Mohamed Khider University, Biskra

Faculty of Sciences and the Natural Sciences and Life

Department of Mathematics

A Thesis Presented for the Degree of

DOCTOR OF SCIENCES

In the Field of Statistics

By

BRAHIM BRAHIMI

Title

Statistics of Bivariate Extreme Values

Examination Committee Members:

Brahim MEZERDI	Professor	U. Mohamed Khider Biskra	Chairman
Abdelhakim NECIR	Professor	U. Mohamed Khider Biskra	Supervisor
Djamel MERAGHNI	M.C.(A)	U. Mohamed Khider Biskra	Examiner
Djamil AÏSSANI	Professor	U. Abderahmane Mira Bejaia	Examiner
Hocine FELLAG	Professor	U. Mouloud Mammeri Tizi-Ouzou	Examiner
Zaher MOHDEB	Professor	U. Mentouri Constantine	Examiner

10/02/2011

A Thesis Presented for the Degree of

DOCTOR OF SCIENCES

In the Field of **Statistics**

By

BRAHIM BRAHIMI

Title

Statistics of Bivariate Extreme Values

10/02/2011

Acknowledgments

In the first place I would like to record my gratitude to Professor Dr. Necir Abdelhakim for his supervision, advice, and guidance from the very early stage of this research as well as giving me extraordinary experiences through out the work. Above all and the most needed, he provided me unflinching encouragement and support in various ways. His truly scientist intuition has made him as a constant oasis of ideas and passions in science, which exceptionally inspire and enrich my growth as a student, a researcher and a scientist want to be. I am indebted to him more than he knows.

This thesis arose in part out of years of research that has been done since I came to INRS's and LSTA's groups. By that time, I have worked with a Professor Taha Ouarda, Professor Fateh chebana, Professor Paul Deheuvels and Professor Daniel Pierre-Loti-Viaud, whose contribution in assorted ways to the research and the making of this thesis deserved special mention. It is a pleasure to convey my gratitude to them all in my humble acknowledgment.

I am very grateful to the members of the supervisory committee, Assoc. Professor Dr. Mezerdi Brahim, Dr. Meraghni Djamel, Professor Dr. Mohdeb Zaher, Professor Dr. Fellag Hocine and Professor. Dr. Aïssani Djamil without whose knowledge and assistance this study would not have been successful.

Special thanks also to all my friends, especially statistical group members; Dr. Meraghni Djamel, Dr. Djabrane Yahia, Mr. Benatia Fateh and Mr. Sayeh Abdellah for sharing the literature and invaluable assistance. Not forgetting to her bestfriends who always been there.

Where would I be without my family? My parents deserve special mention for their inseparable support and prayers. My Father, in the first place is the person who put the fundament my learning character, showing me the joy of intellectual pursuit ever since I was a child. My Mother, is the one who sincerely raised me with her caring and gently love. Words fail me to express my appreciation to my wife Jihane whose dedication, love and persistent confidence in me, has taken the load off my shoulder. I owe her for being unselfishly let her intelligence, passions, and ambitions collide with mine. Sister's and Brather's, thanks for being supportive and caring siblings.

Finally, I would like to thank everybody who was important to the successful realization of this thesis, as well as expressing my apology that I could not mention personally one by one.

To my kid Midou.

Publications Based on This Thesis

1. Brahimi, B., Meraghni, D., Necir, A., 2010. Distortion risk measures for sums of dependent losses. Journal Afrika Statistika, Vol **5**, N° 9. 260-277.

2. Necir, A., Brahimi, B., Meraghni, D., 2010. Erratum to: 'Statistical estimate of the proportional hazard premium of loss'. *Scandinavian Actuarial Journal*, **3**, 246-247.

3. Brahimi, B., Chebana, F., Necir, A., 2010. Copula parameter estimation by bivariate L-moments. submitted.

Contents

List of Figures

List of Tables

Preface

In recent years the field of extreme value theory has been a very active research area. It is of relevance in many practical problems such as the flood frequency analysis, insurance premium, financial area, ... For the literature concerning extreme value theory we refer to Reiss and Thomas (2001, [128]), Coles (2001, [21]) and Beirlant et al. (2005, [8]). The first aim of this thesis to present the investigation in multivariate extreme values distributions (EVDs), and modeling the dependence structure between the extreme data by using copula.

Copula models are becoming increasingly popular for modeling dependencies between random variables (rv). Use of copulas for multidimensional distributions is a powerful method of analyzing this dependence structure. The range of their recent applications include such fields as analysis of extremes in financial assets and returns, failure of paired organs in health science, and human mortality in insurance. The term «copula» was first introduced by Sklar, (1959, [142]) although some of the ideas go back to Hoffding, (1940, [71]). They are useful because they permit us to focus on the dependency structure of the distribution independently of the marginal distributions of the rv's.

The relative mathematical simplicity of copula models and the possibility to construct a diversity of dependence structures based on parametric or non-parametric models of the marginal distributions are one advantages of using this tool in statistical modeling.

The outline of this thesis is as follows: We start as preliminary Chapter 1, and 2, such precise definitions and proprieties of copulas are given, we explores the most important examples of copulas, we describe ranks and dependence measures and we give the most important result about copula parameter estimations.

In the last chapter in preliminary part we discuses Risk measurement wich is a great part of an organization's overall risk management strategy. Risk measurment is a tool to used to assess the probability of a bad event happening. It can be done by businesses as part of disaster recovery planning and as part of the software development lifecycle. The analysis usually involves assessing the expected impact of a bad event such as a hurricane or tornado. Furthermore, risk analysis also involves an assessment of the likelihood of that bad event occurring.

In Chapter 4, we establish a link with the bivariate L-moments (BLM) and the underlying copula functions. This connection provides a new estimate of dependence parameter of bivariate statistical data. We show by simulations that the BLM estimation method is more accurate than the pseudo maximum likelihood for small samples. Consistency and asymptotic normality of this new estimator are also obtained.

We discuss in Chapter 5 two distinct approaches, for distorting risk measures of sums of dependent random variables, that preserve the property of coherence. The first, based on distorted expectations, operates on the survival function of the sum. The second, simultaneously applies the distortion on the survival function of the sum and the dependence structure of risks, represented by copulas. Our goal is to propose an alternative risk measure which takes into account the fluctuations of losses and possible correlations between random variables.

Part I

Preliminary Theory

Chapter 1

Copula Theory

The only real valuable thing is intuition.

Albert Einstein.

The notion of copula was introduced in the seminal paper of Sklar in (1959,[142]), when answering a question proposed by Fréchet about the ties between a multidimensional joint probability function and its margins. At the first, copulas were chiefly used in the development of probabilistic metric spaces theory. Later, it used to define nonparametric measures of dependence between rv's, and since then, they becomes an important tools in probability and mathematical statistics. In this Chapter, a general overview of the theory of copulas will be presented.

1.1 Copulas: Properties and definitions

The standard definition of a copula is a multivariate distribution function defined on the unit cube $[0,1]^d$, with uniformly distributed marginals. This definition is very natural if one considers how a copula is derived from a continuous multivariate distribution function. This definition hides some of the problems of constructing copulas by using other techniques, i.e. it does not say what is meant by a multivariate distribution function. For that reason, we start with a slightly more abstract definition, returning to the standard one later.

First, we focus on general multivariate distributions and then studying the special properties of the copula subset, we refer to Nelsen (2006, [123]) for more details. Throughout this thesis, we denote by $\operatorname{Dom} H$ and $\operatorname{Ran} H$ the domain and range respectively of a function H. Furthermore, a function f will be called increasing[1] whenever $x \leq y$ implies that $f(x) \leq f(y)$.

In this section, we give a general idea about the concept of a multivariate distribution function at first, and then we give an exact definitions and necessary fundamental relationships.

Let $X_1, ..., X_d$ be a random variables, with marginal distribution functions $F_1, ..., F_d$, respectively, and joint distribution function (df.) H. The dependence structure of the variables is represented by the function H as follow

$$H(x_1, ..., x_d) := \mathbb{P}[X_1 \leq x_1, ..., X_d \leq x_d].$$

Any of the marginal distribution functions F_j can be obtained from H, by letting $x_i \to \infty$ for all $i \neq j$. The necessary and sufficient properties of a multivariate distribution function are given in Joe (1997, [85]).

Definition 1.1.1 ***(Multivariate or Joint Distribution Function).*** *A function $H : \mathbb{R}^d \to [0; 1]$ is a multivariate distribution function if the following conditions are satisfied:*

1. *H is right continuous,*
2. *$\lim_{x_i \to -\infty} H(x_1, ..., x_d) = 0$, for $i = 1, ..., d$,*
3. *$\lim_{x_i \to +\infty} H(x_1, ..., x_d) = 1$,*
4. *For all $(a_1, ..., a_d)$ and $(b_1, ..., b_d)$ with $a_i \leq b_i$ for $i = 1, ..., d$ the following inequality holds:*

$$\sum_{i_1=1}^{2} \cdots \sum_{i_d=1}^{2} (-1)^{i_1+\ldots+i_d} H(x_{1_{i_1}}, ..., x_{d_{i_d}}) \geq 0,$$

where $x_{j_1} = a_j$ and $x_{j_2} = b_j$ for all j and $x_{j_k} \in [0, 1]$ for all j and k.

If H has d-th order derivatives, then this condition is equivalent to

$$\frac{\partial^d H}{\partial x_1 \ldots \partial x_d} \geq 0.$$

[1] We can also said that f is nondecreasing.

By the use of a probability integral transformation. Separate the joint distribution function H into its marginal distribution functions, and a part that describes the dependence structure. This can be established by

$$\begin{aligned} H(x_1,...,x_d) &= \mathbb{P}\left[X_1 \leq x_1,...,X_d \leq x_d\right] \\ &= \mathbb{P}\left[F_1(X_1) \leq F_1(x_1),...,F_d(X_d) \leq F_d(x_d)\right] \\ &= C\left(F_1(x_1),...,F_d(x_d)\right), \end{aligned}$$

if the random variable X has a continuous distribution function F, then $F(X)$ has a Uniform $(0,1)$ distribution. Therefore, we define the properties of the function C, called the copula function.

Definition 1.1.2 ***(Copula Function).*** *An d-dimensional copula C is a function from $[0,1]^d$ to $[0,1]$ having the following properties:*

1. *$C(u_1,...,u_d) = 0$ if $u_i = 0$ for some $i = 1,...,d$.*
2. *$C(1,...,1,u_i,1,...,1) = u_i$ for all $u_i \in [0,1]$:*
3. *C is d-increasing, for all $(a_1,...,a_d)$ and $(b_1,...,b_d)$ with $a_i \leq b_i$ for $i = 1,...,d$, the following inequality holds:*

$$\sum_{i_1=1}^{2} \cdots \sum_{i_d=1}^{2} (-1)^{i_1+...+i_d} C(u_{1_{i_1}},...,u_{d_{i_d}}) \geq 0,$$

where $u_{j_1} = a_j$ and $u_{j_2} = b_j$ for all j and $u_{j_k} \in [0,1]$ for all j and k.

Let denote $\mathbf{x} := (x_1,...,x_d)$.

Definition 1.1.3 *Let $S_1,...,S_d$ be nonempty subsets of $\overline{\mathbb{R}}$, where $\overline{\mathbb{R}}$ denotes the extended real line $[-\infty,\infty]$. Let H be a real function of d variables such that* Dom $H = S_1 \times \cdots \times S_d$ *and for $\mathbf{a} \leq \mathbf{b}$ ($a_k \leq b_k$ for all k) let $B = [\mathbf{a},\mathbf{b}]$ $(= [a_1,b_1] \times \cdots \times [a_d,b_d])$ be an d-box whose vertices are in* Dom H. *Then the H-volume of B is given by*

$$V_H(B) = \sum sgn(\mathbf{c})H(\mathbf{c}),$$

where the sum is taken over all vertices $\mathbf{c}$ of B, and $sgn(\mathbf{c})$ is given by

$$sgn(\mathbf{c}) = \begin{cases} 1, & \text{if} \quad c_k = a_k \quad \text{for an even number of } k\text{'s}, \\ -1, & \text{if} \quad c_k = a_k \quad \text{for an odd number of } k\text{'s}. \end{cases}$$

Equivalently, the H-volume of an d-box $B = [\mathbf{a}, \mathbf{b}]$ is the d-th order difference of H on B

$$V_H(B) = \Delta_{\mathbf{a}}^{\mathbf{b}} H(\mathbf{t}) = \Delta_{a_d}^{b_d} ... \Delta_{a_1}^{b_1} H(\mathbf{t}),$$

where the k first order differences are given by

$$\Delta_{a_k}^{b_k} H(\mathbf{t}) = H(t_1, ..., t_{k-1}, b_k, t_{k+1}, ..., t_d) - H(t_1, ..., t_{k-1}, a_k, t_{k+1}, ..., t_d).$$

Definition 1.1.4 *A real function H of d variables is d-increasing if $V_H(B) \geq 0$ for all d-boxes B whose vertices lie in* Dom H.

Let S_k has a smallest element a_k, and suppose that the domain of a real function H of d variables is given by Dom $H = S_1 \times ... \times S_d$. We say that H is grounded[2] if $H(\mathbf{t}) = 0$ for all $\mathbf{t}$ in Dom H such that $t_k = a_k$. If S_k is non empty for each k and has a greatest element b_k, then H has margins H_k, with Dom $H_k = S_k$ and we have

$$H_k(x) = H(b_1, ..., b_{k-1}, x, b_{k+1}, ..., b_d)$$

for all x in S_k. H_k is called one-dimensional margins.

Lemma 1.1.1 *Let $S_1, ..., S_d$ be non empty subsets of $\overline{\mathbb{R}}$, and let H be a grounded d-increasing function with domain $S_1 \times ... \times S_d$. Then H is increasing in each argument, i.e., if $(t_1, ..., t_{k-1}, x, t_{k+1}, ..., t_d)$ and $(t_1, ..., t_{k-1}, y, t_{k+1}, ..., t_d)$ are in* Dom H *and $x \leq y$, then*

$$H(t_1, ..., t_{k-1}, x, t_{k+1}, ..., t_d) \leq H(t_1, ..., t_{k-1}, y, t_{k+1}, ..., t_d).$$

Lemma 1.1.2 *Let $S_1, ..., S_d$ be non empty subsets of $\overline{\mathbb{R}}$, and let H be a grounded d-increasing function with margins and domain $S_1 \times ... \times S_d$. Then, if $\mathbf{x}$ and $\mathbf{y}$ are any points in $S_1 \times ... \times S_d$,*

$$H(\mathbf{x}) - H(\mathbf{y}) \leq \sum_{k=1}^{d} |H_k(x_k) - H_k(y_k)| .$$

For the proof, see Schweizer and Sklar (1983, [136]).

Definition 1.1.5 *An d-dimensional distribution function is a function H with domain $\overline{\mathbb{R}}^d$ such that H is grounded, d-increasing and $H(\infty, ..., \infty) = 1$.*

We say that a function H from $S_1 \times S_2$ into $\mathbb{R}$ is grounded if $H(x, a_2) = 0 = H(a_1, y)$ for all (x, y) in $S_1 \times S_2$.

From Lemma 1.1.1 it follows that the margins of an d-dimensional distribution function are distribution functions, which we denote $F_1, ..., F_d$.

Definition 1.1.6 *An d-dimensional copula is a function C with domain $[0,1]^d$ such that*

1. *C is grounded and d-increasing.*

2. *C has margins $C_k, k = 1, 2, ..., d$, which satisfy $C_k(u) = u$ for all u in $[0,1]$.*

Note that for any d-copula C, $d \geq 3$, each k-dimensional margin of C is a k-copula. Equivalently, an d-copula is a function C from $[0,1]^d$ to $[0,1]$ with the following properties:

1. For every $\mathbf{u}$ in $[0,1]^d$, $C(\mathbf{u}) = 0$ if at least one coordinate of $\mathbf{u}$ is $\mathbf{0}$, and $C(\mathbf{u}) = u_k$ if all coordinates of $\mathbf{u}$ equal 1 except u_k.

2. For every $\mathbf{a}$ and $\mathbf{b}$ in $[0,1]^d$ such that $a_i \leq b_i$ for all i, $V_C([\mathbf{a}, \mathbf{b}]) \geq 0$.

Copulas are joint distribution functions (on $[0,1]^d$), induces that copulas are a probability measures on $[0,1]^d$ via a standard extension to arbitrary Borel subsets of $[0,1]^d$.

$$V_C([0,u_1] \times ... \times [0,u_d]) = C\left(u_1, ..., u_d\right).$$

A standard result from measure theory says that there is a unique probability measure on the Borel subsets of $[0,1]^d$ which coincides with V_C on the set of d-boxes of $[0,1]^d$. This probability measure will also be denoted V_C.

It follows from Definition 1.1.6 that copula C is a distribution function on $[0,1]^d$ with $[0,1]$-uniformly distributed margins. The following theorem follows directly from Lemma 1.1.2.

Theorem 1.1.1 *Let C be an d-copula. Then for every $\mathbf{u}$ and $\mathbf{v}$ in $[0,1]^d$,*

$$C(\mathbf{v}) - C(\mathbf{u}) \leq \sum_{k=1}^{d} v_k - u_k.$$

Hence C is uniformly continuous on $[0,1]^d$.

1.1.1 Sklar's Theorem

The Sklar's theorem is the basic tool of copulas theory and is the foundation of most, of the applications of that theory to statistics. Is show the role that copulas play in the relationship between multivariate distribution functions and their univariate margins.

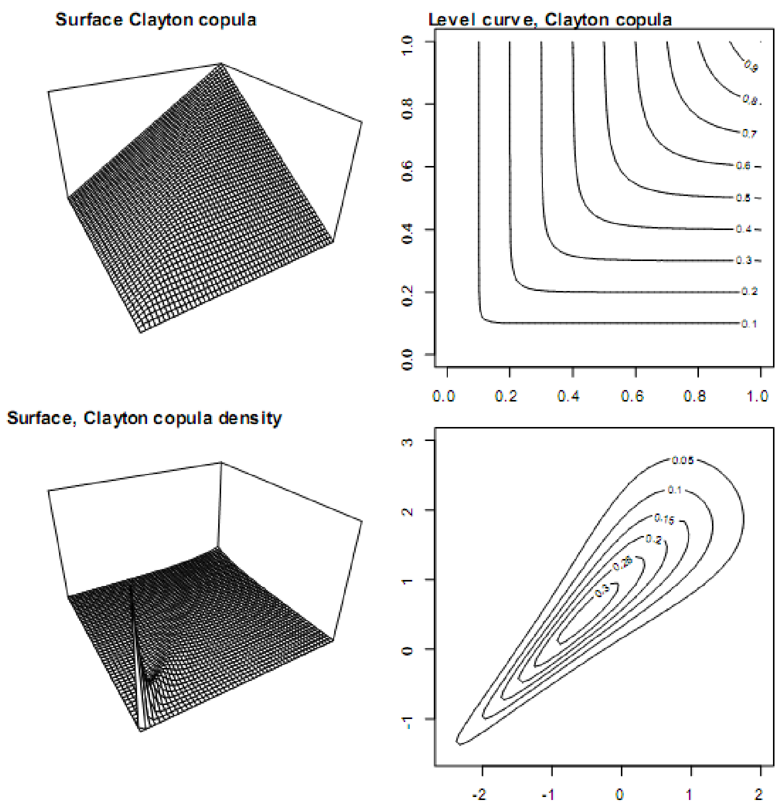

Figure 1.1: Clayton Copula on top (distribution function), and the associated density below.

Let X be a real random variable on the probability space $(\Omega, \mathcal{U}, \mathbb{P})$ with distribution function F and let $V \sim U(0,1)$ be uniformly distributed on $(0,1)$ such that V independent of X. We define the modified distribution function $F(x,\lambda)$ by

$$F(x,\lambda) := \mathbb{P}(X < x) + \lambda \mathbb{P}(X = x),$$

and the (generalized) distributional transform of X by

$$U := F(X, V). \tag{1.1}$$

The representation of the distributional transform is

$$U = F(X^-) + V(F(X) - F(X^-)).$$

A good reference for this transform is the statistics book of Ferguson (1967, [44]). $U = F(X) \stackrel{d}{=} U(0,1)$ is holds true for continuous df's F, $F(x, \lambda)$, in general for the distributional transform and the quantile transform. Here the inverse or the quantile of a df F is defined as usual by

$$F^{-1}(u) = \inf\{x \in \mathbb{R} : F(x) \geq u\}, u \in (0,1).$$

We give a proof of this interesting results.

Proposition 1.1.1 ***(Distributional transform)*** *Let U be the distributional transform of X as defined in (1.1). Then*

$$U \stackrel{d}{=} U(0,1) \quad \text{and} \quad X = F^{-1}(U) \text{ almost surely (a.s.).}$$

Proof. For $0 < \alpha < 1$ let $q_\alpha^-(X)$ denote the lower α-quantile, that is $q_\alpha^-(X) = \sup\{x : \mathbb{P}(X \leq x) < \alpha\}$. Then $F(X, V) \leq \alpha$ if and only if

$$(X, V) \in \{(x, \lambda) : \mathbb{P}(X < x) + \lambda \mathbb{P}(X = x) \leq \alpha\}.$$

If $\beta := \mathbb{P}(X = q_\alpha^-(X)) > 0$ and with $q := \mathbb{P}(X < q_\alpha^-(X))$ this is equivalent to $\{X < q_\alpha^-(X)\} \cup \{X = q_\alpha^-(X), q + V\beta \leq \alpha\}$. Thus we obtain

$$\mathbb{P}(U \leq \alpha) = \mathbb{P}(F(X, V) \leq \alpha) = q + \beta \mathbb{P}(V \leq \frac{\alpha - q}{\beta}) = q + \beta \frac{\alpha - q}{\beta} = \alpha.$$

If $\beta = 0$, then

$$\mathbb{P}(F(X, V) \leq \alpha) = \mathbb{P}(X < q_\alpha^-(X)) = \mathbb{P}(X \leq q_\alpha^-(X)) = \alpha.$$

By definition of U,

$$F(X^-) \leq U \leq F(X). \tag{1.2}$$

For any $u \in (F(x^-), F(x)]$ it holds that $F^{-1}(u) = x$. Thus by (1.2) we obtain that $F^{-1}(U) = X$ a.s. ∎

The distributional quantile transform is a useful tool, in general case it allows to give a simple proof of Sklar's Theorem. The idea of Sklar's Theorem is to represent an n-

dimensional distribution function H in two parts, the marginal distribution functions F_i and the copula C, witch describe the dependence concepts of the distribution.

Theorem 1.1.2 *Let H be an d-dimensional df with margins $F_1, ..., F_d$. Then there exists an d-copula C such that for all x in $\overline{\mathbb{R}}^d$,*

$$H(x_1, ..., x_d) = C\left(F_1\left(x_1\right), ..., F_d\left(x_d\right)\right). \tag{1.3}$$

If $F_1, ..., F_d$ are all continuous, then C is unique; otherwise C is uniquely determined on $\operatorname{Ran} F_1 \times ... \times \operatorname{Ran} F_d$. Conversely, if C is an d-copula and $F_1, ..., F_d$ are df's, then the function H defined above is an d-dimensional df with margins $F_1, ..., F_d$.

Proof. On a probability space $(\Omega, \mathcal{U}, \mathbb{P})$, let $\mathbf{X} = (X_1, ..., X_d)$ be a random vector with df F and let V be independent of X and uniformly distributed on $(0,1), V \sim U(0,1)$. Suppose that distributional transforms $U_i := F_i(X_i, V), 1 \leq i \leq d$, by Proposition 1.1.1 we have $U_i \sim U(0,1)$, and $X_i = F_i^{-1}(U_i)$ a.s., $1 \leq i \leq d$. Thus we define C to be the distribution function of $U = (U_1, ..., U_d)$ as follow

$$\begin{aligned} F(\mathbf{x}) &= \mathbb{P}(\mathbf{X} \leq \mathbf{x}) = \mathbb{P}(F_1^{-1}(U_1) \leq x_1, ..., F_d^{-1}(U_d) \leq x_d) \\ &= \mathbb{P}(U_1 \leq F_1(x_1), ..., U_d \leq F_d(x_d)) \\ &= C(F_1(x_1), ..., F_d(x_d)), \end{aligned}$$

i.e. C is a copula of F. ■

We show from Sklar's Theorem that we can separete univariate margins and the multivariate dependence structure of continuous multivariate distribution functions, and the dependence structure can be represented by a copula function.

Let F be a univariate distribution function and F^{-1} the quantile function of F, using the convention $\inf \varnothing = -\infty$.

Corollary 1.1.1 *Let H be an d-dimensional df with continuous margins $F_1, ..., F_d$ and copula C (where C satisfies (1.3)). Then for any $\mathbf{u}$ in $[0,1]^d$,*

$$C(\mathbf{u}) = H(F_1^{-1}(u_1), ..., F_d^{-1}(u_d)).$$

For the Proof we can see Nelsen (2006, [123]) or Marshall (1996, [113]).

Example 1.1.1 *Let Φ denote the standard univariate normal df and let $\Phi_{\mathcal{R}}^d$ denote*

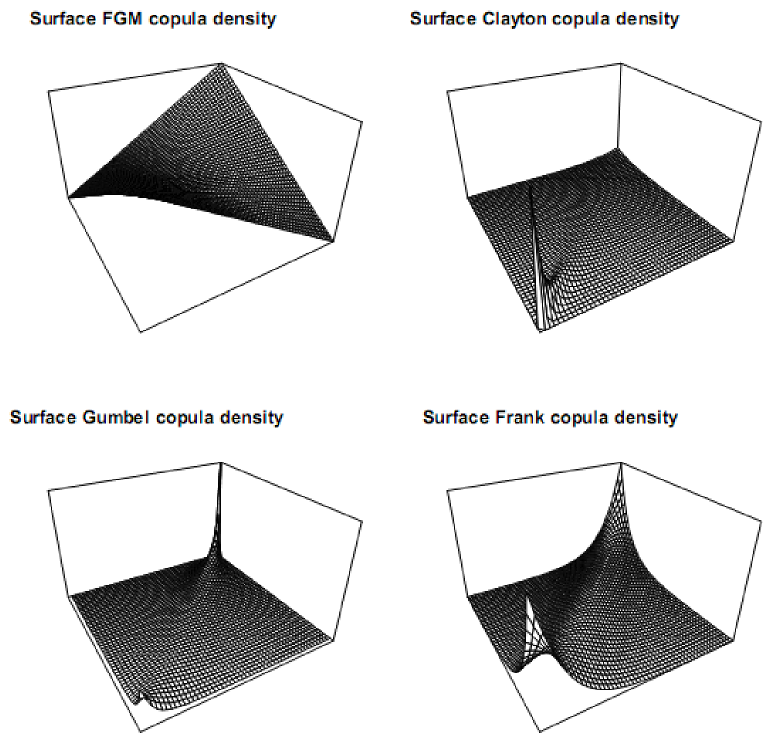

Figure 1.2: Surface FGM (0.3), Clayton (5.2), Gumbel (4) and Frank (8) copulas density.

the standard multivariate normal df with linear correlation matrix $\mathcal{R}$. *Then*

$$C(u_1, ..., u_d) = \Phi^n_{\mathcal{R}}(\Phi^{-1}(u_1), ..., \Phi^{-1}(u_d))$$

is the Gaussian or normal d*-copula.*

The concept of copula density functions play an important role in the parameter estimation methods for copulas that are based on the Maximum Likelihood Principle, and later we give a new principle of L-comoment based estimation.

Corollary 1.1.2 ***(Copula Density).*** *The relationship between the multivariate density function* $h(x_1, ..., x_d)$ *and the copula density* c*, is given by:*

$$h(x_1, ..., x_d) = c(F_1(x_1), ..., F_d(x_d)) \prod_{i=1}^{d} f_i(x_i). \tag{1.4}$$

Proof. Let h be the d-dimensional density function that belongs to the distribution function H. Define h by:

$$h(x_1, ..., x_d) = \frac{\partial^d}{\partial x_1, ..., \partial x_d} H(x_1, ..., x_d).$$

Now substitute (1.3), to obtain:

$$h(x_1, ..., x_d) = \frac{\partial^d}{\partial x_1, ..., \partial x_d} C(F_1(x_1), ..., F_d(x_d)).$$

Using the substitution $u_i = F_i(x_i)$ for $i = 1, ..., d$ we obtain:

$$\begin{aligned} h(x_1, ..., x_d) &= \frac{\partial^d}{\partial u_1, ..., \partial u_d} C(u_1, ..., u_d) \prod_{i=1}^{d} f_i(x_i) \\ &= c(u_1, ..., u_d) \prod_{i=1}^{d} f_i(x_i). \end{aligned}$$

Inserting again $u_i = F_i(x_i)$ yields the relationship given by (1.4). ■

1.1.2 The Fréchet-Hoeffding bounds for joint df's

Hoeffding (1940, [71], 1942, [73]) give an explicit formulation of the statement «there is a functional dependence between random variables X and Y».

Let's the functions M^d, Π^d and W^d defined on $[0,1]^d$ as follows:

$$\begin{aligned} M^d(\mathbf{u}) &= \min(u_1, ..., u_d), \\ \Pi^d(\mathbf{u}) &= u_1 ... u_d, \\ W^d(\mathbf{u}) &= \max(u_1 + ... + u_d - d + 1, 0). \end{aligned}$$

It easy to check that M^d and Π^d are d-copulas for all $d \geq 2$ but the function W^d is not a copula for any $d \geq 3$ as shown in the following example.

Example 1.1.2 *Consider the d-cube* $[1/2,1]^d \subset [0,1]^d$.

$$\begin{aligned} V_{W^d}([1/2,1]^d) &= \max(1+...+1-d+1,0) \\ &\quad -d\max(1/2+1+...+1-d+1,0) \\ &\quad +\binom{d}{2}\max(1/2+1/2+1+...+1-d+1,0) \\ &\quad ... \\ &\quad +\max(1/2+...+1/2-d+1,0) \\ &= 1-d/2+0+...+0. \end{aligned}$$

Hence W^d *is not a copula for* $d \geq 3$.

The following theorem given in Fréchet (1957, [50]) is called the Fréchet–Hoeffding bounds inequality.

Theorem 1.1.3 *If* C *is any d-copula, then for every* $\mathbf{u}$ *in* $[0,1]^d$,

$$W^d(\mathbf{u}) \leq C(\mathbf{u}) \leq M^d(\mathbf{u}).$$

Although the Fréchet–Hoeffding lower bound W^d is never a copula for $d \geq 3$, it is the best possible lower bound in the following sense.

Theorem 1.1.4 *For any* $d \geq 3$ *and any* $\mathbf{u}$ *in* $[0,1]^d$, *there is an d-copula* C *(which depends on* $\mathbf{u}$*) such that*

$$C(\mathbf{u}) = W^d(\mathbf{u}).$$

For the proof, see Nelsen (2006, [123, p. 48]).
Let $(U_1,...,U_d)$ d-random variables with joint distribution function C. We denote by $\overline{C}$ the joint survival function of C, then

$$\overline{C}\,(u_1,...,u_d) = \mathbb{P}\{U_1 > u_1,...,U_d > u_d\}.$$

The following definition give the notion of partial ordering, since not every pair of copulas is comparable in this order. However many important parametric families of copulas are totally ordered.

Definition 1.1.7 *If* C_1 *and* C_2 *are copulas,* C_1 *is smaller than* C_2 *(written* $C_1 \prec C_2$*) if*

$$C_1(\mathbf{u}) \leq C_2(\mathbf{u}) \quad \text{and} \quad \overline{C}_1(\mathbf{u}) \leq \overline{C}_2(\mathbf{u}),$$

for all $\mathbf{u}$ *in* $[0,1]^d$.

In bivariate case,

$$\begin{aligned}\overline{C}_1(u_1,u_2) \leq \overline{C}_2(u_1,u_2) &\iff 1-u_1-u_2+C_1(u_1,u_2)\\ &\leq 1-u_1-u_2+C_2(u_1,u_2)\\ &\iff C_1(u_1,u_2)\leq C_2(u_1,u_2).\end{aligned}$$

The Fréchet–Hoeffding lower bound W^2 is smaller than every 2-copula, and every d-copula is smaller than the Fréchet–Hoeffding upper bound M^d. This partial ordering of the set of copulas is called a concordance ordering.

We denote by C_θ for one-parameter family of copula, we called positively ordered if $C_{\theta_1} \prec C_{\theta_2}$ whenever $\theta_1 \leq \theta_2$. Examples of such one-parameter families will be given later.

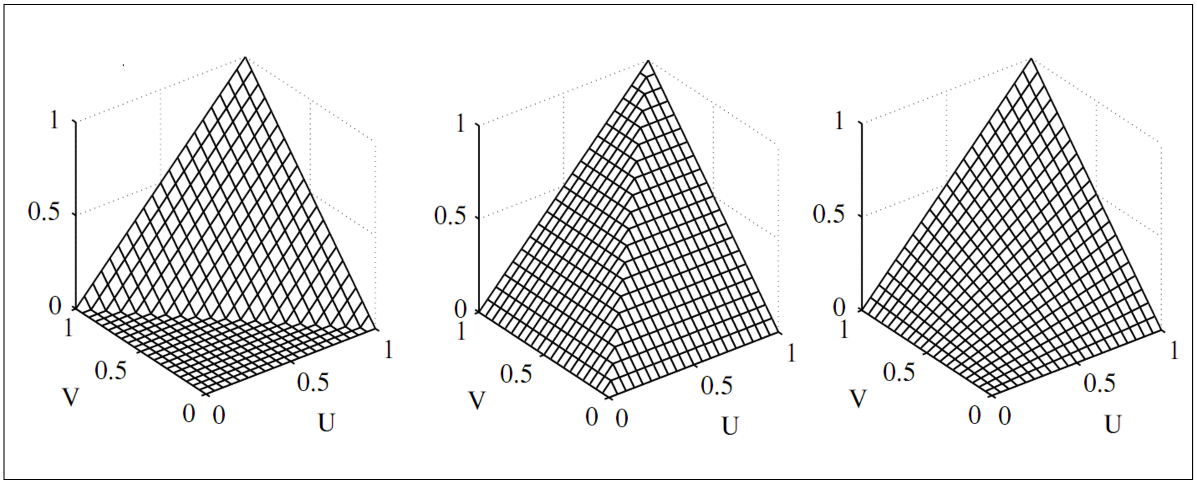

Figure 1.3: The 2-copulas $\mathbf{W}^2$, $\mathbf{M}^2$ and $\mathbf{\Pi}^2$, respectively from the left to the right.

1.1.3 Copulas and random variables

Let $X_1, ..., X_d$ be random variables with continuous distribution functions $F_1, ..., F_d$, respectively, and joint distribution function H. Then $(X_1, ..., X_d)$ has a unique copula C, where C is given by (1.3). The standard copula representation of the distribution of the random vector $(X_1, ..., X_d)$ then becomes:

$$H(x_1, ..., x_d) = \mathbb{P}\{X_1 \leq x_1, ..., X_d \leq x_d\} = C\left(F_1\left(x_1\right), ..., F_d\left(x_d\right)\right).$$

The transformations $X_i \to F_i(X_i)$ used in the above representation are usually referred to as the probability-integral transformations (to uniformity) and form a standard tool in simulation methodology.

Since $X_1, ..., X_d$ are independent if and only if

$$H(x_1, ..., x_d) = \prod_{i=1}^{d} F_i(x_i),$$

for all $x_1, ..., x_d$ in $\mathbb{R}$, the following result follows from Theorem 1.1.2.

Theorem 1.1.5 *Let $(X_1, ..., X_d)$ be a vector of continuous random variables with copula C, then $X_1, ..., X_d$ are independent if and only if $C = \Pi^d$.*

One of the nice property of copulas is that for strictly monotone transformations of the random variables, copulas are either invariant, or change in certain simple ways. Note that if the distribution function of a random variable X is continuous, and if α is a strictly monotone function whose domain contains $\operatorname{Ran} X$, then the distribution function of the random variable $\alpha(X)$ is also continuous.

Theorem 1.1.6 *Let $(X_1, ..., X_d)$ be a vector of continuous random variables with copula C. If $\alpha_1, ..., \alpha_d$ are strictly increasing on* $\operatorname{Ran} X_1, ..., \operatorname{Ran} X_d$, *respectively, then also $(\alpha_1(X_1), ..., \alpha_d(X_d))$ has copula C.*

Proof. Let $F_1, ..., F_d$ denote the distribution functions of $X_1, ..., X_d$ and let $G_1, ..., G_d$ denote the distribution functions of $\alpha_1(X_1), ..., \alpha_d(X_d)$, respectively. Let $(X_1, ..., X_d)$ have copula C, and let $(\alpha_1(X_1), ..., \alpha_d(X_d))$ have copula C_α. Since α_k is strictly increasing for each k,

$$G_k(x) = \mathbb{P}\{\alpha_k(X_k) \leq x\} = \mathbb{P}\{X_k \leq \alpha_k^{-1}(x)\} = F_k(\alpha_k^{-1}(x))$$

for any x in $\mathbb{R}$, hence

$$\begin{aligned} C_\alpha(G_1(x_1), ..., G_d(x_d)) &= \mathbb{P}\{\alpha_1(X_1) \leq x_1, ..., \alpha_d(X_d) \leq x_d\} \\ &= \mathbb{P}\{X_1 \leq \alpha_1^{-1}(x_1), ..., X_d \leq \alpha_d^{-1}(x_d)\} \\ &= C(F_1(\alpha_1^{-1}(x_1)), ..., F_d(\alpha_d^{-1}(x_d))) \\ &= C(G_1(x_1), ..., G_d(x_d)). \end{aligned}$$

Since $X_1, ..., X_d$ are continuous, $\operatorname{Ran} G_1 = ... = \operatorname{Ran} G_d = [0, 1]$. Hence it follows that $C_\alpha = C$ on $[0, 1]^d$. ■

From Theorem 1.1.2 we know that the copula function C «separates» an d-dimensional distribution function from its univariate margins. The next theorem will show that there is also a function, C, that separates an d-dimensional survival function from its univariate survival margins. Furthermore this function can be shown to be a copula, and this survival copula can rather easily be expressed in terms of C and its k-dimensional margins.

Theorem 1.1.7 *Let $(X_1, ..., X_d)$ be a vector of continuous random variables with copula $C_{X_1,...,X_d}$. Let $\alpha_1, ..., \alpha_d$ be strictly monotone on* $\operatorname{Ran} X_1, ...,$ $\operatorname{Ran} X_d$*, respectively, and let $(\alpha_1(X_1), ..., \alpha_d(X_d))$ have copula $C_{\alpha_1(X_1),...,\alpha_d(X_d)}$. Furthermore let α_k be strictly decreasing for some k. Without loss of generality let $k = 1$. Then*

$$\begin{aligned} C_{\alpha_1(X_1),...,\alpha_d(X_d)}(u_1, u_2, ..., u_d) &= C_{\alpha_2(X_2),...,\alpha_d(X_d)}(u_2, ..., u_d) \\ &\quad -C_{X_1,\alpha_2(X_2),...,\alpha_d(X_d)}(1 - u_1, u_2, ..., u_d). \end{aligned}$$

Proof. Let $X_1, ..., X_d$ have distribution functions $F_1, ..., F_d$ and let $\alpha_1(X_1), ...,$ $\alpha_d(X_d)$ have distribution functions $G_1, ..., G_d$. Then

$$\begin{aligned} C_{\alpha_1(X_1),\alpha_2(X_2),...,\alpha_d(X_d)}(G_1(x_1), ..., G_d(x_d)) &= \mathbb{P}\{\alpha_1(X_1) \leq x_1, ..., \alpha_d(X_d) \leq x_d\} \\ &= \mathbb{P}\{X_1 > \alpha_1^{-1}(x_1), \alpha_2(X_2) \leq x_2, ..., \alpha_d(X_d) \leq x_d\} \\ &= \mathbb{P}\{\alpha_2(X_2) \leq x_2, ..., \alpha_d(X_d) \leq x_d\} \\ &\quad -\mathbb{P}\{X_1 \leq \alpha_1^{-1}(x_1), \alpha_2(X_2) \leq x_2, ..., \alpha_d(X_d) \leq x_d\} \\ &= C_{\alpha_2(X_2),...,\alpha_d(X_d)}(G_2(x_2), ..., G_d(x_d)) \\ &\quad -C_{X_1,\alpha_2(X_2),...,\alpha_d(X_d)}(F_1(\alpha_1^{-1}(x_1)), G_2(x_2), ..., G_d(x_d)) \\ &= C_{\alpha_2(X_2),...,\alpha_d(X_d)}(G_2(x_2), ..., G_d(x_d)) \\ &\quad -C_{X_1,\alpha_2(X_2),...,\alpha_d(X_d)}(1 - G_1(x_1), G_2(x_2), ..., G_d(x_d)), \end{aligned}$$

from which the conclusion follows directly. ■

By using the two theorems above recursively it is clear that the copula $C_{\alpha_1(X_1),...,\alpha_d(X_d)}$ can be expressed in terms of the copula $C_{X_1,...,X_d}$ and its lower-dimensional margins. This is exemplified below.

Example 1.1.3 *Consider the bivariate case. Let α_1 be strictly decreasing and let α_2 be strictly increasing. Then*

$$\begin{aligned} C_{\alpha_1(X_1),\alpha_2(X_2)}(u_1, u_2) &= u_2 - C_{X_1,\alpha_2(X_2)}(1 - u_1, u_2) \\ &= u_2 - C_{X_1,X_2}(1 - u_1, u_2). \end{aligned}$$

Let α_1 and α_2 be strictly decreasing. Then

$$\begin{aligned} C_{\alpha_1(X_1),\alpha_2(X_2)}(u_1,u_2) &= u_2 - C_{X_1,\alpha_2(X_2)}(1-u_1,u_2) \\ &= u_2 - (1-u_1-C_{X_1,X_2}(1-u_1,1-u_2)) \\ &= u_1+u_2-1+C_{X_1,X_2}(1-u_1,1-u_2). \end{aligned}$$

Here $C_{\alpha_1(X_1),\alpha_2(X_2)}$ is the survival copula C, of (X_1,X_2), i.e.,

$$H(x_1,x_2) = \mathbb{P}\{X_1 > x_1, X_2 > x_2\} = C(F_1(x_1),F_2(x_2)).$$

Note also that the joint survival function of d $U(0,1)$ random variables whose joint distribution function is the copula C is $C(u_1,...,u_d) = C(1-u_1,...,1-u_d)$. The mixed kth order partial derivatives of a copula C, $\partial^k C(u)/\partial u_1...\partial u_k$, exist for almost all $\mathbf{u}$ in $[0,1]^d$. For such $\mathbf{u}, 0 \leq \partial^k C(u)/\partial u_1...\partial u_k \leq 1$. For details, see Nelsen (2006, [123, p. 26]).

