Eidgenössische
Technische Hochschule
Zürich

Ecole polytechnique fédérale de Zurich
Politecnico federale di Zurigo
Swiss Federal Institute of Technology Zurich

Modelling Dependence with Copulas
and
Applications to Risk Management

Master Thesis by Filip Lindskog
For: Dr. Alexander J. McNeil

Period: September 1999 to January 2000
Version: July 2, 2000

Abstract

The dependence between random variables is completely described by their joint distribution. However, dependence and marginal behavior can be separated. The copula of a multivariate distribution can be considered to be the part describing the dependence structure. Furthermore, strictly increasing transformations of the underlying random variables result in the transformed variables having the same copula. Hence copulas are invariant under strictly increasing transformations of the margins. This provides a way of studying scale-invariant measures of associations and also a starting point for construction of multivariate distributions. Scale-invariant measures of association such as Kendall's tau and Spearman's rho only depend on the copula and are thus invariant under strictly increasing transformations of the margins, which means that we can apply arbitrary continuous margins to our chosen copula leaving among other things the measures of association unchanged.

Tail dependence and Kendall's tau and Spearman's rho are presented and evaluated for a large number of copula families. Among these copula families are families suitable for modelling extreme events, which are highly relevant as a basis for risk models in insurance and finance.

The multivariate normal distribution and linear correlation are the basis of most models used to model dependence. Even though this distribution has a wide range of dependence it is quite seldom suitable for modelling real world situations in insurance and finance. We will show that using a model based on the multivariate normal distribution without knowledge of its limitations can prove very dangerous. Linear correlation is a natural measure of dependence in the context of the normal distribution. However, it should be noted that it is not invariant under strictly increasing transformations of the marginals and can be misleading as a measure of dependence.

The problem of simulating dependent data arises naturally in Monte Carlo approaches to risk management. One main aim of this paper is to show that when addressing this problem knowledge of copulas and copula based dependence concepts is important, and also the usefulness of copula ideas in this approach to risk management. Another main aim of this paper is the construction of multivariate extensions of bivariate copula families. In particular we focus on multivariate extensions with a flexible and wide range of dependence for which efficient algorithms for random variate generation are presented.

Acknowledgements

This thesis was written during my stay at ETH in Zürich the fall 1999. I would like to express my gratitude and appreciation to Alexander McNeil for whom I had the pleasure of working with my thesis. I would also like to thank RiskLab for providing me with an office and Roger Kaufmann at RiskLab for valuable suggestions regarding the layout. Finally, I would like to thank Paul Embrechts at ETH and Jan Grandell, my supervisor at KTH in Stockholm, for giving me the opportunity to visit ETH.

Contents

1 Motivation

Consider a portfolio of n financial risks $X_1, \dots, X_n$. This might be

1. a portfolio of traded financial assets e.g. equities,
2. a credit portfolio of loans to various counterparties,
3. a portfolio of potential insurance losses in various lines of business or geographical areas.

Suppose that we have n-variate samples from this portfolio and that we want to examine the distribution of some pay-off function $f(X_1, \dots, X_n)$ representing the risk of the portfolio or some contract written on the portfolio. In general estimating a joint distribution describing this portfolio is very difficult and requires a lot of data. Since estimation of the univariate marginal distributions is well understood, the following approach may seem natural:

1. Estimate marginal distributions $F_1, \dots, F_n$,
2. Estimate a matrix of pairwise linear correlations, ρ_{ij},
3. Use this information in some Monte Carlo simulation procedure to generate dependent vectors.

However this approach is in general not good. Even if the matrix of estimated linear correlations is a proper linear correlation matrix it is possible that the estimated linear correlations are not consistent with the marginal distributions so that no corresponding multivariate distribution exists. If a multivariate distribution exists, it is in general not unique. If instead of linear correlations we use so called rank correlations, then the problem of consistency with the marginals does not appear. If moreover the matrix of estimated rank correlations is a proper rank correlation matrix then there exists a multivariate distribution having this rank correlation matrix and margins $F_1, \dots, F_n$. However, there is still no unique solution and finding such a multivariate distribution with certain statistical properties indicated by the data is often difficult.

Rather than looking simply at correlations we can consider so called copulas, which are multivariate uniform distributions, to describe dependence. In the case when marginal distributions $F_1, \dots, F_n$ and an n-copula C are specified, a unique multivariate distribution H given by

$$H(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n))$$

exists, with marginal distributions $F_1, \dots, F_n$. Furthermore, given a proper rank correlation matrix it is often not difficult to find a copula with a parametrization that gives this rank correlation matrix. The problem of simulating from the multivariate distribution H is no longer a theoretical problem. The problem is rather finding efficient simulation techniques for the desired copulas. The general approach is the following:

1. Estimate marginal distributions $F_1, \dots, F_n$,
2. Estimate a matrix of pairwise rank correlations,
3. Choose an n-copula C having the wanted rank correlation matrix,
4. Simulate a random vector $(U_1, \dots, U_n)$ having joint distribution C,
5. Apply the transformations $u_i \mapsto F_i^{-1}(u_i)$ for $i = 1, \dots, n$ to the ith component. Then $(F_1^{-1}(U_1), \dots, F_n^{-1}(U_n))$ have joint distribution H with marginal distributions $F_1, \dots, F_n$ and the desired rank correlation matrix.