With this in mind, let

$$C(u_1,...,u_d) = A_C(u_1,...,u_d) + S_C(u_1,...,u_d),$$

where

$$\begin{aligned} A_C(u_1,...,u_d) &= \int_0^{u_1} ... \int_0^{u_d} \frac{\partial^d}{\partial s_1...\partial s_d} C(s_1,...,s_d) ds_1...ds_d, \\ S_C(u_1,...,u_d) &= C(u_1,...,u_d) - A_C(u_1,...,u_d). \end{aligned}$$

Unlike multivariate distributions in general, the margins of a copula are continuous, hence a copula has no individual points $\mathbf{u}$ in $[0,1]^d$ for which $V_C(\mathbf{u}) > 0$. If $C = A_C$ on $[0,1]^d$, then C is said to be absolutely continuous. In this case C has density

$$\frac{\partial^d}{\partial u_1...\partial u_d} C(u_1,...,u_d).$$

If $C = S_C$ on $[0,1]^d$, then C is said to be singular, and

$$\frac{\partial^d}{\partial u_1...\partial u_d} C(u_1,...,u_d) = 0$$

almost everywhere in $[0,1]^d$. The support of a copula is the complement of the union of all open subsets A of $[0,1]^d$ with $V_C(A) = 0$. When C is singular its support has Lebesgue measure zero and conversely. However a copula can have full support without being absolutely continuous. Examples of such copulas are so-called Marshall-Olkin copulas which are presented later.

Example 1.1.4 *Consider the bivariate Fréchet-Hoeffding upper bound M given by*

$$M(u,v) = \min(u,v)$$

on $[0,1]^2$. It follows that

$$\frac{\partial^2}{\partial u \partial v} M(u,v) = 0,$$

everywhere on $[0,1]^2$ except on the main diagonal (which has Lebesgue measure zero), and $V_M(B) = 0$ for every rectangle B in $[0,1]^2$ entirely above or below the main diagonal. Hence M is singular.

Now we present a general algorithm for random variate generation from copulas. Note however that in most cases it is not an efficient one to use.

Consider the general situation of random variate generation from the d-copula C. Let

$$C_k(u_1, ..., u_k) = C(u_1, ..., u_k, 1, ..., 1), k = 2, ..., d-1,$$

denote k-dimensional margins of C, with $C_1(u_1) = u_1$ and $C_d(u_1, ..., u_d) = C(u_1, ..., u_d)$. Let $U_1, ..., U_d$ have joint distribution function C. Then the conditional distribution of U_k given the values of $U_1, ..., U_{k-1}$, is given by

$$\begin{aligned} C_k(u_k|u_1, ..., u_{k-1}) &= \mathbb{P}\{U_k \leq u_k | U_1 = u_1, ..., U_{k-1} = u_{k-1}\} \\ &= \frac{\partial^{k-1}}{\partial u_1 ... \partial u_{k-1}} C_k(u_1, ..., u_k) / \frac{\partial^{k-1}}{\partial u_1 ... \partial u_{k-1}} C_{k-1}(u_1, ..., u_{k-1}), \end{aligned}$$

given that the numerator and denominator exist and that the denominator is not zero. The following algorithm generates a random variate $(u_1, ..., u_d)^T$ from C. As usual, let $U(0,1)$ denote the uniform distribution on $[0,1]$.

Algorithm 1.1.1

1. Simulate a random variate u_1 from $U(0,1)$.
2. Simulate a random variate u$_2$ from $C_2(\cdot|u_1)$.

   ⋮

3. Simulate a random variate un from $C_d(\cdot|u_1, ..., u_{d-1})$.

This algorithm is in fact a particular case of what is called «the standard construction». The correctness of the algorithm can be seen from the fact that for independent $U(0,1)$ random variables $Q_1, ..., Q_d$,

$$(Q_1, C_2^{-1}(Q_2|Q_1), ..., C_d^{-1}(Q_d|Q_1, C_2^{-1}(Q_2|Q_1), ...))$$

has distribution function C. To simulate a value u_k from $C_k(\cdot\ |u_1, ..., u_{k-1})$ in general means simulating q from $U(0,1)$ from which $u_k = C_k^{-1}(q|u_1, ..., u_{k-1})$ can be obtained through the equation $q = C_k(u_k|u_1, ..., u_{k-1})$ by numerical rootfinding. When $C_k^{-1}(q|u_1, ..., u_{k-1})$ has a closed form (and hence there is no need for numerical rootfinding) this algorithm can be recommended.

Example 1.1.5 *Let the copula C be given by*

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

for $\theta > 0$. Then

$$\begin{aligned} C_{2|1}(v|u) &= \frac{\partial C}{\partial u}(u, v) = -\frac{1}{\theta}(u^{-\theta} + v^{-\theta} - 1)^{-1/\theta-1}(-\theta u^{-\theta-1}) \\ &= (u^{\theta})^{\frac{-1-\theta}{\theta}}(u^{-\theta} + v^{-\theta} - 1)^{-1/\theta-1} \\ &= (1 + u^{\theta}(v^{-\theta} - 1))^{\frac{-1-\theta}{\theta}}. \end{aligned}$$

Solving the equation $q = C_{2|1}(v|u)$ for v yields

$$C_{2|1}^{-1}(q|u) = v = \left(\left(q^{\frac{-\theta}{1+\theta}} - 1\right)u^{-\theta} + 1\right)^{-1/\theta}.$$

The following algorithm generates a random variate $(u, v)^T$ from the above copula C.

1. *Simulate two independent random variates u and q from $U(0,1)$.*
2. *Set $v = \left(\left(q^{\frac{-\theta}{1+\theta}} - 1\right)u^{-\theta} + 1\right)^{-1/\theta}$.*

1.2 Measures of association

Generally speaking, random variables X and Y are said to be associated when they are not independent, i.e. when $F(x, y) = F(x)F(y)$. With regards to measuring association, the terms used in the literature are concordance and dependence. Formally, there is a clear distinction between the two, as dependence measure obtains its minimum value when the measured random variables are independent, while the minimum value of concordance indicates that the random variables in question are countermonotonic.

This terminology is not standard however, we tend to use the words interchangeably, meaning association along the definition of concordance. Furthermore, while we examine the two dimensional cases in this section for simplicity, the extensions to higher dimensions are straightforward.

Copulas provide a helpful way of studying and measuring the association between rv's. In the following, we give a formal definition of concordance and some of its properties and present the most crucial measures of concordance that are relevant to this thesis. For completeness, we also shortly reiterate the definition of linear correlation.

1.2.1 Concordance

Loosely speaking, concordance seeks to capture the probability of rv's X and Y both having a «large» or a «small» values versus the probability of one them having a «large» and the other one a «small» value. Formally speaking, a measure $\delta_{X,Y}$ with copula C is a measure of concordance if it satisfies the following properties (Scarsini 1984, [134]).

Definition 1.2.1 ***(Measure of Dependence).*** *A numeric measure δ of association between two continuous rv's X and Y whose copula is C is a measure of dependence if it satisfies the following properties:*

1. *δ is defined for every pair X, Y of continuous rv's.*
2. $\delta_{X,Y} = \delta_{Y,X}$.
3. $0 \leq \delta_{X,Y} \leq 1$.
4. *$\delta_{X,Y} = 0$ if and only if X and Y are independent.*
5. *$\delta_{X,Y} = 1$ if and only if each of X and Y is almost surely a strictly monotone function of the other.*
6. *If α and β are almost surely strictly monotone functions on $\operatorname{Rang} X$ and $\operatorname{Rang} Y$, respectively, then $\delta_{\alpha(X),\beta(Y)} = \delta_{X,Y}$.*

7. *If $\{(X_n, Y_n)\}$ is a sequence of continuous random variables with copulas C_n, and if $\{C_n\}$ converges pointwise to C, then $\lim_{n\to\infty} \delta_{C_n} = \delta_C$.*

The above definition implies invariance with respect to increasing transformations and the existence of bounds for δ. These bounds are referred to as comonotonicity and countermonotonicity and imply $\delta_{X,Y} = 1$ and $\delta_{X,Y} = -1$ respectively. It is also worth noting that independence is a sufficient but not necessary condition for $\delta_{X,Y}$ to be zero.

Below we recall the basic properties of linear correlation, and then continue with some copula based measures of dependence.

1.2.2 Linear Correlation

Linear correlation (or Pearson's correlation) is the measure of association that is by far the most used, but possibly also the most misunderstood. The popularity of linear correlation stems from the ease with which it can be calculated and understood as a scalar representing dependence in elliptical distributions (with often used members such as Gaussian distribution and Student's t-distribution). Most random variables however are not jointly elliptically distributed and using linear correlation for these can be misleading.

The value of Pearson's correlation (ρ) may not be a solid indicator of the strength of dependence. For one thing, independence of two rv's implies they are uncorrelated (i.e., $\rho = 0$) but the converse does not hold in general, as shown in the next example.

Example 1.2.1 ***(Two dependent variables with zero correlation).*** *Let Y be an rv taking the values 0, $\pi/2$ and π with probability $1/3$ each. Then, it is easy to see that $X_1 = \sin Y$ and $X_2 = \cos Y$ are uncorrelated (i.e., corr $(X_1, X_2) = 0$). However, they are not independent since X_1 and X_2 are functionally connected (by the relation $X_1^2 + X_2^2 = 1$).*

Linear correlation is defined in the following way.

Let (X, Y) be a vector of rv's with nonzero finite variances. The linear correlation coefficient for (X, Y) is

$$\rho^L(X,Y) = \frac{Cov(X,Y)}{Var(X)Var(Y)}, \tag{1.5}$$

where $Cov(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$ is the covariance of (X,Y), and $Var(X)$ and $Var(Y)$ are the variances of X and Y, respectively.

Linear correlation satisfies properties 1-6 but not the last property of a concordance measure. Linear correlation is invariant under only linear increasing transformations but not under general increasing transformations.

1.2.3 Kendall's τ and Spearman's ρ

Kendall's τ and Spearman's ρ are probably the two most important measures of concordance, as they provide arguably the best alternatives to linear correlation coefficient as a measure of association for non-elliptical distributions[1]. They are, however, not the only possible scalar valued measures of concordance (see e.g. Schweizer and Wolff (1981, [137]) for a complete list). It can be proved that both Kendall's τ and Spearman's ρ satisfy the properties of the definition of concordance (1.8).

Definition 1.2.2 *Kendal l's τ for random vectors (X_1, Y_1) and (X_2, Y_2) independent and identically distributed (i.i.d.) is defined as*

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

Kendall's τ is simply the difference between the probability of concordance and the probability of discordance.

Theorem 1.2.1 *Let (X,Y) be a vector of continuous random variables with copula C. Then Kendal l's τ can be written as*

$$\tau = 4\int_0^1 \int_0^1 C(u,v)dC(u,v) - 1, \tag{1.6}$$

The double integral in the above is the definition of the expected value for the function $C(U_1, U_2)$ where both U_1 and U_2 are standard uniform and have a joint distribution C. So $\tau = 4\mathbb{E}[C(U_1, U_2)] - 1$ and, $-1 \leq 4\mathbb{E}[C(U_1, U_2)] - 1 \leq 1$ i.e., Kendall's τ is a normalized expected value.

Definition 1.2.3 *Spearman's ρ for i.i.d. random vectors (X_1, Y_1), (X_2, Y_2) and (X_3, Y_3) is defined as*

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

Recall that linear correlation is a problem for non-elliptical distributions for which it often gives misleading results

Spearman's ρ then measures the difference of the probability of concordance and discordance for vectors (X_1, Y_1) and (X_2, Y_2) the latter being made up of independent rv's. Therefore in the ρ case the probabilities of concordance and discordance are measured with regards to the independent case.

Theorem 1.2.2 *Let (X, Y) be a vector of continuous rv's with copula C. Then Spearman's ρ can be written as*

$$\rho = 12 \int_0^1 \int_0^1 \{C(u,v) - uv\} dudv, \tag{1.7}$$

Hence, if $X \sim F$ and $Y \sim G$, and we let $U_1 = F(X)$ and $U_2 = G(Y)$, then

$$\begin{aligned} \rho &= 12 \int_0^1 \int_0^1 uvdC(u,v) - 3 \\ &= 12\mathbb{E}[U_1 U_2] - 3 \\ &= \frac{\mathbb{E}[U_1 U_2] - 1/4}{1/12} \end{aligned}$$

Since 1/2 and 1/12 are resp. the mean and variance of standard uniforms, we have

$$\begin{aligned} \rho &= \frac{\mathbb{E}[U_1 U_2] - (1/2)^2}{1/12} \\ &= \frac{Cov(U_1, U_2)}{\sqrt{Var(U_1)}\sqrt{Var(U_2)}} \\ &= \rho(F(X), G(Y)). \end{aligned}$$

Thus Spearman's ρ is the linear correlation coefficient between the integral transforms of X and Y, $(F(X)$ and $G(Y))$. For this reason it is often denoted as being a rank correlation measure. Because Kendall's τ and Spearman's ρ satisfy the properties of concordance (Definition 1.2.1) and if we recall the Fréchet-Hoeffding bounds (Theorem 1.1.3) it follows that e.g. the maximum concordance is achieved, when using the maximum copula, i.e. $C = C^{\perp} \Longrightarrow \tau = \rho = 1$. The converse is also true, so that

$$\begin{aligned} \rho = \tau = -1 &\quad \text{if and only if} \quad C = \mathbf{W}, \\ \rho = \tau = 1 &\quad \text{if and only if} \quad C = \mathbf{M}. \end{aligned}$$

Although the rank correlation measures have the property of invariance under monotonic transformations and can capture perfect dependence, they are not simple functions of moments and hence computation is more involved; see some examples in Table 1.1.

Table 1.1: Some standard copula functions.

Copulas type	θ domain	Kendall's τ	Spearman's ρ
Product	N. A	0	0
FGM	$-1 \leq \theta \leq 1$	$\frac{2}{9}\theta$	$\frac{1}{3}\theta$
Gaussian	$-1 < \theta < 1$	$\frac{2}{\pi}\arcsin(\theta)$	$\frac{6}{\pi}\arcsin(\theta/2)$
Clayton	$0 < \theta < \infty$	$\frac{\theta}{\theta+2}$	no closed form
Frank	$-\infty < \theta < \infty$	$1-\frac{4(1-D_1(\theta))}{\theta}$	$1-\frac{12(D_1(\theta)-D_2(\theta))}{\theta}$
AMH	$-1 \leq \theta \leq 1$	$\frac{3\theta-2}{\theta} - \frac{2\left(1-\frac{1}{\theta}\right)^2 \ln(1-\theta)}{3}$	no closed form
Plackett	$0 < \theta < \infty$	no closed form	$\frac{\theta+1}{\theta-1} - \frac{2\theta\ln\theta}{(\theta-1)^2}, \theta \neq 1$
Fréchet	$\alpha, \beta \geq 0,\ \alpha+\beta \leq 1$	$\frac{(\alpha-\beta)(\alpha+\beta+2)}{3}$	$\alpha-\beta.$

In some cases one can use (1.6) or (1.7). The relationship between τ and ρ is shown by a pair of inequalities due to Durbin and Stuart (1951) who showed that

$$\frac{3}{2}\tau - \frac{1}{2} \leq \rho \leq \frac{1}{2} + \tau - \frac{1}{2}\tau^2, \quad \text{for} \quad \tau \geq 0,$$

$$\frac{1}{2}\tau^2 + \tau - \frac{1}{2} \leq \rho \leq \frac{3}{2}\tau + \frac{1}{2}, \quad \text{for} \quad \tau \leq 0.$$

These inequalities form the basis of a widely presented 4-quadrant diagram that displays the (ρ, τ)-region; see Figure 1.4. Nelsen (1991, [122]) presents expressions for ρ and τ and their relationship for a number of copula families.

1.2.4 Perfect dependence

For every d-copula C we know from the Fréchet–Hoeffding inequality (Theorem 1.1.3) that

$$\mathbf{W}^d(u_1, ..., u_d) \leq C(u_1, ..., u_d) \leq \mathbf{M}^d(u_1, ..., u_d).$$

Furthermore, for $d = 2$ the upper and lower bounds are themselves copulas and we have seen that $\mathbf{W}$ and $\mathbf{M}$ are the bivariate distributions functions of the random vectors $(U, 1-U)$ and (U, U), respectively, where $U \sim U(0,1)$ (i.e. U is uniformly distributed on $[0,1]$). In this case we say that $\mathbf{W}$ describes perfect negative dependence and $\mathbf{M}$ describes perfect positive dependence.

Theorem 1.2.3 *Let (X, Y) have one of the copulas $\mathbf{W}$ or $\mathbf{M}$. Then there exist two*

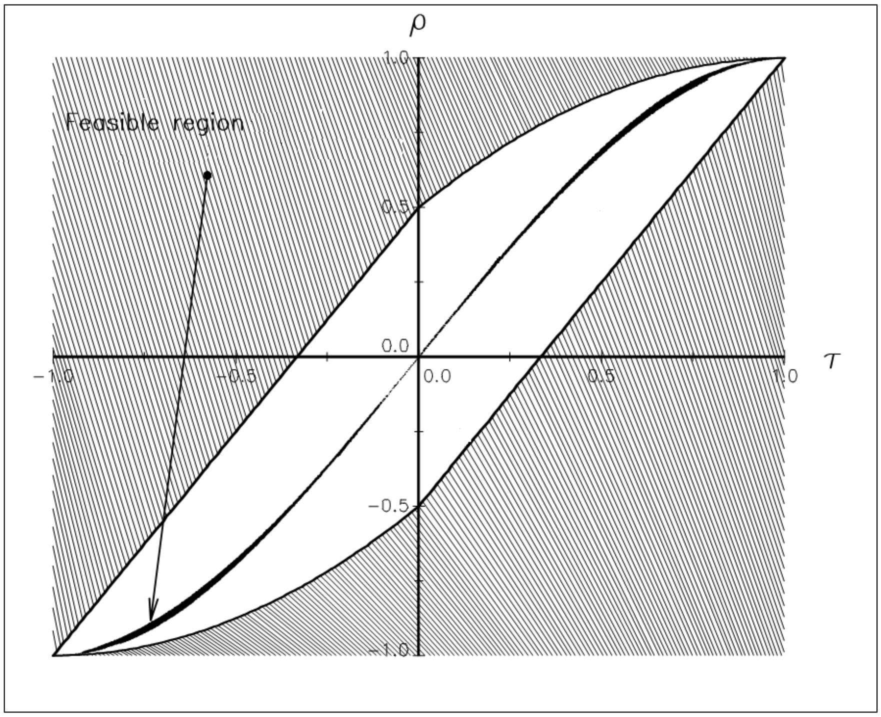

Figure 1.4: Bounding region for ρ and τ.

monotone functions $\alpha, \beta : \mathbb{R} \longrightarrow \mathbb{R}$ and a random variable Z so that

$$(X, Y) \stackrel{d}{=} (\alpha(Z), \beta(Z)),$$

with α increasing and β decreasing in the former case ($\mathbf{W}$) and both α and β increasing in the latter case ($\mathbf{M}$). The converse of this result is also true.

For a proof, see Embrechts *et al.* (2002, [40]). In a different form this result was already in Fréchet (1951, [49]).

Definition 1.2.4 *If (X, Y) has the copula $\mathbf{M}$ then X and Y are said to be comonotonic. If it has the copula $\mathbf{W}$ they are said to be countermonotonic.*

Note that if any of F and G (the df's of X and Y, respectively) have discontinuities, so that the copula is not unique, then $\mathbf{W}$ and $\mathbf{M}$ are possible copulas. In the case of F and G being continuous, a stronger version of the result can be stated:

$$\begin{aligned} C &= \mathbf{W} \Leftrightarrow Y = T(X) \ \text{ a.s., } \ T = G^{-1} \circ (1 - F) \text{ decreasing,} \\ C &= \mathbf{M} \Leftrightarrow Y = T(X) \ \text{ a.s., } \ T = G^{-1} \circ F \text{ increasing.} \end{aligned}$$

Other characterizations of comonotonicity can be found in Denneberg (1994, [30]).

1.2.5 Measures of dependence

Intuitively it is clear that measures of dependence are based on a «distance» between the copula of (X, Y) and the product copula Π.

Definition 1.2.5 ***(Schweizer and Wolff's Sigma).*** *Schweizer and Wolff's Sigma for a vector of continuous rv's (X, Y) with copula C is given by:*

$$\sigma_{X,Y} = 12 \iint_{[0,1]^2} |C(u,v) - uv| \, dudv.$$

Notice from this definition the similarity with Spearman's rho. The difference between the two is that this measure reports the absolute distance between the copula under consideration and the product copula whereas Spearman's rho reports the «signed» distance.

Definition 1.2.6 ***(Hoeffding's Dependence Index).*** *Hoeffding's Dependence Index for a vector of continuous rv's (X, Y) with copula C is given by:*

$$\Phi^2_{X,Y} = 90 \iint_{[0,1]^2} |C(u,v) - uv|^2 \, dudv.$$

1.2.6 Tail Dependence

The concept of tail dependence relates to the amount of dependence in the upper-right quadrant tail or lower-left-quadrant tail of a bivariate distribution. It is a concept that is relevant for the study of dependence between extreme values. It turns out that tail dependence between two continuous random variables X and Y is a copula property and hence the amount of tail dependence is invariant under strictly increasing transformations of X and Y.

Definition 1.2.7 *Let (X, Y) be a vector of continuous rv's with marginal df's F and G. The coefficient of upper tail dependence of (X, Y) is*

$$\lim_{u \nearrow 1} \mathbb{P}\{Y > G^{-1}(u) | X > F^{-1}(u)\} = \lambda_U,$$

provided that the limit $\lambda_U \in [0, 1]$ exists. If $\lambda_U \in (0, 1]$, X and Y are said to be asymptotically dependent in the upper tail; if $\lambda_U = 0$, X and Y are said to be asymptotically independent in the upper tail.

Since $\mathbb{P}\{Y > G^{-1}(u)|X > F^{-1}(u)\}$ can be written as

$$\frac{1 - \mathbb{P}\{X \leq F^{-1}(u)\} - \mathbb{P}\{Y \leq G^{-1}(u)\} + \mathbb{P}\{X \leq F^{-1}(u), Y \leq G^{-1}(u)\}}{1 - \mathbb{P}\{X \leq F^{-1}(u)\}},$$

an alternative and equivalent definition (for continuous rv's), from which it is seen that the concept of tail dependence is indeed a copula property, is the following which can be found in Joe (1997, [85, P. 33]).

Definition 1.2.8 *If a bivariate copula C is such that*

$$\lambda_U = \lim_{u \nearrow 1} \frac{1 - 2u + C(u,u)}{1 - u}$$

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

Example 1.2.2 *Consider the bivariate Gumbel family of copulas given by*

$$C_\theta(u,v) = \exp(-[(-\ln u)^\theta + (-\ln v)^\theta]^{1/\theta}),$$

for $\theta \geq 1$. Then

$$\begin{aligned} \frac{1 - 2u + C(u,u)}{1 - u} &= \frac{1 - 2u + \exp(2^{1/\theta} \ln u)}{1 - u} \\ &= \frac{1 - 2u + u^{2^{1/\theta}}}{1 - u}, \end{aligned}$$

and hence

$$\begin{aligned} \lim_{u \nearrow 1} \frac{(1 - 2u + C(u,u))}{(1 - u)} &= 2 - \lim_{u \nearrow 1} 2^{1/\theta} u^{2^{1/\theta}} - 1 \\ &= 2 - 2^{1/\theta}. \end{aligned}$$

Thus for $\theta > 1$, C_θ has upper tail dependence.

For copulas without a simple closed form an alternative formula for λ_U is more useful. An example is given in the case of the Gaussian copula

$$C_{\mathcal{R}}(u,v) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi-1(v)} \frac{1}{2\pi\sqrt{1 - \mathcal{R}_{12}^2}} \exp\left\{-\frac{s^2 - 2\mathcal{R}_{12} st + t^2}{2(1 - \mathcal{R}_{12}^2)}\right\} dsdt,$$

where $-1 < \mathcal{R}_{12} < 1$ and Φ is the univariate standard normal distribution function. Consider a pair of $U(0,1)$ rv's (U, V) with copula C. First note that $\mathbb{P}\{V \leq v|U =$

$u\} = \partial C(u,v)/\partial u$ and $\mathbb{P}\{V > v|U = u\} = 1 - \partial C(u,v)/\partial u$, and similarly when conditioning on V. Then

$$\begin{aligned}
\lambda_U &= \lim_{u \nearrow 1} \frac{C(u,u)}{(1-u)} \\
&= -\lim_{u \nearrow 1} \frac{dC(u,u)}{du} \\
&= -\lim_{u \nearrow 1} \left(-2 + \frac{\partial}{\partial s} C(s,t)\Big|_{s=t=u} + \frac{\partial}{\partial t} C(s,t)\Big|_{s=t=u} \right) \\
&= \lim_{u \nearrow 1} (\mathbb{P}\{V > u|U = u\} + \mathbb{P}\{U > u|V = u\}).
\end{aligned}$$

Furthermore, if C is an exchangeable copula, i.e. $C(u,v) = C(v,u)$, then the expression for λ_U simplifies to $\lambda_U = 2 \lim_{u \nearrow 1} \mathbb{P}\{V > u|U = u\}$.

Example 1.2.3 *Let (X,Y) have the bivariate standard normal distribution function with linear correlation coefficient ρ. That is $(X,Y) \sim C(\Phi(x), \Phi(y))$, where C is a member of the Gaussian family given above with $\mathcal{R}_{12} = \rho$. Since copulas in this family are exchangeable,*

$$\lambda_U = 2 \lim_{u \nearrow 1} \mathbb{P}\{V > u|U = u\},$$

and because Φ is a distribution function with infinite right endpoint,

$$\begin{aligned}
\lim_{u \nearrow 1} \mathbb{P}\{V > u|U = u\} &= \lim_{x \to \infty} \mathbb{P}\{\Phi^{-1}(V) > x|\Phi^{-1}(U) = x\} \\
&= \lim_{x \to \infty} \mathbb{P}\{X > x|Y = x\}.
\end{aligned}$$

Using the well known fact that $Y|X = x \sim \mathcal{N}(\rho x, 1-\rho^2)$ we obtain

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

from which it follows that $\lambda_U = 0$ for $\mathcal{R}_{12} < 1$. Hence the Gaussian copula C with $\rho < 1$ does not have upper tail dependence.

The concept of lower tail dependence can be defined in a similar way. If the limit

$$\lim_{u \searrow 0} \frac{C(u,u)}{u} = \lambda_L \tag{1.8}$$

exists, then C has lower tail dependence if $\lambda_L \in (0,1]$, and lower tail independence if $\lambda_L = 0$. For copulas without a simple closed form an alternative formula for λ_L is more useful. Consider a random vector (U, V) with copula C. Then

$$\begin{aligned}
\lambda_L &= \lim_{u \searrow 0} \frac{C(u,u)}{u} \\
&= \lim_{u \searrow 0} \frac{dC(u,u)}{du} \\
&= \lim_{u \searrow 0} \left(\frac{\partial}{\partial s} C(s,t) \Big|_{s=t=u} + \frac{\partial}{\partial t} C(s,t) \Big|_{s=t=u} \right) \\
&= \lim_{u \searrow 0} (\mathbb{P}\{V < u | U = u\} + \mathbb{P}\{U < u | V = u\}).
\end{aligned}$$

Furthermore if C is an exchangeable copula, i.e. $C(u,v) = C(v,u)$, then the expression for λ_L simplifies to

$$\lambda_L = 2 \lim_{u \searrow 0} \mathbb{P}\{V < u | U = u\}.$$

Recall that the survival copula of two random variables with copula C is given by

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

and the joint survival function for two $U(0,1)$-rv's whose joint df is C given by

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

Hence it follows that

$$\begin{aligned}
\lim_{u \nearrow 1} \frac{\widehat{C}(u,u)}{1-u} &= \lim_{u \nearrow 1} \frac{\overline{C}(1-u, 1-u)}{1-u} \\
&= \lim_{u \searrow 0} \frac{\overline{C}(u,u)}{u},
\end{aligned}$$

so the coefficient of upper tail dependence of C is the coefficient of lower tail dependence of $\overline{C}$. Similarly the coefficient of lower tail dependence of C is the coefficient of upper tail dependence of $\overline{C}$.

1.3 Copula estimation

There is a growing literature on the estimation of multivariate densities using copulas. Both parametric and nonparametric estimators have been proposed in the literature. Each method may be further divided into two-step and one-step approaches. In the two-step approach, each margin is estimated first and the estimated marginal df's are used to estimate copulas in the second step. The estimated parameters (in the parametric case) are typically ineffcient when estimated in two steps. In principle we can also estimate the joint density in one-step. The margins and the copula are estimated jointly in this approach. Although the estimated parameters (in parametric case) are effcient in this case, the one-step approach is more computationally burdensome than the two-step approach. In empirical work, we may have prior knowledge on the margins but not on the structure of the joint dependence. Therefore, the two-step approach may have an advantage over the one-step approach in terms of the computational consideration, although the estimates may be less effcient.

In practice there is usually little guidance on how to choose the best combination of the margins and the copula in parametric estimation. Therefore, semiparametric and non-parametric estimations have become popular in the literature recently. The main advantage of these estimation methods is to let the data determine the copula without imposing restrictive functional assumptions. In semiparametric estimations, often a parametric copula is specified but not the margins. The parameters in the copula function are estimated by maximum likelihood estimation. See earlier application in Oakes (1986, [125]), Genest and Rivest (1993, [59]), Genest, Ghoudi and Rivest (1995, [60]) and more recently in Liebscher (2005) and Chen et al. (2006, [18]).

Alternatively, nonparametric estimation does not assume any parametric distribution on both the margins and the copula. In this way, nonparametric estimation provides a higher degree of flexibility, since the dependence structure of the copula is not directly observable. It also illustrates a rough picture helpful to researchers for subsequent parametric estimation of the copula. In addition, the problem of misspecification in the copula can be avoided in the context of nonparametric estimation. The earliest nonparametric estimation in copulas is due to Deheuvels (1979, [26]) who estimated the copula density based on the empirical distribution.

Further work using kernel methods have been proposed by Gijbels and Mielnicnuk (1990, [65]), Fermanian and Scaillet (2005, [46]) in a time series framework and Chen and Huang (2007, [19]) with boundary corrections. Recently, Sancetta and Satchell (2004, [133]) use the Bernstein polynomials to approximate the Kimeldorf and Sampson copula. Hall and Neumeyer (2006, [70]) use wavelet estimators to approximate the copula density. Alternatively, Cai et al. (2008, [12]) use a mixture of parametric copulas to estimate unknown copula functions.

The kernel density estimator is one of the mostly popular methods in nonparametric estimations. Li and Racine (2007, [107]) provide a comprehensive review of this method. In spite of its popularity, there are several drawbacks in kernel estimation. If one uses a higher order kernel estimator in order to achieve a faster rate of convergence, it can result in negative density estimates. In addition, the support of data is often bounded with high concentration at or close to the boundaries in application. This boundary bias problem is well known in the univariate case, and can be more severe in the case of multivariate bounded support variables; see Muller (1991, [118]) and Jones (1993, [89]). The log-spline estimators have also drawn considerable attention in the literature[1] and have been studied extensively by Stone (1990, [143]). This estimator has been shown to perform well for density estimations. However, it suffers from a saturation problem. If we denote s the order of the spline and the logarithm of the density defined on a bounded support has r square integrable derivatives, the fastest convergence rate is achieved only if $s > r$. Like the kernel estimator, the log-spline estimator also faces boundary bias problem. It is known that boundary bias exists if the tail has a non-vanishing kth order derivative, while the order the (local) polynomial at the tail is less than k.

1.3.1 Exact Maximum Likelihood (EML)

Let's assume that we have a sample $\{x_1^t, ..., x_d^t\}_{t=1,...,d}$ containing the values of d different variables over T periods. The joint distribution of these variables is described by the copula $C(F_1(x_1^t), ..., F_d(x_d^t))$, with $F_1, ..., F_d$ all the marginal distributions. The expression for the log-likelihood function then becomes:

$$\ell(\boldsymbol{\theta}) = \sum_{t=1}^{T} \ln c(F_1(x_1^t), ..., F_d(x_d^t)) + \sum_{t=1}^{T} \sum_{i=1}^{d} \ln f_i\left(x_i^t\right),$$

with $\boldsymbol{\theta}$ a $k \times 1$ vector of parameters, including both the margin and the copula parameters. By maximizing the log-likelihood with respect to $\boldsymbol{\theta}$, we find the EML estimator, by $\hat{\boldsymbol{\theta}}^{EML}$. For this estimator it holds that:

$$\sqrt{T}\left(\hat{\boldsymbol{\theta}}^{EML} - \boldsymbol{\theta}\right) \to \mathcal{N}\left(0, \mathcal{J}^{-1}(\boldsymbol{\theta})\right)$$

with $\mathcal{J}(\boldsymbol{\theta}_0)$ the Fisher information matrix.

A closely related literature is the bivariate log-spline estimator studied by Stone (1994), Koo (1996) and Kooperberg (1998).

1.3.2 Inference Functions for Margins (IFM)

As stated previously, the ML method does not use the copula idea of a separation of margins from the dependence structure. The IFM method on the other hand is based on that principle. As a consequence of Sklar's theorem we can write the log-likelihood as:

$$\ell(\boldsymbol{\theta}) = \sum_{t=1}^{T} \ln c(F_1(x_1^t, \theta_1), ..., F_n(x_d^t, \theta_d); \boldsymbol{\alpha}) + \sum_{t=1}^{T} \sum_{i=1}^{d} \ln f_i \left(x_i^t, \theta_i\right),$$

with $\boldsymbol{\theta} = (\theta_1, ..., \theta_d; \boldsymbol{\alpha})$. Now we split the estimation of $\boldsymbol{\theta}$ in two steps:

- Estimation of the parameters in the univariate margins:

  $$\hat{\theta}_i = \arg\max_{\theta_i} \ell^i(\theta_i) := \arg\max_{\theta_i} \sum_{t=1}^{T} \ln f_i \left(x_i^t, \theta_i\right),$$

  for $i = 1, ..., d$.

- Estimation of the copula parameter, using the estimates obtained for the margins:

  $$\hat{\boldsymbol{\alpha}} = \arg\max_{\boldsymbol{\alpha}} \ell^C(\boldsymbol{\theta}) := \arg\max_{\boldsymbol{\alpha}} \sum_{t=1}^{T} \ln c(F_1(x_1^t, \hat{\theta}_1), ..., F_d(x_d^t, \hat{\theta}_d); \boldsymbol{\alpha})$$

Now the IFM estimator is defined as:

$$\hat{\boldsymbol{\theta}}^{IFM} = (\hat{\theta}_1, ..., \hat{\theta}_d; \hat{\boldsymbol{\alpha}})$$

It can be shown that the IFM estimator verifies the property of asymptotic normality

$$\sqrt{T}\left(\hat{\boldsymbol{\theta}}^{IFM} - \boldsymbol{\theta}\right) \to \mathcal{N}\left(0, \mathcal{V}^{-1}(\boldsymbol{\theta})\right),$$

with $\mathcal{V}(\theta)$ the information matrix of Godambe. Define the score function as

$$g(\boldsymbol{\theta}) = \left(\frac{\partial}{\partial\theta_1}\ell^1, ..., \frac{\partial}{\partial\theta_d}\ell^d, \frac{\partial}{\partial\boldsymbol{\alpha}}\ell^C\right),$$

then the Godambe information matrix can be written as (see Joe 1997, [85]):

$$\mathcal{V}(\boldsymbol{\theta}) = \left(\mathbb{E}\left(\frac{\partial}{\partial\theta}g\left(\theta\right)^{T}\right)\right)^{-1}\mathbb{E}\left(g\left(\theta\right)^{T}g\left(\theta\right)\right)\left(\left(\mathbb{E}\left(\frac{\partial}{\partial\theta}g\left(\theta\right)^{T}\right)\right)^{-1}\right)^{T},$$

where A^T denote the transpose of A.

1.3.3 Canonical Maximum Likelihood (CML)

When we use the IFM method we will end up with an estimate for $\boldsymbol{\alpha}$ that depends on the distributional assumptions that we made about the margins. How can we change our approach and get an estimate for $\boldsymbol{\alpha}$ that is margin independent?

First transform the data $(x_1^t, ..., x_d^t)$ into uniform variates $(\hat{u}_1^t, ..., \hat{u}_d^t)$ using the empirical distribution functions of the univariate margins. After that we estimate the copula parameters $\boldsymbol{\alpha}$ in the following way:

$$\hat{\boldsymbol{\alpha}} = \arg\max\sum_{t=1}^{T}\ln c(\hat{u}_1^t, ..., \hat{u}_d^t; \boldsymbol{\alpha}).$$

In practice transforming the data into uniform variates is performed using a transformation of the empirical distribution function. Define $F_1^T, ..., F_d^T$, the empirical distribution functions, in the following way:

$$F_j^T\left(x\right) = \frac{1}{T}\sum_{t=1}^{T}\mathbf{1}_{\left\{X_j^t \leq x_j\right\}},$$

where $\mathbf{1}(\cdot)$ is the indicator function. In the estimation procedure we will use

$$\frac{T}{T+1}F_j^T\left(x\right),$$

to avoid possible problems with unboundedness of the copula density if some of the u_i's are equal to 1. This means that the transformation of the data into standard uniform random variates is performed by:

$$\hat{u}_j^t = \frac{T}{T+1}F_j^T\left(x_j^t\right), \quad \text{for } j = 1, ..., d. \tag{1.9}$$

The resulting estimator for $\boldsymbol{\alpha}$ is also called a Pseudo-Likelihood Estimator or Omnibus Estimator. More information about this procedure and the properties of the estimator can be found in Genest *et al.* (1995, [60]).

Note that the difference between IFM and CML is that IFM first estimates the marginal distributions under distributional assumptions and then transforms the data into uniform variates while the CML method directly transforms the data into uniform variates using the empirical distribution of the data. Another point worth noticing is that the empirical distribution function can also be described in terms of rank numbers. For the definition of the rank

$$\text{Rank}(X_i) := 1 + \# \{ j | \, X_j < X_i \} + \frac{1}{2} \# \{ j | \, j \neq i \text{ and } X_j = X_i \} ,$$

Write R_j^t to denote the rank of observation t for variable j. Now the data transformation can also be written in the following way:

$$\hat{u}_j^t = \frac{R_j^t}{T+1},$$

In Genest *et al.* (1995, [60]) it is proven that (under certain regularity conditions):

$$\sqrt{T} \left(\hat{\boldsymbol{\alpha}}^{CML} - \boldsymbol{\alpha} \right) \to \mathcal{N} \left(0, \mathcal{V} \right),$$

where $\boldsymbol{\alpha}$ denotes the true copula parameter (under the assumption that the data generating process is driven by this specific copula) and $\mathcal{V}$ is not dependent upon T. In Genest and Werker (2002, [64]) conditions are given under which CML estimators are semiparametrically efficient in large samples, and it is argued that for most copulas these requirements are not satisfied.

1.3.4 The Empirical Copula Function

Here we give a non parametric method for getting a bivariate copula. Consider a sample $(X_1, Y_1), (X_2, Y_2), ..., (X_n, Y_n)$, i.i.d n copies of a random vector (X, Y). The bivariate empirical distribution function (see Embrechts *et al.* 1997, [42, p182]) associated with (X, Y) is

$$\mathbb{H}_n(x, y) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{\{X_i \leq x, Y_i \leq y\}},$$

with marginals

$$\mathbb{F}_n(x) = \mathbb{H}_n(x, -\infty) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{\{X_i \leq x\}},$$

and

$$\mathbb{G}_n(y) = \mathbb{H}_n(-\infty, y) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{Y_i \le y\}}.$$

Then (see Veraverbeke 2005, [148]) the empirical copula function is given by

$$\begin{aligned}\mathbb{C}_n(u,v) &= \mathbb{H}_n(\mathbb{F}_n^{-1}(u), \mathbb{G}_n^{-1}(v)) \\ &= \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{X_i \le \mathbb{F}_n^{-1}(u), Y_i \le \mathbb{G}_n^{-1}(v)\}}.\end{aligned}$$

(see Nelsen 2006, [123, p 219]) defined this copula as

$$\mathbb{C}_n(\frac{i}{n}, \frac{j}{n}) = \frac{\text{number of pairs } (x,y) \text{ in the sample with } x \le X_{i:n}, y \le Y_{j:n}}{n}$$

where $X_{i:n}$ and $Y_{j:n}$, $1 \le i, j \le n$, denote the order statistics of the sample. Note that the empirical copula function based on $(X_1, Y_1), (X_2, Y_2), ..., (X_n, Y_n)$, is the same as that based on uniform $[0,1]$ rv's $(U_1, V_1), (U_2, V_2), ..., (U_n, V_n)$, where $U_i = F(X_i)$ and $V_i = G(Y_i)$, $i \in \{1, 2, ..., n\}$ (see Veraverbeke 2005, [148]).

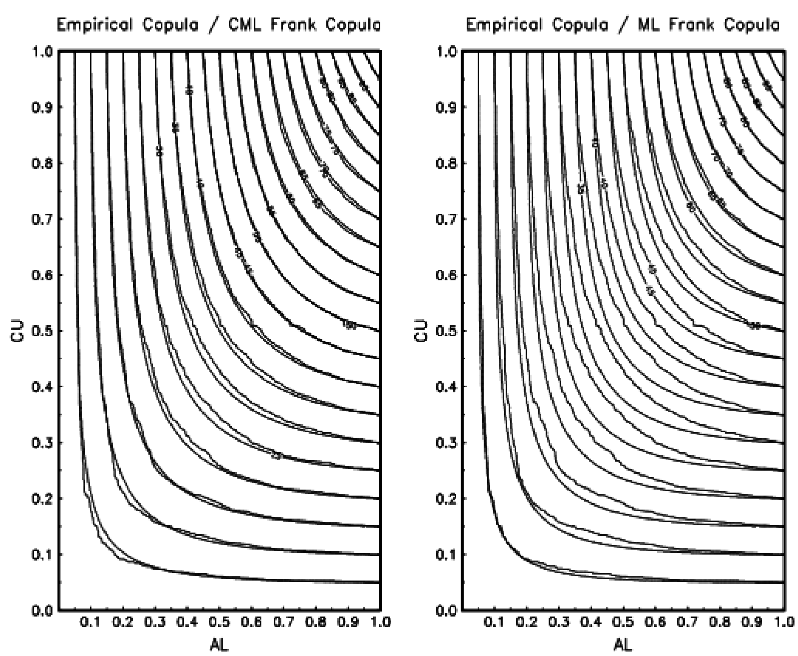

Figure 1.5: Comparison of empirical copula, CML Frank copula and ML Frank copula

Chapter 2

Copula families

«The essence of mathematics is not to make simple things complicated,

but to make complicated things simple.»

S. Gudder

In This Chapter will deal mainly with parametric copulas. The word «parametric» suggests that there is also something like non-parametric copulas, and indeed this is the case. Three different forms of non-parametric copulas are distinguished in the current literature. The so-called Deheuvels or empirical copula (1979, [26]). Is defined similar to an empirical distribution function, by counting the number of outcomes below certain values and dividing that number by the total number of outcomes. Empirical copulas are used in practice, but have the disadvantage that they are discontinuous functions. Kernel approximations to copula functions were introduced very recently in Fermanian and Scaillet (2003, [45]). The kernel approximations to copulas, were introduced in Li et *al.* (1997, [106]) . In this approach an approximating functional form for the copula is specified. Two important approximating forms are called the Bernstein polynomial, and the checkerboard copula, these approximations converge to the true copula when the number of observations increases, the tail dependence will always equal zero for the approximations, as proven in Durrleman et al. (2000, [37]).

Note that we will only look at copula families that are parametric. That is the most commonly used approach in the literature. Because of the specific nature of this research, multivariate distributions of order higher than two, only two specific classes of parametric copulas will be treated. The extension of twodimensional copulas to copulas of higher dimension is certainly not a trivial one, but for the three classes treated here it is feasible. Other classes of copulas can be found in the literature, such as the Fréchet Family (See Hürlimann 2004, [83]) and Extreme Value copulas (See Joe 1997, [85]).

The first class of copulas that will be treated is the Marshall-Olkin copulas. The elliptical copulas, specifically the Gaussian and t-copulas. (for definitions see, Embrechts *et al.* 2000, [39]) and in the end the Archimedean copulas.

2.1 Marshall-Olkin Copulas

In this section we discuss a class of copulas called Marshall-Olkin copulas. To be able to derive these copulas and present explicit expressions for rank correlation and tail dependence coefficients without tedious calculations, we gives with bivariate Marshall-Olkin copulas and we then suggest applications of Marshall-Olkin copulas to the modelling of dependent risks. For further details about Marshall-Olkin distributions we refer to Marshall and Olkin (1967, [114]). Similar ideas are contained in Muliere and Scarsini (1987, [119]).

Consider a two-component system where the components are subject to shocks, which are fatal to one or both components. Let X and Y denote the lifetimes of the two components. Furthermore assume that the shocks follow three independent Poisson processes with parameters λ_1, λ_2, $\lambda_{12} \geq 0$, where the index indicates whether the shocks effect only component 1, only component 2 or both. Then the times Z_1, Z_2 and Z_{12} of occurrence of these shocks are independent exponential rv's with parameters λ_1, λ_2 and λ_{12} respectively. Hence

$$\begin{aligned} \overline{H}(x,y) &= \mathbb{P}\{X > x, Y > y\} \\ &= \mathbb{P}\{Z_1 > x\}\mathbb{P}\{Z_2 > y\}\mathbb{P}\{Z_{12} > \max(x,y)\}. \end{aligned}$$

The univariate survival functions for X and Y are $\overline{F}(x) = \exp(-(\lambda_1 + \lambda_{12})x)$ and $\overline{G}(y) = \exp(-(\lambda_2 + \lambda_{12})y)$. Furthermore, since $\max(x,y) = x + y - \min(x,y)$,

$$\begin{aligned} \overline{H}(x,y) &= \exp(-(\lambda_1 + \lambda_{12})x - (\lambda_2 + \lambda_{12})y + \lambda_{12}\min(x,y)) \\ &= \overline{F}(x)\overline{G}(y)\min(\exp(\lambda_{12}x), \exp(\lambda_{12}y)). \end{aligned}$$

Let $\alpha = \lambda_{12}/(\lambda_1 + \lambda_{12})$ and $\beta = \lambda_{12}/(\lambda_2 + \lambda_{12})$. Then $\exp(\lambda_{12}x) = \overline{F}_1(x)^{-\alpha}$ and

$\exp(\lambda_{12}y) = \overline{F}_2(y)^{-\beta}$, and hence the survival copula of (X, Y) is given by

$$\overline{C}\,(u,v) = uv\min(u^{-\alpha}, v^{-\beta}) = \min(u^{1-\alpha}v, uv^{1-\beta}).$$

This construction leads to a copula family given by

$$\begin{aligned} C_{\alpha,\beta}\,(u,v) &= \min(u^{1-\alpha}v, uv^{1-\beta}) \\ &= \begin{cases} u^{1-\alpha}v, & u^\alpha \geq v^\beta, \\ uv^{1-\beta}, & u^\alpha \leq v^\beta. \end{cases} \end{aligned}$$

This family is known as the Marshall-Olkin family. Marshall-Olkin copulas have both an absolutely continuous and a singular component. Since

$$\frac{\partial^2}{\partial u \partial v} C_{\alpha,\beta}\,(u,v) = \begin{cases} u^{-\alpha}, & u^\alpha > v^\beta, \\ v^{-\beta}, & u^\alpha < v^\beta, \end{cases}$$

the mass of the singular component is concentrated on the curve $u^\alpha = v^\beta$ in $[0,1]^2$.

Kendall's tau and Spearman's rho are easily evaluated for this copula family. For Spearman's rho, applying Theorem 1.2.2 yields:

$$\begin{aligned} \rho(C_{\alpha,\beta}) &= 12 \iint_{[0,1]^2} C_{\alpha,\beta}(u,v)dudv - 3 \\ &= 12 \int_0^1 \left(\int_0^{u^{\alpha/\beta}} u^{1-\alpha}vdv + \int_{u^{\alpha/\beta}}^1 uv^{1-\alpha}dv \right) du - 3 \\ &= \frac{3\alpha\beta}{2\alpha + 2\beta - \alpha\beta}. \end{aligned}$$

To evaluate Kendall's tau we use the following theorem, a proof of which is found in Nelsen (2006, [123, p.168]).

Theorem 2.1.1 *Let C be a copula such that the product $(\partial C/\partial u)\,(\partial C/\partial v)$ is integrable on $[0,1]^2$. Then*

$$\iint_{[0,1]^2} C(u,v)dC(u,v) = \frac{1}{2} - \iint_{[0,1]^2} \frac{\partial}{\partial u} C(u,v) \frac{\partial}{\partial v} C(u,v)dudv.$$

Using Theorems 1.2.1 and 2.1.1 we obtain

$$\begin{aligned}\tau(C_{\alpha,\beta}) &= 4\iint_{[0,1]^2} C_{\alpha,\beta}(u,v)dC_{\alpha,\beta}(u,v) - 1\\ &= 4\left(\frac{1}{2} - \iint_{[0,1]^2} \frac{\partial}{\partial u}C_{\alpha,\beta}(u,v)\frac{\partial}{\partial v}C_{\alpha,\beta}(u,v)dudv\right) - 1\\ &= \frac{\alpha\beta}{\alpha+\beta-\alpha\beta}.\end{aligned}$$

Thus, all values in the interval $[0,1]$ can be obtained for $\rho(C_{\alpha,\beta})$ and $\tau(C_{\alpha,\beta})$. The Marshall-Olkin copulas have upper tail dependence. Without loss of generality assume that $\alpha > \beta$, then

$$\begin{aligned}\lim_{u\nearrow 1}\frac{\overline{C}(u,u)}{1-u} &= \lim_{u\nearrow 1}\frac{1-2u+u^2\min(u^{-\alpha},u^{-\beta})}{1-u}\\ &= \lim_{u\nearrow 1}\frac{1-2u+u^2u^{-\beta}}{1-u}\\ &= \lim_{u\nearrow 1}(2-2u^{1-\beta}+\beta u^{1-\beta})\\ &= \beta,\end{aligned}$$

and hence $\lambda_U = \min(\alpha,\beta)$ is the coefficient of upper tail dependence.

2.2 Elliptical Copulas

The class of elliptical copulas has an unfavorable property when talking about application in the field of finance. The dependence structure in financial data cannot be represented correctly. For instance, the asymmetry of the lower and upper tail of a distribution cannot be described properly by an elliptical copula. This is because elliptical copulas exhibit «radial symmetry» which has the property that $C(u,v) = C(v,u)$ and $C(u,v) = \overline{C}(u,v) = u+v-1+C(1-u,1-v)$. Another fact that elliptical copulas do not have closed form expressions.

The class of elliptical distributions provides a rich source of multivariate distributions which share many of the tractable properties of the multivariate normal distribution and enables modelling of multivariate extremes and other forms of nonnormal dependences.

Elliptical copulas are simply the copulas of elliptical distributions. Simulation from elliptical distributions is easy, and as a consequence of Sklar's Theorem so is simulation from elliptical copulas. Furthermore, we will show that rank correlation and tail dependence coefficients can be easily calculated. For further details on elliptical distributions we refer to Fang, Kotz, and Ng (1990, [43]) and Cambanis, Huang, and Simons (1981, [13]).

2.2.1 Elliptical distributions

Definition 2.2.1 *Let* $\mathbf{X}$ *is a* d*-dimensional random vector and, for some* $\boldsymbol{\mu} \in \mathbb{R}^n$ *and some* $d \times d$ *nonnegative definite, symmetric matrix* Σ*, the characteristic function* $\varphi_{\mathbf{X}-\boldsymbol{\mu}}(\mathbf{t})$ *of* $\mathbf{X} - \boldsymbol{\mu}$ *is a function of the quadratic form* $\mathbf{t}^T \Sigma \mathbf{t}$*, ie.*

$$\varphi_{\mathbf{X}-\boldsymbol{\mu}}(\mathbf{t}) = \phi(\mathbf{t}^T \Sigma \mathbf{t}),$$

we say that $\mathbf{X}$ *has an elliptical distribution with parameters* $\boldsymbol{\mu}$, Σ *and* ϕ*, and we write* $\mathbf{X} \sim E_n(\boldsymbol{\mu}, \Sigma, \phi)$.

When $d = 1$, the class of elliptical distributions coincides with the class of onedimensional symmetric distributions. A function ϕ as in Definition 2.2.1 is called a characteristic generator.

Theorem 2.2.1 $\mathbf{X} \sim E_d(\mu, \Sigma, \phi)$ *with* $\mathrm{Rank}(\Sigma) = k$ *if and only if there exist a rv* $\mathcal{R} \geq 0$ *independent of* $\mathbf{U}$*, a* k*-dimensional random vector uniformly distributed on the unit hypersphere* $\{\mathbf{z} \in \mathbb{R}^k | \mathbf{z}^T \mathbf{z} = 1\}$*, and an* $d \times k$ *matrix* A *with* $AA^T = \Sigma$*, such that*

$$X \stackrel{d}{=} \boldsymbol{\mu} + \mathcal{R} A \mathbf{U}.$$

For the proof of Theorem 2.2.1 and the relation between $\mathcal{R}$ and ϕ see Fang *et al.* (1990, [43]) or Cambanis, Huang and Simons (1981, [13]).

Example 2.2.1 *Let* $\mathbf{X} \sim \mathcal{N}_d(\mathbf{0}, \mathbf{I}_d)$*. Since the components* $X_i \sim \mathcal{N}(0,1)$*,* $i = 1, ..., d$*, are independent and the characteristic function of* X_i *is* $\varphi_{X_i}(t_i) = \exp(-t_i^2/2)$*, the characteristic function of* $\mathbf{X}$ *is*

$$\begin{aligned} \phi_{\mathbf{X}}(\mathbf{t}) &= \exp\left\{-\frac{1}{2}(t_1^2 + ... + t_n^2)\right\} \\ &= \exp\left\{-\frac{1}{2}\mathbf{t}^T\mathbf{t}\right\}. \end{aligned}$$

From Theorem 2.2.1 it then follows that $\mathbf{X} \sim E_d(\mathbf{0}, \mathbf{I}_d, \phi)$*, where* $\phi(\mathbf{u}) = \exp(-\mathbf{u}/2)$.

If $\mathbf{X} \sim E_d(\boldsymbol{\mu}, \Sigma, \phi)$, where Σ is a diagonal matrix, then $\mathbf{X}$ has uncorrelated components (if $0 < Var(X_i) < \infty$). If $\mathbf{X}$ has independent components, then $\mathbf{X} \sim \mathcal{N}_d(\boldsymbol{\mu}, \Sigma)$. Note that the multivariate normal distribution is the only one among the elliptical distributions where uncorrelated components imply independent components.

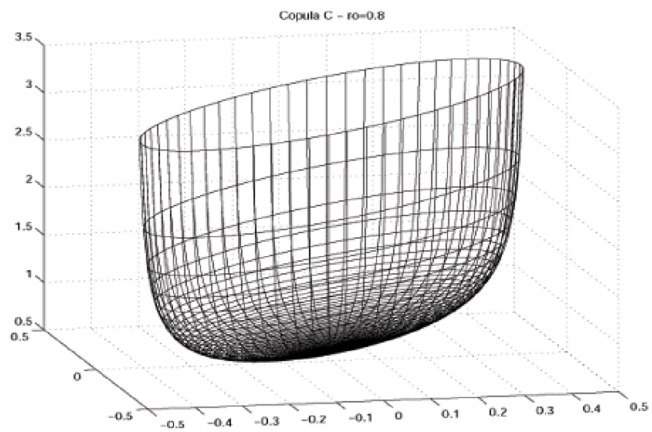

Figure 2.1: A density function of an elliptical copula with correlation 0.8.

A random vector $\mathbf{X} \sim E_d(\boldsymbol{\mu}, \Sigma, \phi)$ does not necessarily have a density. If $\mathbf{X}$ has a density it must be of the form

$$|\Sigma|^{-1/2} g((\mathbf{X} - \boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{X} - \boldsymbol{\mu})),$$

for some nonnegative function g of one scalar variable.
Hence the contours of equal density form ellipsoids in $\mathbb{R}^d$. Given the distribution of $\mathbf{X}$, the representation $E_d(\boldsymbol{\mu}, \Sigma, \phi)$ is not unique. It uniquely determines $\boldsymbol{\mu}$ but Σ and ϕ are only determined up to a positive constant. More precisely, if $\mathbf{X} \sim E_d(\boldsymbol{\mu}, \Sigma, \phi)$ and $\mathbf{X} \sim E_d(\boldsymbol{\mu}^*, \Sigma^*, \phi^*)$, then

$$\boldsymbol{\mu}^* = \boldsymbol{\mu}, \quad \Sigma^* = c\Sigma, \quad \phi^*(\cdot) = \phi(\cdot/c),$$

for some constant $c > 0$. In order to find a representation such that $Cov(\mathbf{X}) = \Sigma$, we use Theorem 2.2.1 to obtain

$$Cov(\mathbf{X}) = Cov(\boldsymbol{\mu} + \mathcal{R}A\mathbf{U}) = A\mathbb{E}(\mathcal{R}^2)Cov(\mathbf{U})A^T,$$

provided that(1) $\mathbb{E}(\mathcal{R}^2) < \infty$. Let $\mathbf{Y} \sim \mathcal{N}_d(\mathbf{0}, \mathbf{I}_d)$. Then $Y \stackrel{d}{=} \|\mathbf{Y}\| \, \mathbf{U}$, where $\|\mathbf{Y}\|$ is independent of $\mathbf{U}$. Furthermore $\|\mathbf{Y}\|^2 \sim \chi_d^2$, so $\mathbb{E}(\|\mathbf{Y}\|^2) = d$. Since $Cov(\mathbf{Y}) = \mathbf{I}_d$ we see that if $\mathbf{U}$ is uniformly distributed on the unit hypersphere in $\mathbb{R}^d$, then $Cov(\mathbf{U}) = \mathbf{I}_d/d$.

Thus $Cov(\mathbf{X}) = AA^T\mathbb{E}(\mathcal{R}^2)/d$. By choosing the characteristic generator $\phi^*(s) = \phi(s/c)$, where $c = \mathbb{E}(\mathcal{R}^2)/d$, we get $Cov(\mathbf{X}) = \Sigma$.

Hence an elliptical distribution is fully described by $\boldsymbol{\mu}$, Σ and ϕ, where ϕ can be chosen so that([2]) $Cov(\mathbf{X}) = \Sigma$. If $Cov(\mathbf{X})$ is obtained as above, then the distribution of $\mathbf{X}$ is uniquely determined by $\mathbb{E}(\mathbf{X})$, $Cov(\mathbf{X})$ and the type of its univariate margins.

As usual, let $\mathbf{X} \sim E_d(\boldsymbol{\mu}, \Sigma, \phi)$. Whenever $0 < Var(X_i), Var(X_j) < \infty$,

$$\rho(X_i, X_j) := Cov(X_i, X_j)/\sqrt{Var(X_i)Var(X_j)} = \Sigma_{ij}/\sqrt{\Sigma_{ii}\Sigma_{jj}}.$$

This explains why linear correlation is a natural measure of dependence between rv's with a joint nondegenerate ($\Sigma_{ii} > 0$ for all i) elliptical distribution. Throughout this section we call the matrix $\mathcal{R}$, with

$$\mathcal{R}_{ij} = \frac{\Sigma_{ij}}{\Sigma_{ii}\Sigma_{jj}},$$

the linear correlation matrix of $\mathbf{X}$.

Note that this definition is more general than the usual one and in this situation (elliptical distributions) makes more sense. Since an elliptical distribution is uniquely determined by $\boldsymbol{\mu}$, Σ and ϕ, the copula of a nondegenerate elliptically distributed random vector is uniquely determined by $\mathcal{R}$ and ϕ.

One practical problem with elliptical distributions in multivariate risk modelling is that all margins are of the same type. To construct a realistic multivariate distribution for some given risks, it may be reasonable to choose a copula of an elliptical distribution but different types of margins (not necessarily elliptical), such a model seems to be that the copula parameter $\mathcal{R}$ can not be estimated directly from data[3]. In such cases, $\mathcal{R}$ can be estimated using (robust) linear correlation estimators. The Kendall's tau rank correlation matrix for a random vector is invariant under strictly increasing transformations of the vector components, and the next theorem provides

$\mathbb{E}(X)$ denote the expectation of X.

If $Cov(\mathbf{X})$ is defined

Recall that for nondegenerate elliptical distributions with finite variances, $\mathcal{R}$ is just the usual linear correlation matrix.

a relation between the Kendall's tau rank correlation matrix and $\mathcal{R}$ for nondegenerate elliptical distributions, then $\mathcal{R}$ can in fact easily be estimated from data.

Theorem 2.2.2 *Let* $\mathbf{X} \sim E_d(\boldsymbol{\mu}, \Sigma, \phi)$ *with* $\mathbb{P}\{X_i = \mu_i\} < 1$ *and* $\mathbb{P}\{X_j = \mu_j\} < 1$. *Then*

$$\tau(X_i, X_j) = (1 - \sum_{x \in \mathbb{R}} (\mathbb{P}\{X_i = x\})^2) \frac{2}{\pi} \arcsin(\mathcal{R}_{ij}), \tag{2.1}$$

where the sum extends over all atoms of the distribution of X_i. *If* $\mathrm{Rank}(\Sigma) \geq 2$, *then (2.1) simplifies to*

$$\tau(X_i, X_j) = (1 - (\mathbb{P}\{X_i = \mu_i\})^2) \frac{2}{\pi} \arcsin(\mathcal{R}_{ij}).$$

For a proof, see Lindskog, McNeil, and Schmock (2003, [110]). Note that if $\mathbb{P}\{X_i = \mu_i\} = 0$ for all i, which is true for e.g. multivariate t or normal distributions with strictly positive definite dispersion matrices Σ, then

$$\tau(X_i, X_j) = \frac{2}{\pi} \arcsin(\mathcal{R}_{ij}),$$

for all i and j.

The nonparametric estimator of $\mathcal{R}$, $\sin(\pi\hat{\tau}/2)$ (dropping the subscript for simplicity), provided by the above theorem, inherits the robustness properties of the Kendall's tau estimator and is an efficient (low variance) estimator of $\mathcal{R}$ for both elliptical distributions and nonelliptical distributions with elliptical copulas.

2.2.2 Gaussian or Normal Copulas

The copula of the d-variate normal distribution with linear correlation matrix $\mathcal{R}$ is

$$C_{\mathcal{R}}^{Ga}(u) = \Phi_{\mathcal{R}}^d(\Phi^{-1}(u_1), ..., \Phi^{-1}(u_d)),$$

where $\Phi_{\mathcal{R}}^d$ denotes the joint distribution function of the d-variate standard normal distribution function with linear correlation matrix $\mathcal{R}$, and Φ^{-1} denotes the inverse of the distribution function of the univariate standard normal distribution. Copulas of the above form are called *Gaussian copulas.* In the bivariate case the copula expression can be written as

$$C_{\mathcal{R}}^{Ga}(u, v) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi(1 - \mathcal{R}_{12}^2)^{1/2}} \exp\left\{ -\frac{s^2 - 2\mathcal{R}_{12}st + t^2}{2(1 - \mathcal{R}_{12}^2)} \right\} dsdt.$$

Note that $\mathcal{R}_{12}$ is simply the usual linear correlation coefficient of the corresponding bivariate normal distribution. Example 1.2.3 shows that Gaussian copulas do not have upper tail dependence. Since elliptical distributions are radially symmetric, the coefficient of upper and lower tail dependence are equal. Hence Gaussian copulas do not have lower tail dependence.

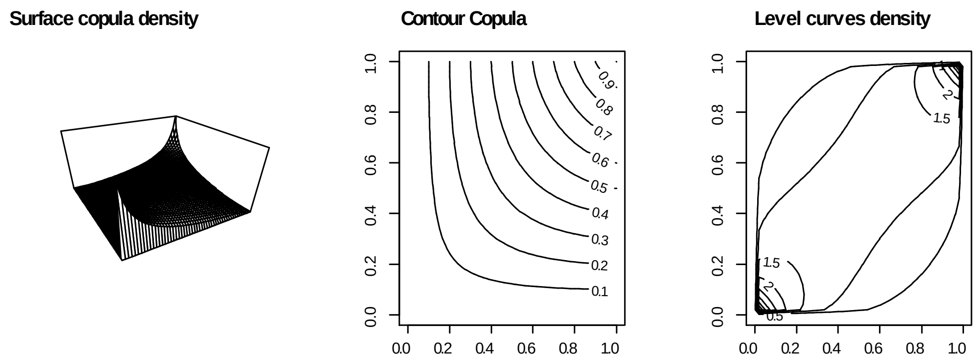

Figure 2.2: The Normal 2-copula density and the corresponding density contour and level curves. Here $\mathcal{R}_{12} = 0.5$.

We now address the question of random variate generation from the Gaussian copula $C_{\mathcal{R}}^{Ga}$. For our purpose, it is sufficient to consider only strictly positive definite matrices $\mathcal{R}$. Write $\mathcal{R} = AA^T$ for some $d \times d$ matrix A, and if $Z_1, ..., Z_d \sim \mathcal{N}(0,1)$ are independent, then

$$\boldsymbol{\mu} + AZ \sim \mathcal{N}_d(\boldsymbol{\mu}, \mathcal{R}).$$

One natural choice of A is the Cholesky decomposition of $\mathcal{R}$. The Cholesky decomposition of $\mathcal{R}$ is the unique lower-triangular matrix L with $LL^T = \mathcal{R}$. Furthermore Cholesky decomposition routines are implemented in most mathematical software. This provides an easy algorithm for random variate generation from the Gaussian d-copula $C_{\mathcal{R}}^{Ga}$.

Algorithm 2.2.1

1. Find the Cholesky decomposition A of $\mathcal{R}$.

2. Simulate d independent random variates $z_1, ..., z_d$ from $\mathcal{N}(0,1)$.

3. Set $\mathbf{x} = A\mathbf{z}$.

4. Set $u_i = \Phi(x_i), i = 1, ..., d$.

5. $(u_1, ..., u_d) \sim C_{\mathcal{R}}^{Ga}$.

As usual Φ denotes the univariate standard normal distribution function.

2.2.3 t-copulas

The d-dimensional random vector $\mathbf{X} = (X_1, ..., X_d)$ is said to have a (non-singular) multivariate t-distribution with ν degrees of freedom, mean vector $\boldsymbol{\mu}$ and positive-definite dispersion or scatter matrix Σ, if its density is given by

$$f(\mathbf{x}) = \frac{\Gamma\left(\frac{\nu+d}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{(\nu\pi)^d |\Sigma|}} \left(1 + \frac{(\mathbf{x}-\boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{x}-\boldsymbol{\mu})}{\nu}\right)^{-\frac{\nu+d}{2}}.$$

It is well-known that the multivariate t belongs to the class of multivariate normal variance mixtures and has the stochastic representation

$$X \overset{d}{=} \boldsymbol{\mu} + \frac{\sqrt{\nu}}{\sqrt{S}}\mathbf{Z}, \tag{2.2}$$

where $\boldsymbol{\mu} \in \mathbb{R}^d$, $S \sim \chi_\nu^2$ and $\mathbf{Z} \sim \mathcal{N}_d(\mathbf{0}, \Sigma)$ are independent, then $\mathbf{X}$ has an d-variate t_ν-distribution with mean $\boldsymbol{\mu}$ (for $\nu > 1$) and covariance matrix $\frac{\nu}{\nu-2}\Sigma$ (for $\nu > 2$). If $\nu \leq 2$ then $Cov(\mathbf{X})$ is not defined. In this case we just interpret Σ as being the shape parameter of the distribution of $\mathbf{X}$.

The t-copula is obtained after filtering out all univariate t_ν-distributions from the multivariate $t_{\Sigma,\nu}$-distribution. Again, this follows directly from Sklar's theorem.

Definition 2.2.2 *Let t_ν denote the univariate t-distribution function with ν degrees of freedom, and let $t_{\Sigma,\nu}$ denote the standard multivariate d-dimensional t-distribution function with linear correlation matrix Σ and degrees of freedom ν. The d-dimensional t-copula is then defined as:*

$$C_{\Sigma,\nu}^t(u_1, ..., u_d) = t_{\Sigma,\nu}\left(t_\nu^{-1}(u_1), ..., t_\nu^{-1}(u_d)\right).$$

Equivalently, the following definition is often used:

Definition 2.2.3 ***(Multivariate t-Copula).*** *The d-dimensional t-copula with lin-*

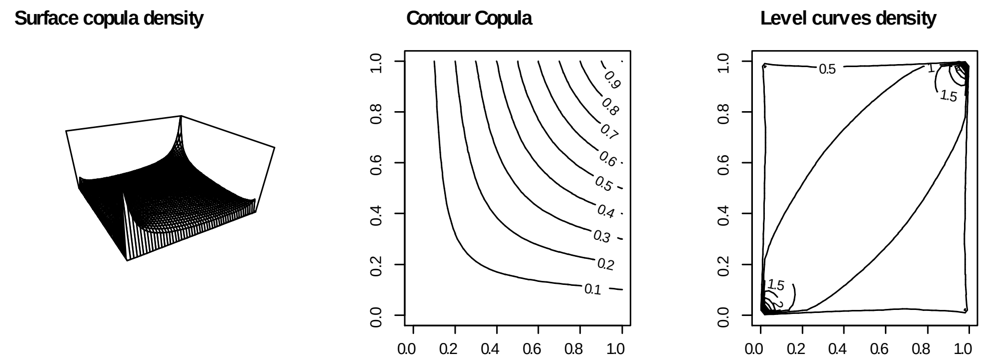

Figure 2.3: The Student's t-copula density and the corresponding density contour and level curves. Here $\rho = 0.71$, $\nu = 3$.

ear correlation matrix Σ and degrees of freedom ν, is defined as:

$$C^t_{\Sigma,\nu}(u_1,...,u_d) = \int_{-\infty}^{t_\nu^{-1}(u_1)} ... \int_{-\infty}^{t_\nu^{-1}(u_d)} \frac{\Gamma((\nu+d)/2)\left(1+\mathbf{w}^T\Sigma^{-1}\mathbf{w}/\nu\right)^{-(\nu+d)/2}}{|\Sigma|^{1/2}\Gamma(\nu/2)(\nu\pi)^{d/2}} dw_1...dw_d,$$

where $|\Sigma|$ stands for the determinant of Σ, $\mathbf{w} = (w_1,...,w_d)$ and $\Gamma(\cdot)$ is the Gamma function.

For more details, see Hult and Lindskog (2002, [81]), and for details about regular variation in general see Resnick (1987, [129]) or Embrechts et *al.* (1997, [42]).

Equation (2.2) provides an easy algorithm for random variate generation from the t-copula, $C^t_{\Sigma,\nu}$.

Algorithm 2.2.2

1. Find the Cholesky decomposition A of Σ.

2. Simulate d independent random variates $z_1,...,z_d$ from $\mathcal{N}(0,1)$.

3. Simulate a random variate s from χ^2_ν independent of $z_1,...,z_d$.

4. Set $\mathbf{y} = A\mathbf{z}$.

5. Set $\mathbf{x} = \frac{\sqrt{\nu}}{\sqrt{s}}\mathbf{y}$.

6. Set $u_i = t_\nu(x_i), i = 1,...,d$.

7. $(u_1,...,u_d) \sim C^t_{\Sigma,\nu}$.

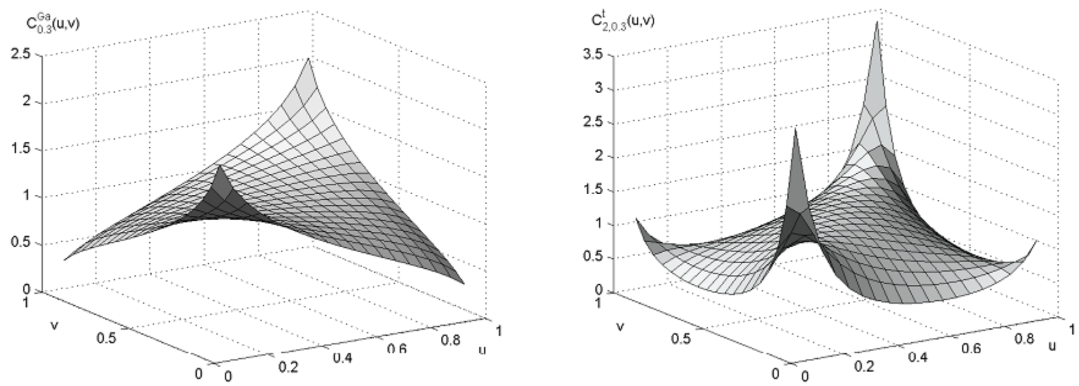

Figure 2.4: The copula densities of a Gaussian copula (left) and a Student t-copula (right). Both copulas have correlation coefficient $\rho = 0.3$ and the t-copula has 2 degrees of freedom.

2.3 Archimedean Copulas

In this Section we focus on a very important class of copulas called Archimedean copulas. The adventages of this class are

- Ease in construction.
- Rich of great variety of families of copulas belonging to this class.
- Nice properties of copula belonging to this class.
- Reduce the study of a multivariate copula to a single univariate function.

The word Archimedean was employed the first time by Ling in 1965 for Archimedean t-norms (every Archimedean copula is also an Archimedean t-norm). And the term «Archimedean copula» was first appeared in the statistical literature in two papers by Genest and Mackay (1986a, [57], 1986b, [58]). Archimedean copulas also appear in Schweizer and Sklar (1983, [136]) but without the name.

For some background on bivariate Archimedean copulas and a discussion on other statistical questions we refer to Genest and MacKay (1986a, [57]), Genest and Rivest (1993, [59]), Joe (1997, [85]), and Nelsen (2006, [123]).

2.3.1 Definitions and properties

We begin with a general definition of Archimedean copulas, which can be found in Nelsen (2006, [123, p. 109]). As our aim is the construction of multivariate extensions of Archimedean 2-copulas, this general definition will later prove to be a bit more general than needed.

Definition 2.3.1 *Let φ be a continuous, strictly decreasing function from $[0,1]$ to $[0,\infty]$ such that $\varphi(1)=0$. The pseudo-inverse of φ is the function $\varphi^{[-1]}:[0,\infty]\to[0,1]$ given by*

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

Note that $\varphi^{[-1]}$ is continuous and decreasing on $[0,\infty]$, and strictly decreasing on $[0,\varphi(0)]$. Furthermore, $\varphi^{[-1]}(\varphi(u))=u$ on $[0,1]$, and

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

Finally, if $\varphi(0)=\infty$, then $\varphi^{[-1]}=\varphi^{-1}$.

Theorem 2.3.1 *Let φ be a continuous, strictly decreasing function from $[0,1]$ to $[0,\infty]$ such that $\varphi(1)=0$, and let $\varphi^{[-1]}$ be the pseudo-inverse of φ. Let C be the function from $[0,1]^2$ to $[0,1]$ given by*

$$C(u,v)=\varphi^{[-1]}(\varphi(u)+\varphi(v)). \tag{2.3}$$

Then C is a copula if and only if φ is convex.

For a proof, see Nelsen (2006, [123, p. 111]).

Copulas of the form (2.3) are called Archimedean copulas. The function φ is called a generator of the copula. If $\varphi(0)=\infty$, we say that φ is a strict generator. In this case, $\varphi^{[-1]}=\varphi^{-1}$ and

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

is said to be a strict Archimedean copula.

The results in the following theorem will enable us to formulate multivariate extensions of Archimedean copulas.

Theorem 2.3.2 *Let C be an Archimedean copula with generator φ. Then*

1. *C is symmetric, i.e. $C(u,v) = C(v,u)$ for all u, v in $[0,1]$.*

2. *C is associative, i.e. $C(C(u,v),w) = C(u,C(v,w))$ for all u, v,w in $[0,1]$.*

Proof. The first part follows directly from (2.3). For 2.,

$$\begin{aligned} C(C(u,v),w) &= \varphi^{[-1]}(\varphi(\varphi^{[-1]}(\varphi(u)+\varphi(v)))+\varphi(w)) \\ &= \varphi^{[-1]}(\varphi(u)+\varphi(v)+\varphi(w)) \\ &= \varphi^{[-1]}(\varphi(u)+\varphi(\varphi^{[-1]}(\varphi(v)+\varphi(w)))) \\ &= C(u,C(v,w)). \end{aligned}$$

■

The associativity property of Archimedean copulas is not shared by copulas in general as shown by the following example.

Example 2.3.1 *Let C_θ be a member of the bivariate Farlie-Gumbel-Morgenstern family of copulas, i.e. $C_\theta(u,v) = uv + \theta uv(1-u)(1-v)$, for $\theta \in [-1,1]$. Then*

$$C_\theta\left(\frac{1}{4}, C_\theta\left(\frac{1}{2},\frac{1}{3}\right)\right) \neq C_\theta\left(\left(C_\theta\frac{1}{4},\frac{1}{2}\right),\frac{1}{3}\right)$$

for all $\theta \in [-1,1]\backslash\{0\}$. Hence the only member of the bivariate Farlie-Gumbel-Morgenstern family of copulas that is Archimedean is Π.

Theorem 2.3.3 *Let C be an Archimedean copula generated by φ and let*

$$K_C(t) = V_C(\{(u,v) \in [0,1]^2 | C(u,v) \leq t\}).$$

Then for any t in $[0,1]$,

$$K_C(t) = t - \frac{\varphi(t)}{\varphi'(t^+)}. \tag{2.4}$$

For a proof, see Nelsen (2006, [123, p. 127]).

Corollary 2.3.1 *If (U,V) has joint distribution function C, where C is an Archimedean copula generated by φ, then the function K_C given by (2.4) is the distribution function of the random variable $C(U,V)$.*

The next theorem will provide the basis for a general algorithm for random variate generation from Archimedean copulas. Before the theorem can be stated we need an expression for the density of an absolutely continuous Archimedean copula. From (2.3) it follows that

$$\begin{aligned}\varphi'(C(u,v))\frac{\partial}{\partial u}C(u,v) &= \varphi'(u),\\ \varphi'(C(u,v))\frac{\partial}{\partial v}C(u,v) &= \varphi'(v),\end{aligned}$$

and

$$\varphi''(C(u,v))\frac{\partial}{\partial u}C(u,v)\frac{\partial}{\partial v}C(u,v) + \varphi'(C(u,v))\frac{\partial^2}{\partial u \partial v}C(u,v) = 0,$$

when C is absolutely continuous, its density is given by

$$\begin{aligned}\frac{\partial^2}{\partial u \partial v}C(u,v) &= -\frac{\varphi''(C(u,v))\frac{\partial}{\partial u}C(u,v)\frac{\partial}{\partial v}C(u,v)}{\varphi'(C(u,v))}\\ &= -\frac{\varphi''(C(u,v))\varphi'(u)\varphi'(v)}{[\varphi(C(u,v))]^3}. \qquad (2.5)\end{aligned}$$

Theorem 2.3.4 *Under the hypotheses of Corollary 2.3.1, the joint distribution function $H(s,t)$ of the random variables $S = \varphi(U)/[\varphi(U)+\varphi(V)]$ and $T = C(U,V)$ is given by $H(s,t) = sK_C(t)$ for all (s,t) in $[0,1]^2$. Hence S and T are independent, and S is uniformly distributed on $[0,1]$.*

Proof. (This proof, for the case when C is absolutely continuous, can be found in Nelsen (2006, [123, p. 104]). For the general case, see Genest and Rivest (1993, [59])). The joint density $h(s,t)$ of S and T is given by

$$h(s,t) = \frac{\partial^2}{\partial u \partial v}C(u,v)\left|\frac{\partial(u,v)}{\partial(s,t)}\right|,$$

where $\partial^2 C(u,v)/\partial u \partial v$ is given by (2.5) and $|\partial(u,v)/\partial(s,t)|$ denotes the Jacobian of the transformation $\varphi(u) = s\varphi(t)$, $\varphi(v) = (1-s)\varphi(t)$. But

$$\frac{\partial(u,v)}{\partial(s,t)} = \frac{\varphi(t)\varphi'(t)}{\varphi'(u)\varphi'(v)},$$

and hence

$$h(s,t) = -\frac{\varphi''(t)\varphi'(u)\varphi'(v)}{[\varphi(t)]^3} - \frac{\varphi(t)\varphi'(t)}{\varphi'(u)\varphi'(v)} = \frac{\varphi''(t)\varphi(t)}{[\varphi'(t)]^2}.$$

Therefore

$$H(s,t) = \int_0^s\int_0^t \frac{\varphi''(y)\varphi(y)}{[\varphi'(y)]^2}dydx = s\left[y - \frac{\varphi(y)}{\varphi'(y)}\right]_0^t = sK_C(t),$$

from which the conclusion follows. ■

An application of Theorem 2.3.4 is the following algorithm for generating random variates (u, v) whose joint distribution is an Archimedean copula C with generator φ.

Algorithm 2.3.1

1. Simulate two independent $U(0,1)$ random variates s and q.
2. Set $t = K_C^{-1}(q)$, where K_C is the distribution function of $C(U,V)$.
3. Set $u = \varphi^{[-1]}(s\varphi(t))$ and $v = \varphi^{[-1]}((1-s)\varphi(t))$.

Note that the variates s and t correspond to the random variables S and T in Theorem 2.3.4 and from the proof it follows that this algorithm yields the desired result.

2.3.2 Kendall's tau of Archimedean Copula

Recall that Kendall's tau for a copula C can be expressed as a double integral of C. This double integral is in most cases not straightforward to evaluate. However for an Archimedean copula, Kendall's tau can be expressed as an (one-dimensional) integral of the generator and its derivative, as shown in the following theorem from Genest and MacKay (1986a, [57]).

Theorem 2.3.5 *Let X and Y be rv's with an Archimedean copula C generated by φ. Kendall's tau of X and Y is given by*

$$\tau_C = 1 + 4\int_0^1 \frac{\varphi(t)}{\varphi'(t)}dt. \tag{2.6}$$

Proof. Let U and V be $U(0,1)$ rv's with joint df C, and let K_C denote the distribution function of $C(U,V)$. Then from Theorem 1.2.1 we have

$$\begin{aligned}
\tau_C &= 4\mathbb{E}(C(U,V)) - 1 \\
&= 4\int_0^1 t dK_C(t) - 1 \\
&= 4\left([tK_C(t)]_0^1 - \int_0^1 K_C(t)dt\right) - 1 \\
&= 3 - 4\int_0^1 K_C(t)dt.
\end{aligned}$$

From Theorem 2.3.3 and Corollary 2.3.1 it follows that $K_C(t) = t - \frac{\varphi(t)}{\varphi'(t^+)}$. Since φ is convex, $\varphi(t^+)$ and $\varphi(t^-)$ exist for all t in $(0,1)$ and the set $\{t \in (0,1) | \varphi(t^+) = \varphi(t^-)\}$ is at most countable (i.e. it has Lebesgue measure zero). Hence

$$\begin{aligned} \tau_C &= 3 - 4\int_0^1 \left(t - \frac{\varphi(t)}{\varphi'(t^+)}\right) dt \\ &= 1 + 4\int_0^1 \frac{\varphi(t)}{\varphi'(t)} dt. \end{aligned}$$

■

2.3.3 Tail dependence of Archimedean Copula

The following Theorem give an expression of the tail dependence for Archimedean copulas, in terms of the generators.

Theorem 2.3.6 *Let φ be a strict generator such that φ^{-1} belongs to the class of Laplace transforms of strictly positive random variables. If $\varphi^{-1\prime}(0)$ 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 $\varphi^{-1\prime}(0) = -\infty$ and the coefficient of upper tail dependence is given by

$$\lambda_U = 2 - 2 \lim_{s \searrow 0} \left[\frac{\varphi^{-1\prime}(2s)}{\varphi^{-1\prime}(s)} \right].$$

Proof. See Joe (1997, [85, p. 103]). ■

The additional condition on the generator φ might seem somewhat strange. It will however prove quite natural when we turn to the construction of multivariate Archimedean copulas. Furthermore, the condition is satisfied by the majority of the commonly encountered Archimedean copulas.

Theorem 2.3.7 *Let φ be as in Theorem 2.3.6. The coefficient of lower tail dependence for the copula C is equal to*

$$\lambda_L = 2 \lim_{s \to \infty} \left[\frac{\varphi^{-1\prime}(2s)}{\varphi^{-1\prime}(s)} \right].$$

The proof is similar to that of Theorem 2.3.6.