This is one example showing that copulas are useful when modelling dependent risks. More generally, copulas represent the dependence structure of multivariate distributions. Hence when modelling dependence, knowledge of copulas is essential. In this paper we discuss a large number of copula families and suggest efficient simulation techniques. The results provide a basis for those interested in modelling dependence in theory and practice.

2 Copulas

2.1 Mathematical Introduction

Definition 1. Let S_1, $S_2, \ldots S_n$ 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 n variables such that $\mathrm{Dom}H = S_1 \times S_2 \times \ldots \times S_n$ and let $B = [\mathbf{a}, \mathbf{b}]$ be an n-box all of whose vertices are in $\mathrm{Dom}H$. Then the H-volume of B is given by

$$V_H(B) = \sum \operatorname{sgn}(\mathbf{c}) H(\mathbf{c}), \tag{2.1.1}$$

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

$$\operatorname{sgn}(\mathbf{c}) = \begin{cases} 1, & \text{if } c_k = a_k \text{ for an even number of } k\text{'s}, \\ -1, & \text{if } c_k = a_k \text{ for an odd number of } k\text{'s}. \end{cases} \tag{2.1.2}$$

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

$$V_H(B) = \triangle_{\mathbf{a}}^{\mathbf{b}} H(\mathbf{t}) = \triangle_{a_n}^{b_n} \triangle_{a_{n-1}}^{b_{n-1}} \ldots \triangle_{a_1}^{b_1} H(\mathbf{t}),$$

where we define the n first order differences as

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

Definition 2. A real function H of n variables is n-increasing if $V_H(B) \geq 0$ for all n-boxes B whose vertices lies in $\mathrm{Dom}H$.

Suppose that the domain of a real function H of n variables is given by $\mathrm{Dom}H = S_1 \times S_2 \times \ldots \times S_n$ where each S_k has a least element a_k. We say that H is grounded if $H(\mathbf{t}) = 0$ for all $\mathbf{t}$ in $\mathrm{Dom}H$ such that $t_k = a_k$ for at least one k. If each S_k is nonempty and has a greatest element b_k, then we say that H has margins, and the one-dimensional margins of H are the functions H_k given by $\mathrm{Dom}H_k = S_k$ and

$$H_k(x) = H(b_1, \ldots, b_{k-1}, x, b_{k+1}, \ldots, b_n) \text{ for all } x \text{ in } S_k. \tag{2.1.3}$$

Higher dimensional margins are defined by fixing fewer places in H. One-dimensional margins will be called simply "margins" and for $k \geq 2$ k-dimensional margins will be called "k-margins".

Lemma 2.1. *Let $S_1, S_2, \ldots, S_n$ be nonempty subsets of $\overline{\mathbb{R}}$, and let H be a grounded n-increasing function with domain $S_1 \times S_2 \times \ldots \times S_n$. Then H is nondecreasing in each argument, that is, if $(t_1, \ldots, t_{k-1}, x, t_{k+1}, \ldots, t_n)$ and $(t_1, \ldots, t_{k-1}, y, t_{k+1}, \ldots, t_n)$ are in DomH and $x < y$, then $H(t_1, \ldots, t_{k-1}, x, t_{k+1}, \ldots, t_n) \leq H(t_1, \ldots, t_{k-1}, y, t_{k+1}, \ldots, t_n)$.*

Lemma 2.2. *Let $S_1, S_2, \ldots, S_n$ be nonempty subsets of $\overline{\mathbb{R}}$, and let H be a grounded n-increasing function with margins whose domain is $S_1 \times S_2 \times \ldots \times S_n$. Let $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, and $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ be any points in $S_1 \times S_2 \times \ldots \times S_n$. Then*

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

For the proof, see Schweizer and Sklar (1983) [15].

Definition 3. An n-dimensional subcopula is a function C' with the following properties:

1. Dom$C' = S_1 \times S_2 \times \ldots \times S_n$, where each S_k is a subset of $\mathbf{I}$ containing 0 and 1;
2. C' is grounded and n-increasing;
3. C' has margins C'_k, $k = 1, 2, \ldots, n$, which satisfy
   $C'_k(u) = u$ for all u in S_k.