2.3.4 Examples of Archimedean Copulas

Clayton copula

The Clayton (1978) copula, also referred to as the Cook and Johnson (1981) copula, originally studied by Kimeldorf and Sampson (1975), takes the form:

$$C(u,v) = (u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}, \ \theta \in [-1, \infty)\backslash\{0\},$$

with the dependence parameter θ restricted on the region $(0, \infty)$. The limiting case $\theta = 0$ corresponds to the independent case, i.e. $C = \Pi^2$. In particular, this family is positively ordered, and its members are absolutely continuous for $\theta > 0$.

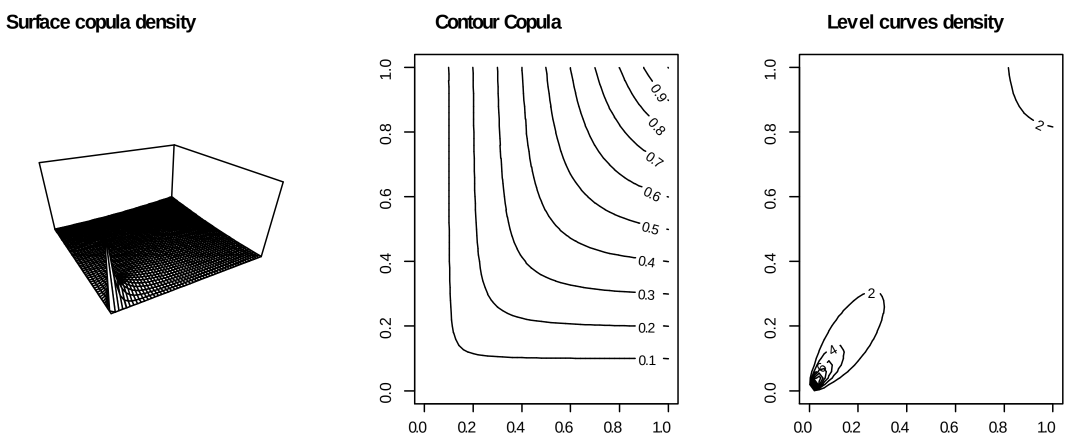

Figure 2.5: The Clayton 2-copula density and the corresponding density contour and level curves. Here the parameter is $\theta = 2$.

This strict copula family has generator

$$\varphi(t) = (t^{-\theta} - 1)/\theta.$$

It follows that

$$\varphi^{-1}(s) = (1 + \theta s)^{-1/\theta}.$$

Using Theorem 2.3.6 and 2.3.7 shows that $\lambda_U = 0$ and that the coefficient of lower tail dependence given by

$$\begin{aligned}\lambda_L &= 2\lim_{s\to\infty}[\varphi^{-1\prime}(2s)/\varphi^{-1\prime}(s)]\\ &= 2\lim_{s\to\infty}\left[\frac{(1+2\theta s)^{-1/\theta-1}}{(1+\theta s)^{-1/\theta-1}}\right]\\ &= 2\times 2^{-1/\theta-1}\\ &= 2^{-1/\theta}.\end{aligned}$$

As approaches zero, the marginals become independent. As approaches infinity, the copula attains the Fréchet upper bound, but for no value does it attain the Fréchet lower bound. The Clayton copula cannot account for negative dependence. It has been used to study correlated risks because it exhibits strong left tail dependence and relatively weak right tail dependence.

Using Theorem 2.3.5 we can calculate Kendall's tau for the Clayton family

$$\begin{aligned}\tau_\theta &= 1+4\int_0^1\frac{t^{\theta+1}-t}{\theta}dt\\ &= 1+\frac{4}{\theta}\left(\frac{1}{\theta+2}-\frac{1}{2}\right)\\ &= \frac{\theta}{\theta+2}.\end{aligned}$$

which may provide a way to fit a Clayton 2-copula to the available data.

Gumbel family

The Gumbel copula (1960) takes the form:

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

The dependence parameter is restricted to the interval $[1,\infty)$. Values of 1 and ∞ correspond to independence and the Fréchet upper bound, but this copula does not attain the Fréchet lower bound for any value of θ. The Gumbel copulas are strict Archimedean with generator

$$\varphi(t)=(-\ln t)^\theta.$$

Hence $\varphi^{-1}(s)=\exp(-s^{1/\theta})$ and

$$\varphi^{-1\prime}(s)=-s^{1/\theta-1}\exp(-s^{1/\theta})/\theta.$$

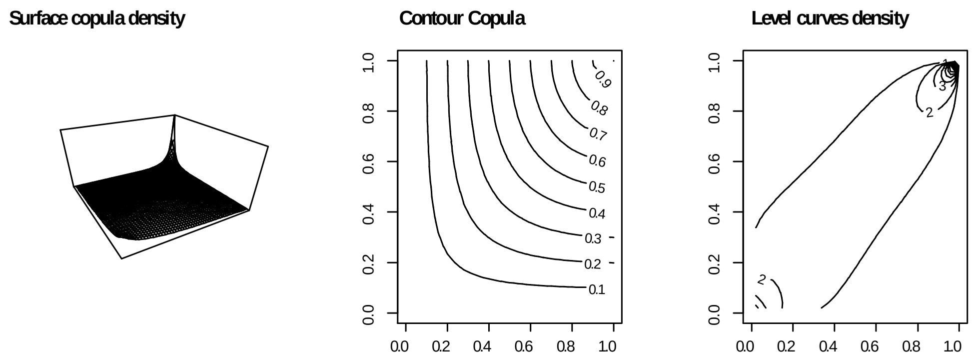

Figure 2.6: The Gumbel 2-copula density and the corresponding density contour and level curves. Here the parameter is $\theta = 2$.

Using Theorem 2.3.6 we get

$$\begin{aligned}\lambda_U &= 2 - 2\lim_{s\searrow 0}[\varphi^{-1\prime}(2s)/\varphi^{-1\prime}(s)] \\ &= 2 - 2^{1/\theta}\lim_{s\searrow 0}\left[\frac{\exp(-(2s)^{1/\theta})}{\exp(-s^{1/\theta})}\right] \\ &= 2 - 2^{1/\theta}.\end{aligned}$$

Similar to the Clayton copula, Gumbel does not allow negative dependence, but it contrast to Clayton, Gumbel exhibits strong right tail dependence and relatively weak left tail dependence. If outcomes are known to be strongly correlated at high values but less correlated at low values, then the Gumbel copula is an appropriate choice.

Using Theorem 2.3.5 we can calculate Kendall's tau for the Gumbel family

$$\begin{aligned}\tau_\theta &= 1 + 4\int_0^1 \frac{t\ln t}{\theta}dt \\ &= 1 + \frac{4}{\theta}\left(\left[\frac{t^2}{2}\ln t\right]_0^1 - \int_0^1 \frac{t}{2}dt\right) \\ &= 1 + \frac{4}{\theta}(0 - 1/4) \\ &= 1 - 1/\theta.\end{aligned}$$

As a consequence, in order to have Kendall's tau equal to 0.5, we put $\theta = 2$.

Frank family

Consider the Frank family given by

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

for $\theta \in \mathbb{R}\backslash\{0\}$. This strict copula family has generator

$$\varphi(t) = -\ln\frac{e^{-\theta t}-1}{e^{-\theta}-1}.$$

It follows that

$$\varphi^{-1}(s) = -\frac{1}{\theta}\ln\left(1-(1-e^{-\theta})e^{-s}\right),$$

and

$$\varphi^{-1\prime}(s) = -\frac{1}{\theta}\frac{(1-e^{-\theta})e^{-s}}{1-(1-e^{-\theta})e^{-s}}.$$

Since

$$\varphi^{-1\prime}(0) = -\frac{e^{\theta}-1}{\theta}$$

is finite, the Frank family does not have upper tail dependence according to Theorem 2.3.6. Furthermore, members of the Frank family are radially symmetric, i.e. $C = \overline{C}$, and hence the Frank family does not have lower tail dependence.

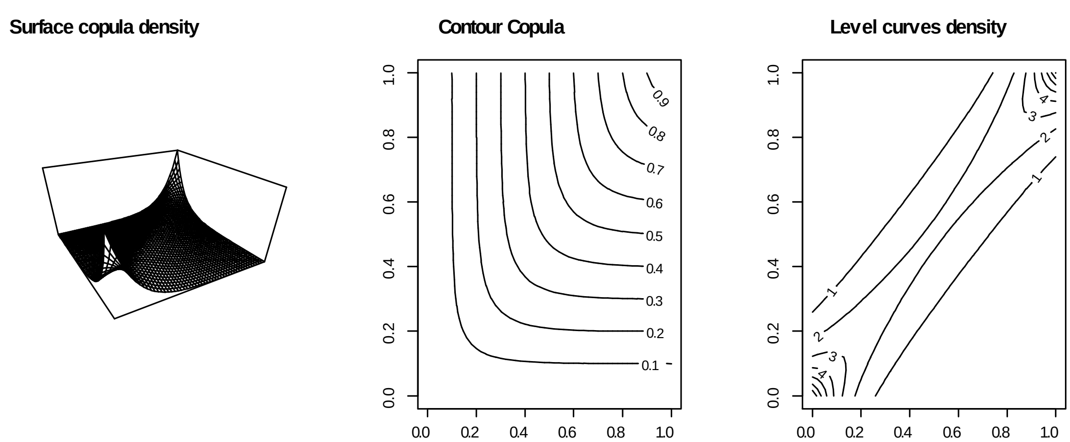

Figure 2.7: The Frank 2-copula density and the corresponding density contour and level curves. Here the parameter is $\theta = 12.825$.

It can be shown that (see, e.g. Genest, 1987, [56]) Kendall's tau is

$$\tau_\theta = 1 - \frac{4(1 - D_1(\theta))}{\theta},$$

where $D_k(x)$ is the Debye function, given by

$$D_k(x) = \frac{k}{x^k} \int_0^x \frac{t^k}{e^t - 1} dt,$$

for any positive integer k.

2.3.5 Simulation

In this section we present two algorithms to generate an observation (u, v) from an Archimedean copula C with generator φ.

Algorithm 2.3.1

1. Generate two independent uniform $(0, 1)$ variates s and t.
2. Set $w = K^{(-1)}(t)$, where
$$K(t) = t - \frac{\varphi(t)}{\varphi'(t^+)}$$
and $K^{(-1)}(t) = \sup\{x | K(t) \leq x\}$.
3. Set $w = \varphi^{[-1]}(s\varphi(w))$ and $v = \varphi^{[-1]}((1 - s)\varphi(w))$.

Algorithm 2.3.1 is a consequence of the fact that if U and V are uniform random variables with an Archimedean copula C, then $W = C(U, V)$ and $S = \varphi(U)/(\varphi(U) + \varphi(V))$ are independent, S is uniform $(0, 1)$, and the distribution function of W is K (Genest and Rivest, 1993, [59]).

Algorithm 2.3.2

1. Generate two independent uniform $(0, 1)$ variates u and t.
2. Set $w = \varphi'^{(-1)}\left(\frac{\varphi'(u)}{t}\right)$.
3. Set $v = \varphi^{[-1]}(\varphi(w) - \varphi(u))$.

Algorithm 2.3.2 is the «conditional distribution function» method, where $v = c_u^{(-1)}(t)$ for

$$c_u(t) = \frac{\partial C(u, v)}{\partial u} = \mathbb{P}[V \leq v | U = u].$$

2.4 Copulas with two dependence parameters

In many instances, a single parameter does not provide sufficient flexibility for modeling purposes. In this section we discuss methods to add a parameter.

Although they are rarely used in empirical applications, it is possible to construct Archimedean copulas with two dependence parameters, each of which measures a different dependence feature. For example, one parameter might measure left tail dependence while the other might measure right tail dependence. The bivariate Student's t-distribution was mentioned earlier as an example of a two-parameter copula. Transformation copulas in Section 2.3 are a second example. Two parameter Archimedean copulas take the form:

$$C(u,v) = \varphi\left(-\ln \mathbb{K}\left(e^{-\varphi(u)}, e^{-\varphi(v)}\right)\right), \tag{2.7}$$

where $\mathbb{K}$ is max-stable and φ is a Laplace transformation. If $\mathbb{K}$ assumes an Archimedean form, and if $\mathbb{K}$ has dependence parameter θ_1, and φ is parameterized by θ_2, then $C(u,v;\theta_1,\theta_2)$ assumes an Archimedean form with two dependence parameters. Joe (1997, [85]) discusses this calculation in more detail. As an example, we present one Archimedean copula with two parameters, and we direct the interested reader to Joe (1997, [85]) for more examples. If $\mathbb{K}$ assumes the Gumbel form, then (2.7) takes the form:

$$C(u,v;\theta_1,\theta_2) = \varphi\left(\varphi^{-1}(u) + \varphi^{-1}(v)\right),$$

where $\varphi(t) = (1 + t^{1/\theta_1})^{-1/\theta_2}$, $\theta_2 > 0$ and $\theta_1 \geq 1$. For this copula, $2^{-1/(\theta_1\theta_2)}$ measures lower tail dependence and $2 - 2^{1/\theta_1}$ captures upper tail dependence.

2.4.1 Interior and exterior power

In this section we discuss methods to add a parameter at Archimedean copula withe single parameter. Let φ_θ be a generator of an Archimedean copula, and define $\varphi_{\theta,\alpha}(t) = \varphi_\theta(t^\alpha)$ and $\varphi_{\theta,\beta}(t) = [\varphi(t)]^\beta$. Then:

1. $\varphi_{\theta,\beta}$ is a generator for all $\beta \geq 1$,
2. $\varphi_{\theta,\alpha}$ is a generator for all α in $(0,1]$,
3. If φ_θ is twice differentiable and $\varphi_\theta(t)$ is nondecreasing on $(0,1)$, then $\varphi_{\theta,\alpha}$ is a generator for all $\alpha > 0$.

The set $\{\varphi_{\theta,\alpha}\}$ is the interior power family of generators associated with φ_θ and $\{\varphi_{\theta,\beta}\}$ is the exterior power family of generators associated with φ_θ (see, Oakes, 1994, [126]). We note that we can create a three-parameter family of generators from a single generator φ_θ by $\varphi_{\theta,\alpha,\beta}(t) = [\varphi_\theta(t^\alpha)]^\beta$ for appropriate values of α and β.

Example 2.4.1 *The single Gumbel-Hougaard Copula is given by*

$$C_{\beta_1}(u,v) = \exp\left\{-\left[(-\ln u)^{\beta_1} + (-\ln v)^{\beta_1}\right]^{1/\beta_1}\right\},$$

withe generator $\varphi_{\beta_1} = (-\ln t)^{\beta_1}$. *This copula has more probability concentrated in the tails. It is also asymmetric, with more weight in the right tail (Nelsen 2006, [123]). The dependence parameter* β_1 *is restricted to the interval* $[1,\infty)$. *Let* $\psi_{\beta_1,\beta_2}(t) = \left(-\ln t^{\beta_2}\right)^{\beta_1}$, *the interior power family of Gumbel-Hougaard Copula generators, then for* β_1 *in* $[1,\infty)$ *and* β_2 *in* $(0,1]$

$$C_{\beta_1,\beta_2}(u,v) = \left(\exp\left(-\left(\left(-\ln u^{\beta_2}\right)^{\beta_1} + \left(-\ln v^{\beta_2}\right)^{\beta_1}\right)^{1/\beta_1}\right)\right)^{1/\beta_2}.$$

Note that the exterior power family of generators φ_{β_1} *is the single Gumbel-Hougaard Copula withe parameter* $\beta_1\beta_2$.

Example 2.4.2 *The single Clayton Copula is given by:*

$$C_{\gamma_1}(u,v) = \left(\max\left\{\left(u^{-\gamma_1} + v^{-\gamma_1} - 1\right), 0\right\}\right)^{-1/\gamma_1},$$

withe generator $\varphi_{\gamma_1}(t) = (t^{-\gamma_1} - 1)/\gamma_1$, *and* γ_1 *in* $[-1,\infty)\setminus\{0\}$. *The Clayton copula is an asymmetric Archimedean. copula, exhibiting greater dependence in the left tail than in the right (Salvadori et al. 2007). The exterior power family of generators associated with* φ_{γ_1} *is given by* $\psi_{\gamma_1,\gamma_2}(t) = \left(\varphi_{\gamma_1}(t)\right)^{\gamma_2}$. *Then a two parameter Clayton copula is given for* $\gamma_2 \geq 1$ *by*

$$C_{\gamma_1,\gamma_2}(u,v) = \left(\max\left\{\left(\left(\left(u^{-\gamma_1} - 1\right)^{\gamma_2} + \left(v^{-\gamma_1} - 1\right)^{\gamma_2}\right)^{1/\gamma_2} + 1\right), 0\right\}\right)^{-1/\gamma_1}.$$

Note that the interior power family of Clayton copula is a single Clyton copula with parameter $\gamma_1\gamma_2$.

Example 2.4.3 *(Fang et al., 2000)* $\varphi_\theta(t) = \ln([1-\theta(1-t)]/t)$ *generates the Ali-Mikhail-Haq family (see Table 4.1 of Nelsen 2006, [123]), and since* $t\varphi'_\theta(t)$ *is nondecreasing for* θ *in* $[0,1]$, *the interior power family of copulas associated with* φ_θ *is, for*

$u,\ v,\ \theta\ in\ [0,1],\ \alpha > 0$

$$C_{\theta,\alpha,1}(u,v) = \frac{uv}{[1-\theta(1-u^{1/\alpha})(1-v^{1/\alpha})]^{\alpha}}.$$

2.4.2 Method of nesting

Archimedean copulas can be extended to include additional marginal distributions. We Focus on the trivariate case, it easy to include a third marginal by

$$C(u,v,w) = \varphi\left(\varphi^{-1}(u) + \varphi^{-1}(v) + \varphi^{-1}(w)\right). \tag{2.8}$$

This construction can be readily used in empirical applications, but it is necessary to assume that φ^{-1} is completely monotonic (Cherubini et *al.*, 2004, [20, p. 149]). The forme (2.8) implies symmetric dependence between the three pairs (u,v), (v,w), and (u,w), due to having a single dependence parameter. This restriction becomes more onerous as the number of marginals increases. It is not possible to model separately the dependence between all pairs.

The functional form of an Archimedean copula will be recognized by those familiar with the theory of separable and additively separable functions. Note that many Archimedean copulas are additively separable[1].

Remark 2.4.1 *A function may be separable but not additively separable.*

Under separability variables can be nested. For example if $d = 3$, then the following groupings are possible: $(u,v,w;\theta)$, $(u,[v,w;\theta_2];\theta_1)$, $(v,[u,w;\theta_2];\theta_1)$ and $(w,[u,v;\theta_2];\theta_1)$.

When fitting copulas to data, the alternative groupings have different implications and interpretations. Presumably each grouping is justified by some set of assumptions about dependence. The first grouping restricts the dependence parameter θ to be the same for all pairs. The remaining three groupings allow for two dependence parameters, first one, θ_1, for a pair and a second one, θ_2, for dependence between the singleton and the pair.

The existence of generators ϕ^t that lead to flexible forms of Archimedean copulas seems to be an open question. Certain types of extensions to multivariate copulas

A function $f(\mathbf{u})$ is (weakly) separable if it can be written as $f(\mathbf{u}) = \phi\{\phi^1(u_1),...,\phi^k(u_k)\}$,and additively separable if $f(\mathbf{u}) = \phi\left\{\sum_{t=1}^{k}\phi^t(\mathbf{u}_t)\right\}$,where $(\mathbf{u}_1,...,\mathbf{u}_t)$ is a separation of the set of variables $(u_1,u_2,...,u_d)$ into t nonoverlapping groups.

are not possible. For example, Genest et al. (1995, [60]) considered a copula C such that

$$H(x_1, ..., x_i, y_1, ..., y_j) = C(F(x_1, ..., x_i), G(y_1, ..., y_j))$$

defines a $(i+j)$-dimensional distribution function with marginals F and G, $i+j \geq 3$. They found that the only copula consistent with these marginals is the independence copula. Multivariate Archimedean copulas with a single dependence parameter can be obtained if restrictions are placed on the generator. For multivariate generalizations of Gumbel, Frank and Clayton (see Cherubini et *al.*, 2004, [20, p. 150-151]). Copula densities for dimensions higher than 2 are tedious to derive; however, Cherubini et *al.* 2004, [20, Section 7.5] gives a general expression for the Clayton copula density, and for the Frank copula density for the special case of four variables.

We exploit the mixtures of powers method to extend Archimedean copulas to include a third marginal. For a more detailed exposition of this method, see Joe (1997, [85, CH. 5]) and Zimmer and Trivedi (2006). The trivariate mixtures of powers representation is

$$C(u, v, w) = \int_0^\infty \int_0^\infty G^\beta(u) G^\beta(v) dM_2(\beta; \alpha) G^\alpha(w) dM_1(\alpha), \tag{2.9}$$

where $G(u) = \exp(-\phi^{-1}(u))$, $G(v) = \exp(-\phi^{-1}(v))$, $G(w) = \exp(-\varphi^{-1}(w))$, and φ is a Laplace transformation. In this formulation, the power term α affects u, v, and w, and a second power term β affects u and v. The distribution M_1 has Laplace transformation $\varphi(\cdot)$, and M_2 has Laplace transformation $(\varphi^{-1} \circ \phi)^{-1}(-\alpha^{-1} log(\cdot))^{-1}$. When $\phi = \varphi$, expression (2.9) simplifies to expression (2.8). (The mathematical notation $f \circ g$ denotes the functional operation $f(g(x))$).
When $\phi = \varphi$, the trivariate extension of Archimedean copula corresponding to (2.9) is

$$C(u, v, w) = \varphi\left(\varphi^{-1} \circ \phi[\phi^{-1}(u) + \phi^{-1}(v)] + \varphi^{-1}(w)\right). \tag{2.10}$$

Therefore, different Laplace transformations produce different families of trivariate copulas. Expression (2.8) has symmetric dependence in the sense that it produces one dependence parameter $\theta = \theta_{uv} = \theta_{uw} = \theta_{vw}$.

But the dependence properties of three different marginals are rarely symmetric in empirical applications. The trivariate representation of expression (2.10) is symmetric with respect to (u, v) but not with respect to w. Therefore, (2.10) is less restrictive than (2.8). The partially symmetric formulation of expression (2.10) yields two dependence parameters, θ_1 and θ_2, such that $\theta_1 \leq \theta_2$. The parameter $\theta_2 = \theta_{uv}$ measures dependence between u and v. The parameter $\theta_1 = \theta_{uw} = \theta_{vw}$ measures dependence between u and w as well as between v and w, and the two must be equal. Distributions greater than three dimensions also have a mixtures of powers representations, but this technique yields only $d-1$ dependence parameters for an d-variate distribution function. Therefore, the mixtures of powers approach is more restrictive for higher dimensions. While this restriction constitutes a potential weakness of the approach, it is less restrictive than formulation (2.8) which yields only one dependence parameter. Moreover, the multivariate representation in Eq. (2.10) allows a researcher to explore several dependence patterns by changing the ordering of the marginals. For example, instead of (u, v, w), one could order the marginals (w, v, u), which provides a different interpretation for the two dependence parameters.

As an example, we demonstrate how the Frank copula is extended to include a third marginal. If $\phi(s) = exp(-s^{1/\theta})$ and $(\varphi^{-1} \circ \phi)(s) = s^{\theta_1/\theta_2}$, then expression (2.10) becomes

$$\begin{aligned} & C(u, v, w; \theta_1, \theta_2) \\ = \; & -\theta_1 \ln \left\{ 1 - \frac{(1-e^{-\theta_1 w})}{1-e^{-\theta_1}} \left(1 - \left[1 - \frac{(1-e^{-\theta_2 u})(1-e^{-\theta_2 v})}{1-e^{-\theta_2}} \right]^{\theta_1/\theta_2} \right) \right\}, \end{aligned} \tag{2.11}$$

where $\theta_1 \leq \theta_2$. (The proof is complicated; see Joe (1993, [84])). Despite the ability of some bivariate Archimedean copulas to accommodate negative dependence, trivariate Archimedean copulas derived from mixtures of powers restrict θ_1 and θ_2 to be greater than zero, which implies positive dependence. This reflects an important property of mixtures of powers. In order for the integrals in Eq. (2.9) to have closed form solutions, then power terms α and β, which are imbedded within θ_1 and θ_2, must be positive.

2.4.3 Distortion of copula

In this section, we limit ourselves to the case of 2-dimensional copula. Let C be a copula. Given a bijection $\Gamma : [0,1] \to [0,1]$, we can now define $C^\Gamma : [0,1]^2 \to [0,1]$ by

$$C^\Gamma(u, v) = \Gamma^{-1}(C(\Gamma(u), \Gamma(v))). \tag{2.12}$$

We can find necessary and sufficient conditions for C^Γ being a copula by introducing strong conditions for Γ.

Theorem 2.4.1 *Assume that* Γ *is a concave,* $\mathcal{C}^1$*-diffeomorphism from* $]0,1[$ *to* $]0,1[$, *twice differentiable and continuous from* $[0,1]$ *to* $[0,1]$, *such that* $\Gamma(0) = 0$ *and* $\Gamma(1) = 1$, *then* C^Γ *is a copula.*

Theorem 2.4.2 *Under the assumptions of Theorem 2.4.1, the function* C^Γ *is a copula if and only if*

$$\frac{\partial^2 C}{\partial u \partial v}(u,v) \geq \frac{\Gamma''\left(\Gamma^{-1}\left(C\left(u,v\right)\right)\right)}{\left[\Gamma'\left(\Gamma^{-1}\left(C\left(u,v\right)\right)\right)\right]^2} \frac{\partial C}{\partial u}(u,v) \frac{\partial C}{\partial v}(u,v),$$

for every (u,v) *where the derivatives of* C *exist.*

Proof. It is sufficient to prove the second result because in the case where Γ is concave, we trivially have

$$\frac{\Gamma''\left(\Gamma^{-1}\left(C\left(u,v\right)\right)\right)}{\left[\Gamma'\left(\Gamma^{-1}\left(C\left(u,v\right)\right)\right)\right]^2} \frac{\partial C}{\partial u}(u,v) \frac{\partial C}{\partial v}(u,v) \leq 0$$

and therefore the condition is match as soon as C is a copula. In order to prove the second Theorem, assume that C^Γ is a copula. By definition, we have

$$\Gamma\left(C^\Gamma(u,v)\right) = (C(\Gamma(u),\Gamma(v))),$$

and by doing $(u,v) = (0,0)$ and $(u,v) = (1,1)$, we obtain the necessary condition that $\Gamma(0) = 0$ and $\Gamma(1) = 1$. The continuity of Γ and the assuption that Γ be a $\mathcal{C}^1$-diffeomorphism imply that $\Gamma' > 0$. At every (u,v)

$$\begin{aligned}\frac{\partial^2 C^\Gamma}{\partial u \partial v}(u,v) &= \frac{\Gamma'(u)\,\Gamma'(v)}{\Gamma'\left(\Gamma^{-1}\left(C\left(\Gamma(u),\Gamma(v)\right)\right)\right)} \times \left[\frac{\partial^2 C}{\partial u \partial v}\left(\Gamma(u),\Gamma(v)\right)\right. \\ &\left. - \frac{\Gamma''\left(\Gamma^{-1}\left(C\left(\Gamma(u),\Gamma(v)\right)\right)\right)}{\left[\Gamma'\left(\Gamma^{-1}\left(C\left(\Gamma(u),\Gamma(v)\right)\right)\right)\right]^2} \frac{\partial C}{\partial u}\left(\Gamma(u),\Gamma(v)\right) \frac{\partial C}{\partial v}\left(\Gamma(u),\Gamma(v)\right)\right].\end{aligned}$$

The condition is then necessary. On the other hand, by integrating the cross-derivative between $0 < u_1 \leq u_2 < 1$ and $0 < v_1 \leq v_2 < 1$, we get

$$C^\Gamma(u_2,v_2) - C^\Gamma(u_1,v_2) - C^\Gamma(u_2,v_1) + C^\Gamma(u_1,v_1) \geq 0. \tag{2.13}$$

Since $\Gamma(0) = 0$ and $\Gamma(1) = 1$, we obtain

$$\begin{aligned} C^{\Gamma}(u,0) &= 0, \\ C^{\Gamma}(u,1) &= u, \\ C^{\Gamma}(0,v) &= 0, \\ C^{\Gamma}(1,v) &= v, \end{aligned}$$

this shows us that inequality (2.13) holds for evry $0 < u_1 \leq u_2 < 1$ and $0 < v_1 \leq v_2 < 1$, and this completes the proof ■

Example 2.4.4 ***(Frank copula).*** *The Frank copula has the following form*

$$C(u,v) = -\frac{1}{\theta}\ln\left\{\frac{(1-e^{-\theta}) - (1-e^{-\theta u})(1-e^{-\theta v})}{1-e^{-\theta}}\right\}.$$

Assume that $\Gamma(x) = x^{1/\beta}$ *with* $\beta \geq 1$. *verifies the Theorem 2.4.1 and it comes that* C^{Γ} *is a copula. The corresponding density function is then*

$$\begin{aligned} c^{\Gamma}(u,v) &= \frac{\bar{u}\bar{v}C^{\Gamma}(u,v)\,e^{-\theta(\bar{u}+\bar{v})}}{\beta uv\left(\bar{C}^{\Gamma}(u,v)\right)^2}\left(\frac{1}{(1-e^{-\theta})\exp\left(-\theta\bar{C}^{\Gamma}(u,v)\right)}\right)^2 \times \\ &\quad \left(\theta(1-e^{-\theta})\bar{C}^{\Gamma}(u,v) - (1-\beta)(1-e^{-\theta\bar{u}})(1-e^{-\theta\bar{v}})\right), \end{aligned}$$

with

$$C^{\Gamma}(u,v) = \left(-\frac{1}{\theta}\ln\left\{\frac{(1-e^{-\theta}) - (1-e^{-\theta\bar{u}})(1-e^{-\theta\bar{v}})}{1-e^{-\theta}}\right\}\right)^{\beta}.$$

where $\bar{u} = u^{1/\beta}$, $\bar{v} = v^{1/\beta}$ *and* $\bar{C}^{\Gamma}(u,v) = \left(C^{\Gamma}(u,v)\right)^{1/\beta}$. *In Figure 2.8, we represented the density contours of different multivariate distributions generated by Frank Copula. We have set* $\theta = 5.736$ *(the corresponding Kendall's tau is then equal to* 0.5*). When the margins are uniform (left and top quadrant), we obtain directly the density of the copula. In the oyher quadrant, the margins are Gaussian, Student or* α*-stable distributions.*

In Figure 2.9 and 2.10 we give the contour plots of transformed copula in the cases $\beta = 3$ and $\beta = 7$. We remark clearly that this transform function has an important impact on the dependence structure. Note moreover that the transformed copula belongs to two-parameter family.

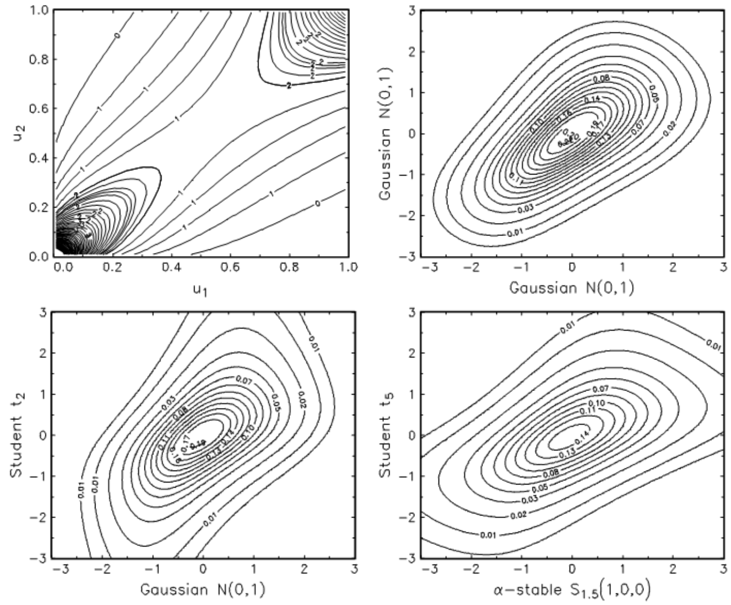

Figure 2.8: Countours of density for Frank copula with $\theta = 5.736$.

2.5 Choosing the right copula

To assess the fit of copula in the dependency structure of a sample, we can use graphical tools include:

1. Empirical comparison of densities (estimated nonparametrically from the sample by kernel method) with theoretical 3-dimensional or as contour lines.

2. The dependogramme.

3. Kendall-plot or K-plot (see Genest and Boies, 2003, [55]).

These graphs can be specially observed if the tail dependence is present in the data and if it is well modeled by the copula chosen.

2.5.1 Empirical comparison of densities

In the first stage, we estimate marginals by

$$\hat{F}_n(x) = \frac{1}{n}\sum_{i=1}^{n} K\left(\frac{x - X_i}{b_{n1}}\right), \quad \hat{G}_n(Y_i) = \frac{1}{n}\sum_{i=1}^{n} K\left(\frac{y - Y_i}{b_{n2}}\right)$$

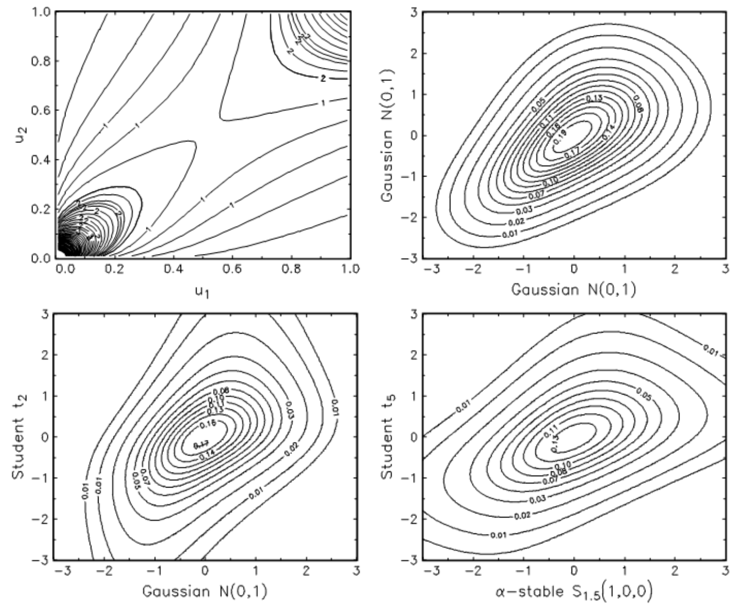

Figure 2.9: Contours of density for the transformed Frank copula withe $\theta = 5.736$ and $\beta = 3$.

with a bandwidth b_{n1} and b_{n2}, see Bowman, Hall and Prvan (1998) for more details on this kernel distribution function estimator. and K the integral of a symmetric bounded kernel function k supported on $[-1, 1]$. In the second stage, the pseudo-observations $\hat{U}_i = \hat{F}_n(X_i)$ and $\hat{V}_i = \hat{G}_n(Y_i)$ are used to estimate the joint distribution function of the unobserved $F(X_i)$ and $G(Y_i)$, which gives the estimate of the unknown copula C. To prevent boundary bias, Chen and Huang (2007, [19]) suggested using a local linear version of the kernel k given by

$$k_{u,h}(x) = \frac{k(x)(a_1 - a_2 x)}{a_0 a_2 - a_1^2} \mathbf{1}_{\left\{\frac{u-1}{h} < x < \frac{u}{h}\right\}},$$

a local linear version of K,to smooth at a $u \in [0, 1]$ with a bandwidth $h > 0$, where

$$a_i = \int_{(u-1)/h}^{u/h} t^i k(t)\, dt, \text{ for } i = 0, 1, 2.$$

Finally, the local linear type estimator of the copula density is given by

$$\hat{c}_h(u, v) = \frac{1}{Th^2} \sum_{i=1}^{T} K_{u,h_n}\left(\frac{u - \hat{U}_i}{h}\right) K_{v,h_n}\left(\frac{v - \hat{V}_i}{h}\right),$$

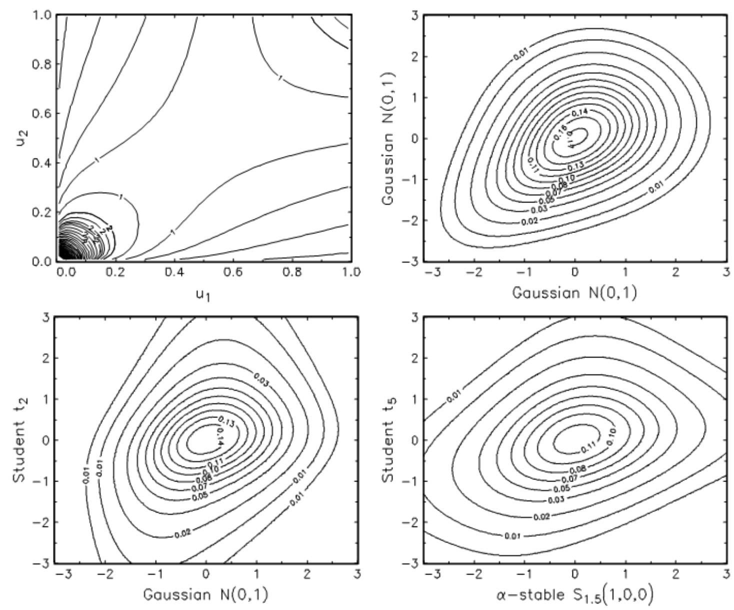

Figure 2.10: Contours of density for the transformed Frank copula withe $\theta = 5.736$ and $\beta = 7$.

where $K_{u,h}(x) = \int_{-\infty}^{x} k_{u,h}(s)ds$, and we compare the empirical estimator with the theoretical 3-dimensional densities see Figure 2.11.

2.5.2 Dependogramme

The dependogramme represents the dependence structure in the form of scatter uniform margins (u, v) extracted from the sample or simulation of a copula theory. Note that the pairs (u, v) from the sample built the empirical copula. It is simply defined by the rank statistics from the sample. We compare the empirical dependogramme copula with the other theoretical copula estimated on the sample.

The dependogramme can also observe the more or less simultaneous creations of the sample. More accurately, in the tails, it is useful to analyze whether simultaneity is high and therefore it is necessary to calibrate our sample to a copula with tail dependence see Figure 2.12.

2.5.3 Kendall plot

The Kendall-plot allow more direct comparison between the empirical copula and copula theory. The algorithm of constructing a Kendall-plot is as follows:

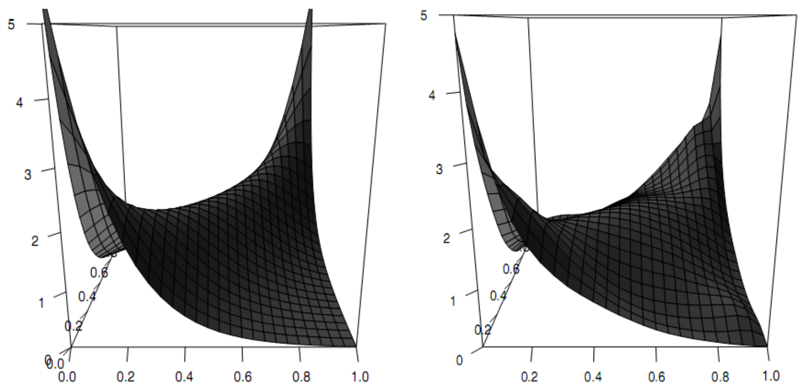

Figure 2.11: (Left) The density of the Frank copula with a $\tau = 0.5$. (Rihgt) Estimation of the copula density using a Gaussian kernel and Gaussian transformations with 1,000 observations drawn from the Frank copula.

1. Compute the function H_i for each pair of a ranks observations of the sample (u_i, v_i)
$$H_i = \frac{1}{n-1} card\{j \neq i : u_j \leq u_i, v_j \leq v_i\},$$
for $1 \leq i \leq n$, where n is the length of the sample. They must then be ordered for $H_{(1)} \leq ... \leq H_{(n)}$. We thus obtain the empirical part of K-plot from the sample.

2. It should be compared with the theoretical copula, For this, we determine $H_{i:n}^{th} = \mathbb{E}\left(H_{(i)}^{th}\right)$, where
$$H_{(i)}^{th} = \frac{1}{n-1} card\left\{j \neq i : u_j^{th} \leq u_i^{th}, v_j^{th} \leq v_i^{th}\right\}.$$
for each pair (u_i^{th}, v_i^{th}) from the copula theory. We determine these couples using Monte Carlo simulations of the copula theory. Thus, we simulate n realizations of the theoretical copula m times, determining $H_{(i)}^{th}$ for each m-th simulation, which is then ordered, and finally we calculate the expectation of m $H_{(i)}^{th}$ for each i. And we obtain $H_{i:n}^{th}$.

3. It remains to graphing pairs $\left(H_{(i)}, H_{i:n}^{th}\right)$ to obtain the K-plot

More Kendall-plot approximates a straight line, plus the adjustment between the dependency structure of the sample and the estimated copula on the same sample is good see Figure 2.13.