Note that for every u in DomC', $0 \leq C'(\mathbf{u}) \leq 1$, so that Ran$C'$ is also a subset of $\mathbf{I}$.

Definition 4. An n-dimensional copula is an n-subcopula C whose domain is $\mathbf{I}^n$.

Equivalently, an n-copula is a function C from $\mathbf{I}^n$ to $\mathbf{I}$ with the following properties:

1. For every $\mathbf{u}$ in $\mathbf{I}^n$,
   $C(\mathbf{u}) = 0$ if at least one coordinate of $\mathbf{u}$ is 0
   and
   if all coordinates of $\mathbf{u}$ is 1 except u_k, then $C(\mathbf{u}) = u_k$;
2. For every $\mathbf{a}$ and $\mathbf{b}$ in $\mathbf{I}^n$ such that $\mathbf{a} \leq \mathbf{b}$,
   $V_C([\mathbf{a}, \mathbf{b}]) \geq 0$.

Note that for any n-copula C, $n \geq 3$, each k-margin of C is a k-copula, $2 \leq k < n$. The following theorem follows directly from Lemma 2.2.

Theorem 2.1. *Let C' be an n-subcopula. Then for every $\mathbf{u}$ and $\mathbf{v}$ in DomC',*

$$\left|C'(\mathbf{v}) - C'(\mathbf{u})\right| \leq \sum_{k=1}^{n} \left|v_k - u_k\right|. \tag{2.1.4}$$

Hence C' is uniformly continuous on its domain.

2.2 Sklar's Theorem

Definition 5. An n-dimensional distribution function is a function H with domain $\overline{\mathbb{R}}^n$ such that

1. H is n-increasing,
2. $H(\mathbf{t}) = 0$ for all $\mathbf{t}$ in $\overline{\mathbb{R}}^n$ such that $t_k = -\infty$ for at least one k, and
   $H(\infty, \infty, \ldots, \infty) = 1$.

Thus H is grounded, and since Dom$H = \overline{\mathbb{R}}^n$, it follows from Lemma 2.1 that the margins of an n-dimensional distribution function are distribution functions, which we will denote $F_1, F_2, \ldots, F_n$.

Theorem 2.2. *Sklar's theorem in n-dimensions. Let H be an n-dimensional distribution function with margins $F_1, F_2, \ldots, F_n$. Then there exists an n-copula C such that for all $\mathbf{x}$ in $\overline{\mathbb{R}}^n$,*

$$H(x_1, x_2, \ldots, x_n) = C(F_1(x_1), F_2(x_2), \ldots, F_n(x_n)). \tag{2.2.1}$$

If $F_1, F_2, \ldots, F_n$ are all continuous, then C is unique; otherwise C is uniquely determined on Ran$F_1 \times$Ran$F_2 \times \ldots \times$ RanF_n. Conversely, if C is an n-copula and $F_1, F_2, \ldots, F_n$ are distribution functions, then the function H defined above is an n-dimensional distribution function with margins $F_1, F_2, \ldots, F_n$.

For the proof, see Sklar (1996) [16].

From Sklar's theorem we see that for continuous multivariate distributions, the univariate margins and the multivariate dependence structure can be separated, and the dependence structure can be represented by a copula.

Definition 6. Let F be a distribution function. Then the quasi-inverse of F is any function $F^{(-1)}$ with domain $\mathbf{I}$ such that

1. if t is in Ran$F$, then $F^{(-1)}(t)$ is any number x in $\overline{\mathbb{R}}$ such that $F(x) = t$ i.e., for all t in Ran$F$, $F(F^{(-1)}(t)) = t$;
2. if t is not in Ran$F$, then $F^{(-1)}(t) = \inf\{x|F(x) \geq t\} = \sup\{x|F(x) \leq t\}$.

If F is strictly increasing, then the quasi-inverse is the ordinary inverse, which we denote F^{-1}.

Corollary 2.1. *Let H, C, $F_1, F_2, \dots, F_n$ be as in Theorem 2.2, and let $F_1^{(-1)}, F_2^{(-1)}, \dots, F_n^{(-1)}$ be quasi-inverses of $F_1, F_2, \dots, F_n$, respectively. Then for any $\mathbf{u}$ in $\mathbf{I}^n$,*

$$C(u_1, u_2, \dots, u_n) = H(F_1^{(-1)}(u_1), F_2^{(-1)}(u_2), \dots, F_n^{(-1)}(u_n)). \tag{2.2.2}$$

Example 2.1. Let Φ denote the standard univariate normal distribution function and let Φ_ρ^n denote the standard multivariate normal distribution function with linear correlation matrix ρ. Then

$$C(u_1, u_2, \dots, u_n) = \Phi_\rho^n(\Phi^{-1}(u_1), \Phi^{-1}(u_2), \dots, \Phi^{-1}(u_n))$$

is the Gaussian or normal n-copula. This copula is used very often in practice because it has some nice properties as a member of the family of elliptical copulas, and as it is easy to simulate from as well as practitioners knowledge of copulas is quite limited. That the linear correlation matrix is used as parameterization of the copula is not standard for copulas. The reasons for this follow from results in the preceding chapters.