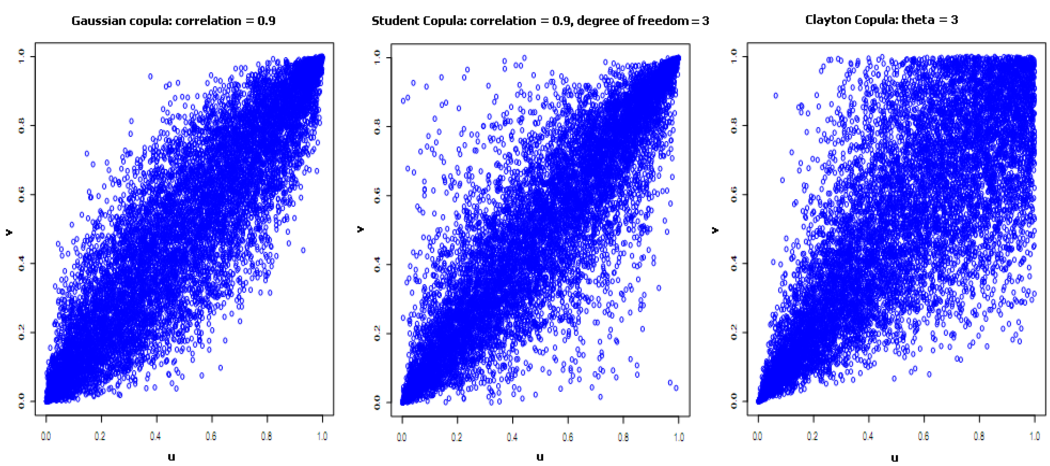

Figure 2.12: Dependogrammes for simulated data from three different copulas

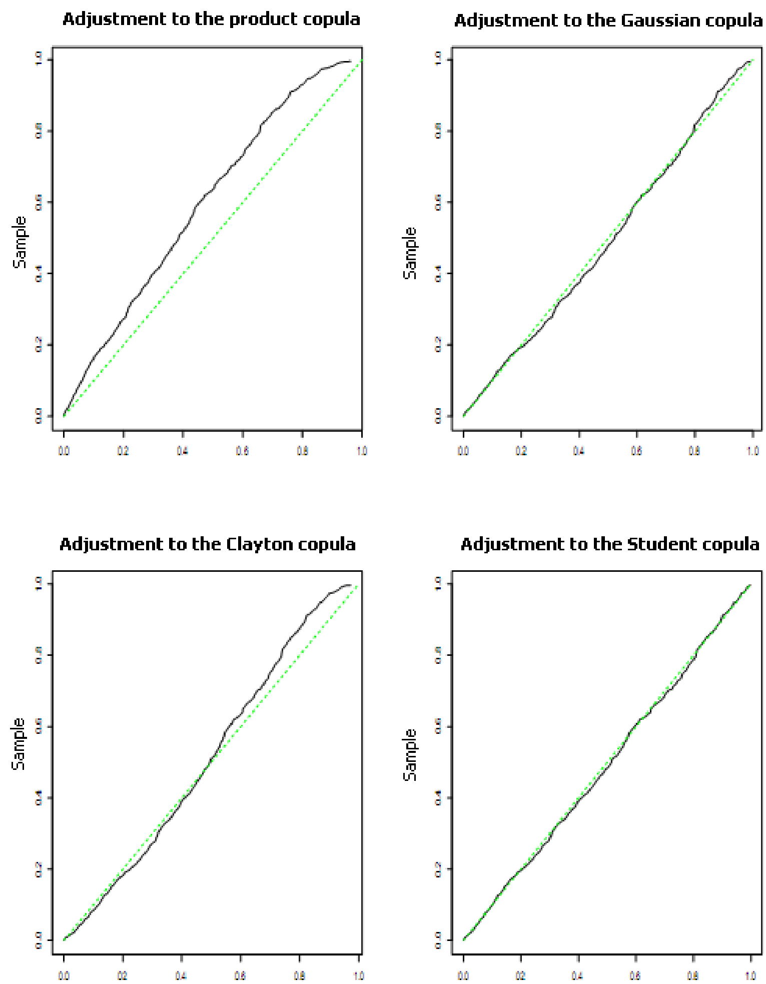

Figure 2.13: Kendall plots comparing a sample of simulated data from a Student copula (correlation 0.5, 3 degree of freedom) several copulas estimated on the same sample.

Chapter 3

Measuring Risk

Lord Kelvin once said «Anything that exists,
exists in some quantity and can therefore be measured»
(quoted in Beer 1967).

Quantifying and placing risks in some order of priority is an important activity to the farmer and his enterprise. There are two elements of each risk which need to be quantified for controlling it reliably:

- the frequency of the risk occurring,
- the cost and economic consequences of it occurring.

This quantification of risk is fundamental to all the commercial decisions which may be taken about an enterprise, then the initial investment capital must be sufficient to start and operate the business and to cover the risks it faces or to divert the costs of the risks elsewhere.

The principal decisions facing the farmer or the investor, can be sub-divided into three categories, namely:

1. Commercial decisions. These are the basic decisions about business, and are made through financial comparison of the anticipated return on investment with the cost of any risk if it occurs.
2. Reduction and control decisions. The decisions specific for each risk which must be made if its impact is to be reduced or eliminated altogether. If the

risk is only to be reduced, then it is important to decide to what acceptable level, and at what cost.

3. Financing decisions. These are the decisions which deal with ways of financing the risk and their acceptability.

A good reference resources it is still very necessary to make the quantification specific to the farm in question. As the risks to the farm operations are site related, it is necessary that the quantification of those risks is also site related. It is therefore important that as much information as possible is assembled for each particular farming enterprise, and its local environment.

3.1 Risk Measures

3.1.1 Definition

Since risks are modelled as non-negative rv's, measuring risk is equivalent to establishing a correspondence between the space of rv's and non-negative real numbers $\mathbb{R}^+$. The real number denoting a general risk measure associated with the risk X will henceforth be denoted as X. Thus, a risk measure is nothing but a functional that assigns a non-negative real number to a risk. See Szegö (2004) for an overview. It is essential to understand which aspect of the riskiness associated with the uncertain outcome the risk measure attempts to quantify.

In this chapter, we will focus on risk measures that can be used for determining provisions and capital requirements in order to avoid insolvency. In that respect, we will concentrate on risk measures that measure upper tails of distribution functions. We are now ready to state the definition of a risk measure.

Definition 3.1.1 *A risk measure is a functional ϱ mapping a risk X to a non-negative real number $\varrho[X]$, possibly infinite, representing the extra cash which has to be added to X to make it acceptable.*

The idea is that ϱ quantifies the riskiness of X: large values of $\varrho[X]$ tell us that X is 'dangerous'. Specifically, if X is a possible loss of some financial portfolio over a time horizon, we interpret $\varrho[X]$ as the amount of capital that should be added as a buffer to this portfolio so that it becomes acceptable to an internal or external risk controller. In such a case, $\varrho[X]$ is the risk capital of the portfolio.

Such risk measures are used for determining provisions and capital requirements in order to avoid insolvency; see Panjer (1998, [127]).

Another function that is useful in analysing the thickness of tails is the mean-excess loss, whose definition by.

Definition 3.1.2 *Given a non-negative rv X, the associated mean-excess function (mef) e_X is defined as*

$$e_X(x) = \mathbb{E}[X - x | X > x], x > 0.$$

The mef corresponds to the well-known expected remaining lifetime in life insurance. In reliability theory, when X is a non-negative rv, X can be thought of as the lifetime of a device and $e_X(x)$ then expresses the conditional expected residual life of the device at time x given that the device is still alive at time x.

3.2 Premium Principles

A premium principle is a rule for assigning a premium to an insurance risk. In this Section, we focus on the premium that accounts for the monetary payout by the insurer in connection with insurable losses plus the risk loading that the insurer imposes to reflect the fact that experienced losses rarely.

We list and discuss desirable properties of premium principles. First, we present some notation that we use throughout afterwards. Let χ denote the set of nonnegative random variables on the probability space $(\Omega, F, \mathbb{P})$, this is our collection of insurance-loss rv's, also called insurance risks. Let X, Y, Z, etc. denote typical members of χ. Finally, let ϱ denote the premium principle, or function, from χ to the set of (extended) non-negative real numbers. Thus, it is possible that $\varrho[X]$ takes the value ∞. It is possible to extend the domain of a premium principle ϱ to include possibly negative rv's. That might be necessary if we were considering a general loss rv of an insurer, namely, the payout minus the premium (Bowers *et al.,* [9]). However, in this thesis, we consider only the insurance payout and refer to that as the insurance loss rv.

3.2.1 Properties of premium calculation principles

1) Independence

$\varrho[X]$ depends only on the df of X, namely S_X, inwhich

$$S_X(t) = \mathbb{P}\{\omega \in : X(\omega) > t\}.$$

That is, the premium of X depends only on the tail probabilities of X. This property states that the premium depends only on the monetary loss of the insurable event and the probability that a given monetary loss occurs, not the cause of the monetary loss.

2) Risk loading

$\varrho[X] \geq \mathbb{E}[X]$ for all $X \in \chi$. Loading for risk is desirable because one generally requires a premium rule to charge at least the expected payout of the risk X, namely $\mathbb{E}[X]$, in exchange for insuring the risk. Otherwise, the insurer will lose money on average.

3) No unjustified risk loading

If a risk $X \in \chi$ is identically equal to a constant $c \geq 0$ (almost everywhere), then $\varrho[X] = c$. In contrast to Property 2 (Risk loading), if we know for certain (with probability 1) that the insurance payout is c, then we have no reason to charge a risk loading because there is no uncertainty as to the payout.

4) Maximal loss (or no rip-off)

$$\varrho[X] \leq \max[X], \text{ for all } X \in \chi.$$

5) Translation equivariance (or translation invariance)

That means that adding (resp. subtracting) the sure initial amount a to the initial position and investing it in the reference instrument, simply decreases (resp. increases) the risk measure by a. Translation invariance: for all $X \in \chi$ and all real numbers $a \geq 0$, we have

$$\varrho[X \pm a] = \varrho[X] \pm a.$$

If we increase a risk X by a fixed amount a, then Property 5 states that the premium for $X + a$ should be the premium for X increased by that fixed amount a.

6) Subadditivity

One can argue that subadditivity is a reasonable property because the no-arbitrage argument works well to ensure that the premium for the sum of two risks is not greater than the sum of the individual premiums

$$\varrho[X+Y] \leq \varrho[X] + \varrho[Y] \text{ for all } X,\ Y \in \chi.$$

Otherwise, the buyer of insurance would simply insure the two risks separately. However, the no-arbitrage argument that asserts that $\varrho[X+Y]$ cannot be less than $\varrho[X]+\varrho[Y]$ fails because it is generally not possible for the buyer of insurance to sell insurance for the two risks separately.

7) Positive homogeneity

Axiom 6 implies that $\varrho(nX) \leq n\varrho(X)$ for $n = 1, 2, ...$ In Axiom 7 we have imposed the reverse inequality (and require equality for all positive λ)

$$\varrho[\lambda X] = \lambda\varrho[X] \text{ for all } X \in \chi \text{ and all } \lambda \geq 0,$$

to model what a government or an exchange might impose in a situation where no netting or diversification occurs, in particular because the government does no prevent many firms to all take the same position.

Remark 3.2.1 *If position size directly influences risk (for example, if positions are large enough that the time required to liquidate them depend on their sizes) then we should consider the consequences of lack of liquidity when computing the future net worth of a position. With this in mind, Axioms 6 and 7 about mappings from random variables into the reals, remain reasonable.*

8) Additivity

This Property is a stronger form of Property 7 (Positive homogeneity). One can use a similar no-arbitrage argument to justify the additivity property

$$\varrho[X+Y] = \varrho[X] + \varrho[Y] \text{ for all } X,\ Y \in \chi,$$

(see, Albrecht 1992, [2]).

9) Superadditivity

Might be a reasonable property of a premium principle if there are surplus constraints that require that an insurer charge a greater risk load for insuring larger risks

$$\varrho[X+Y] \geq \varrho[X] + \varrho[Y] \text{ for all } X,\ Y \in \chi.$$

For example, we might observe in the market that $\varrho[2X] > 2\varrho[X]$ because of such surplus constraints. Note that both Properties 8 and 9 can be weakened by requiring only $\varrho[\lambda X] \leq \lambda\varrho[X]$ or $\varrho[\lambda X] \geq \lambda\varrho[X]$ for $\lambda > 0$, respectively. Next, we weaken the additivity property by requiring additivity only for certain insurance risks.

10) Additivity for independent risks

Some actuaries might feel that Property 7 (Additivity) is too strong and that the no-arbitrage argument only applies to risks that are independent

$$\varrho[X+Y] = \varrho[X] + \varrho[Y] \text{ for all } X,\ Y \in \chi,$$

such that X and Y are independent. They, there by, avoid the problem of surplus constraints for dependent risks.

11) Additivity for comonotonic risks

Is desirable because if one adopts subadditivity as a general rule

$$\varrho[X+Y] = \varrho[X] + \varrho[Y] \text{ for all } X,\ Y \in \chi,$$

such that X and Y are comonotonic (see Comonotonicity). Then it is unreasonable to have

$$\varrho[X+Y] < \varrho[X] + \varrho[Y]$$

because neither risk is a hedge[1] against the other, that is, they move together (Yaari 1987, [154]). If a premium principle is additive for comonotonic risks, then is it layer additive (Wang 1996, [150]). Note that Property 11 implies Property 6, (Scale equivariance), if ϱ additionally satisfies a continuity condition. Next, we consider properties of premium rules that require that they preserve common orderings of

See the book of Jon Gregory and Angelo Arvanitis, 2004, Credit: The Complete Guide to Pricing, Hedging and Risk Management risks.

12) Monotonicity

If

$$X(\omega) \le Y(\omega) \text{ for all } \omega \in \Omega,$$

then $\varrho[X] \le \varrho[Y]$.

13) Preserves first stochastic dominance (FSD) ordering

If

$$S_X(t) \le S_Y(t) \text{ for all } t \ge 0,$$

then $\varrho[X] \le \varrho[Y]$.

14) Preserves stop-loss ordering (SL) ordering

Property 1, (Independence), together with Property 12, (Monotonicity), imply Property 13, (Preserves FSD ordering, Wang et al. 1997, [152]). Also, if ϱ preserves SL ordering, then ϱ preserves FSD ordering because stop-loss ordering is weaker (see, Rothschild et *al.* 1970, [131]), if

$$\mathbb{E}[X - d]_+ \le \mathbb{E}[Y - d]_+ \text{ for all } d \ge 0,$$

then $\varrho[X] \le \varrho[Y]$. These orderings are commonly used in actuarial science to order risks (partially) because they represent the common orderings of groups of decision makers (see, Kaas et *al.* 1994, [93], Van Heerwaarden, 1991, [145]), for example. Finally, we present a technical property that is useful in characterizing certain premium principles.

15) Continuity

Let $X \in \chi$, then,

$$\lim_{a \to 0^+} \varrho[\max(X - a, 0)] = \varrho[X],$$

and

$$\lim_{a \to \infty} \varrho[\min(X, a)] = \varrho[X].$$

3.2.2 Coherent risk measures

Several authors have selected some of these conditions to form a set of requirements that any risk measure should satisfy. The following definition is taken from the seminal paper of Artzner et al. (1999, [5]).

Definition 3.2.1 *A risk measure that is translative, positive homogeneous, subadditive and monotone is called coherent.*

It is worth mentioning that coherence is defined with respect to a set of axioms, and no set is universally accepted. Modifying the set of axioms regarded as desirable leads to other 'coherent' risk measures.

3.3 Value-at-Risk

Moste part of practitioners interest in quantiles of probability distributions. Since quantiles have a simple interpretation in terms of over or undershoot probabilities they have found their way into current risk management practice in the form of the concept of value-at-risk (VaR). This concept was introduced to answer the following question: how much can we expect to lose in one day, week, year, with a given probability?

Recently, VaR has become the benchmark risk measure: its importance is unquestion since regulators accept this model as the basis for setting capital requirements for market risk exposure. A textbook treatment of VaR is given in Jorion (2000, [91]).

Non-subadditivity. VaR has been fundamentally criticized as a riskmeasure on the grounds that is has poor aggregation properties. This critique has its origins in the work of Artzner et al. (1997 [4], 1999 [5]), who showed that VaR is not a coherent risk measure, since it violates the property of subadditivity that they believe reasonable risk measures should have.

VaR is defined as follows.

Definition 3.3.1 *Given a risk* X *and a probability level* $p \in [0,1]$, *the corresponding VaR, denoted by* $VaR(X,p)$, *is defined as*

$$VaR(X,p) = F_X^{-1}(p).$$

Note that the VaR risk measure reduces to the percentile principle of Goovaerts *et al.*, (1984, [67]).

It is worth mentioning that VaR's always exist and are expressed in the proper unit of measure, namely in lost money. Since VaR is defined with the help of the quantile function F_X^{-1}. We have the following equivalence relation, which holds for all $x \in \mathbb{R}$ and $p \in [0,1]$,

$$VaR(X,p) \leq x \Longleftrightarrow p \leq F_X(x).$$

VaR fails to be subadditive (except in some very special cases, such as when the X_i are multivariate normal). Thus, in general, VaR has the surprising property that the VaR of a sum may be higher than the sum of the VaR's. In such a case, diversification will lead to more risk being reported. Consider two independent Pareto risks of parametre 1, X and Y. Show that the inequality

$$VaR(X,p) + VaR(Y,p) < VaR(X+Y,p)$$

holds for any p, so that VaR cannot be subadditive in this simple case.

A possible harmful aspect of the lack of subadditivity is that a decentralized risk management system may fail because VaR's calculated for individual portfolios may not be summed to produce an upper bound for the VaR of the combined portfolio.

3.4 Tail Value-at-Risk

A single VaR at a predetermined level p does not give any information about the thickness of the upper tail of the distribution function. This is a considerable shortcoming since in practice a regulator is not only concerned with the frequency of default, but also with the severity of default. Also shareholders and management should be concerned with the question «how bad is bad?» when they want to evaluate the risks at hand in a consistent way.

Therefore, one often uses another risk measure, which is called the tail value-at-risk (TVaR) and defined by

Definition 3.4.1 *Given a risk X and a probability level p, the corresponding TVaR, denoted by $TVaR(X,p)$, is defined as*

$$TVaR(X,p) = \frac{1}{1-p}\int_1^p VaR[X,\xi]\,d\xi,\ \ 0<p<1.$$

We thus see that $TVaR(X,p)$ can be viewed as the arithmetic average of the VaR's of X, from p on $[0,1]$.

3.5 Some related risk measures

3.5.1 Conditional tail expectation

The conditional tail expectation (CTE) represents the conditional expected loss given that the loss exceeds its VaR:

$$\mathbb{CTE}(X,p) = \mathbb{E}\left(X \mid X > VaR(X,p)\right).$$

Thus the CTE is nothing but the mathematical transcription of the concept of 'average loss in the worst $100(1-p)\%$ cases'. Defining by $c =$VaR(X,p) a critical loss threshold corresponding to some confidence level p, $\mathbb{CTE}(X,p)$ provides a cushion against the mean value of losses exceeding the critical treshold c.

3.5.2 Conditional VaR

An alternative to CTE is the conditional VaR (or CVaR). The CVaR is the expected value of the losses exceeding VaR

$$\begin{aligned} CVaR(X,p) &= \mathbb{E}\left(X - VaR(X,p) \mid X > VaR(X,p)\right) \\ &= \mathbb{CTE}(X,p) - VaR(X,p). \end{aligned}$$

It is easy to see from Definition 3.1.2 that CVaR is related to the mean-excess function through

$$CVaR(X,p) = e_X\left(VaR(X,p)\right).$$

Therefore, evaluating the mef at quantiles yields CVaR.

3.5.3 Expected shortfall

As the VaR at a fixed level only gives local information about the underlying distribution, a promising way to escape from this shortcoming is to consider the so-called expected shortfall over some quantile. Expected shortfall at probability level p is the stop-loss premium with retention $\mathrm{VaR}(X,p)$. Specifically,

$$\begin{aligned}\mathbb{ES}(X,p) &= \mathbb{E}\left((X - VaR(X,p))_+\right) \\ &= \pi_X(VaR(X,p)).\end{aligned}$$

3.5.4 Relationships between risk measures

The following relation holds between the three risk measures defined above.

Proposition 3.5.1 *For any $p \in [0,1]$, the following identities are valid:*

$$TVaR(X,p) = VaR(X,p) + \frac{1}{1-p}\mathbb{ES}(X,p), \tag{3.1}$$

$$\mathbb{CTE}(X,p) = VaR(X,p) + \frac{1}{\overline{F}_X(VaR(X,p))}\mathbb{ES}(X,p), \tag{3.2}$$

$$CVaR(X,p) = \frac{\mathbb{ES}(X,p)}{\overline{F}_X(VaR(X,p))}. \tag{3.3}$$

Proof. See Denuit *et al. (2005, [31]).* ∎

Corollary 3.5.1 *Note that if F_X is continuous then by combining (3.1) and (3.2) we find*

$$\mathbb{CTE}(X,p) = TVaR(X,p), \ p \in [0,1]. \tag{3.4}$$

so that $\mathbb{CTE}$ and $TVaR$ coincide for all p in this special case. In general, however, we only have

$$TVaR(X,p) = \mathbb{CTE}(X,p) + \left(\frac{1}{1-p} - \frac{1}{\overline{F}_X(VaR(X,p))}\right)\mathbb{ES}(X,p).$$

Since the quantity between the brackets can be different from 0 for some values of p, TVaR and $\mathbb{CTE}$ are not always equal. Looking back at Figure 3.1, we see that the values of p for which the quantity between the brackets does not vanish correspond to jumps in the df (e.g., for p_3, $F_X(F_x^{-1}(p_3)) > p_3$). See also Acerbi and Tasche (2002, [1]).

Remark 3.5.1 *From (3.1) in Proposition 3.5.1 it follows that the minimal value of the cost function in (3.6) can be expressed as*

$$\begin{aligned}CVaR[X, VaR(X,1-\epsilon)] &= \mathbb{E}\left[(X - VaR(X,1-\epsilon))_+\right] + \epsilon VaR(X,1-\epsilon) \\ &= \epsilon TVaR(X,1-\epsilon). \end{aligned} \tag{3.5}$$

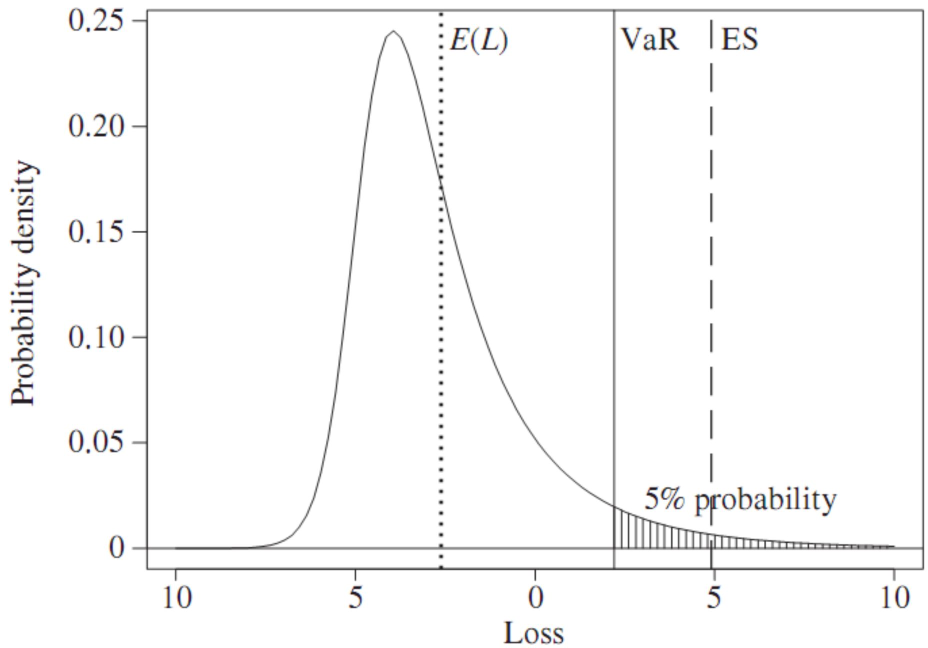

Figure 3.1: An example of a loss distribution with the 95%VaR marked as a vertical line; the mean loss is shown with a dotted line and an alternative risk measure known as the 95% expected shortfall is marked with a dashed line.

A more general version of the minimization problem

$$\min_{\varrho[X]} \left\{ \mathbb{E}\left[(X - \varrho[X])_{+} \right] + \varrho[X]\,\epsilon \right\}, \qquad 0 < \epsilon < 1, \tag{3.6}$$

involving a distortion risk measure, is considered in Dhaene, Goovaerts and Kaas (2003, [68]), Laeven and Goovaerts (2004, [104]) and Goovaerts, Van den Borre and Laeven (2004, [69]).

3.6 Risk measures based on Distorted Expectation Theory

Consider a decision-maker with a future random fortune equal to X. Using integration by parts, the expectation of X can be written as

$$\mathbb{E}[X] = -\int_{-\infty}^{0} \left(1 - \overline{F}_X(x)\right) dx + \int_{0}^{+\infty} \overline{F}_X(x)\, dx.$$

Under the 'distorted expectations hypotheses' it is assumed that each decision-maker has a non-decreasing function $g\left[0,1\right] \rightarrow \left[0,1\right]$ with $g\left(0\right) = 0$ and $g\left(1\right) = 1$ (called a distortion function) and that he values a fortune X at its 'distorted expectation' $\pi_g\left(X\right)$ defined as

$$\pi_g\left[X\right] := -\int_{-\infty}^{0} \left(1 - g\left(\overline{F}_X\left(x\right)\right)\right) dx + \int_{0}^{+\infty} g\left(\overline{F}_X\left(x\right)\right) dx. \tag{3.7}$$

The function g is called a distortion because it distorts the probabilities $\overline{F}_X\left(x\right)$ before calculating a generalized expected value. As $g\left(\overline{F}_X\left(x\right)\right)$ is a non-decreasing function of $\overline{F}_X\left(x\right)$, where $\overline{F}_X\left(x\right)$ is a non-increasing function of x, $g\left(\overline{F}_X\left(x\right)\right)$ is also a non-increasing function of x, and can be thought of as the risk-adjusted tail function. Note that $g\left(\overline{F}_X\left(x\right)\right)$ is not necessarily a tail function (indeed the right-continuity condition involved in properties satisfied by all dfs is not always fulfilled). Hence $\pi_g\left[X\right]$ is not necessarily the expectation of some transformed rv.

We will see that under additional assumptions on g, this will be the case. If X is non-negative, then we find from (3.7) that

$$\pi_g\left[X\right] = \int_{0}^{+\infty} g\left(\overline{F}_X\left(x\right)\right) dx.$$

Note that $g\left(0\right) = 0$ implies $\pi_g\left[0\right] = 0$ and that $g\left(1\right) = 1$ implies $\pi_g\left[1\right] = 1$.

A decision-maker is said to base his preferences on the «distorted expectations hypotheses» if he acts in order to maximize the distorted expectation of his wealth. This means that there exists a distortion function g such that the decision-maker prefers Y to a fortune X if, and only if, $\pi_g\left[X\right] \leq \pi_g\left[Y\right]$.

Part II

Main results

Chapter 4

Copula parameter estimation by bivariate *L*-moments

«Mathematics is the art of giving the same name to different things.»

Henri POINCARE, 1854-1912

Recently, Serfling and Xiao (2007, [138]) extended the L-moment theory (Hosking, 1990, [75]) to the multivariate setting. In the present paper, we focus on the two-dimension random vectors to establish a link between the bivariate L-moments (BLM) and the underlying bivariate copula functions. This connection provides a new estimate of dependence parameters of bivariate statistical data. Extensive simulation study is carried out to compare estimators based on the BLM, the maximum likelihood, the minimum distance and the rank approximate Z-estimation. The obtained results show that, when the sample size increases, BLM's based estimation performs better as far as the bias and computation time are concerned. Moreover, the root means squared error (RMSE) is quite reasonable and less sensitive in general to outliers than those of the above cited methods. Further, we expect that BLM's method will be an easy-to-use tool for the estimation of multiparameter copula models.

4.1 Introduction and motivation

The copula method is a tool to construct multivariate distributions and describe the dependence structure in multivariate data sets (e.g., Joe, 1997, [85] or Nelsen, 2006,

[123]). Modelling dependence structures by copulas is a topic of current research and of recent use in several areas, such as financial assessments (e.g., Malevergne and Sornette, 2003, [112]), insurance (e.g., Drees and Müller, 2008, [34]) and hydrology (e.g., Dupuis, 2007, [35]). For the sake of simplicity, throughout the paper, we restrict ourself to the two-dimensional case. Let $\left(X^{(1)}, X^{(2)}\right)$ be a bivariate random variable with joint distribution function

$$F\left(x_1, x_2\right) = \mathbb{P}\left(X^{(1)} \leq x_1, X^{(2)} \leq x_2\right), \quad \left(x_1, x_2\right) \in \mathbb{R}^2,$$

and marginal df $F_j\left(x_j\right) = \mathbb{P}\left(X^{(j)} \leq x_j\right)$ for $x_j \in \mathbb{R}$ and $j = 1, 2$. If not stated otherwise, we assume that the F_j are continuous functions. According to Sklar's theorem (Sklar, 1959, [142]) there exists a unique copula $C : \mathbb{I}^2 \to \mathbb{I}$, with $\mathbb{I} = [0, 1]$, such that

$$F\left(x_1, x_2\right) = C\left(F_1\left(x_1\right), F_2\left(x_2\right)\right), \text{ for } \left(x_1, x_2\right) \in \mathbb{R}^2.$$

The copula C is the joint df of the uniform random variables (r.v.'s) $U_j = F_j\left(X^{(j)}\right)$, $j = 1, 2$, defined for $\left(u_1, u_2\right) \in \mathbb{I}^2$, by

$$C\left(u_1, u_2\right) = F\left(F_1^{-1}\left(u_1\right), F_2^{-1}\left(u_2\right)\right),$$

where G^{-1} is the generalized inverse function (or the quantile function) of a df G.

A parametric copula model arises for $\left(X^{(1)}, X^{(2)}\right)$ when C is unknown but assumed to belong to a class $\mathcal{C} := \{C_{\boldsymbol{\theta}}, \ \boldsymbol{\theta} \in \mathcal{O}\}$, where $\mathcal{O}$ is an open subset of $\mathbb{R}^r$ for some integer $r \geq 1$. Statistical inference on the dependence parameter $\boldsymbol{\theta}$ is one of the main topics in multivariate statistical analysis. Several methods of copula parameter estimation have been developed, included the pseudo maximum likelihood (PML), inference of margins, minimum distance and others, see for instance Genest *et al.* (2009, [62]). All these methods use in general some optimization technics under constraints, this in general require enough computational time to run calculations. In this paper, we present a new estimation method of θ based on the bivariate L-moments that may be serve as alternative in front of computation's time issue and produces estimation results reasonable enough. The multivariate L-moments have been introduced by Serfling and Xiao (2007, [138]) as an extension of the univariate L-moments introduced by Hosking (1990, [75]). The L-comoments have interpretations similar to the classical central moment covariance, coskewness, and cokurtosis that also possess the features of the L-moments. This extension is useful to solve some problems in connection with multivariate heavy-tailed distributions and small samples. As mentioned, for instance, in Hosking (1990, [75]) and recently in Delicado and Goria (2008, [29]), the main advantage of L-moments vis-a-vis of classical estimation methods (e.g. least squares, moments and maximum likelihood) is their relative slight sensitivity to outlying data and their performance in statistical inference with small samples. In this paper we establish a functional representation of bivariate L-moments (BLM) by the underlying copula function and propose a new estimation method of the parameters of copula models. By considering multiparameter Farlie-Gumbel-Morgenstern (FGM) and Archimedean copulas, simulation studies are carried out to compare the performance of this method with those of the PML, minimum distance (MD) and rank approximate Z-estimation. The rest of the paper is organized as follows. In Section 4.2, we briefly introduce the univariate and bivariate L-moment approaches. We present, in Section 4.3, functional representations of the bivariate L-moments by copula functions and give some examples. A new estimator of copula parameter and its asymptotic behavior are given in Sections 4.4. In Section 4.5, a simulation study evaluates the BLM performance is given.

4.2 Bivariate L-moments

First we begin with a brief introduction on the univariate L-moments. Hosking (1990, [75]) introduced L-moments λ_k as an alternative to the classical central moments $\mu_k = \mathbb{E}\left[(Y-\mu)^k\right]$ determined by the df F_Y of the underlying r.v. Y. An L-moment λ_k is defined as a specific linear combination of the expectations of the order statistics $Y_{1:k} \leq ... \leq Y_{k:k}$. More precisely, the kth L-moment is defined by

$$\lambda_k = \frac{1}{k}\sum_{\ell=0}^{k-1}\frac{(-1)^{\ell}(k-1)!}{\ell!\,(k-1-\ell)!}\mathbb{E}\left[Y_{k-\ell:k}\right],\ \ k=1,2,...$$

By analogy with the classical moments, the first four L-moments λ_1, λ_2, λ_3 and λ_4 measure location, scale, skewness and kurtosis features respectively. The L-functional representation of λ_k is terms of the quantile function F_Y^{-1} is given by (see Hosking, 1998, [77]):

$$\lambda_k = \int_{\mathbb{I}} F_Y^{-1}(u)\, P_{k-1}(u)\, du, \tag{4.1}$$

where $P_k(u) := \sum_{\ell=0}^{k} p_{k,\ell} u^{\ell}$, with $p_{k,\ell} = (-1)^{k+\ell}(k+\ell)!/[(\ell^2)!(k-\ell)!]$ is the shifted Legendre polynomials (SLP). In the sequel, we will make use of the three first SLP

$$P_0(u) = 1,\ P_1(u) = 2u - 1,\ P_2(u) = 6u^2 - 6u + 1.$$

A straightforward transformation in (4.1) using $P_0 \equiv 1$ and the orthogonality of P_{k-1} leads to a representation in terms of covariance, that is

$$\lambda_k = \begin{cases} \mathbb{E}[Y] & k = 1; \\ Cov(Y, P_{k-1}(F_Y(Y))) & k \geq 2. \end{cases} \tag{4.2}$$

L-moments may be used as summary statistics for data samples, to identify probability distributions and fit them to data. A brief description of these methods is given in Hosking (1998, [77]). L-moments are now widely used in water sciences especially in flood frequency analysis. Recent studies include Kjeldsen *et al.* (2002, [100]), Kroll and Vogel (2002, [102]), Lim and Lye (2003, [109]), Chebana and Ouarda (2007, [16]) and Chebana *et al.* (2009, [17]). In other recent work, Karvanen *et al.* (2002, [95]) used L-moments for fitting distributions in independent component analysis in signal processing, and Jones and Balakrishnan (2002, [90]) pointed out some relationships between integrals occurring in the definition of moments and L-moments. Hosking (2006, [78]) showed that, for a wide range of distributions, the characterization of a distribution by its L-moments is non-redundant. That is, if one L-moment is dropped, the remaining L-moments no longer suffice to determine the entire distribution. Recently, Serfling and Xiao (2007, [138]) extended this approach to the multivariate case, this has already begun to be developed and applied in statistical hydrology by Chebana and Ouarda (2007, [16]) and Chebana *et al.* (2009, [17]).

Next we present basic notations and definitions of the bivariate L-moments. Let $X^{(1)}$ and $X^{(2)}$ be two r.v.'s with finite means, margins F_1 and F_2 and L-moments sequences $\lambda_k^{(1)}$ and $\lambda_k^{(2)}$, respectively. By analogy with the covariance representation (4.2) for L-moments, and the central comoments, Serfling and Xiao (2007, [138]) defined the kth L-comoment of $X^{(1)}$ with respect to $X^{(2)}$ by the covariance of the couple of r.v.'s $X^{(1)}$ and $P_k(F_2(X^{(2)}))$, for every $k \geq 1$, as

$$\lambda_{k[12]} = Cov\left(X^{(1)}, P_k\left(F_2\left(X^{(2)}\right)\right)\right). \tag{4.3}$$

So, the kth L-comoment of $X^{(2)}$ with respect to $X^{(1)}$ is defined by

$$\lambda_{k[21]} = Cov\left(X^{(2)}, P_k\left(F_1\left(X^{(1)}\right)\right)\right).$$

If we suppose that F belongs to a parametric family of df's, then the set of parameters define the models of margins and the dependence structure between r.v.'s $X^{(1)}$ and $X^{(2)}$. Since we focus only on the estimation of copula parameters, then it is convenient to use the kth L-comoment of $F_1\left(X^{(1)}\right)$ with respect to $X^{(2)}$ instead of $\lambda_{k[12]}$, that is

$$\delta_{k[12]} := Cov\left(F_1\left(X^{(1)}\right), P_k\left(F_2\left(X^{(2)}\right)\right)\right). \tag{4.4}$$

So that the kth L-comoment of $F_2\left(X^{(2)}\right)$ with respect to $X^{(1)}$ is given by

$$\delta_{k[21]} = Cov\left(F_2\left(X^{(2)}\right), P_k\left(F_1\left(X^{(1)}\right)\right)\right).$$

If the copula C is symmetric in the sense that $C(u,v) = C(v,u)$, then $\delta_{k[12]} = \delta_{k[21]} := \delta_k$, for each $k = 1, 2, ...$ The quantity $\delta_{k[12]}$ will be called "*the kth bivariate copula L-moment*" of $X^{(1)}$ with respect to $X^{(2)}$, so $\delta_{k[21]}$ is the kth copula L-moment of $X^{(2)}$ with respect to $X^{(1)}$.

In application, we will often make use of the three first bivariate copula L-moments, that is:

$$\delta_1 = 2Cov\left(F_1\left(X^{(1)}\right), F_2\left(X^{(2)}\right)\right)$$

$$\delta_2 = -6Cov\left(F_1\left(X^{(1)}\right), F_2\left(X^{(2)}\right)\left(1 - F_2\left(X^{(2)}\right)\right)\right)$$

$$\delta_3 = Cov\left(F_1\left(X^{(1)}\right), 20F_2^3\left(X^{(2)}\right) - 30F_2^2\left(X^{(2)}\right) + 12F_2\left(X^{(2)}\right) - 1\right).$$

4.3 Bivariate copula representation of kth copula L-moment

Theorem 4.3.1 below gives a representation of the kth bivariate L-moment in terms of the underlying copula function. This result provides a new estimate of bivariate copula parameters.

Theorem 4.3.1 *The kth bivariate copula L-moment of $X^{(1)}$ with respect to $X^{(2)}$ may be rewritten, for each $k \geq 1$, as*

$$\delta_{k[12]} = \int_{\mathbb{I}^2} \left(C\left(u_1, u_2\right) - u_1 u_2 \right) du_1 dP_k\left(u_2\right), \tag{4.5}$$

or

$$\delta_{k[12]} = \int_{\mathbb{I}^2} u_1 P_k\left(u_2\right) dC\left(u_1, u_2\right).$$

Observe that $\delta_{1[12]} = \delta_{1[21]} = \rho/6$ where ρ is the Spearman rho ρ, defined in term of copula C by

$$\rho = 12 \int_{\mathbb{I}^2} u_1 u_2 dC\left(u_1, u_2\right) - 3, \tag{4.6}$$

(see Nelsen, 2006, [123, page 167]).

In view of Theorem 4.3.1, according to our needs, we may construct a system of equations that will serve to the estimation of multiparameter copula models. For this reason, the proposed estimator is more likely to be used for the multiparameter copulas. In the case of the one-parameter copulas, it is equivalent to the rho-inversion method (see (4.17)). Indeed, suppose that we are dealing with the estimation of one dimension parameter of a copula model, then it suffices to use one of the kth bivariate copula L-moment, says $\delta_{1[12]}$. In the case of $d-$dimension parameters we have to take the d first bivariate copula L-moment, so we obtain a system of d equations with d unknown parameters. Then, by replacing the coefficients $\delta_{k[12]}$, $k = 1, ..., d$ by their empirical counterparts, we obtain estimators of the d parameters. Indeed, suppose that $d = 3$ and $C = C_\theta$, $\theta = (\theta_1, \theta_2, \theta_3)$, then from Theorem 4.3.1, the first three bivariate copula L-moments of $X^{(1)}$ with respect to $X^{(2)}$ are

$$\begin{aligned}
\delta_{1[12]} &= 2 \int_{\mathbb{I}^2} C_\theta\left(u_1, u_2\right) du_1 du_2 - \frac{1}{2} \\
\delta_{2[12]} &= 6 \int_{\mathbb{I}^2} \left(2u_2 - 1\right) C_\theta\left(u_1, u_2\right) du_1 du_2 - \frac{1}{2} \\
\delta_{3[12]} &= \int_{\mathbb{I}^2} \left(60u_2^2 - 60u_2 + 12\right) C_\theta\left(u_1, u_2\right) du_1 du_2 - \frac{1}{2}.
\end{aligned}$$

Next we present applications of Theorem 4.3.1 to parameter estimation of two popular families of copula, namely the FGM and Archimedean copulas.

4.3.1 FGM families

One of the most popular parametric family of copulas is the FGM family defined for $|\alpha| \leq 1$ by

$$C_\alpha(u_1, u_2) = u_1 u_2 + \alpha u_1 u_2 \overline{u}_1 \overline{u}_2, \ 0 \leq u_1, u_2 \leq 1, \tag{4.7}$$

with $\overline{u}_j := 1 - u_j$, $j = 1, 2$. The model is useful for the moderate correlation which occurs in engineering and medical applications (see, e.g., Blischke and Prabhaker Murthy, 2000 and Chalabian and Dunnington, 1998). The Pearson correlation coefficient ρ corresponds to the model (4.7) can never exceed 1/3, (see, e.g., Huang and Kotz, 1984, [80]). In order to increase the dependence between two random variables obeying the type of FGM distribution, Johnson and Kotz (1977, [87]) introduced the $(r-1)$-iterated FGM family with r-dimensional parameter $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_r)$:

$$C_{\boldsymbol{\alpha}}(u_1, u_2) = u_1 u_2 + \sum_{j=1}^{r} \alpha_j (u_1 u_2)^{[j/2]+1} (\overline{u}_1 \overline{u}_2)^{[j/2+1/2]},$$

where $[z]$ denotes the greatest integer less than or equal to z. For example, the one-iterated FGM family (Huang and Kotz, 1984) is a two-parameter copula model:

$$C_{\alpha_1,\alpha_2}(u_1, u_2) = u_1 u_2 \{1 + \alpha_1 \overline{u}_1 \overline{u}_2 + \alpha_2 u_1 u_2 \overline{u}_1 \overline{u}_2\}. \tag{4.8}$$

The range of parameters (α_1, α_2) is given by the region

$$\mathcal{R} := \left\{ (\alpha_1, \alpha_2), |\alpha_1| \leq 1, \alpha_1 + \alpha_2 \geq -1, \ \alpha_2 \leq \frac{1}{2} \left[3 - \alpha_1 + \left(9 - 6\alpha_1 - 3\alpha_1^2 \right)^{1/2} \right] \right\}. \tag{4.9}$$

The maximal reached correlation for this family is

$$\rho_{FGM}^{\max} = 0.42721, \quad \text{for} \ (\alpha_1, \alpha_2) = \left(-1 + 7/\sqrt{13}, 2 - 2/\sqrt{13} \right). \tag{4.10}$$

and the minimal correlation is $\rho_{FGM}^{\min} = -1/3$ for $(\alpha_1, \alpha_2) = (-1, 0)$. The two-iterated FGM family is given by

$$C_{\alpha_1,\alpha_2,\alpha_3}(u_1, u_2) = u_1 u_2 \left\{ 1 + \alpha_1 \overline{u}_1 \overline{u}_2 + \alpha_2 u_1 u_2 \overline{u}_1 \overline{u}_2 + \alpha_3 u_1 u_2 (\overline{u}_1 \overline{u}_2)^2 \right\},$$

and it has been discussed by Lin (1987, [111]).