2.3 The Fréchet-Hoeffding Bounds for Joint Distribution Functions

Consider the functions M^n, Π^n and W^n given by

$$\begin{aligned} M^n(\mathbf{u}) &= \min(u_1, u_2, \dots, u_n); \\ \Pi^n(\mathbf{u}) &= u_1 u_2 \dots u_n; \\ W^n(\mathbf{u}) &= \max(u_1 + u_2 + \dots + u_n - n + 1, 0). \end{aligned}$$

The functions M^n and Π^n are n-copulas for all $n \geq 2$ whereas the function W^n is not a copula for any $n > 2$ as shown in the following example.

Example 2.2. Consider the n-cube $[1/2, 1]^n \subset \mathbf{I}^n$. Because W is symmetric the W-volume of $[1/2, 1]^n$ is given by

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

Hence, W^n is not a copula for $n \geq 3$.

The following theorem is the n-dimensional version of the Fréchet-Hoeffding bounds inequality.

Theorem 2.3. *If C' is any n-subcopula, then for every* $\mathbf{u}$ *in DomC′,*

$$W^n(\mathbf{u}) \leq C'(\mathbf{u}) \leq M^n(\mathbf{u}). \tag{2.3.1}$$

For the proof, see Nelsen (1999) [13].

Although the Fréchet-Hoeffding lower bound W^n is never a copula for $n > 2$, the left-hand side in (2.3.1) is the best possible in the sense that for all $n \geq 3$ and any $\mathbf{u}$ in $\mathbf{I}^n$, there is an n-copula C such that $C(\mathbf{u}) = W^n(\mathbf{u})$.

Theorem 2.4. *For any $n \geq 3$ and any* $\mathbf{u}$ *in* $\boldsymbol{I}^n$*, there is an n-copula C (which depends on* $\mathbf{u}$*) such that*

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

For the proof, see Nelsen (1999) [13].

Definition 7. If C_1 and C_2 are copulas, we say that C_1 is smaller than C_2 (or C_2 is larger than C_1), and write $C_1 \prec C_2$ (or $C_2 \succ C_1$) if $C_1(u_1, u_2, \dots, u_n) \leq C_2(u_1, u_2, \dots, u_n)$ for all $u_1, u_2, \dots, u_n$ in $\mathbf{I}$.

Thus the Fréchet-Hoeffding lower bound W^n is smaller than every n-copula C and the Fréchet-Hoeffding upper bound M^n is larger than every n-copula C. This partial ordering of the set of copulas is called a concordance ordering. It is a partial ordering since not every pair of copulas is comparable. However many important parametric families of copulas are totally ordered. We call the one-parameter family $\{C_\theta\}$ positively ordered if $C_{\theta_1} \prec C_{\theta_2}$ whenever $\theta_1 \leq \theta_2$ and negatively ordered if $C_{\theta_1} \succ C_{\theta_2}$ whenever $\theta_1 \leq \theta_2$.

2.4 Copulas and Random Variables

Let $X_1, X_2, \dots, X_n$ be continuous random variables with distribution functions $F_1, F_2, \dots, F_n$, respectively, and joint distribution function H. Then we say that $X_1, X_2, \dots, X_n$ have copula C, where C is given by (2.2.1).

Theorem 2.5. *Let $X_1, X_2, \dots, X_n$ be continuous random variables with copula C. Then $X_1, X_2, \dots, X_n$ are independent if and only if $C = \Pi^n$.*

One 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 Ran$X$, then the distribution function of the random variable $\alpha(X)$ is also continuous.

Theorem 2.6. *Let $X_1, X_2, \dots, X_n$ be continuous random variables with copula C. If $\alpha_1, \alpha_2, \dots, \alpha_n$ are strictly increasing on RanX$_1$, RanX$_2, \dots,$ RanX$_n$ respectively, then $\alpha_1(X_1), \alpha_2(X_2), \dots, \alpha_n(X_n)$ have copula C. Thus C is invariant under strictly increasing transformations of $X_1, X_2, \dots, X_n$.*

Proof. Let $F_1, F_2, \dots, F_n$ denote the distribution functions of $X_1, X_2, \dots, X_n$ respectively, and let $G_1, G_2, \dots, G_n$ denote the distribution functions of $\alpha_1(X_1)$, $\alpha_2(X_2)$, $\dots, \alpha_n(X_n)$ respectively. Let $X_1, X_2, \dots, X_n$ have copula C, and let $\alpha_1(X_1)$, $\alpha_2(X_2)$, $\dots, \alpha_n(X_n)$ 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 $\overline{\mathbb{R}}$.