According to the Theorem 4.3.1, we may give explicit formulas of bivariate copula L-moments for the FGM, the one-iterated FGM and the two-iterated FGM. Since the number of parameters equals $k \in \{1, ..., r\}$, then we are dealing with first k bivariate copula L-moments that will provide a system of k equations and therefore a tool for the estimation of the parameters of the copulas. Next we give the first bivariate copula L-moments for the FGM family C_α, the one-iterated FGM copula C_{α_1,α_2} and the two-iterated FGM copula $C_{\alpha_1,\alpha_2,\alpha_3}$.

- The first bivariate copula L-moment of FGM family C_α is

$$\delta_{1[12]} = \alpha_1/18$$

- The two first bivariate copula L-moments of one-iterated FGM copula C_{α_1,α_2} are:

$$\begin{cases} \delta_{1[12]} = \alpha_1/18 + \alpha_2/72 \\ \delta_{2[12]} = \alpha_2/120 \end{cases} \tag{4.11}$$

- The three first bivariate copula L-moments of two-iterated FGM copula $C_{\alpha_1,\alpha_2,\alpha_3}$ are:

$$\begin{cases} \delta_{1[12]} = \alpha_1/18 + \alpha_2/72 + \alpha_3/450 \\ \delta_{2[12]} = \alpha_2/120 \\ \delta_{3[12]} = -\alpha_3/1050 \end{cases}$$

4.3.2 Archimedean copula families

The Archimedean copula family is one of important class of copula models that contains the Gumbel, Clayton, Frank, ... (see, Table 4.1 in Nelsen, 2006, [123, page 116]). In the bivariate case, an Archimedean copula is defined by

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

where $\varphi : \mathbb{I} \to \mathbb{R}$ is a twice differentiable function called the generator, satisfying: $\varphi(1) = 0$, $\varphi'(x) < 0$, $\varphi''(x) \geq 0$ for any $x \in \mathbb{I}/\{0, 1\}$. The notation φ^{-1} stands for the inverse function of φ. For examples, the three generators $\varphi_\theta(t) = (-\ln((1-\theta(1-t))/t))$, $\varphi_\alpha(t) = (t^{-\alpha} - 1)/\alpha$ and $\varphi_\beta(t) = (-\ln t)^\beta$ define, respectively, the one parameter Frank, Clayton and Gumbel copula families. For more flexibility in fitting data, it is better to use the multi-parameters copula models than those of one parameter. To have a copula with more one parameter, we use, for instance, the distorted copula defined by $C_\Gamma(u, v) = \Gamma^{-1}(C(\Gamma(u), \Gamma(v)))$, where

$\Gamma : \mathbb{I} \rightarrow \mathbb{I}$ is a continuous, concave and strictly increasing function with $\Gamma(0) = 0$ and $\Gamma(1) = 1$. Note that if C is an Archimedean copula with generator φ, then C_Γ is also Archimedean copula with generator generator $\varphi \circ \Gamma$. For more details see Nelsen (2006, [123, 96]). As example, suppose that $\Gamma = \Gamma_{\beta_2}$, with $\Gamma_{\beta_2}(t) = \exp\left(t^{-\beta_2} - 1\right)$, $\beta_2 > 0$ and consider a Gumbel copula C_{β_1} with generator $\varphi_{\beta 1}(t) = (-\ln t)^{\beta_1}$, $\beta_1 \geq 1$. Then the copula $C_{\beta_1,\beta_2}(u,v) = \Gamma_{\beta_2}^{-1}\left(C_{\beta_1}\left(\Gamma_{\beta_2}(u), \Gamma_{\beta_2}(v)\right)\right)$ given by

$$C_{\beta_1,\beta_2}(u,v) := \left(\left(\left(u^{-\beta_2} - 1\right)^{\beta_1} + \left(v^{-\beta_2} - 1\right)^{\beta_1}\right)^{1/\beta_1} + 1\right)^{1/\beta_2}, \tag{4.12}$$

is a two-parameter Archimedean copula with generator $\varphi_{\beta_1,\beta_2}(t) := \left(t^{-\beta_2} - 1\right)^{\beta_1}$.

To have the two first bivariate copula L-moments correspond to C_{β_1,β_2}, we apply the Theorem 4.3.1 to get the following system of equations:

$$\begin{cases} \delta_{1[12]} = 2\int_0^1\int_0^1 \left(C_{\beta_1,\beta_2}(u,v) - uv\right) dudv, \\ \delta_{2[12]} = 6\int_0^1\int_0^1 (2v-1)\left(C_{\beta_1,\beta_2}(u,v) - uv\right) dudv. \end{cases} \tag{4.13}$$

In this case we cannot give explicit formulas, in terms of $\left\{\delta_{1[12]}, \delta_{2[12]}\right\}$, for the parameters $\{\beta_1, \beta_2\}$, however for a given values of the bivariate copula L-moments, we can solving the previous system by numerical methods and obtain the corresponding values of $\{\beta_1, \beta_2\}$.

Remark 4.3.1 *The previous system provides estimators for copula parameters by replacing the bivariate copula L-moments by their sample counterparts. This is similar to the method of moments (see Section 4.4).*

4.4 Semi-parametric BLM-based estimation

The aim of the present section is to provide a semi-parametric estimation for bivariate copula parameters on the basis of results of Section 4.3. Suppose that the underlying copula C belongs to a parametric family C_θ with $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_r)$, $r \geq 1$, and consider a random sample $\left(X_i^{(1)}, X_i^{(2)}\right)_{i=1,n}$, from the bivariate r.v. $\left(X^{(1)}, X^{(2)}\right)$. For each $j = 1, 2$, let $\mathbb{F}_{j:n}^* := n\mathbb{F}_{j:n}/(n+1)$ denotes the rescaled empirical df corresponds to the empirical df

$$\mathbb{F}_{n:j}(x_j) = n^{-1}\sum_{i=1}^n \mathbf{1}\left\{X_i^{(j)} \leq x_j\right\}.$$

For the sake of simplicity we suppose that the underlying copula C is symmetric and therefore $\delta_{k[12]} = \delta_{k[21]} = \delta_k$. The asymmetric case for the object of the future work.

We are now in position to present, in three steps, the semi-parametric BLM-based estimation:

- *Step* 1: For each $k = 1, ..., r$, compute

$$\widehat{\delta}_k = n^{-1} \sum_{i=1}^{n} \mathbb{F}^*_{1:n} \left(X_i^{(1)} \right) P_k \left(\mathbb{F}^*_{2:n} \left(X_i^{(2)} \right) \right). \tag{4.14}$$

  given in equation (4.4).

- *Step 2*: Using Theorem 4.3.1 to generate a system of r equations given by equation (4.5), for $k = 1, ..., r$.

- *Step 3:* Solve the system

$$\begin{cases} \delta_1 \left(\theta_1, ..., \theta_r \right) = \widehat{\delta}_1 \\ \delta_2 \left(\theta_1, ..., \theta_r \right) = \widehat{\delta}_2 \\ \vdots \\ \delta_r \left(\theta_1, ..., \theta_r \right) = \widehat{\delta}_r. \end{cases} \tag{4.15}$$

  The obtained solution $\hat{\boldsymbol{\theta}}^{BLM} := \left(\hat{\theta}_1, ..., \hat{\theta}_r \right)$ is called a BLM estimator for $\boldsymbol{\theta} = \left(\theta_1, ..., \theta_r \right)$.

The existence and the convergence of a solution of the previous system are established in Theorem 4.4.1, (see Section 4.4.2).

As an application of the BLM based estimation, we choose the one-iterated FGM copula C_{α_1, α_2} given in (4.8) and propose estimators for the parameters (α_1, α_2) noted $(\widehat{\alpha}_1, \widehat{\alpha}_2)$. For this family, recall (4.11), the system (4.15) becomes

$$\begin{cases} \alpha_1/18 + \alpha_2/72 = \widehat{\delta}_{1[12]}, \\ \alpha_2/120 = \widehat{\delta}_{2[12]}, \end{cases}$$

where $\widehat{\delta}_k, k = 1, 2$ are given in (4.14). Therefore

$$\begin{cases} \widehat{\alpha}_1 = 18\widehat{\delta}_{1[12]} - 30\widehat{\delta}_{2[12]}, \\ \widehat{\alpha}_2 = 120\widehat{\delta}_{2[12]}. \end{cases}$$

4.4.1 BLM as a rank approximate Z-estimation

Tsukahara (2005, [144]) introduced a new estimation method for copula models called the rank approximate Z-estimation (RAZ) that generalizes the PML one. The BLM method may be interpreted as a RAZ estimation. Indeed, let $\Psi(\cdot;\boldsymbol{\theta})$ be an $\mathbb{R}^r$-valued function on $\mathbb{I}^2$, called "score function", whose components $\Psi_j(\cdot;\boldsymbol{\theta})$ satisfy the condition

$$\int_{\mathbb{I}^2} \Psi_j(u_1, u_2; \boldsymbol{\theta})\, dC(u_1, u_2) = 0,\ \ j = 1, ..., r.$$

Any solution $\widehat{\boldsymbol{\theta}}^{RAZ}$ of the following equation

$$\sum_{i=1}^{n} \Psi\left(\mathbb{F}^*_{1:n}\left(X_i^{(1)}\right), \mathbb{F}^*_{2:n}\left(X_i^{(2)}\right); \widehat{\boldsymbol{\theta}}^{RAZ}\right) = 0, \tag{4.16}$$

is called a RAZ estimator. There may not be an exact solution to equation (4.16) in general, so in practice, we should choose $\widehat{\boldsymbol{\theta}}^{RAZ}$ to be any value of θ which minimizes the absolute value of the left-hand side of equation (4.16). It is worth mentioning that if the copula $C_{\boldsymbol{\theta}}$ is absolutely continuous with density $c_{\boldsymbol{\theta}}$, then the function $\Psi = \overset{\bullet}{c}_{\boldsymbol{\theta}}/c_{\boldsymbol{\theta}}$, with $\overset{\bullet}{c}_{\boldsymbol{\theta}} = (\partial c_{\boldsymbol{\theta}}/\partial\theta_j)_{j=1,...,r}$, leads the PML based estimation, see for instance Genest *et al.* (1995, [60]). Note in passing that the existence of a sequence of consistent roots of Z-estimation in this context is discussed in Theorem 1 in Tsukahara (2005, [144]). In our case $\Psi(u_1, u_2; \boldsymbol{\theta})$ corresponding to L_k (see (4.19)).

On the other hand, the population measures of concordance produce also a Z-estimation for copulas models. Indeed, the most popular population measures of concordance (see, Nelsen, 2006, [123, page 182]) are Kendall's tau (τ), Spearman's rho (ρ), Gini's gamma (γ) and Spearman's foot-rule phi (φ), given respectively by

$$\begin{aligned}
\tau(\boldsymbol{\theta}) &= 4\int_{\mathbb{I}^2} C_\theta(u_1, u_2)\, dC_{\boldsymbol{\theta}}(u_1, u_2) - 1,\\
\rho(\boldsymbol{\theta}) &= 12\int_{\mathbb{I}^2} u_1 u_2 dC_{\boldsymbol{\theta}}(u_1, u_2) - 3,\\
\gamma(\boldsymbol{\theta}) &= 4\int_{\mathbb{I}} C_{\boldsymbol{\theta}}(u_1, 1-u_1)\, du_1 - \int_{\mathbb{I}} (u_1 - C_{\boldsymbol{\theta}}(u_1, u_1))\, du_1,\\
\varphi(\boldsymbol{\theta}) &= 1 - 3\int_{\mathbb{I}^2} |u_1 - u_2|\, dC_{\boldsymbol{\theta}}(u_1, u_2).
\end{aligned}$$

It follows that the concordance score (CS) functions associated to τ, ρ, γ and φ respectively are

$$\Psi_1(u_1, u_2; \boldsymbol{\theta}) := 4C_{\boldsymbol{\theta}}(u_1, u_2) - \tau(\boldsymbol{\theta}),$$
$$\Psi_2(u_1, u_2; \boldsymbol{\theta}) := 12u_1u_2 - 3 - \rho(\theta),$$
$$\Psi_3(u_1, u_2; \boldsymbol{\theta}) := 4C_{\boldsymbol{\theta}}(u_1, 1-u_1) - u_1 + C_{\boldsymbol{\theta}}(u_1, u_1) - \gamma(\boldsymbol{\theta}),$$
$$\Psi_4(u_1, u_2; \boldsymbol{\theta}) := 1 - 3|u_1 - u_2| - \varphi(\boldsymbol{\theta}).$$

It is now clear that $\int_{\mathbb{I}^2} \Psi_j(u_1, u_2; \boldsymbol{\theta})\, dC_{\boldsymbol{\theta}}(u_1, u_2) = 0$, $j = 1, ..., 4$, then whenever the dimension of parameters $r = 4$, the function $\Psi = (\Psi_1, ..., \Psi_4)$ provides Z-estimators for copula models. If the dimension of parameters $r < 4$, then we may choose any r functions from $\Psi_1, ..., \Psi_4$ to have a system of r equations that provides estimators of the r parameters.

Tsukahara (2005, [144]) also discussed the RAZ-estimators based on Kendall's tau (τ) and Spearman's rho (ρ), called τ-score and ρ-score RAZ-estimators. Suppose that $r = 1$ and let $\widehat{\tau}_n$ and $\widehat{\rho}_n$ be the sample versions of Kendall's tau (τ) and Spearman's rho (ρ). By using the same idea as the method of moments, the τ-inversion $\widehat{\theta}_\tau$ estimator and the ρ-inversion $\widehat{\theta}_\rho$ estimator of θ are defined by

$$\widehat{\theta}_\tau = \tau^{-1}(\widehat{\tau}_n) \text{ and } \widehat{\theta}_\rho = \rho^{-1}(\widehat{\rho}_n). \tag{4.17}$$

In the case when $r = 2$, we may also estimate $\boldsymbol{\theta} = (\theta_1, \theta_2)$ by solving the system

$$\begin{cases} \tau(\theta_1, \theta_2) = \widehat{\tau}_n \\ \rho(\theta_1, \theta_2) = \widehat{\rho}_n. \end{cases}$$

Suppose that we are dealing with the estimation of parameters (α_1, α_2) of the one-iterated FGM copula C_{α_1,α_2} in (4.8). Then, the associated Kendall's tau (τ) and Spearman's rho (ρ) are

$$\begin{cases} \tau(\alpha_1, \alpha_2) = 2\alpha_1/9 + \alpha_2/18 + \alpha_1\alpha_2/450 \\ \rho(\alpha_1, \alpha_2) = \alpha_1/3 + \alpha_2/12. \end{cases} \tag{4.18}$$

We call (τ, ρ)-inversion estimator of parameters (α_1, α_2) a solution of the system

$$\begin{cases} \tau(\widehat{\alpha}_1, \widehat{\alpha}_2) = \widehat{\tau}_n \\ \rho(\widehat{\alpha}_1, \widehat{\alpha}_2) = \widehat{\rho}_n. \end{cases}$$

Similarly, if we consider the FGM family $C_{\alpha_1,\alpha_2,\alpha_3}$ we have to add $\gamma-$score and $\varphi-$score to have a system of four equations, we omit details.

The kth bivariate copula L-moments $\delta_k(\boldsymbol{\theta})$, $k = 1, ..., r$, may also generate score functions. Indeed, recall that from Theorem 4.3.1 we have

$$\delta_k(\boldsymbol{\theta}) = \int_{\mathbb{I}^2} u_1 P_k(u_2)\, dC_{\boldsymbol{\theta}}(u_1, u_2),\ k = 1, 2...,$$

and define the copula L-moment score (CLS) functions by

$$L_k(u_1, u_2; \boldsymbol{\theta}) := u_1 P_k(u_2) - \delta_k(\boldsymbol{\theta}),\ k = 1, ..., r, \tag{4.19}$$

satisfying $\int_{\mathbb{I}^2} L_k(u_1, u_2; \boldsymbol{\theta})\, dC_{\boldsymbol{\theta}}(u_1, u_2) = 0,\ k = 1, ..., r$. Then the RAZ estimator corresponding to the CLS function $\mathbf{L} = (L_1, ..., L_r)$ is a solution in $\boldsymbol{\theta}$ of the system

$$\sum_{i=1}^{n} \mathbf{L}\left(F_{1:n}^*\left(X_i^{(1)}\right), F_{2:n}^*\left(X_i^{(2)}\right); \boldsymbol{\theta}\right) = 0, \tag{4.20}$$

that is

$$\sum_{i=1}^{n} F_{1:n}^*\left(X_i^{(1)}\right) P_k\left(F_{2:n}^*\left(X_i^{(2)}\right)\right) - n\delta_k(\boldsymbol{\theta}) = 0,\ k = 1, ..., r, \tag{4.21}$$

therefore $\delta_k(\boldsymbol{\theta}) = \widehat{\delta}_k$, $k = 1, ..., r$, which in fact the system of bivariate copula L-moments given in system (4.15).

One of the main question of RAZ-estimation is the choice of the score function Ψ producing, in a certain sense, the best estimator. In Section (4.5), we show that the CLS functions improve the concordance score functions in terms of bias and root mean scare error (RMSE).

4.4.2 Asymptotic behavior of the BLM estimator

By considering BLM's estimator as a RAZ-estimator, a straight application of Theorem 1 in Tsukahara (2005, [144]) leads to the consistency and asymptotic normality of the considered estimator. Then we will state the following Theorem 4.4.1 without giving proofs. Indeed, let $\boldsymbol{\theta}_0$ be the true value of $\boldsymbol{\theta}$ and assume that the assumptions $[H.1] - [H.3]$ listed below are required.

- $[H.1]$ $\boldsymbol{\theta}_0$ is the unique zero of the mapping $\boldsymbol{\theta} \to \int_{\mathbb{I}^2} \mathbf{L}(u_1, u_2; \boldsymbol{\theta})\, dC_{\boldsymbol{\theta}}(u, v)$.

- $[H.2]$ $\mathbf{L}(\cdot;\boldsymbol{\theta})$ is differentiable with respect to $\boldsymbol{\theta}$ with the Jacobian matrix denoted by
$$\overset{\bullet}{\mathbf{L}}(u_1,u_2;\boldsymbol{\theta}) := \left[\frac{\partial L_k(u_1,u_2;\boldsymbol{\theta})}{\partial \theta_k}\right]_{r\times r},$$
$\overset{\bullet}{\mathbf{L}}(u_1,u_2;\boldsymbol{\theta})$ is continuous in $\boldsymbol{\theta}$, and the Euclidian norm $\left\|\overset{\bullet}{\mathbf{L}}(u_1,u_2;\boldsymbol{\theta})\right\|$ is dominated by a $dC_{\boldsymbol{\theta}_0}$-integrable function $h(u_1,u_2)$.

- $[H.3]$ The $r\times r$ matrix $A_0 := \int_{\mathbb{I}^2} \overset{\bullet}{\mathbf{L}}(u_1,u_2;\boldsymbol{\theta}_0)\, dC_{\boldsymbol{\theta}_0}(u_1,u_2)$ is nonsingular.

Theorem 4.4.1 *Assume that the assumptions $[H.1]-[H.3]$ hold. Then with probability tending to one as $n\to\infty$, there exists a solution $\widehat{\boldsymbol{\theta}}^{BLM}$ to the equation (4.21) which converges to $\boldsymbol{\theta}_0$. Moreover*

$$\sqrt{n}\left(\widehat{\boldsymbol{\theta}}^{BLM}-\boldsymbol{\theta}_0\right)\xrightarrow{\mathcal{D}}\mathcal{N}\left(0, A_0 \textstyle\sum_0 A_0^{-1}\right),\ \text{as } n\to\infty,$$

with

$$\textstyle\sum_0 := var\left\{\mathbf{L}(\xi_1,\xi_2;\boldsymbol{\theta}_0)+\sum_{j=1}^{2}\int_{\mathbb{I}^2}\mathbf{M}_j(u_1,u_2)\left(\mathbf{1}\{\xi_j\le u_j\}-u_j\right)dC_{\boldsymbol{\theta}_0}(u_1,u_2)\right\},$$

where (ξ_1,ξ_2) is a bivariate r.v. with joint distribution function C_{θ_0},

$$\mathbf{M}_1(u_1,u_2) := \{P_k(u_2)\}_{k=1,r}\ \text{and}\ \mathbf{M}_2(u_1,u_2) := \{u_1 P_k'(u_2)\}_{k=1,r},$$

where P_k' denotes the derivative of the polynomials P_k.

4.4.3 A discussion on Theorem 4.1

Notice that assumption $[H.1]$ is verified for any parametric copula $C_{\boldsymbol{\theta}}$ satisfying the concordance ordering condition of copulas (**??**). Indeed, suppose that there exists $\boldsymbol{\theta}_1 \neq \boldsymbol{\theta}_0$, such that

$$\int_{\mathbb{I}^2} L_k(u_1,u_2;\boldsymbol{\theta}_1)\, dC_{\boldsymbol{\theta}_0}(u_1,u_2)=0,\ \text{for every } k\in\{1,...,r\}. \tag{4.22}$$

Recall that $\delta_{k[12]}(\boldsymbol{\theta}_0)=\int_{\mathbb{I}^2}u_1P_k(u_2)\,dC_{\boldsymbol{\theta}_0}(u_1,u_2)$ and $\delta_{k[12]}(\boldsymbol{\theta}_1)=\int_{\mathbb{I}^2}u_1P_k(u_2)\,dC_{\boldsymbol{\theta}_1}(u_1,u_2)$ and, from assumption (**??**), $C_{\boldsymbol{\theta}_0}$ ($>$ or $<$) $C_{\boldsymbol{\theta}_1}$. It follows, by monotonicity of the integral, that $\delta_{k[12]}(\boldsymbol{\theta}_1)$ ($>$ or $<$) $\delta_{k[12]}(\boldsymbol{\theta}_0)$, this implies that $\delta_{k[12]}(\boldsymbol{\theta}_1)-\delta_{k[12]}(\boldsymbol{\theta}_0)\neq$

0, for every $k \in \{1,...,r\}$. Observe that $\int_{\mathbb{I}^2} dC_{\boldsymbol{\theta}_0}(u_1,u_2) = 1$ then $\delta_{k[12]}(\boldsymbol{\theta}_1) = \int_{\mathbb{I}^2} \delta_{k[12]}(\boldsymbol{\theta}_1)\, dC_{\boldsymbol{\theta}_0}(u_1,u_2)$, consequently, since $L_k(u_1,u_2;\boldsymbol{\theta}_1) = u_1 P_k(u_2) - \delta_{k[12]}(\boldsymbol{\theta}_1)$, that

$$\int_{\mathbb{I}^2} L_k(u_1,u_2;\boldsymbol{\theta}_1)\, dC_{\boldsymbol{\theta}_0}(u_1,u_2) = \int_{\mathbb{I}^2} u_1 P_k(u_2)\, dC_{\boldsymbol{\theta}_0}(u_1,u_2) - \int_{\mathbb{I}^2} \delta_{k[12]}(\boldsymbol{\theta}_1)\, dC_{\boldsymbol{\theta}_0}(u_1,u_2)$$
$$= \delta_{k[12]}(\boldsymbol{\theta}_0) - \delta_{k[12]}(\boldsymbol{\theta}_1) \neq 0, \text{ for every } k \in \{1,...,r\},$$

which is a contradiction with equation (4.22), as sought. Let's now discuss the rest of assumptions. In $[H.2]$, the continuity and the differentiability with respect to $\boldsymbol{\theta}$ and (u_1,u_2) of $\mathbf{L}(\cdot;\boldsymbol{\theta})$ and $\dot{\mathbf{L}}(\cdot;\boldsymbol{\theta})$ are lie with that of copula $C_{\boldsymbol{\theta}}$, which are natural assumptions in parametric copula models. Some examples on this issue are illustrated in Fredricks *et al.* (2007). The second part of $[H.2]$ and $[H.3]$ may be checked for a given copula model. For example, if we consider the FGM family (see (4.7)) we get $L_1(u_1,u_2;\alpha) = u_1(2u_2-1) - \alpha/18$ and $L_k(u_1,u_2;\alpha) = 0$, for $k = 2,3,...$Then $dL_1(u_1,u_2;\alpha)/d\alpha = -1/18$ and $dL_k(u_1,u_2;\alpha)/d\alpha = 0$, for $k = 2,3,...$ Let α_0 denote the true value of parameter α. It is clear that each compound of $\dot{\mathbf{L}}$ is continuous with respect to α and (u_1,u_2), $\left|\dot{\mathbf{L}}(u_1,u_2;\alpha)\right| = 1/18$, which is $C_\alpha-$integrable function and, $A_0 = \int_{\mathbb{I}^2} \dot{\mathbf{L}}(u_1,u_2;\alpha_0)\, dC_{\alpha_0} = -1/18$ which is nonsingular matrix, then the assumptions $[H.2]$ and $[H.3]$ are well verified. By a little algebra we get to the corresponding value of $\sum_0$ that is defined in (**??**), and by Theorem 4.4.1 we get

$$\sqrt{n}\left(\widehat{\alpha}^{BLM} - \alpha_0\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \alpha_0^2/270 + 1/5\right), \text{ as } n \to \infty.$$

For the one-iterated FGM family (see (4.8)), by letting $\alpha_1 = \alpha$ and $\alpha_2 = \beta$, it is readily to verify that

$$L_1(u_1,u_2;\alpha,\beta) = u_1(2u_2-1) - \alpha/18 - \beta/72,$$
$$L_2(u_1,u_2;\alpha,\beta) = u_1\left(6u_2^2 - 6u_2 + 1\right) - \beta/120,$$

which, obviously, are continuous with respect to (α,β) and (u_1,u_2) and $C_{\alpha,\beta}-$integrable function, and

$$\dot{\mathbf{L}}(u_1,u_2;\alpha,\beta) = \begin{bmatrix} -1/18 & -1/72 \\ 0 & -1/120 \end{bmatrix}.$$

Let (α_0, β_0) denote the true value of parameter (α, β) and by calculating the elements of the matrix

$$A_0 := \int_{\mathbb{I}^2} \overset{\bullet}{\mathbf{L}} (u_1, u_2; \alpha_0, \beta_0)\, dC_{\alpha_0,\beta_0},$$

we get

$$A_0 = \begin{bmatrix} -1/18 & 1/72 \\ 0 & -1/120 \end{bmatrix},$$

which is nonsingular because its determinant equals $1/2160 \neq 0$, therefore $[H.2]$ and $[H.3]$ are also verified. Then, in view of Theorem 4.4.1, we have

$$\sqrt{n}\left\{\begin{pmatrix} \widehat{\alpha}^{BLM} \\ \widehat{\beta}^{BLM} \end{pmatrix} - \begin{pmatrix} \alpha_0 \\ \beta_0 \end{pmatrix}\right\} \overset{\mathcal{D}}{\to} \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \sum\nolimits^2\right), \text{ as } n \to \infty,$$

where $\sum^2 := A_0^{-1} \sum_0 \left(A_0^{-1}\right)^T$. After a tedious computation we get

$$\sum\nolimits_0 = \begin{bmatrix} \dfrac{\alpha_0^2}{270} + \dfrac{\alpha_0\beta_0}{540} + \dfrac{\beta_0^2}{3780} + \dfrac{1}{5} & \dfrac{\beta_0^2}{8640} + \dfrac{\alpha_0\beta_0}{2160} \\ \dfrac{\beta_0^2}{8640} + \dfrac{\alpha_0\beta_0}{2160} & \dfrac{\alpha_0^2}{105} + \dfrac{\alpha_0\beta_0}{252} + \dfrac{17\beta_0^2}{21000} + \dfrac{1}{15} \end{bmatrix},$$

it follows that

$$\sum\nolimits^2 = \begin{bmatrix} \dfrac{342\alpha_0^2}{35} + \dfrac{327\alpha_0\beta_0}{70} + \dfrac{263\beta_0^2}{280} + \dfrac{624}{5} & \dfrac{240\alpha_0^2}{7} + \dfrac{107\alpha_0\beta_0}{7} + \dfrac{443\beta_0^2}{140} + 240 \\ \dfrac{240\alpha_0^2}{7} + \dfrac{107\alpha_0\beta_0}{7} + \dfrac{443\beta_0^2}{140} + 240 & \dfrac{960\alpha_0^2}{7} + \dfrac{400\alpha_0\beta_0}{7} + \dfrac{408\beta_0^2}{35} + 960 \end{bmatrix}.$$

Finally, we note that assumptions $[H.1] - [H.3]$ may be also verified for one and two parameters copula families given in (**??**) and (4.12), respectively, but that requires tedious calculations which would get us out of the context of the paper.

4.5 Simulation study

To check and compare the performance of BLM's estimator with PML, (τ, ρ) $-$inversion (RAZ) and the minimum distance (MD) methods (see the Appendix for MD method), a simulation study is carried out with $r = 2$ by considering C_{α_1,α_2} (the one iterated

FGM family) and C_{β_1,β_2} (the two parameters Gumbel family) given in (4.11) and (4.12) respectively. The evaluation of the performance is based on the bias and the RMSE defined as follows:

$$\text{Bias} = \frac{1}{N}\sum_{i=1}^{N}\left(\hat{\theta}_i - \theta\right), \ \text{RMSE} = \left(\frac{1}{N}\sum_{i=1}^{N}\left(\hat{\theta}_i - \theta\right)^2\right)^{1/2}, \tag{4.23}$$

where $\hat{\theta}_i$ is an estimator (from the considered method) of θ from the ith samples for N generated samples from the underlying copula. In both parts, we selected $N = 1000$. We compare the BLM estimator with the PML, RAZ and MD estimators. The procedure outlined in Section (4.4) is repeated for different sample sizes n with $n = 30, 50, 100, 500$ to assess the improvement in the bias and RMSE of the estimators with increasing sample size. Furthermore, the simulation procedure is repeated for a large set of parameters of the true copulas C_{α_1,α_2} and C_{β_1,β_2}. For each sample, we solve systems (4.11) and (4.13) to obtain, respectively, the BLM-estimators $(\widehat{\alpha}_{1,i}, \widehat{\alpha}_{2,i})$ and $\left(\widehat{\beta}_{1,i}, \widehat{\beta}_{2,i}\right)$ of (α_1, α_2) and (β_1, β_2) for $i = 1, ..., N$, and the estimators $\widehat{\alpha}_k$, $\widehat{\beta}_k$ for $k = 1, 2$ are given by $\widehat{\alpha}_k = \frac{1}{N}\sum_{i=1}^{N}\widehat{\alpha}_{k,i}$ and $\widehat{\beta}_k = \frac{1}{N}\sum_{i=1}^{N}\widehat{\beta}_{k,i}$.

4.5.1 Performance of the BLM-based estimation

We first select parameters, as the true values of the parameters, of Gumbel and FGM copula models. The choice of the parameters have to be meaningful, in the sense that each couple of parameters assigns a value of one of the dependence measure, that is weak, moderate and strong dependence. In other words, if we consider Spearman's rho ρ as a dependence measure, then we should select values for copula parameters that correspond to specified values of ρ by using equation (4.6). Recall that for the FGM family C_{α_1,α_2}, the dependence reaches the maximum $\rho_{FGM}^{\max} = 0.42721$ in $\alpha_1 = -1+7/\sqrt{13} \approx 0.941$ and $\alpha_2 = 2-2/\sqrt{13} \approx 1.445$ (see (4.10)). So, we may chose $(\alpha_1, \alpha_2) = (0.941, 1.445)$ as the true parameters of FGM family that correspond to the strong dependence. For the true values of (α_1, α_2) corresponding to the weak and the moderate dependence, we proceed as follows. We assign a value to the couple (ρ, α_1) such that $|\alpha_1| \leq 1$, then we solve by numerical methods the equation (4.6) in the region (4.9) and get the corresponding value to α_2. We summarize the results in the following table:

By the same procedure, we select the true parameters (β_1, β_2) of the Gumbel copula C_{β_1,β_2} and get:

To evaluate the performance of the BLM estimators, we proceed as follows:

ρ	α_1	α_2
0.001	0.100	0
0.208	0.400	0.900
0.427	0.941	1.445

Table 4.1: The true parameters of FGM copula used for the simulation study .

ρ	β_1	β_2
0.001	1	0.001
0.500	1.400	0.200
0.900	2.500	1

Table 4.2: The true parameters of Gumbel copula used for the simulation study.

1. By using the Algorithm in Nelsen 2006, [123, page 41] and the Theorem 4.3.7 in Nelsen 2006, [123, page 129], respectively, we generate twice N samples of size n from each one the considered copulas C_{α_1,α_2} and C_{β_1,β_2}.
2. Obtain the BML estimators $(\widehat{\alpha}_1, \widehat{\alpha}_2)$ of (α_1, α_2) and $\left(\widehat{\beta}_1, \widehat{\beta}_2\right)$ of (β_1, β_2) .
3. By computing, for each estimator, the appropriate Bias and RMSE, we compare $(\widehat{\alpha}_1, \widehat{\alpha}_2)$ and $\left(\widehat{\beta}_1, \widehat{\beta}_2\right)$, respectively, with the true parameters (α_1, α_2) and (β_1, β_2) .

All computations were performed in R Software version 2.10.1. The results of the simulation study are summarized in Tables 4.3 and 4.4. We observed that BLM's method product, in terms of bias and RMSE, reasonable results, notably when the sample size increases. However, in the case of strong dependence for FGM's family when the sample size is small and less than 30, the estimation of the first parameter α_1 is better that of the second one α_2. However, for the sample sizes greater than 100 the results become reasonable and more better for sample sizes greater than $n = 500$. For Gumbel family the performance of BLM's method looks good even for small samples.

4.5.2 Comparative study: BLM, RAZ, MD and PML

As the previous Subsection, we consider the bivariate two-parameter FGM and Gumbel copula families with the trues parameters those given in Tables 4.1 and 4.2 respectively. The simulation study proceeds as follows:

1. Generate N samples of size $n = 30$, 50, 100, 500 from the copula C_θ.

2. Assess the performances of the BLM, RAZ, MD and PML estimators.

3. Compare the BLM, RAZ, MD and PML estimators with the true parameter $\boldsymbol{\theta}$ by computing, for each estimator, the appropriate criteria given by (4.23) .

It is clear, from Tables 4.5 to 4.10, that the BLM estimate preforms better than the RAZ, MD and PML ones as far as the Bias is concerned. On the other hand, in the case of small samples the RAZ, MD and PLM methods give better RMSE than the BLM one. However, when the sample size increases, the RMSE of the BLM estimator becomes reasonable. Moreover, for the computation time point of view, we observed that the RAZ, MD and PLM estimates require hours to be obtained, notably when the sample size becomes large, whereas the BLM estimate execution time is in terms of minutes. This is a natural conclusion, because the RAZ, MD and PLM methods use the optimization problem under constraints, while the BLM method uses systems of equations.

4.5.3 Comparative robustness study: BLM, RAZ, MD and PML

In this subsection we study the sensitivity to outliers of BLM's estimator and compare with those of the RAZ, MD and PML ones. We consider an $\epsilon-$contaminated model for two-parameters FGM familly by means of a copula from the same family. In other terms, we are dealing with the following mixture copula model:

$$C_{\alpha_1,\alpha_2}(\epsilon) := (1-\epsilon)\, C_{\alpha_1,\alpha_2} + \epsilon C_{\alpha_1^*,\alpha_2^*}, \tag{4.24}$$

where $0 < \epsilon < 1$ is the amount of contamination. For the implementation of mixtures models to the study outliers one refers, for instance, to Barnett and Lewis (1994[**?**]), page 43. In this context, we proceed our study as follows. First, we select $(\alpha_1, \alpha_2) = (0.4, 0.9)$ corresponds to Spearman's Rho $\rho = 0.208$ (see Table 4.1) and chose $(\alpha_1^*, \alpha_2^*) = (0, 0)$ to have the contamination model as the product copula that is $C_{\alpha_1^*,\alpha_2^*}(u, v) = uv$. Then we consider four contamination scenarios according to $\epsilon = 5\%, 10\%, 20\%, 30\%$. For each value ϵ, we generate 1000 samples of size $n = 40$ from the copula $C_{\alpha_1,\alpha_2}(\epsilon)$. Finally, we compare the BLM, RAZ, MD and PML estimators with the true parameter (α_1, α_2) by computing, for each estimator, the appropriate Bias and RMSE and summarize the results in Table 4.11. We observed that, for example, in 0% contamination the (Bias, RMSE) of $\widehat{\alpha}_1$ equals $(0.044, 0.832)$, while for 30% contamination is $(-0.165, 0.835)$. We may conclude that the RMSE of BLM's estimation in less sensitive (or robust) to outliers, however the Bias is not. The same conclusion is for the RAZ method but the BLM's one is better. For PLM's estimation both the Bias and the RMSE are sensitive, indeed for 0% contamination the (Bias, RMSE) of $\widehat{\alpha}_1$ equals $(-0.238, 0.440)$, while for 30% contamination is $(-0.328, 0.589)$. Both the bias and the RMSE of MD's estimation are note sensitive to outliers, then we may conclude that is the better among the four estimation methods. However, the computation time cost in MD's method is important which is considered as an handicap from practitioners.

	$\rho = 0.001$				$\rho = 0.208$				$\rho = 0.427$			
	$\alpha_1 = 0.1$		$\alpha_2 = 0$		$\alpha_1 = 0.4$		$\alpha_2 = 0.9$		$\alpha_1 = 0.941$		$\alpha_2 = 1.445$	
n	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
30	0.197	0.882	−0.089	2.417	0.113	0.931	−0.210	2.804	0.098	0.823	−0.312	2.861
50	0.133	0.672	−0.050	1.781	0.042	0.703	−0.065	2.074	−0.051	0.712	0.276	2.197
100	0.065	0.456	−0.040	1.105	0.026	0.498	0.048	1.408	0.041	0.513	−0.055	1.572
500	−0.017	0.206	0.041	0.639	0.021	0.215	0.031	0.659	−0.020	0.308	−0.031	0.692

Table 4.3: Bias and RMSE of BLM's estimator of two-parameters FGM copula.

	$\rho = 0.001$				$\rho = 0.5$				$\rho = 0.9$			
	$\beta_1 = 1$		$\beta_2 = 0.001$		$\beta_1 = 1.4$		$\beta_2 = 0.2$		$\beta_1 = 2.5$		$\beta_2 = 1$	
n	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
30	0.162	0.994	0.428	1.945	0.214	1.002	0.549	1.421	0.404	0.920	−0.653	1.109
50	0.134	0.725	0.294	1.107	0.187	0.695	0.498	0.999	0.350	0.854	−0.550	0.835
100	−0.094	0.697	0.219	0.804	−0.136	0.619	0.287	0.665	0.183	0.597	−0.536	0.526
500	−0.071	0.597	0.107	0.358	−0.081	0.489	0.148	0.477	−0.096	0.395	−0.340	0.480

Table 4.4: Bias and RMSE of BLM's estimator of two-parameters FGM copula.

	$\alpha_1 = 0.1$		$\alpha_2 = 0$		
	Bias	RMSE	Bias	RMSE	Time (h)
		$n = 30$			
BLM	0.227	0.952	−0.194	2.882	0.640
RAZ	0.458	1.005	0.782	2.157	0.978
MD	0.575	0.571	0.494	0.851	1.566
PML	0.550	0.552	0.424	0.872	1.033
		$n = 50$			
BLM	−0.140	0.702	−0.112	2.193	1.215
RAZ	0.358	0.958	0.558	1.428	1.856
MD	0.468	0.559	0.238	0.846	3.455
PML	0.444	0.546	−0.237	0.840	2.421
		$n = 100$			
BLM	−0.039	0.565	0.082	1.364	1.847
RAZ	0.229	0.664	0.195	0.985	2.548
MD	0.125	0.521	0.145	0.684	6.888
PML	0.121	0.520	0.131	0.673	4.107
		$n = 500$			
BLM	0.021	0.417	0.071	0.634	8.963
RAZ	0.084	0.588	0.118	0.748	11.548
MD	0.077	0.504	0.086	0.640	19.598
PML	0.076	0.502	0.081	0.639	17.073

Table 4.5: Bias and RMSE of the BLM, RAZ, MD and PML estimators for two-parameters of FGM copula for weak dependence ($\rho = 0.001$).

4.6 Conclusions

In this paper, a formula of the bivariate L-moments in terms of copulas is given. This formula leads to introduce a new estimation method for bivariate copula parameters, that we called the BLM based estimation. The limiting distribution of the estimators given by the BLM method are established. Moreover, we compared by simulations the BLM method with the well-known (τ, ρ)-inversion (RAZ), the minimum distance (MD) and the pseudo maximum likelihood (PML) estimators by focusing on the Bias and the RMSE. We conclude that the BLM based estimation performs well the Bias and reasonably the RMSE. However, BLM's method may be an alternative robust method as far as the RMSE is concerned. As finale conclusion, it is worth noting that computation's time of the proposed method is quite small compared to MD and PML ones.

	$\alpha_1 = 0.4$		$\alpha_2 = 0.9$		
	Bias	RMSE	Bias	RMSE	Time (h)
		$n = 30$			
BLM	0.127	0.855	−0.297	2.668	0.011
RAZ	0.102	1.322	−0.290	1.383	0.016
MD	−0.174	0.777	−0.322	1.058	3.583
PML	−0.191	0.906	−0.372	1.261	0.954
		$n = 50$			
BLM	−0.059	0.755	0.123	2.001	1.035
RAZ	0.091	0.892	−0.141	1.272	1.101
MD	−0.173	0.730	−0.223	1.010	6.428
PML	0.122	0.775	−0.200	0.853	1.823
		$n = 100$			
BLM	0.031	0.715	0.060	1.404	1.920
RAZ	0.082	0.791	−0.130	0.942	1.037
MD	−0.130	0.652	−0.121	0.919	11.217
PML	0.090	0.599	0.100	0.794	2.652
		$n = 500$			
BLM	−0.025	0.300	0.049	0.629	9.205
RAZ	0.054	0.393	−0.087	0.701	8.285
MD	−0.071	0.602	−0.061	0.742	19.210
PML	0.047	0.573	0.056	0.632	16.458

Table 4.6: Bias and RMSE of the BLM, RAZ, MD and PML estimators for two-parameters of FGM copula for moderate dependence ($\rho = 0.208$).