$$\begin{aligned} C_\alpha(G_1(x_1), \dots, G_n(x_n)) &= \mathbb{P}[\alpha_1(X_1) \leq x_1, \dots, \alpha_n(X_n) \leq x_n] \\ &= \mathbb{P}[X_1 \leq \alpha_1^{-1}(x_1), \dots, X_n \leq \alpha_n^{-1}(x_n)] \\ &= C(F_1(\alpha_1^{-1}(x_1)), \dots, F_n(\alpha_n^{-1}(x_n))) \\ &= C(G_1(x_1), \dots, G_n(x_n)). \end{aligned}$$

Since $X_1, X_2, \dots, X_n$ are continuous, $\text{Ran}G_1 = \text{Ran}G_2 = \dots = \text{Ran}G_n = \mathbf{I}$. Hence it follows that $C_\alpha = C$ in $\mathbf{I}^n$. □

From Sklar's theorem we know that the copula function, C, separates an n-dimensional distribution function from its univariate margins. The next theorem will show that there is also a function, $\hat{C}$, that separates an n-dimensional survival function from its univariate survival margins. Furthermore this function can be shown to be a copula, and this survival copula can quite easily be expressed in terms of C and its k-margins.

Theorem 2.7. *Let $X_1, X_2, \dots, X_n$ be continuous random variables with copula $C_{X_1,X_2,\dots,X_n}$. Let $\alpha_1, \alpha_2, \dots, \alpha_n$ be strictly monotone on RanX_1, RanX_2,..., RanX_n, respectively, and let $\alpha_1(X_1), \alpha_2(X_2), \dots, \alpha_n(X_n)$ have copula $C_{\alpha_1(X_1),\alpha_2(X_2),\dots,\alpha_n(X_n)}$. Furthermore let α_k be strictly decreasing for some k, where $1 \le k \le n$. Without loss of generality let $k = 1$. Then*

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

Proof. Let $X_1, \dots, X_n$ have margins $F_1, \dots, F_n$ and let $\alpha_1(X_1), \dots, \alpha_n(X_n)$ have margins $G_1, \dots, G_n$. Then

$$\begin{aligned} & C_{\alpha_1(X_1),\alpha_2(X_2),\dots,\alpha_n(X_n)}(G_1(x_1), \dots, G_n(x_n)) = \\ = \; & \mathbb{P}[\alpha_1(X_1) \le x_1, \dots, \alpha_n(X_n) \le x_n] \\ = \; & \mathbb{P}[X_1 > \alpha_1^{-1}(x_1), \alpha_2(X_2) \le x_2, \dots, \alpha_n(X_n) \le x_n] \\ = \; & \{\mathbb{P}[A^c \cap B] = \mathbb{P}[B] - \mathbb{P}[A \cap B]\} \\ = \; & \mathbb{P}[\alpha_2(X_2) \le x_2, \dots, \alpha_n(X_n) \le x_n] - \\ & \mathbb{P}[X_1 \le \alpha_1^{-1}(x_1), \alpha_2(X_2) \le x_2, \dots, \alpha_n(X_n) \le x_n] \\ = \; & C_{\alpha_2(X_2),\dots,\alpha_n(X_n)}(G_2(x_2), \dots, G_n(x_n)) - \\ & C_{X_1,\alpha_2(X_2),\dots,\alpha_n(X_n)}(F_1(\alpha_1^{-1}(x_1), G_2(x_2), \dots, G_n(x_n)) \\ = \; & \{G_1(x) = \mathbb{P}[\alpha_1(X_1) \le x] = \mathbb{P}[X_1 > \alpha_1^{-1}(x)] = 1 - F_1(\alpha_1^{-1}(x))\} \\ = \; & C_{\alpha_2(X_2),\dots,\alpha_n(X_n)}(G_2(x_2), \dots, G_n(x_n)) - \\ & C_{X_1,\alpha_2(X_2),\dots,\alpha_n(X_n)}(1 - G_1(x_1), G_2(x_2), \dots, G_n(x_n)), \end{aligned}$$

from which the conclusion follows directly. □

By using the result of the two theorems above recursively it is clear that the copula $C_{\alpha_1(X_1),\alpha_2(X_2),\dots,\alpha_n(X_n)}$ can be expressed in terms of the copula $C_{X_1,X_2,\dots,X_n}$. This is done in the following example.