	$\alpha_1 = 0.941$		$\alpha_2 = 1.445$		
	Bias	RMSE	Bias	RMSE	Time (h)
		$n = 30$			
BLM	0.091	0.832	0.402	2.715	0.017
RAZ	0.171	1.142	−0.471	1.229	0.057
MD	−0.142	0.871	−0.420	1.025	2.083
PML	−0.121	0.927	−0.415	1.061	0.781
		$n = 50$			
BLM	0.054	0.641	0.300	1.982	1.020
RAZ	0.157	0.997	−0.321	1.120	1.021
MD	−0.135	0.753	0.351	0.940	6.633
PML	0.092	0.892	−0.307	1.150	1.754
		$n = 100$			
BLM	0.030	0.449	0.090	1.391	1.620
RAZ	0.081	0.463	−0.153	0.931	1.037
MD	0.070	0.743	−0.114	0.904	9.217
PML	0.050	0.712	0.102	0.800	2.652
		$n = 500$			
BLM	0.021	0.315	0.046	0.602	9.205
RAZ	0.071	0.357	−0.098	0.765	8.285
MD	−0.064	0.541	−0.054	0.782	19.210
PML	0.052	0.472	0.076	0.699	17.458

Table 4.7: Bias and RMSE of the BLM, RAZ, MD and PML estimators for two-parameters of FGM copula for strong dependence ($\rho = 0.427$).

	$\beta_1 = 1$		$\beta_2 = 0.001$		
	Bias	RMSE	Bias	RMSE	Time (h)
		$n = 30$			
BLM	0.174	0.941	0.453	1.854	1.121
RAZ	0.181	0.782	0.532	1.186	1.021
MD	−0.274	0.546	−0.698	1.243	4.691
PML	0.310	0.335	−0.593	0.910	1.065
		$n = 50$			
BLM	−0.157	0.897	0.289	0.977	1.026
RAZ	0.184	0.539	0.476	0.629	1.265
MD	0.262	0.448	−0.310	0.759	3.633
PML	0.250	0.303	−0.302	0.815	2.754
		$n = 100$			
BLM	−0.126	0.530	0.193	0.824	1.920
RAZ	−0.177	0.523	0.250	0.619	2.248
MD	−0.161	0.420	−0.201	0.521	6.285
PML	0.151	0.272	−0.197	0.810	4.153
		$n = 500$			
BLM	−0.098	0.411	0.114	0.324	9.010
RAZ	−0.235	0.502	0.136	0.503	7.149
MD	−0.181	0.409	0.116	0.376	14.984
PML	0.170	0.205	−0.115	0.619	13.147

Table 4.8: Bias and RMSE of the BLM, RAZ, MD and PML estimators for two-parameters of Gumbel copula for weak dependence ($\rho = 0.001$).

	$\beta_1 = 1.4$		$\beta_2 = 0.2$		
	Bias	RMSE	Bias	RMSE	Time (h)
		$n = 30$			
BLM	−0.182	0.989	0.593	1.317	1.024
RAZ	0.191	0.885	0.655	1.215	1.042
MD	0.214	0.985	−0.525	1.056	6.485
PML	0.195	0.524	−0.423	1.051	1.125
		$n = 50$			
BLM	−0.134	0.594	0.526	0.994	1.058
RAZ	0.187	0.512	0.555	0.972	1.012
MD	0.181	0.423	−0.461	0.853	2.588
PML	0.177	0.318	−0.413	0.916	2.859
		$n = 100$			
BLM	−0.122	0.482	0.272	0.492	2.247
RAZ	−0.150	0.421	0.291	0.712	3.153
MD	−0.170	0.439	−0.269	0.474	6.256
PML	0.152	0.293	−0.275	0.471	5.254
		$n = 500$			
BLM	−0.101	0.223	0.135	0.312	11.587
RAZ	−0.149	0.400	0.221	0.655	8.145
MD	−0.106	0.306	−0.212	0.355	14.445
PML	0.102	0.221	−0.200	0.317	13.157

Table 4.9: Bias and RMSE of the BLM, RAZ, MD and PML estimators for two-parameters of Gumbel copula for moderate dependence ($\rho = 0.5$).

	$\beta_1 = 2.5$		$\beta_2 = 1$		
	Bias	RMSE	Bias	RMSE	Time (h)
		$n = 30$			
BLM	0.422	0.954	−0.740	1.119	0.955
RAZ	0.786	1.125	0.782	1.175	1.054
MD	0.546	0.546	0.592	0.563	1.245
PML	0.553	0.551	0.723	0.522	1.165
		$n = 50$			
BLM	0.329	0.817	−0.635	0.852	1.021
RAZ	0.586	0.983	0.745	0.972	1.245
MD	0.321	0.522	0.582	0.552	2.265
PML	0.292	0.512	0.551	0.514	2.255
		$n = 100$			
BLM	0.107	0.584	−0.592	0.713	1.920
RAZ	0.425	0.812	0.611	0.902	2.153
MD	−0.181	0.501	−0.578	0.488	5.544
PML	0.172	0.482	−0.545	0.472	5.458
		$n = 500$			
BLM	−0.066	0.456	−0.367	0.478	9.205
RAZ	0.123	0.757	0.501	0.694	8.789
MD	0.094	0.469	0.408	0.495	14.565
PML	0.084	0.465	0.375	0.482	13.425

Table 4.10: Bias and RMSE of the BLM, RAZ, MD and PML estimators for two parameters of Gumbel copula for strong dependence ($\rho = 0.9$).

	α_1= 0.4		α_2= 0.9	
0% contamination				
	Bias	RMSE	Bias	RMSE
BLM	0.044	0.832	−0.141	2.650
RAZ	0.053	0.432	−0.440	0.711
MD	0.267	0.270	−0.456	0.461
PML	−0.238	0.440	−0.472	0.627
5% contamination				
BLM	0.046	0.833	−0.137	2.662
RAZ	−0.069	0.431	−0.479	0.738
MD	0.254	0.257	−0.472	0.475
PML	−0.274	0.407	0.432	0.613
10% contamination				
BLM	−0.082	0.811	−0.155	2.641
RAZ	−0.090	0.393	−0.461	0.695
MD	0.279	0.281	−0.464	0.468
PML	−0.267	0.506	−0.429	0.637
20% contamination				
BLM	−0.100	0.802	−0.188	2.585
RAZ	−0.130	0.423	−0.537	0.786
MD	0.280	0.282	−0.472	0.477
PML	−0.268	0.524	−0.500	0.639
30% contamination				
BLM	−0.165	0.835	−0.280	2.627
RAZ	−0.179	0.480	−0.619	0.909
MD	0.293	0.266	−0.458	0.465
PML	−0.328	0.589	−0.515	0.641

Table 4.11: Bias and RMSE of the BLM, RAZ, MD and PML estimators for ϵ-contaminated two-parameters of FGM copula by product copula.

Chapter 5

Distortion risk measures for sums of dependent losses

«A mathematician is a machine for turning coffee into theorems.»

Paul Erdös, (1913-1996)

We discuss two distinct approaches, for distorting risk measures of sums of dependent random variables, which preserve the property of coherence. The first, based on distorted expectations, operates on the survival function of the sum. The second, simultaneously applies the distortion on the survival function of the sum and the dependence structure of risks, represented by copulas. Our goal is to propose risk measures that take into account the fluctuations of losses and possible correlations between random variables. For more detail see Brahimi et al. (2010, [10]).

Keywords: Coherence, Distortion parameter, Dependence structure, Heavy-tailed risks, Insurance premium, Wang transform.

5.1 Introduction

Risk measures are used to quantify insurance losses and measuring financial risk assessments. Several risk measures have been proposed in actuarial science literature, namely: the Value-at-Risk (VaR), the expected shortfall or the conditional tail expectation (CTE), and the distorted risk measures (DRM). Before introducing and interpreting the DRM, it is necessary to fix a convention of profit and loss appropriate to the application to market finance, credit risk and insurance. Let X be a random variable (rv), representing losses (or gains) of a company, with a continuous distribution function (df) F. The DRM of rv X, due to Wang (1995,[149]), is defined as follows:

$$\pi_\psi [X] := \int_0^\infty \psi (1 - F(x))\, dx, \qquad (5.1)$$

where ψ is a non-decreasing function, called distortion function, satisfying $\psi(0) = 0$ and $\psi(1) = 1$. In the actuarial literature the following functions are frequently used:

$$\psi_\rho (s) = s^\rho, \qquad \text{for} \quad 0 < \rho \leq 1,$$

$$\psi_\kappa (s) = \phi\left(\phi^{-1}(s) + \kappa\right), \qquad \text{for} \quad 0 \leq \kappa < \infty,$$

$$\psi_\zeta (s) = \min\left(s/(1-\zeta), 1\right) \quad \text{for} \quad 0 \leq \zeta < 1,$$

$$\psi_\alpha (s) = s^\alpha (1 - \alpha \ln s), \qquad \text{for} \quad 0 < \alpha \leq 1,$$

where $\phi^{-1}(u) := \inf\{x : \phi(x) \geq u\}$ is the quantile function of the standard normal distribution ϕ. Constants ρ, κ, ζ and α are called distortion parameters. The functions ψ_ρ, ψ_κ, ψ_ζ and ψ_α respectively give rise to the so-called proportional hazard transform (PHT) (Wang, 1995), the normal transform (Wang, 2000, [151]), the CTE and the look-back distortion (Hürlimann, 1998, [82]). When $\rho = 1$ and $\kappa = \zeta = 0$, there is no distortion and the corresponding DRM is equal to the expectation of X. For recent literature on risk measures one refers to Denuit *et al.* (2005, [31]) and Furman and Zitikis (2008a[52] , 2008b[53]).

The problem of the axiomatic foundation of risk measures has received much attention starting with the seminal paper of Artzner *et al.* (1999, [5]), where the definition of coherent risk measure was first provided. A coherent risk measure is a real functional μ, defined on a space of rv's, satisfying the following axioms:

H1. bounded from above by the maximum loss: $\mu(X) \leq \max(X)$.
H2. bounded from below by the mean loss: $\mu(X) \geq \mathbb{E}(X)$.
H3. scalar additive and multiplicative: $\mu(aX + b) = a\mu(X) + b$, for $a, b \geq 0$.
H4. subadditivity: $\mu(X + Y) \leq \mu(X) + \mu(Y)$.

The only axiom that a DRM may lack in order to be a coherent risk measure in the sense of Artzner *et al.* (1999, [5]) is H4. However, the subadditivity theorem of Choquet integrals (Denneberg, 1994, [30]) guarantees that $\mu(X+Y) \leq \mu(X) + \mu(Y)$ if and only if the distortion function ψ is concave. Hence, the DRM $\pi_\psi[X]$ defined in (5.1) with a concave distortion ψ is coherent. It is well known that the CTE and the PHT are examples of concave distortion risk measures, whereas the VaR is not.

In traditional risk theory, individual risks have been usually assumed to be independent. Traceability for this assumption is very convenient, but not realistic. Recently in the actuarial science, the study of the impact of dependence among risks has become a major and flourishing topic. Several notions of dependence were introduced to model the fact that larger values of one component of a multivariate risk tend to be associated with larger values of the others. In this paper, we deal with a vector of risk losses $\mathbf{X} = (X^{(1)}, ..., X^{(d)})$, $d \geq 2$ and we discuss the computation of the DRM of the sum Z of its components. When $X^{(1)}, ..., X^{(d)}$ are independent and identically distributed, their sum is considered as a rv whose df G is the convolution of the marginal distributions of $\mathbf{X}$. In this case, the DRM value of Z, for a given distortion function ψ may be obtained via formula (5.1), that is

$$\pi_{\psi}[Z] := \int_{0}^{\infty} \psi(1 - G(z))\, dz. \tag{5.2}$$

Now, assume that $X^{(1)}, ..., X^{(d)}$ are dependent with joint df H and continuous margins F_i, $i = 1, ..., d.$ In this case, the problem becomes different and its resolution requires more than the usual background. Several authors discussed the DRM, when applied to sums of rv's, against some classical dependency measures such as Person's r, Spearman's ρ and Kendall's τ, see for instance, Darkiewicz *et al.* (2004, [23]) and Burgert and Rüschendorf (2006, [11]). Our contribution is to introduce the copula notion to provide more flexibility to the DRM of sums of rv's in terms of loss and dependence structure. For comprehensive details on copulas one may consult the textbook of Nelsen (2006, [123]). According to Sklar's Theorem (Sklar, 1959[142]), there exists a unique copula $C : [0,1]^d \to [0,1]$ such that

$$H(x_1, ..., x_d) = C(F_1(x_1), ..., F_d(x_d)). \tag{5.3}$$

Copula C is the joint df of rv's $U_i = F_i\left(X^{(i)}\right)$, $i = 1, ..., d.$ It is defined on $[0,1]^d$ by $C(u_1, ..., u_d) = H\left(F_1^{-1}(u_1), ..., F_d^{-1}(u_d)\right)$, where F_i^{-1} denotes the quantile function of F_i. This means that the DRM of the sum is a functional of both copula C and margins F_i. Therefore, one must take into account the dependence structure and the behavior of margin tails. These two aspects have an important influence when quantifying risks. If the correlation factor is neglected, the calculation of the DRM follows formula (5.2), which only focuses on distorting the tail. In order to highlight the dependence structure, we add a distortion on the copula as well. The notion of distorted copula has recently been considered by several authors, see for instance

Frees and Valdez (1998), Genest and Rivest (2001, [63]), Morillas (2005, [117]) and Valdez and Xiao (2010, [147]). Given a copula C and a non-decreasing bijection $\Gamma : [0,1] \to [0,1]$, the distorted copula C^{Γ} is defined by

$$C^{\Gamma}(u_1, ..., u_d) := \Gamma^{-1}(C(\Gamma(u_1), ..., \Gamma(u_d))).$$

This transformation will affect the joint df H and consequently the df G of the sum Z. Their new forms will be denoted by H^{Γ} and G^{Γ} respectively. Morillas (2005, [117]) describes some of the existing families of distortion functions, among which the following are frequently used:

$$\begin{aligned}
\Gamma_r(s) &= s^r, && \text{for} \quad 0 < r \leq 1, \\
\Gamma_\delta(s) &= \frac{\ln(\delta s + 1)}{\ln(\delta + 1)}, && \text{for} \quad \delta > 0, \\
\Gamma_{\xi,\vartheta}(s) &= \frac{(\xi + \vartheta) s}{\xi s + \vartheta}, && \text{for} \quad \xi, \vartheta > 0, \\
\Gamma_\nu(s) &= \frac{s^\nu}{2 - s^\nu}, && \text{for} \quad 0 < \nu \leq 1/3.
\end{aligned}$$

We call the corresponding distorted risk measures by *copula distorted risk measure* (CDRM) defined as

$$\pi_\psi^{\Gamma}[Z] = \int_0^\infty \psi\left(1 - G^{\Gamma}(z)\right) dz.$$

It is worth mentioning that if $X^{(1)}, ..., X^{(d)}$ are independent, the corresponding copula function $C(u_1, ..., u_d) = \prod_{i=1}^{d} u_i$ is called the product copula and denoted by $C^{\perp}$. In this case, we have $C^{\Gamma} = C$ and therefore $\pi_\psi^{\Gamma}[Z] = \pi_\psi[Z]$.

The remainder of this paper is organized as follows. In Section 5.2, we give a copula representation of the DRM's. In Section 5.3, we present a more flexible class of copula given by the notion of distorted Archimedean copulas. By the nice properties of this class and the copula representation of the DRM, we introduce in Section 5.4 of the CDRM's. Finally, an illustrative example explaining the CDRM computation is given in Section 5.5.

5.2 Copula representation of the DRM

Given a vector of risk losses $\mathbf{X} = (X^{(1)}, ..., X^{(d)})$, $d \geq 2$, with joint df H and continuous margins F_i, $i = 1, ..., d$. The df of the rv $Z = \sum_{i=1}^{d} X_i$, is given by

$$G(t) = \int_{A(t)} dH\left(x_1, .., x_d\right), \text{ for any } t \geq 0,$$

where $A\left(t\right) := \left\{ \left(x_1, .., x_d\right) : 0 \leq \sum_{i=1}^{d} x_i \leq t \right\}$. Using the representation (5.3), we get

$$G(t) = \int_{A(t)} dC(F_1(x_2), ..., F_d(x_d)).$$

If we suppose that the copula C and margins F_i are differentiable with densities c and f_i, respectively, then

$$G(t) = \int_{A(t)} c\left(F_1\left(x_1\right), ..., F_d\left(x_d\right)\right) \prod_{i=1}^{d} f_i\left(x_i\right) dx_1, ...dx_d.$$

The change of variables $F_i\left(x_i\right) = u_i$, $i = 1, ..., d$, yields

$$G(t) = \int_0^{F_d(t)} \int_0^{F_{d-1}\left(t - F_d^{-1}(u_d)\right)} \cdots \int_0^{F_1\left(t - \sum_{i=0}^{d-2} F_{d-i}^{-1}(u_{d-i})\right)} c(u_1, ..., u_d) du_1 ... du_d. \quad (5.4)$$

According to (5.4), the computation of the DRM corresponding to Z, given in (5.2), requires the knowledge of the copula density and the margins of vector $\mathbf{X}$. In particular, for the bivariate case $(d = 2)$, we have

$$G(t) = \int_0^{F_2(t)} \int_0^{F_1\left(t - F_2^{-1}(u_2)\right)} c(u_1, u_2) du_1 du_2.$$

Whenever X_1 and X_2 are independent, we have $c(u_1, u_2) = 1$, and therefore

$$G(t) = \int_0^{F_2(t)} F_1\left(t - F_2^{-1}\left(u_2\right)\right) du_2 = \int_0^t F_1\left(t - x\right) dF_2\left(x\right),$$

which is the usual convolution of the F_i's.

5.3 Distorted Archimedean copulas

In this paper, we focus on one important class of copulas called: Archimedean copulas. This class contains several copula families useful in dependence modelling. Their nice properties are captured by an additive generator function $\varphi : [0,1] \to [0,\infty]$, which is continuous, strictly decreasing and convex with $\varphi(1) = 0$. The main advantage of the Archimedean copulas is the achievement of the reduction in dimensionality of a d-variate distribution in a single argument. In econometrics, this property has the potential to be of use in models of limited dependent variables, especially those requiring some probabilistic enumeration on high-dimensional subspaces. In the bivariate case, an Archimedean copula is defined by

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

where

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

Note that $\varphi^{[-1]}$ is continuous and non-increasing on $[0,\infty]$ and φ is the unique generator up to a scaling constant. If the terminal $\varphi(0) = \infty$, the generator is termed strict and $\varphi^{[-1]} = \varphi^{-1}$. Numerous single-parameter families of Archimedean copulas are listed in Table 4.1 in Nelsen (2006, [123]). Particular examples are $\varphi_\theta(t) = \left(t^{-\theta} - 1\right)/\theta$, $\varphi_\alpha(t) = (-\ln t)^\alpha$ and $\varphi_\beta(t) = -\ln(\left(e^{-\beta t} - 1\right) / \left(e^{-\beta} - 1\right))$ which are, respectively, the generators of the Clayton family

$$C_\theta(u,v) = \left(u^{-\theta} + v^{-\theta} - 1\right)^{-1/\theta}, \ \theta \ge 0,$$

the Gumbel family

$$C_\alpha(u,v) = \exp\left\{-[(-\ln u)^\alpha + (-\ln v)^\alpha]^{1/\alpha}\right\}, \ \alpha \ge 1,$$

and the Frank family

$$C_\beta(u,v) = -\frac{1}{\beta}\ln\left[1 + \frac{\left(e^{\beta u} - 1\right)\left(e^{\beta v} - 1\right)}{e^\beta - 1}\right], \ \beta \in \mathbb{R}\setminus\{0\}.$$

The generators φ_θ, φ_α and φ_β are strict and therefore their corresponding copulas C_θ, C_α and C_β verify

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

Next, we discuss some properties of distortion functions acting on bivariate Archimedean copulas. Given an Archimedean copula C and a strictly increasing bijection Γ : $[0,1] \rightarrow [0,1]$, we consider the function $C^{\Gamma} : [0,1]^2 \rightarrow [0,1]$ defined by

$$C^{\Gamma}(u,v) = \Gamma^{-1}(C(\Gamma(u),\Gamma(v))).$$

Under what conditions on Γ, the function C^{Γ} is an Archimedean copula?

First, from Theorem 3.3.3. in Nelsen (2006, [123]), C^{Γ} is a copula if Γ is concave and continuous on $[0,1]$ with $\Gamma(0) = 0$ and $\Gamma(1) = 1$. The following Theorem gives an additional condition so that the copula C^{Γ} remains Archimedean. For convenience, let $\mathbb{K}$ represents the set of the functions Γ verifying the assumptions above.

Theorem 5.3.1 *Let C be an Archimedean copula with generator φ and suppose that $\Gamma \in \mathbb{K}$, then the copula C^{Γ} is Archimedean if and only if $\varphi \circ \Gamma$ is convex.*

Proof. Indeed, let φ be the generator of the copula C and let $\Gamma \in \mathbb{K}$, then

$$C^{\Gamma}(u_1,...,u_d) = \Gamma^{-1}(C(\Gamma(u_1),...,\Gamma(u_d))).$$

We have $\Gamma^{[-1]} = \Gamma^{-1}$, then

$$C^{\Gamma}(u_1,...,u_d) = \Gamma^{[-1]}\varphi^{[-1]}\left(\varphi\left(\Gamma\left(u_1\right)\right) + ... + \varphi\left(\Gamma\left(u_d\right)\right)\right).$$

It is easy to show that $\Gamma^{[-1]}\varphi^{[-1]} = (\varphi \circ \Gamma)^{[-1]}$, it follows that

$$C^{\Gamma}(u_1,...,u_d) = \mathcal{T}^{[-1]}\left(\mathcal{T}\left(u_1\right) + ... + \mathcal{T}\left(u_d\right)\right), \tag{5.5}$$

with $\mathcal{T} := \varphi \circ \Gamma$. From Theorem 4.1.4. Nelsen 2006, [123], C^{Γ} is Archimedean if and only if $\mathcal{T}$ is convex. Notice that $\varphi \circ \Gamma$ is the generator of the copula C^{Γ}. ■

Corollary 5.3.1 *The crucial distortion function $t \rightarrow \Gamma^{\perp}(t) := \exp(-\varphi(t))$ transforms any Archimedean copula C in the product copula $C^{\perp}$.*

Proof. Straightforward. ■

Next, we see the influence of the distortion of copulas on the association measures. Kendall's tau and Spearman's rho are the most popular measures of association, their representations in terms of the copula C are given by

$$\tau = 4\int_0^1\int_0^1 C(u,v)dC(u,v) - 1 \text{ and } \rho = 12\int_0^1\int_0^1 (C(u,v) - uv)\,dudv,$$

respectively. Let τ^{Γ} and ρ^{Γ}, respectively, denote Kendall's tau and Spearman's rho of copula C^{Γ}. According to Theorem 10 in Durrleman et al. (2000, [37]), we have under suitable assumptions

$$1+\frac{\tau-1}{a^2}\leq\tau^{\Gamma}\leq 1+\frac{\tau-1}{b^2},$$

and

$$\frac{\rho+3}{a^3}-3\leq\rho^{\Gamma}\leq\frac{\rho+3}{b^3}-3,$$

where $0<a\leq b<\infty$ are bounds for the derivative of Γ.

5.4 Risk measures for sums of losses

It may happen that the model (represented by the copula C) chosen, to fit the data, does not provide enough information. This leads us to transform C to a more flexible copula C^{Γ} of the same class. Consequently, the joint df of $\mathbf{X}$ may be represented, via Sklar's Theorem, as

$$H\left(x_1,...,x_d\right)=C^{\Gamma}\left(F_1(x_1),...,F_d(x_d)\right).$$

Suppose the copula C is Archimedean with generator φ, then from Theorem 5.3.1, the copula C^{Γ} defined in (5.5) is also Archimedean. Assume that C^{Γ} has a density function c^{Γ}, then in view of the representation (5.4) the df G^{Γ} of the sum Z may be written as

$$\begin{aligned}G^{\Gamma}\left(t\right) &:= \int_0^{F_d(t)}\int_0^{F_{d-1}\left(t-F_d^{-1}(u_d)\right)} \\ &\quad ...\int_0^{F_1\left(t-\sum_{i=0}^{d-2}F_{d-i}^{-1}(u_{d-i})\right)} c_{\Gamma}(u_1,...,u_d)du_1...du_d.\end{aligned}$$

Applying Wang's principle (5.1) to the loss distribution G^{Γ}, we have

$$\pi_{\psi}^{\Gamma}\left[Z\right]:=\int_0^{+\infty}\psi\left(1-G^{\Gamma}\left(t\right)\right)dt,$$

which we call the CDRM. This may be considered as manner of measuring the risk Z by distorting both the dependence structure and the distribution tail, without losing the coherence feature. The CDRM adjusts the true probability measure to give more weight to higher risk events and less weight to dependence structure. In other words, the simultaneous transformations yield a new risk measure bounded by the expectation and Wang's measure, that is

$$\mathbb{E}[Z] \leq \pi_{\psi}^{\Gamma}[Z] \leq \pi_{\psi}[Z]. \tag{5.6}$$

In the following example, we verify the previous inequalities on a selected model.

5.5 Illustrative example

Let X_1 and X_2 be two risks with joint df represented by the Clayton copula C_θ, $\theta > 0$ and Pareto-distributed margins F_1 and F_2 with parameters $0 < \alpha_1, \alpha_2 < 1$, that is $F_i(x_i) = 1 - x_i^{-1/\alpha_i}$, $x_i > 1$, $i = 1, 2, .$ The corresponding Kendall tau of Clayton copula is $\tau = \theta/(\theta + 2)$. Let $\psi(x) = x^{1/\rho}$, $\rho \geq 1$, and $\Gamma(t) = t^{1/\delta}$, $\delta \geq 1$. The distorted copula C_θ^Γ, denoted by C_θ^δ, is of Clayton type with generator $(\varphi \circ \Gamma)(t) = \left(t^{-\theta/\delta} - 1\right)/\theta$ and the corresponding Kendall's tau is $\tau^\Gamma = (\theta/\delta)/(\theta/\delta + 2)$. The df of the sum $Z = X_1 + X_2$ is

$$G^\delta(t; \theta, \alpha_1, \alpha_2) = \int_1^{1-t^{-1/\alpha_2}} \left(\int_1^{1-(t-(1-v)^{-\alpha_2})^{-1/\alpha_1}} c_\theta^\delta(u, v) du \right) dv,$$

where

$$c_\theta^\delta(u, v) = (\theta/\delta + 1) u^{-\theta/\delta - 1} v^{-\theta/\delta - 1} \left(u^{-\theta/\delta} + v^{-\theta/\delta} - 1\right)^{-\delta/\theta - 2},$$

is the density of C_θ^δ. Figures 5.1 and 5.2 gives a preview of the effect of the copula distortion.

The DRM and the CDRM of Z are repetitively denoted by

$$\pi_\rho[Z] = \int_2^\infty (1 - G(t))^{1/\rho} dt,$$

and

$$\pi_\rho^\delta[Z] = \int_2^{+\infty} \left(1 - G^\delta(t)\right)^{1/\rho} dt.$$

We select Pareto model with $\theta = 3/2$, $\alpha_1 = 1/3$, $\alpha_2 = 1/5$. We obtain $\mathbb{E}(Z) = 0.750$, $\tau = 0.428$. For two different tail distortion parameters $\rho = 1.2$ and $\rho = 1.4$ the respective DRM's are 1.225 and 2.091. The CDRM's for distinct values of the copula distortion parameter δ are summarized Tables 5.1 and 5.2, where we see that the inequalities (5.6) are satisfied for any value of the copula distortion parameter. This is well shown graphically in Figure 5.3 in which the three risk measures of (5.6) are

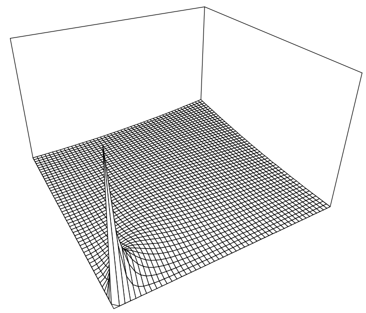

Figure 5.1: Clayton copula density with $\theta = 2$.

plotted as functions of δ.

δ	1	1.5	2	2.5	3	3.5	4	5	6
τ^{δ}	0.428	0.333	0.272	0.230	0.200	0.176	0.157	0.130	0.111
$\pi_{\rho}^{\delta}[Z]$	1.225	1.030	0.988	0.969	0.964	0.961	0.958	0.953	0.950

Table 5.1: CDRM's and transformed Kendall tau of the sum of two Pareto-distributed risks with tail distortion parameter $\rho = 1.2$.

δ	1	1.5	2	2.5	3	3.5	4	5	6
τ^{δ}	0.428	0.333	0.272	0.230	0.200	0.176	0.157	0.130	0.111
$\pi_{\rho}^{\delta}[Z]$	2.091	1.801	1.736	1.712	1.703	1.699	1.694	1.685	1.680

Table 5.2: CDRM's and transformed Kendall tau of the sum of two Pareto-distributed risks with tail distortion parameter $\rho = 1.4$.

Taking $\delta = 1$ means that we make no distortion on the dependence structure, that is $C^1 = C$, and $\pi_{\rho}^{1}[Z] = \pi_{\rho}[Z]$. In other words, the CDRM with $\delta = 1$ reduces to Wang's DRM, which can be seen in the second columns of Tables 5.1 and 5.2. This fact is also clear in Figure 5.3. On the other hand, as δ increases, the transformed Kendall's tau decreases meaning that the dependence gets weaker (see the second lines of Tables 5.1 and 5.2). Moreover, starting from some δ the CDRM values (see the third lines of Tables 5.1 and 5.2) become roughly constant while being always greater than the expectation.

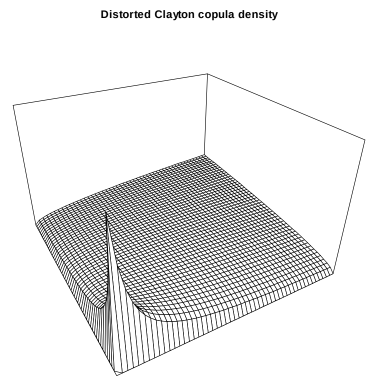

Figure 5.2: Distorted Clayton copula density with $\theta = 2$, $\delta = 4$.

5.6 Concluding remarks

In portfolio analysis, the dependence structure has a major role to play in quantifying risks. This led us to think of risk measure taking into account this fact in addition to the tail behavior. In this paper, we proposed a risk measure for the sum of two losses by simultaneously transforming the distribution tail and the copula, which represents the dependence between the margins, by means of two distortion functions. We obtained a coherent measure that we called the *Copula Distorted Risk Measure.* This new measure has the characteristic to be greater than the expectation and less than the popular Wang's distorted risk measure. In the insurance business, the main advantage of this property is to reduce Wang's premium while respecting the standard axioms of the premium principle.

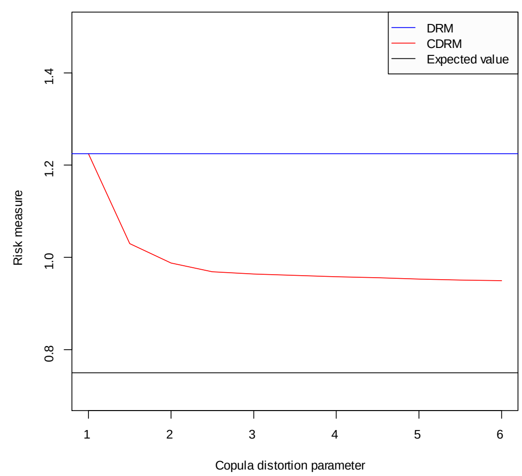

Figure 5.3: Risk measures of the sum of two Pareto-distributed risks with tail distortion parameter $\rho = 1.2$.

General conclusion

One of the most widely used tools to study multivariate outcomes is the copula function. In the case of dependent multivariate data, multivariate copulas provide a useful tool to assist in the process of model building. In this essay we have shown how to analyse some real-world scenarios using copulas.

We started with a discussion of the significance of the copula function in the first chapter. A copula is a function that relates a multivariate distribution function to its one-dimensional marginal distribution functions. We reviewed its properties, such as the invariance under strictly monotone transformations. We also looked at some methods of constructing bivariate copulas and how copulas could be used to simulate multivariate outcomes, an important tool for applied work, where many variables need to be considered. We also provided an example where we showed how one can estimate a copula where bivariate data is at hand, for purposes of simplicity.

Since copulas are parametric families, standard techniques such as the maximum likelihood and inference functions for margins (IFM) methods, are useful for estimating their parameters. All these methods are described in Chapter one.

There are many families of copulas which differ in the detail of the dependence they represent. A family will typically have several parameters which relate to the strength and form of the dependence. Some families of copulas are outlined in Chapter two. A typical use for copulas is to choose one such family and use it to define the multivariate distribution to be used, typically in fitting a distribution to a sample of data. However, it is possible to derive the copula corresponding to any given multivariate distribution.

In Chapter three, we focused on risk measures that can be used to understand which aspect of the riskiness associated with the uncertain outcome the risk measure attempts to quantify. In that respect, we concentrate on risk measures that measure upper tails of distribution functions.

Finally in Chapter four and five, we give the main results in this thesis. The first one is giving a new procedure estimate of copula parameters using the bivariate L-moments, we have already proved performances of this approach, we compared, by simulations, the new method with the well-known (τ, ρ)-inversion, the minimum distance and the pseudo maximum likelihood estimators, which is recommended by some authors. We have show that BLM estimator is good candidate to construct goodness-of-fit tests of copula models. The second one is proposing of a risk measure for the sum of two losses by simultaneously transforming the distribution tail and the copula, which represents the dependence between the margins, by means of two distortion functions. We obtained a coherent Copula Distorted Risk Measure. This new measure has the characteristic to be greater than the expectation and less than the popular Wang's distorted risk measure, the main advantage of this property is to reduce Wang's premium while respecting the standard axioms of the premium principle.

Appendix A

Proofs

A.1 Proof of Theorem 4.3.1

Since $\left(F_2(X^{(2)})\right)$ is $(0,1)$ −uniform r.v., then copula's representation of the joint df of the pair of r.v.'s $\left(X^{(1)}, \left(F_2(X^{(2)})\right)^j\right)$ is

$$D(u,v) := C\left(F_1\left(u\right), v^{1/j}\right),\ j = 1, 2, ..., k-1,$$

it follows that, the covariance of $\left(X^{(1)}, \left(F_2(X^{(2)})\right)^j\right)$ is

$$\begin{aligned} Cov\left(X^{(1)}, \left(F_2\left(X^{(2)}\right)\right)^j\right) &= \int_{\mathbb{R}} \int_{\mathbb{I}} \left(D(u,v) - v^{1/j} F_1(u)\right) dv du \\ &= \int_{\mathbb{R}} \int_{\mathbb{I}} \left(C\left(F_1(u), v^{1/j}\right) - v^{1/j} F_1(u)\right) dv du \\ &= \int_{\mathbb{I}^2} \left(C\left(u,v\right) - uv\right) dv^j dF_1^{-1}(u). \end{aligned}$$

Therefore

$$\begin{aligned} \lambda_{k[12]} &= Cov\left(X^{(1)}, P_{k-1}\left(F_2\left(X^{(2)}\right)\right)\right) \\ &= \sum_{j=0}^{k-1} p_{j,k-1} Cov\left(X^{(1)}, \left(F_2\left(X^{(2)}\right)\right)^j\right). \end{aligned}$$

Since $Cov\left(X^{(1)}, \left(F_2\left(X^{(2)}\right)\right)^j\right) = 0$ for $j = 0$, then

$$\begin{aligned}\lambda_{k[12]} &= \sum_{j=1}^{k-1} p_{j,k-1} Cov\left(X^{(1)}, \left(F_2\left(X^{(2)}\right)\right)^j\right) \\ &= \sum_{j=1}^{k-1} p_{j,k-1} \int_{\mathbb{I}^2} \left(C\left(u,v\right) - uv\right) dv^j dF_1^{-1}(u) \\ &= \int_{\mathbb{I}^2} \left(C\left(u,v\right) - uv\right) dP_{k-1}\left(v\right) dF_1^{-1}(u),\end{aligned}$$

as sought. □

A.2 Proof of Theorem 4.4.1

The existence of a sequence of consistent roots $\widehat{\boldsymbol{\theta}}^{BLM}$ to (4.15) or (4.21), may be checked by using a similar argument as the proof of Theorem 1 in Tsukahara (2005). Indeed, we have only to check the conditions in Theorem A.10.2 in Bickel *et al.* (1993). Since we are dealing with an asymptotic result, we may consider that, for all large n, without loss of generality, that the empirical df $F_{j:n}$ and their rescaled version $F_{j:n}^*$ have a same effect. Therefore throughout the proof, we will make use of $F_{j:n}$ instead of $F_{j:n}^*$. For convenience we set

$$\Phi_n\left(\boldsymbol{\theta}\right) := \frac{1}{n}\sum_{i=1}^{n} \mathbf{L}\left(F_{1:n}\left(X_i^{(1)}\right), F_{2:n}\left(X_i^{(2)}\right); \boldsymbol{\theta}\right) \text{ and } \Phi\left(\boldsymbol{\theta}\right) := \int_{\mathbb{I}^2} \mathbf{L}\left(u_1, u_2; \boldsymbol{\theta}\right) dC_{\boldsymbol{\theta}_0}\left(u_1, u_2\right).$$

By assumption $[H.2]$, it is clear that the following derivatives exist:

$$\begin{aligned}\overset{\bullet}{\Phi}_n\left(\boldsymbol{\theta}\right) &= \frac{\partial \Phi_n\left(\boldsymbol{\theta}\right)}{\partial \boldsymbol{\theta}} = \frac{1}{n}\sum_{i=1}^{n} \overset{\bullet}{\mathbf{L}}\left(F_{1:n}\left(X_i^{(1)}\right), F_{2:n}\left(X_i^{(2)}\right); \boldsymbol{\theta}\right), \\ \overset{\bullet}{\Phi}\left(\boldsymbol{\theta}\right) &= \frac{\partial \Phi\left(\boldsymbol{\theta}\right)}{\partial \boldsymbol{\theta}} = \int_{\mathbb{I}^2} \overset{\bullet}{\mathbf{L}}\left(u_1, u_2; \boldsymbol{\theta}\right) dC_{\boldsymbol{\theta}_0}\left(u_1, u_2\right).\end{aligned}$$

Next, we verify that

$$\sup\left\{\left|\overset{\bullet}{\Phi}_n\left(\boldsymbol{\theta}\right) - \overset{\bullet}{\Phi}\left(\boldsymbol{\theta}\right)\right| : \left|\boldsymbol{\theta} - \boldsymbol{\theta}_0\right| < \epsilon_n\right\} \xrightarrow{\mathbf{P}} 0, \text{ as } n \to \infty, \tag{A.1}$$

for any real sequence $\epsilon_n \to 0$. By using the triangular inequality we get

$$\left|\overset{\bullet}{\Phi}_n(\boldsymbol{\theta}) - \overset{\bullet}{\Phi}_n(\boldsymbol{\theta}_0)\right| \leq \frac{1}{n}\sum_{i=1}^{n}\left|\overset{\bullet}{\mathbf{L}}\left(F_{1:n}\left(X_i^{(1)}\right), F_{2:n}\left(X_i^{(2)}\right); \boldsymbol{\theta}\right) - \overset{\bullet}{\mathbf{L}}\left(F_{1:n}\left(X_i^{(1)}\right), F_{2:n}\left(X_i^{(2)}\right); \boldsymbol{\theta}_0\right)\right|.$$

Since $\overset{\bullet}{\mathbf{L}}$ is continuous in $\boldsymbol{\theta}$, then

$$\sup\left\{\left|\overset{\bullet}{\mathbf{L}}\left(F_{1:n}\left(X_i^{(1)}\right), F_{2:n}\left(X_i^{(2)}\right); \boldsymbol{\theta}\right) - \overset{\bullet}{\mathbf{L}}\left(F_{1:n}\left(X_i^{(1)}\right), F_{2:n}\left(X_i^{(2)}\right); \boldsymbol{\theta}_0\right)\right| : |\boldsymbol{\theta} - \boldsymbol{\theta}_0| < \epsilon_n\right\} = o_{\mathbf{P}}(1),$$

therefore

$$\sup\left\{\left|\overset{\bullet}{\Phi}_n(\boldsymbol{\theta}) - \overset{\bullet}{\Phi}_n(\boldsymbol{\theta}_0)\right| : |\boldsymbol{\theta} - \boldsymbol{\theta}_0| < \epsilon_n\right\} \overset{\mathbf{P}}{\to} 0, \text{ as } n \to \infty. \tag{A.2}$$

On the other hand, from the law of the large number, we infer that

$$\frac{1}{n}\sum_{i=1}^{n}\overset{\bullet}{\mathbf{L}}\left(F_1\left(X_i^{(1)}\right), F_2\left(X_i^{(2)}\right); \boldsymbol{\theta}_0\right) \overset{\mathbf{P}}{\to} \overset{\bullet}{\Phi}(\boldsymbol{\theta}_0), \text{ as } n \to \infty.$$

Moreover, in view of the continuity of $\overset{\bullet}{\mathbf{L}}$ in u and since $\sup_{x^{(j)}}\left|F_{i:n}\left(x^{(j)}\right) - F_i\left(x^{(j)}\right)\right| \to 0$, $j = 1, 2$, almost surely, $n \to \infty$ (Glivenko-Cantelli theorem), we have

$$\frac{1}{n}\sum_{i=1}^{n}\left|\overset{\bullet}{\mathbf{L}}\left(F_{1:n}\left(X_i^{(1)}\right), F_{2:n}\left(X_i^{(2)}\right); \boldsymbol{\theta}_0\right) - \overset{\bullet}{\mathbf{L}}\left(F_1\left(X_i^{(1)}\right), F_2\left(X_i^{(2)}\right); \boldsymbol{\theta}_0\right)\right| \overset{\mathbf{P}}{\to} 0.$$

It follows that $\left|\overset{\bullet}{\Phi}_n(\boldsymbol{\theta}_0) - \overset{\bullet}{\Phi}(\boldsymbol{\theta}_0)\right| \overset{\mathbf{P}}{\to} 0$, which together with (A.2), implies (A.1).

Conditions (MG0) and (MG3) in Theorem A.10.2 in Bickel *et al.* (1993) are trivially satisfied by our assumptions $[H1] - [H3]$. In view of the general theorem for Z-estimators (see, van der Vaart and Wellner, 1996, Th. 3.3.1), it remains to prove that $\sqrt{n}\left(\overset{\bullet}{\Phi}_n - \overset{\bullet}{\Phi}\right)(\boldsymbol{\theta}_0)$ converges in law to the appropriate limit. But this follows from Proposition 3 in Tsukahara (2005), which achieves the proof of Theorem 4.4.1. □

A.3 Minimum distance based estimation

We briefly present the minimum distance (MD) base estimation for copula models that posseses a qualitative robusteness (Genest and Rémillard, 2008), this will be compared with the BLM method (see Subsection 5.2). Let C be the true copula

associated to the df of $\left(X^{(1)}, X^{(2)}\right)$ and suppose that we have a given parametric family of copula $\mathcal{C} := \{C_{\boldsymbol{\theta}},\ \boldsymbol{\theta} \in \mathcal{O}\}$ to fit data. Let us define the minimum distance functional T on the space of the copula by

$$T\left(C\right) := \arg\min_{\boldsymbol{\theta} \in \mathcal{O}} \mu\left(C, C_{\boldsymbol{\theta}}\right).$$

Here μ is a distance between probabilities on $\mathbb{I}^2$. In the present paper, we consider the Cramér-von Mises distance difined by

$$\mu^{CVM}\left(C, C_{\boldsymbol{\theta}}\right) := \int_{\mathbb{I}^2} \left\{C\left(u_1, u_2\right) - C_{\boldsymbol{\theta}}\left(u_1, u_2\right)\right\}^2 dC\left(u_1, u_2\right).$$

Consider now a random sample $\left(X_i^{(1)}, X_i^{(2)}\right)_{i=1,n}$, from the bivariate random variables $\left(X^{(1)}, X^{(2)}\right)$. The joint empirical distribution functions is given by

$$\mathbb{F}_n\left(x_1, x_2\right) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}\left\{X_i^{(1)} \leq x_1, X_i^{(2)} \leq x_2\right\}.$$

Following Deheuvels (1979), we define the empirical copula by

$$\mathbb{C}_n\left(u_1, u_2\right) := \mathbb{F}_n\left(\mathbb{F}_{n:1}^{-1}\left(u_1\right), \mathbb{F}_{n:2}^{-1}\left(u_2\right)\right),\ 0 \leq u_1, u_2 \leq 1.$$

The corresponding Cramér-von Mises statistics is

$$\mu^{CVM}\left(\mathbb{C}_n, C_{\boldsymbol{\theta}}\right) = \int_{\mathbb{I}^2} \left\{\mathbb{C}_n\left(u_1, u_2\right) - C_{\boldsymbol{\theta}}\left(u_1, u_2\right)\right\}^2 d\mathbb{C}_n\left(u_1, u_2\right).$$

This may be rewritten into

$$\mu^{CVM}\left(\mathbb{C}_n, C_{\boldsymbol{\theta}}\right) = n^{-1}\sum_{i=1}^{n}\left(\mathbb{C}_n\left(\widehat{U}_i^{(1)}, \widehat{U}_i^{(2)}\right) - C_{\boldsymbol{\theta}}\left(\widehat{U}_i^{(1)}, \widehat{U}_i^{(2)}\right)\right)^2,$$

where $\widehat{U}_i^{(j)} := \mathbb{F}_{j:n}^* \left(X_i^{(j)} \right)$, $i = 1, ..., n$, for each $j = 1, 2$ (see, Genest and Rémillard, 2008, eq. 31). The MD estimator of the parameter $\boldsymbol{\theta}$ is defined by

$$\widehat{\boldsymbol{\theta}} = T\left(\mathbb{C}_n\right) := \arg\min_{\boldsymbol{\theta}\in\mathcal{O}} \mu^{CVM}\left(\mathbb{C}_n, C_{\boldsymbol{\theta}}\right).$$

Note that we may also use the Kolmogorov-Smirnov distance but this is awkward in practice due to the supremum norm uses. Also since the Hellinger distance is defined by copula densities, other nonparametric estimators of the underling copula are needed (see, Biau and Begkamp, 2005) and therefore non-standard computational procedures are required.

Suppose now that we are dealing with the estimation of parameters of one iterated FGM copula family C_{α_1,α_2} in (4.8) by means of the MD method. The MD estimator for $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)$ noted $\widehat{\boldsymbol{\alpha}}^{MD}$ results by minimizing the function $(\alpha_1, \alpha_2) \rightarrow \rho\left(\mathbb{C}_n, C_{\alpha_1,\alpha_2}\right)$ over the region $\mathcal{R}$ given in (4.9). Then to solve the previous optimization problem, we will introduce the Lagrange multiplier principle, that is we have to rewrite the region $\mathcal{R}$ into

$$\mathcal{R} = \left\{(\alpha_1, \alpha_2),\ \ell_j\left(\alpha_1, \alpha_2\right) \geq 0,\ j = 1, 2, 3\right\},$$

where $\ell_1\left(\alpha_1, \alpha_2\right) := 1 - \alpha_1^2$, $\ell_2\left(\alpha_1, \alpha_2\right) := \alpha_1 + \alpha_2 + 1$ and

$$\ell_3\left(\alpha_1, \alpha_2\right) := \frac{1}{2}\left[3 - \alpha_1 + \left(9 - 6\alpha_1 - 3\alpha_1^2\right)^{1/2}\right] - \alpha_2,$$

and then minimize the function

$$\mathbb{K}_n\left(\boldsymbol{\alpha}, \boldsymbol{\nu}\right) := \rho\left(\mathbb{C}_n, C_{\alpha_1,\alpha_2}\right) - \sum_{j=1}^{3} \nu_j \ell_j\left(\alpha_1, \alpha_2\right),$$

over the whole $\mathbb{R}^5$, with $\boldsymbol{\alpha} = (\alpha_1, \alpha_2) \in \mathbb{R}^2$ and $\boldsymbol{\nu} = (\nu_1, \nu_2, \nu_3) \in \mathbb{R}^3$. So, the new formulation of the MD estimator of parameter $\boldsymbol{\alpha}$ is

$$\widehat{\boldsymbol{\alpha}}^{MD} = \arg\min_{(\boldsymbol{\alpha},\boldsymbol{\nu})\in\mathbb{R}^5} \mathbb{K}_n\left(\boldsymbol{\alpha}, \boldsymbol{\nu}\right).$$

We note here that it is difficult, in general, to have an explicit form for $\widehat{\boldsymbol{\alpha}}^{MD}$, then only the numerical computation can solve this issue. This is observed for the one-iterated FGM family, that the optimization problem requires tedious tools.

Appendix B

Code R

B.1 Simulation datas from two parameters FGM copula

Let C 2-parameters FGM copula defined by

$$C_{\alpha_1,\alpha_2}(u,v) = uv\left\{1 + \alpha(1-u)(1-v) + \beta uv(1-u)(1-v)\right\}.$$

the «conditional distribution function» method, where $v = c_u^{-1}(t)$ for

$$\begin{aligned} c_u(t) &= \frac{\partial C(u,v)}{\partial u} \\ &= v + (\alpha v(1-v)(1-u) - \alpha vu(1-v)) \times \\ &\quad (\beta(2(v(uv)))(1-v)(1-u) - \beta((uv)^2)(1-v)) \end{aligned} \tag{B.1}$$

Algorithm (Nelsen 2006, [123, p. 41]).

1. Generate two independent uniform $(0,1)$ variates u and t.
2. Set $v = c_u^{[-1]}(t)$ where $c_u^{[-1]}$ denotes a quasi-inverse of c_u.
3. The desired pair is (u,v).

The code R for this Algorithm for 2-parameters FGM copula is:

```
library(copula, mvtnorm, scatterplot3d)
library(sn)
```

```
library(mnormt)
# load required packages
P <- function(u,v){
                 v+(a*v*(1-v)*(1-u)-a*v*u*(1-v))
                +(b*(2*(v*(u*v)))*(1-v)*(1-u)-b*((u*v)^2)*(1-v))
                } # the formula (B.1).
a <- 0.941     # first copula parameter α
b <- 1.445     # second copula parameter β
n <- 10        # sample size.
F <- function(x) rank(x)/n
# First step we generate 2 independent uniform (0,1) variates u and t.
   teta <- 0   # teta=0 gives 2 independent rv's.
   fgm.cop <- fgmCopula(teta)
  s <- mvdc(fgm.cop, c("unif", "unif"), list(list( 0, 1), list(0,1)))
   x <- rmvdc(s, n)
   u <- x[,1]
   t <- x[,2]
```

Second step we calculate $v = c_u^{[-1]}(t)$

```
r <- numeric(n)
Kinv <- function(x,r) optimize(function(y) (r-P(x,y))^2,c(0, 1))$minimum
       for (i in 1:n){
              r[i] <- Kinv(u[i],t[i])
                     }
v <- (n/(n+1))*F(r) # v is (0,1) uniform.
```

The fact `(n/(n+1))` to avoid possible problems with unboundedness of the copula density, see (1.9).

```
# Third step we have The desired pair is (u,v) by
D <- data.frame(u,v)
D
```

n	u	v
1	0.7685587	0.2727273
2	0.6812183	0.7272727
3	0.8452455	0.9090909
4	0.5175548	0.1818182
5	0.6071752	0.0909091
6	0.2067431	0.5454545
7	0.4336877	0.3636364
8	0.4732063	0.8181818
9	0.8707530	0.6363636
10	0.3043744	0.4545455

B.2 Simulation datas from two parameters of Archimedean copula

B.2.1 Two parameter Clayton Copula

The two parameter Clayton copula is given for $\alpha > 0$, $\beta > 1$ by

$$C_{\alpha,\beta}(u,v) = \left(\left(\left(u^{-\alpha}-1\right)^{\beta} + \left(v^{-\alpha}-1\right)^{\beta}\right)^{1/\beta} + 1\right)^{-1/\alpha}, \tag{B.2}$$

with generator

$$\varphi_{\alpha,\beta}(t) = \frac{\left(t^{-\alpha}-1\right)^{\beta}}{\alpha^{\beta}}, \tag{B.3}$$

and

$$\varphi_{\alpha,\beta}^{-1}(t) = \left(\alpha t^{1/\beta} + 1\right)^{-1/\alpha}. \tag{B.4}$$

Recall Algorithm 2.3.1 page 62.

Code R of 2-parameter Clayton Copula

```
rclayton2 <- function (n, alpha, beta, dim = 2)
  {
 if (alpha < 0)
 stop("invalid argument :  alpha\n")
 if (beta < 1)
 stop("invalid argument :  beta\n")
     {
```

```
v2 <- runif(n)
T <- runif(n)
```

#Generator of 2-parameters Clayton copula $\varphi_{\alpha,\beta}$ see (B.3)

```
phicla <- function(t, alpha, beta){
          (((t^(-alpha)-1)^(beta))/(alpha^(beta)))}
```

Inverse of generator $\varphi^{-1}_{\alpha,\beta}$ see (B.4)

```
invphicla <- function(t, alpha, beta){
          (alpha*(t^(1/beta))+1)^(-1/alpha)}
```

$K(t) = t - \varphi_{\alpha,\beta}(t)/\varphi'_{\alpha,\beta}(t)$

```
K <- function(t) {(t/(alpha*beta))*(alpha*beta-t^(alpha)+1)}
Kinv <- function(x)
          optimize(function(y)(x - K(y))^2, c(0, 1))$minimum
```

$v_1 = K^{(-1)}(t)$

```
v1 <- sapply(T, Kinv)
unifrand <- matrix(0, n, 2)
```

$u = \varphi^{-1}_{\alpha,\beta}\left(v_2\varphi_{\alpha,\beta}(v_1)\right).$

```
unifrand[, 1] <- (invphicla(phicla(v1, alpha, beta)*v2, alpha, beta))
```

$v = \varphi^{-1}_{\alpha,\beta}\left((1-v_2)\varphi_{\alpha,\beta}(v_1)\right).$

```
unifrand[, 2] <- (invphicla(phicla(v1, alpha, beta)*(1-v2), alpha, beta))
}
```

the desired pair (u, v).

```
return(unifrand)
}
```

#$n = 10$, $\alpha = 0.9$, $\beta = 2$.

```
x <- rclayton2(10, 0.9, 2, dim = 2)
x
```

```
          [ ,1]       [ ,2]
[1, ]   0.26274964  0.30944445
[2, ]   0.64953271  0.54324545
[3, ]   0.99727901  0.99468328
[4, ]   0.02599535  0.04773039
[5, ]   0.06511849  0.05682143
[6, ]   0.22730428  0.24055199
[7, ]   0.42627442  0.41271378
[8, ]   0.56402174  0.69592321
[9, ]   0.15663856  0.33747976
[10, ]  0.43824152  0.36357056
```

B.2.2 Two parameter Gumbel-Hougaard Copula

The two parameter Gumbel-Hougaard Copula is given for α in $(0,1]$, β in $[1,\infty)$ by

$$C_{\alpha,\beta}(u,v) = \left(\exp\left(-\left((-\ln u^{\alpha})^{\beta} + (-\ln v^{\alpha})^{\beta}\right)^{1/\beta}\right)\right)^{1/\alpha},$$

with generator

$$\varphi_{\alpha,\beta}(t) = (-\ln t^{\alpha})^{\beta},$$

and

$$\varphi^{-1}_{\alpha,\beta}(t) = \left(\exp\left(-t^{1/\beta}\right)\right)^{1/\alpha}.$$

Code R of 2-parameter Clayton Copula

```
rgumbel2 <- function (n, alpha, beta, dim = 2)
{
    if (alpha > 1)
    stop("invalid argument :  alpha\n")
    if (beta < 1)
    stop("invalid argument :  beta\n")
    {
       v2 <- runif(n)
       T <- runif(n)
       phi <- function(t, alpha, beta){
             (-log(t^(alpha)))^(beta)}
       invphi <- function(t, alpha, beta){
             (exp(-(t^(1/beta))))^(1/alpha)}
```

```
        K <- function(t) {
              t-((t*log(t^(alpha)))/(alpha*beta))}
        Kinv <- function(x)
             optimize(function(y)(x-K(y))^2, c(0, 1))$minimum
        v1 <- sapply(T, Kinv)
        unifrand <- matrix(0, n, 2)
       unifrand[, 1] <- (invphi(phi(v1, alpha, beta)*v2, alpha, beta))
          unifrand[, 2] <- (invphi(phi(v1, alpha, beta)*(1-v2), alpha,
beta))
        }
          return(unifrand)
}
```

#$n = 10$, $\alpha = 0.8$, $\beta = 2$.

```
x <- rgumbel2(10, 0.8, 2, dim = 2)
x
          [ ,1]        [ ,2]
 [1, ]    0.59997968  0.75029541
 [2, ]    0.37215098  0.20827958
 [3, ]    0.81012388  0.63842415
 [4, ]    0.68762034  0.16793805
 [5, ]    0.06777043  0.34035730
 [6, ]    0.21531555  0.02491624
 [7, ]    0.08653391  0.64281437
 [8, ]    0.58736176  0.30007750
 [9, ]    0.77637822  0.95797613
 [10, ]   0.82762130  0.40694409
```

Bibliography

[1] Acerbi, C., Tasche, D., 2002. On the coherence of expected shortfall. Journal of Banking and Finance 26, 1487-1503.

[2] Albrecht, P., 1992. Premium calculation without arbitrage? A note on a contribution by G. Venter, ASTIN Bulletin 22, 247-254.

[3] Arnold, B.C., Balakrishnan, N., Nagaraja, H.N., 1992. A first course in order statistics. Wiley, New York.

[4] Artzner, P.H., Delbaen, F., Eber, J.M., Heath, D., 1997. Thinking Coherently, RISK 10, 68-71.

[5] Artzner, P.H., Delbaen, F., Eber, J.M., Heath, D., 1999. Coherent measures of risk. Mathematical Finance, 203-28.

[6] Barnett, V., Lewis, T., 1995. *Outliers in statistical data,* Third ed. John Willey & Sons, . New York.

[7] Barlow, R., Proschan, F., 1975. Statistical Theory of Reliability and Life Testing. Holt, Rinehart & Winston, New York.

[8] Beirlant, J., Goegebeur, Y.,Segers, J., Teugels, J., 2005. Statistics of Extremes, Theory and Applications. Jon Wiley & Sons Chicester.

[9] Bowers, N.L., Gerber, H.U., Hickman, J.C., Jones, D.A., Nesbitt, C.J., 1997. Actuarial Mathematics, 2nd edition, Society of Actuaries, Schaumburg, IL

[10] Brahimi, B., Meraghni, D., Necir, A., 2010. Distortion risk measures for sums of dependent losses. Journal Afrika Statistika, Vol 5 N° 9. 260-277.

[11] Burgert, C., Rüschendorf, L., 2006. Consistent risk measures for portfolio vectors. Insurance: Mathematics and Economics 38, 289-297.

[12] Cai, Z., Chen, X., Fan, Y., Wang, X., 2008. Selection of Copulas with Applications in Finance. Unpublished manuscript, University of North Carolina at Charlotte.

[13] Cambanis, S., Huang, S., Simons G., 1981. On the theory of elliptically contoured distributions. Journal of Multivariate Analysis, 11, 368-385.

[14] Campana, A., Ferretti, P., 2007. On bounds for concave distortion risk measures for sums of risks, Mathematical and Statistical Methods in Insurance and Finance, Springer-Verlag, 43-51.

[15] Capéraà, P., Genest C., 1993. Spearman's rho is larger than Kendall's tau for positively dependent random variables. Journal of Nonparametric Statistics, 2, 183-194.

[16] Chebana, F., Ouarda, T.B.M.J., 2007. Multivariate L-moment homogeneity test. Water resources research 43: W08406. DOI: 10-1029/2006WR005639.

[17] Chebana, F., Ouarda, T.B.M.J., Bruneau, B., Barbet, M., El Adlouni, S. Latraverse, M., 2009. Multivariate homogeneity testing in a northern case study in the province of Quebec, Hydrological Processes, 23, 1690-1700, DOI: 10.1002/hyp.7304.

[18] Chen, X. H., Fan, Y. Q., Tsyrennikov, V., 2006. Efficient Estimation of Semiparametric Multivariate Copula Models. Journal of American Statistical Association, Vol. 101, 1228-1241.

[19] Chen, S. X., Huang, T., 2007. Nonparametric Estimation of Copula Functions for Dependence Modeling. Canadian Journal of Statistics Volume 35, Issue 2, pages 265–282.

[20] Cherbini, U., Luciano, E., Vecchiato, W., 2004. Copula Methods in Finance, Wiley.

[21] Coles, S., 2001. An Introduction to Statistical Modeling of Extreme Values. Springer Series in Statistics, Springer New York.

[22] Crane, G., van der Hoek, J., 2008. Using Distortions of Copulas to Price Synthetic CDOs' Insurance: Mathematics and Economics, 42, 903-908.

[23] Darkiewicz, G., Dhaene, J., Goovaerts, M.J., 2004. Distortion risk measures for sums of random variables. Blätter der DGVFM, Springer Berlin/Heidelberg. Volume 26, 631-641.

[24] David, H.A., Nagaraja, H.N., 2003. Order Statistics (3 ed), Wiley, New York.

[25] de Haan, L., Ferreira, A., 2006. Extreme value theory: An Introduction. Springer, New York.

[26] Deheuvels, P., 1979. La fonction de dépendance empirique et ses propriétés. Acad. Roy. Belg. Bull. Cl. Sci. 65, 274-292.

[27] Deheuvels, P., 1981b. A nonparametric test for independence. Publ. Inst. Statist. Univ. Paris., 26, 29-50.

[28] Deheuvels, P., Tiago de Oliveira J., 1989. On the Non-Parametric Estimation of the Bivariate Extreme-Value Distributions Statistics & Probability Letters 8, 315-323.

[29] Delicado, P., Goria, M.N., 2008. A small sample comparison of maximum likelihood, moments and L-moments methods for the asymmetric exponential power distribution. Comput. Statist. Data Anal. 52, 1661-1673.

[30] Denneberg, D., 1994. Non-Additive Measure and Integral. Theory and Decision Library. Series B, vol. 27. Kluwer Academic Publilshers.

[31] Denuit, M., Scaillet, O., 2004. Nonparametric tests for positive quadrant dependence. Journal of Financial Econometrics 2, 422-450.

[32] Denuit, M., Lambert, P., 2005. Constraints on concordance measures in bivariate discrete data. Journal of Multivariate Analysis 93, 40-57.

[33] Dhaene, J., Wang, S.S., Young, V. and Goovaerts, M.J., 2000. Comonotonicity and maximal stop-loss premiums, Schweiz. Aktuarver. Mitt., 99-113.

[34] Drees, H., Müller, P., 2008. Fitting and validation of a bivariate model for large claims. Insurance Math. Econom. 42, 638-650.

[35] Dupuis, D.J., 2007. Using copulas in hydrology: Benefits, cautions, and issues. Journal of Hydrologic Engineering, 12, 381-393.

[36] Durante, F., Sempi, C., 2004. Semi-copulae. Kybernetika 41, 315-328.

[37] Durrleman, V., Nikeghbali, A., Roncalli, T., 2000. A simple transformation of copulas, working paper.

[38] Elamir, E.A.H., Seheult, A.H., 2003. Trimmed L-moments. Comput. Statist. Data Anal. 43, 299-314.

[39] Embrechts, P., 2000. Extreme Value Theory: Potential and Limitations as an Integrated Risk Management Tool Derivatives Use, Trading & Regulation 6, 449-456 .

[40] Embrechts, P., Hoeing, A., Juri., A., 2003. Using Copulae to bound the Value-at-Risk for functions of dependent risks Finance & Stochastics 7, 145-167

[41] Embrechts, P., McNeil, A., and Straumann D., 2002. Correlation and Dependence in Risk Management: Properties and Pitfalls To appear in Risk Management: Value at Risk and Beyond. Ed. by M., Dempster and H., K., Moffatt, Cambridge University Press, 176-223.

[42] Embrechts, P., Mikosch, T., Klüppelberg, C., 1997. Modelling Extremal Events for Insurance and Finance. Springer, Berlin.

[43] Fang, K.T., Kotz, S., Ng, K.W., 1990. Symmetric multivariate and related distributions. London: Chapman and Hall.

[44] Ferguson, T.S., 1967. Mathematical Statistics: A Decision Theoretic Approach. Probability and Mathematical Statistics, Vol. 1. New York: Academic Press.

[45] Fermanian, J.D., Radulovic, D., Wegkamp, M.H., 2004. Weak convergence of empirical copula processes. Bernoulli 10, 847-860.

[46] Fermanian, J.D., Scaillet, O., 2005. Sensitivity analysis of VaR and Expected shortfall for portfolios under netting agreements. Journal of Banking & Finance 29, 927-958.

[47] Fermanian, J.D. and Scaillet, O., 2004. Some Statistical Pitfalls in Copula Modeling for Financial Applications. FAME Working Paper, 108.

[48] Frank, M., 1979. On the simultaneous associativity of $F(x,y)$ and $x+y-F(x,y)$. Aequationes Mathematicae, 19, 194-226.

[49] Fréchet, M., 1951. Sur les tableaux de corrélation dont les marges sont données. Annales de l'Université de Lyon. Sciences Mathématiques et Astronomie, 14, 53-77.

[50] Fréchet, M., 1957. Les tableaux de corrélation dont les marges et des bornes sont données. Annales de l'Université de Lyon. Sciences Mathématiques et Astronomie, 20, 13-31.

[51] Frees, W.E., Valdez, E.A., 1998. Understanding relationships using copulas. North American Actuarial Journal 2, 1-25.

[52] Furman, E., Zitikis, R., 2008a. Weighted premium calculation principles. Insurance: Mathematics and Economics, 42, 459-465.

[53] Furman, E., Zitikis, R., 2008b. Weighted risk capital allocations. Insurance: Mathematics and Economics, 43, 263-269.

[54] Gaenssler, P., Stute, W., 1987. Seminar on Empirical Processes, DMV Sem. 9. Basel: Birkhaüser.

[55] Genest, C., Boies, J.C., 2003. Detecting Dependance with Kendall Plots, American Statistical Association.

[56] Genest, C., 1987. Frank's family of bivariate distributions. Biometrika 74, 549-555.

[57] Genest, C., MacKay, R., 1986a. The joy of copulas: Bivariate distributions with uniform marginals. The American Statistician, 40, 280-283.

[58] Genest, C., MacKay, R., 1986b. Copules archimédiennes et familles de lois bidimensionell dont les marges sont données. The Canadian Journal of Statistics, 14, 145-159.

[59] Genest, C., Rivest L.P., 1993. Statistical inference procedures for bivariate Archimedean copulas. Journal of the American Statistical Association, 88, 1034-1043.

[60] Genest, C., Ghoudi, K., Rivest, L.P., 1995. A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. Biometrika 82, 543-552.

[61] Genest, C., Rémillard, B., 2008. Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models. Ann. Inst. Henri Poincaré Probab. Stat. 44, 1096-1127.

[62] Genest, C., Rémillard, B., Beaudoin, D., 2009. Goodness-of-fit tests for copuluas: a review and a power study. Insurance Math. Econom. 44, 199-213.

[63] Genest, C., Rivest, L.P., 2001. On the multivariate probability integral transformation. Statist. Probab. Lett. 53, 391-399.

[64] Genest, C., Werker, B.J.M., 2002. Conditions for the asymptotic semiparametric efficiency of an omnibus estimator of dependence parameters in copula models. In: Distributions with Given Marginals and Statistical Modelling. Kluwer, Dordrecht, The Netherlands, 103-112.

[65] Gijbels, I., Mielniczuk, J., 1990. Estimating the density of a copula function. Comm. Statist. Theory Methods, 19, 445-464.

[66] Gregor, N.F.W., 2009. Copula Parameter Estimation by Maximum Likelihood and Minimum Distance Estimators: A Simulation Study. Social Science Research Network.

[67] Goovaerts, M.J., De Vijlder, F.E., Haezendonck, J., 1984. Insurance Premiums: Theory and Applications. North-Holland, Amsterdam.

[68] Goovaerts, M.J., Kaas, R., Dhaene, J., Tang, Q., 2003. A unified approach to generate risk measures. ASTIN Bulletin 33, 173-191.

[69] Goovaerts, M.J., Van den Borre, E., Laeven, R.J.A., 2004. Manageing economic and virtual economic capital within financial conglomerates. North American Actuarial Journal, in press.

[70] Hall, P., Neumeyer, N., 2006. Estimating a Bivariate Density when there are Extra Data on One or Both Components. Biometrika, Vol. 93, 439-450.

[71] Hoeffding, W., 1940. Massstabinvariante Korrelationstheorie. Schriften des Mathematischen Seminars und des Instituts für Angewandte Mathematik der Universität Berlin, 5, 181-233.

[72] Hoeffding, W., 1941. Masstabinvariante Korrelationsmasse für Diskontinuierliche Verteilungen. Archiv für Mathematische Wirtscafts. 7, 49-70,

reprinted as Scale- Invariant correlations for discontinuous distributions in The collected works of Wassily Hoeffding, Fisher and Sen eds., Springer Verlag, (1994), 109-132.

[73] Hoeffding, W., 1942. Stochastische Abhängigkeit und Funktionaler Zusammenhang. Skandinavisk Aktuarietidskrift. 25, 200-227, reprinted as Stochastic dependence and functional relationships in The collected works of Wassily Hoeffding, Fisher and Seneds., Springer Verlag, (1994), 109-132.

[74] Hoeffding, W., 1948. A class of statistics with asymptotically normal distribution. Annals of Mathematical Statistics. 19, 293-325.

[75] Hosking, J.R.M., 1990. L-moments: analysis and estimation of distributions using linear combinations of order statistics. J. Royal Statist. Soc. Sev. B 52, 105-124.

[76] Hosking, J.R.M., Wallis, J.R., 1997. Regional Frequency Analysis. An Approach Based on L-moments. Cambridge University Press, Cambridge, New York. Melbourne.

[77] Hosking, J.R.M., 1998. L-moments. In: Kotz, S., Read, C., Banks, D.L., (Eds.), Encyclopedia of Statistical Sciences, vol. 2. Wiley, New York, pp. 357-362.

[78] Hosking, J.R.M., 2006. On the characterization of distributions by their L-moments. J. Statist. Plann. Inference 136, 193-198.

[79] Hosking, J.R.M., 2007. Some theory and practical uses of trimmed L-moments. J. Statist. Plann. Inference 137, 3024-3039.

[80] Huang, J.S., Kotz, S.,1984. Correlation structure in iterated Farlie-Gumbel-Morgenstern distributions. Biometrika 71, 633-636.

[81] Hult, H., Lindskog F., 2002. Multivariate extremes, aggregation and dependencein elliptical distributions. Advances in Applied Probability, 34, 587-608.

[82] Hürlimann, W., 1998. On stop-loss order and the distortion pricing principle. ASTIN Bulletin 28, 119-134.

[83] Hürlimann, W., 2004. Fitting bivariate cumulative returns with copulas. Computational Statistics & Data Analysis 45, 355-372.

[84] Joe, H., 1993. Parametric families of bivariate distributions with given marginals. Journal of Multivariate Analysis 46, 262-282.

[85] Joe, H., 1997. Multivariate Models and Dependence Concepts. Chapman & Hall, London.

[86] Joe, H., 2005. Asymptotic efficiency of the two-stage estimation method for copula-based models. J. Multivariate Anal. 94, 401-419.

[87] Johnson, N.L., Kotz, S., 1972. Distributions in statistics: Continuous multivariate distributions, John Wiley and Sons, New York.

[88] Johnson, N. L., Kotz, S., 1977. On some generalized Farlie- Gumbel-Morgenstern distributions-II: Regression, correlation and further generalizations, Communications in Statistics. Theor. Meth., 6,pp. 485-496.

[89] Jones, M., 1993. Simple Boundary Correction for Kernel Density Estimation. Statistical Computing, 3, 135-146.

[90] Jones, M., Balakrishnan, N., 2002. How are moments and moments of spacings related to distribution functions? J. Statist. Plann. Inference 103, 377-390.

[91] Jorion, P., 2000. Value at Risk. McGraw-Hill, New York.

[92] Kaas, R., Goovaerts, M.J., Dhaene, J., Denuit, M., 2001. Modern Actuarial Risk Theory. Kluwer Academic Publishers, Dordrecht.

[93] Kaas, R., van Heerwaarden, A.E., Goovaerts, M.J., 1994. Ordering of Actuarial Risks, Education Series 1, CAIRE, Brussels.

[94] Karvanen, J., 2006. Estimation of quantile mixtures via L-moments and trimmed L-moments. Comput. Statist. Data Anal. 51, 947-959.

[95] Karvanen, J., Eriksson, J., Koivunen, V., 2002. Adaptive score functions for maximum likelihood ICA. J. VLSI Signal Process. 32, 83-92.

[96] Kendall, M., 1938. A New Measure of Rank Correlation. Biometrika, 30, 81-89.

[97] Kendall, M., Stuart, A., 1979. Handbook of Statistics. Griffin & Company, London.

[98] Kim, G., Silvapulle, M.J., Silvapulle, P., 2007. Comparison of semiparametric and parametric methods for estimating copulas. Comm. Statist. Simulation Comput. 51, 2836-2850.

[99] Kimberling, C., 1974. A probabilistic interpretation of complete monotonicity," Aequationes Mathematicae, 10, 152-164.

[100] Kjeldsen, T.R., Smithers, J.C., Schulze, R.E., 2002. Regional flood frequency analysis in the KwaZulu-Natal province, South Africa, using the index-flood method. J. Hydrol. 255, 194-211.

[101] Kotz, S., Nadarajah, S., 2004. Multivariate t distributions and their applications. Cambridge University Press.

[102] Kroll, C.N., Vogel, R.M., 2002. Probability distribution of low streamflow series in the United States. J. Hydrol. Eng. 7, 137-146.

[103] Kruskal, W., 1958. Ordinal measures of association. Journal of the American Statistical Association, 53, 814-861.

[104] Laeven, R.J.A., and Goovaerts, M.J., 2004. An optimization approach to the dynamic allocation of economic capital. Insurance: Mathematics & Economics, 35, 299-320.

[105] Lehmann, E., 1975. Nonparametrics: Statistical Methods Based on Ranks. Holden-Day, Inc., San Francisco.

[106] Li, X., Mikusiński, P., Sherwood H., Taylor M.D., 1997. On approximation of copulas, in V. Beneš and J. Štěpán (Eds.), Distributions with Given Marginals and Moment Problems, Kluwer Academic Publishers, Dordrecht.

[107] Li, Q., Racine, J., 2007. Nonparametric Econometrics: Theory and Practice, Princeton University Press, New Jersey.

[108] Liebscher, E., 2005. Semiparametric Density Estimators Using Copulas. Communications in Statistics A, Vol. 34, 59-71.

[109] Lim,Y.H., Lye, L.M., 2003. Regional flood estimation for ungauged basins in Sarawak, Malaysia. Hydrol. Sci. J. 48, 79-94.

[110] Lindskog, F., McNeil, A., Schmock, U., 2003, Kendall's tau for elliptical distributions. In: G. Bohl, G. Nakhaeizadeh, S.T. Rachev, T. Ridder and K.H.

Vollmer (Eds.), Credit Risk: Measurement, Evaluation and Management, Physica-Verlag Heidelberg, 649–676.

[111] Lin, G.D., 1987. Relationships between two extensions of Farlie–Gumbel–Morgenstern distribution. Annals of the Institute of Statistical Mathematics **39**, 129-140.

[112] Malevergne, Y., Sornette, D., 2003. Testing the Gaussian copula hypotheses for financial assets dependences. Quant. Finance 3, 231-250.

[113] Marshall, A., 1996. Copulas, marginals and joint distributions. In Distributions with Fixed Marginals and Related Topics. Ed. by L. Rüschendorff, B. Schweizer, and M. Taylor, pp. 213-222. Institute of Mathematical Statistics, Hayward, CA.

[114] Marshall, A., and Olkin I., 1967. A multivariate exponential distribution. Journal of the American Statistical Association, 62, 30-44.

[115] Mendes, B., de Melo, E. and Nelsen, R., 2007. Robust Fits for Copula Models. Comm. Statist. Simulation Comput. 36, 997-1017.

[116] Mikusinski, P., Sherwood H., Taylor, M., 1992. The Fréchet bounds revisited. Real Analysis Exchange, 17, 759-764.

[117] Morillas, P. M., 2005. A method to obtain new copulas from a given one. Metrika, 61, 169-184.

[118] Muller, H., 1991. Smooth Optimum Kernel Estimators Near Endpoints. Biometrika, 78, 521-530.

[119] Muliere, P., Scarsini, M., 1987. Cheracterization of a Marshall-Olkin type class of distributions. Annals of the Institute of Statistical Mathematics, 39, 429-441.

[120] Necir, A., Brahimi, B., Meraghni, D., 2010. Erratum to: ‘Statistical estimate of the proportional hazard premium of loss’. *Scandinavian Actuarial Journal*, **3**, 246-247.

[121] Necir, A., Meraghni, D., 2009. Empirical estimation of the proportional hazard premium for heavy-tailed claim amounts, Insurance: Mathematics and Economics 45, 49-58.

[122] Nelsen R.B., 1991. Copulas and association. In: Dall'Aglio G, Kotz S, Salinetti G (eds) Advances in Probability Distributions with Given Marginals. Kluwer, Dordrecht, 51-74.

[123] Nelsen, R.B., 2006. An Introduction to Copulas, second ed. Springer, New York.

[124] Oakes, D., 1982. A model for association in bivariate survival data. J. Roy. Statist. Soc. Ser., 414-422.

[125] Oakes, D., 1986. Semiparametric Inference in a Model for Association in Bivariate Survival Data. Biometrika, 73, 353-361.

[126] Oakes, D., 1994. Multivariate survival distributions. J. Nonparametr. Statist. 3, 343-354.

[127] Panjer, H.H., 1998. Financial Economics, with Applications to Investments, Insurance and Pensions. Actuarial Foundation, Schaumburg, IL.

[128] Reiss, R.D., Thomas M., 2001. Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields, 2nd edition, BirkhÄauser Basel.

[129] Resnick, S., 1987. Extreme values, Regular variation and Point processes. Springer, New York.

[130] Resnick, S., 2007. Heavy-Tail phenomina. Probabilistic and statistical modeling. Springer, New York.

[131] Rothschild, M., Stiglitz, J., 1970. Increasing risk: I. A definition, Journal of Economic Theory 2, 225-243.

[132] Rüschendorf, L., 1976. Asymptotic distributions of multivariate rank order statistics. Ann. Statist. 4, 912-923.

[133] Sancetta, A., Satchell, S., 2004. The Bernstein Copula and Its Applications to Modeling and Approximations of Multivariate Distributions", Econometric Theory, Vol. 20, 535-562.

[134] Scarsini, M., 1984. On measures of concordance. Stochastica, 8, 201-218.

[135] Schmid, F., Schmidt, R., 2007. Multivariate extensions of Spearman's rho and related statistics. Statist. Probab. Lett. 77, 407-416.

[136] Schweizer, B., Sklar, A., 1983. Probabilistic Metric Spaces. North-Holland, New York.

[137] Schweizer, B., Wolff, E., 1981. On nonparametric measures of dependence for random variables. Annals of Statistics, 9, 879-885.

[138] Serfling, R., Xiao, P., 2007. A Contribution to Multivariate L-Moments, L-comoment Matrices. J. Multivariate Anal. 98, 1765-1781.

[139] Sillitto, G.P., 1951. Interrelations between certain linear systematic statistics of samples from any continuous population. Biometrika 38, 377-382.

[140] Sillitto, G.P., 1969. Derivation of approximants to the inverse distribution function of a continuous univariate population from the order statistics of a sample. Biometrika 56, 641-650.

[141] Sklar, A., 1996. Random Variables, Distribution Functions, and Copulas - a Personal Look Backward and Forward published in Distributions with Fixed Marginals and Related Topics, edited by L. Rüschendorf, B. Schweizer, and M.D. Taylor, Institute of Mathematical Statistics, Hayward, CA, 1-14.

[142] Sklar, A., 1959. Fonctions de répartition à n dimensions et leurs marges, Publ. Inst. Statist. Univ. Paris 8, 229-231.

[143] Stone, C., 1990. Large-sample Inference for Log-spline Models. The Annals of Statistics, 18, 717-741.

[144] Tsukahara, H., 2005. Semiparametric estimation in copula models. Canad. J. Statist. 33, 357-375.

[145] Van Heerwaarden, A., 1991. Ordering of Risks: Theory and Actuarial Applications, Thesis Publishers, Amsterdam.

[146] Varadhan, R., Gilbert, P.D., 2009. BB: An R Package for Solving a Large System of Nonlinear Equations and for Optimizing a High-Dimensional Nonlinear Objective Function, J. Statistical Software, 32:4, http://www.jstatsoft.org/v32/i04/.

[147] Valdez, E.A., Xiao, Y., 2010. On the distortion of a copula and its margins. Scandinavian Actuarial Journal (in press).

[148] Veraverbeke, N., 2005. Multivariate Data and Copulas. Hasselt University, Belgium, Lectures notes.

[149] Wang, S.S., 1995. Insurance Pricing And Increased Lmuts Ratemakmg By Proporltional Hazards Transforms Insurance: Mathematics and Economics, 17, 43-54.

[150] Wang, S.S., 1996. Premium calculation by transforming the layer premium density. ASTIN Bulletin, 26, 71-92.

[151] Wang, S.S., 2000. A Class of distortion operators for pricing financial and insurance risks. Journal of Risk and Insurance, 67, 15-36.

[152] Wang, S.S., Young, V.R., Panjer, H.H., 1997. Axiomatic characterization of insurance prices Insurance: Mathematics and Economics, 21, 173-183.

[153] Wirch, J.L., and Hardy, M.R., 2000. A synthesis of risk measures. Insurance: Mathematics and Economics. 25, 337-347.

[154] Yaari, M.E. (1987). The dual theory of choice under risk, Econometrica 55, 95-115.

[155] Zaidman, M.D., Keller, V., Young, A.R., Cadman, D., 2003. Flow-duration-frequency behaviour of British rivers based on annual minima data. J. Hydrol. 277, 195-213.

Abstract

The main motivation of my Doctorate thesis is to give a new procedure estimate of copula parameters using the bivariate L-moments, we have already proved performances of this approach, we compared, by simulations, the new method with the well-known (τ,ρ)-inversion, the minimum distance and the pseudo maximum likelihood estimators, which is recommended by some authors. We have show that BLM estimator is good candidate to construct goodness-of-fit tests of copula models. The second one is proposing of a risk measure for the sum of two losses by simultaneously transforming the distribution tail and the copula, which represents the dependence between the margins, by means of two distortion functions. We obtained a coherent Copula Distorted Risk Measure. This new measure has the characteristic to be greater than the expectation and less than the popular Wang's distorted risk measure, the main advantage of this property is to reduce Wang's premium while respecting the standard axioms of the premium principle.

Résumé

La principale motivation de ma thèse de doctorat est de donner une nouvelle procédure d'estimation des paramètres de la copule on utilisant les L-moments bivariées, nous avons prouvé les performances de cette approche, nous avons comparé, par simulations, la nouvelle méthode avec le (τ, ρ)-inversion, la distance minimale et l'estimateur du pseudo maximum de vraisemblance, ce qui est recommandé par certains auteurs. Nous avons montré que l'estimateur BLM est bon candidat pour construire des tests de qualité de l'ajustement des modèles de copules. La deuxième chose est de proposer une mesure de risque pour la somme de deux pertes en même temps de transformer la queue de distribution et la copule, qui représente la dépendance entre les marges, à l'aide de deux fonctions de distorsion. Nous avons obtenu une mesure cohérente de risque de copule déformée. Cette nouvelle mesure a la particularité d'être plus grande que l'espérance et moins que la mesure de risque de Wang, le principal avantage de cette propriété est de réduire la prime de Wang, tout en respectant les axiomes standards du principe de prime.

الخلاصة

الهد ف الرئيسي لهذه الرسالة هو تقديم طريقة جديدة لتقدير عوامل الربط باستخدام الخطيات الوقتية ذات متغيرين ، لقد أثبتنا دقة هذه الطريقة ، من خلال المحاكاة. لقد أثبتنا أن هذا المقدر هو مرشح جيد لبناء نموذج اختبار جودة. الشيء الثاني هو طرح اجراء خطر لمجموع الخسائربمراقبة الذيل والرابط في نفس الوقت ، وحصلنا على مقياس ثابت من خطر الروابط المشوهة. هذا الاجراء الجديد قد تميز بأنه أكبر من الأمل الرياضي وأقل من مقياس المخاطر المشوهة لوانغ ، والميزة الرئيسية لهذه الخاصية هي للتقليل من قسط وانغ ، مع احترام مبدأ البديهيات القياسية .