Example 2.3. 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

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

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

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

Consider the trivariate case.
Let α_1, α_2 and α_3 be strictly decreasing. Then

$$\begin{aligned}
&C_{\alpha_1(X_1),\alpha_2(X_2),\alpha_3(X_3)}(u_1,u_2,u_3)=\\
&\quad=C_{\alpha_2(X_2),\alpha_3(X_3)}(u_2,u_3)-C_{X_1,\alpha_2(X_2),\alpha_3(X_3)}(1-u_1,u_2,u_3)\\
&\quad=\dots\\
&\quad=u_1+u_2+u_3-2+C_{X_1,X_2}(1-u_1,1-u_2)+C_{X_1,X_3}(1-u_1,1-u_3)\\
&\qquad+C_{X_2,X_3}(1-u_2,1-u_3)-C_{X_1,X_2,X_3}(1-u_1,1-u_2,1-u_3).
\end{aligned}$$

Here $C_{\alpha_1(X_1),\alpha_2(X_2),\alpha_3(X_3)}$ is the survival copula of X_1,X_2,X_3.

Note also that the joint survival function of n uniform $(0,1)$ random variables whose joint distribution function is the copula C is

$$\overline{C}(u_1,u_2,\dots,u_n)=\hat{C}(1-u_1,1-u_2,\dots,1-u_n).$$

Although most results could be presented without needing measure theory, there are results in the remaining sections that are best presented and understood using some terminology and results from measure theory. Each joint distribution H induces a probability measure on $\mathbb{R}^n$ via $V_H((-\infty,x_1]\times(-\infty,x_2]\times\dots\times(-\infty,x_n])=H(x_1,x_2,\dots,x_n)$. Since copulas are joint distribution functions, each copula induces a probability measure on $\mathbf{I}^n$ via $V_C([0,u_1]\times[0,u_2]\dots\times[0,u_n])=C(u_1,u_2,\dots,u_n)$. Hence, at an intuitive level, the C-measure of a subset of $\mathbf{I}^n$ is the probability that n uniform $(0,1)$ random variables with joint distribution C assume values in that subset.
For any copula C, let

$$C(u_1,u_2,\dots,u_n)=A_C(u_1,u_2,\dots,u_n)+S_C(u_1,u_2,\dots,u_n),$$

where

$$\begin{aligned}
A_C(u_1,\dots,u_n)&=\int_0^{u_1}\dots\int_0^{u_n}\frac{\partial^n}{\partial u_1\dots\partial u_n}C(u_1,\dots,u_n)\,\mathrm{d}s_1\dots\mathrm{d}s_n,\\
S_C(u_1,\dots,u_n)&=C(u_1,\dots,u_n)-A_C(u_1,\dots,u_n).
\end{aligned}$$

Unlike multivariate distributions in general, the margins of a copula are continuous, hence a copula has no individual points in $\mathbf{I}^n$ whose C-measure is positive.
If $C=A_C$ on $\mathbf{I}^n$ then C is said to be absolutely continuous. In this case C has density $\frac{\partial^n}{\partial u_1\dots\partial u_n}C(u_1,\dots,u_n)$.
If $C=S_C$ on $\mathbf{I}^n$ then C is said to be singular, and $\frac{\partial^n}{\partial u_1\dots\partial u_n}C(u_1,\dots,u_n)=0$ almost everywhere in $\mathbf{I}^n$.

The support of a copula is the complement of the union of all open subsets of $\mathbf{I}^n$ with C-measure 0. When the support of a copula C is $\mathbf{I}^n$, it is said to have "full support". When C is singular its support has Lebesgue measure zero and conversely. However a copula can have full support without being absolutely continuous.

Example 2.4. Consider the bivariate Fréchet-Hoeffding upper bound M. Since $\frac{\partial^2}{\partial u\partial v}M(u,v)=0$ everywhere on $\mathbf{I}^2$ except on the main diagonal (which has Lebesgue measure zero), and the M-measure of every rectangle in $\mathbf{I}^2$ entirely above or below the main diagonal is zero, M is singular. The support of M is the main diagonal of $\mathbf{I}^2$ as shown in figure 2.1.
Similarly the support of the bivariate Fréchet-Hoeffding lower bound W is the secondary diagonal of $\mathbf{I}^2$ as shown in figure 2.1.

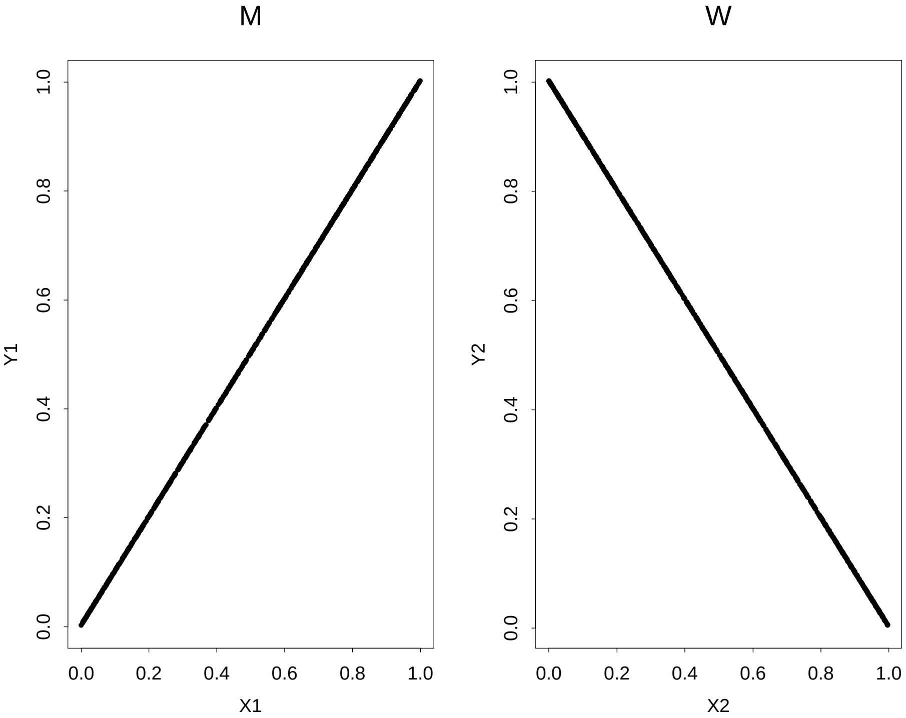

Figure 2.1: The support of the Fréchet-Hoeffding upper and lower bounds.

One of the main aims of this paper is to present effective algorithms for random variate generation from the various copula families studied. The properties of the specific copula family is often essential for the efficiency of the corresponding algorithm. We now present a general algorithm for random variate generation from copulas. However it is in most cases not a suitable one to use.

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

$$C_k(u_1, \dots, u_n) = C(u_1, \dots, u_k, 1, \dots, 1), \quad k = 2, \dots, n-1$$

denote k-margins of $C(u_1, \dots, u_n)$. Furthermore $C_1(u_1) = u_1$ and $C_n(u_1, \dots, u_n) = C(u_1, \dots, u_n)$. Let $(U_1, \dots, U_n)$ have joint distribution function C. Then the conditional distribution of U_k given the values of the first $k-1$ components of $(U_1, \dots, U_n)$ is given by

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

Note that since the increasing functions of one variable given by $u_j \mapsto C(\mathbf{u})$ for $j = 1, \dots, n$ are differentiable almost everywhere, the first order partial derivatives exists for almost all $(u_1, \dots, u_n)$ in $\mathbf{I}^n$. The following algorithm generates a random variate $(u_1, \dots, u_n)$ from $C(u_1, \dots, u_n)$:

Algorithm 1.

- Simulate a value u_1 from $U(0,1)$,
- Simulate a value u_2 from $C_2(u_2|u_1)$,

$\vdots$

- Simulate a value u_n from $C_n(u_n|u_1, \dots, u_{n-1})$.

The correctness of the algorithm follows from the fact that for independent random $U(0,1)$ variables $Q_1, \dots, Q_n$,

$$(Q_1, C_2^{-1}(Q_2|Q_1), \dots, C_n^{-1}(Q_n|Q_1, C_2^{-1}(Q_2|Q_1), \dots))$$

has joint distribution function C.

To simulate a value from $C_k(u_k|u_1, \dots, u_{k-1})$ in general means simulating q from $U(0,1)$ from which $u_k = C_k^{-1}(q|u_1, \dots, u_{k-1})$ can be obtained from the equation $q = C_k(u_k|u_1, \dots, u_{k-1})$ by numeric rootfinding. However we will present situations where $C_k^{-1}(\cdot|u_1, \dots, u_{k-1})$ exists in closed form and hence there is no need for numeric rootfinding. In these situation this algorithm can be recommended. We will refer to this method as the conditional method for random variate generation.

Example 2.5. Consider the bivariate copula family given by

$$C(u,v) = (u^{-\theta} + v^{-\theta} - 1)^{-1/\theta} \tag{2.4.1}$$

for $\theta > 0$.

$$\begin{aligned} C_2(v|u) &= \frac{\partial C}{\partial u}(u,v) = -\frac{1}{\theta}(u^{-\theta} + v^{-\theta} - 1)^{-1/\theta - 1}(-\theta u^{-\theta-1}) \\ &= (u^{\theta})^{\frac{-1-\theta}{\theta}}(u^{-\theta} + v^{-\theta} - 1)^{-1/\theta-1} = (1 + u^{\theta}(v^{-\theta} - 1))^{\frac{-1-\theta}{\theta}}. \end{aligned}$$

Solving the equation $q = C_2(v|u)$ for v yields

$$C_2^{-1}(q|u) = v = \left((q^{\frac{-\theta}{1+\theta}} - 1)u^{-\theta} + 1\right)^{-1/\theta}.$$

Thus the following algorithm generates a random variate from the copula given by (2.4.1).

- Generate a value u from $U(0,1)$,
- Generate a value q from $U(0,1)$,
  Set $v = ((q^{\frac{-\theta}{1+\theta}} - 1)u^{-\theta} + 1)^{-1/\theta}$.

(u,v) is the desired random variate.

3 Dependence

Since the dependence structure among random variables is represented by copulas it provides a natural way to study and measure dependence and association between random variables. Many of these properties and measures are invariant under strictly increasing transforms (as a direct consequence of Theorem 2.6). Linear correlation (or Pearson's correlation) is often used in practice as a measure of dependence. However since linear correlation is not a copula based dependence measure, it is often quite misleading and should not be taken as the canonical dependence measure.
We begin by presenting linear correlation, and then we continue with some copula based measures of dependence.

3.1 Linear Correlation

Definition 8. Let X and Y be two real valued random variables with finite variances. The linear correlation coefficient between X and Y is

$$\rho_l(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}},$$

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

Linear correlation is a measure of linear dependence. In the case of perfect linear dependence, i.e., $Y = aX + b$ almost surely for $a \in \mathbb{R}\backslash\{0\}, b \in \mathbb{R}$, we have $\rho_l(X,Y) = \pm 1$. Otherwise, $-1 < \rho_l(X,Y) < 1$. Furthermore linear correlation has the property that

$$\rho_l(\alpha X + \beta, \gamma Y + \delta) = \mathrm{sgn}(\alpha\gamma)\rho_l(X,Y),$$

for $\alpha, \gamma \in \mathbb{R}\backslash\{0\}, \beta, \delta \in \mathbb{R}$. Hence linear correlation is invariant under strictly increasing *linear* transformations.

Linear correlation is easily manipulated under linear operations. Let $A, B \in \mathbb{R}^{m\times n}; a, b \in \mathbb{R}^m$ and let $\mathbf{X}, \mathbf{Y}$ be random n-vectors. Then

$$\mathrm{Cov}(A\mathbf{X} + a, B\mathbf{Y} + b) = A\mathrm{Cov}(\mathbf{X}, \mathbf{Y})B^T.$$

From this it follows that for $\alpha \in \mathbb{R}^n$,

$$\mathrm{Var}(\alpha^T \mathbf{X}) = \alpha^T \mathrm{Cov}(\mathbf{X}, \mathbf{X})\alpha,$$

and hence the variance of a linear combination is fully determined by pairwise covariances between the components.

Linear correlation is a popular but also often misunderstood measure of dependence. Two reasons for the popularity of linear correlation is that it is often easy to calculate and it is a natural measure of dependence in elliptical distributions (with often used members such as the multivariate normal and the multivariate t-distribution). However often things are not elliptically distributed and using linear correlation as a measure of dependence in such situations might prove very misleading.

Even in the world of elliptical distributions there are situations where using linear correlation does not make sense. We might choose to model some scenario using heavy-tailed distributions such as t-distributions for low degrees of freedom. In such cases the linear correlation may not even be defined because of infinite second order moments.

3.2 Perfect Dependence

For every n-copula C we know from the Fréchet-Hoeffding inequality that

$$W^n(u_1, u_2, \ldots, u_n) \leq C(u_1, u_2, \ldots, u_n) \leq M^n(u_1, u_2, \ldots, u_n).$$

Furthermore, for $n = 2$ the upper and lower bounds are themselves copulas and we have seen that W and M are the bivariate distributions functions of the random vectors $(U, 1-U)$ and (U, U) respectively where U is uniform $(0, 1)$. In this case we say that W describes perfect negative dependence and M describes perfect positive dependence.

Theorem 3.1. *Let (X, Y) have one of the copulas W or M. Then there exists two monotonic functions $\alpha, \beta : \mathbb{R} \longrightarrow \mathbb{R}$ and a real-valued random variable Z so that*

$$(X, Y) =_d (\alpha(Z), \beta(Z)),$$

with α increasing and β decreasing in the former case and both α and β increasing in the latter case. The converse of this result is also true.

For the proof, see Embrechts, McNeil and Straumann (1999) [3].

Definition 9. If (X, Y) has the copula M then X and Y are said to be comonotonic; if it has the copula W they are said to be countermonotonic.

Note that if any of F_1 and F_2 (the distribution functions of X and Y respectively) have discontinuities so that the copula is not unique, then W and M are possible copulas.

In the case of F_1 and F_2 being continuous a stronger version of the result can be stated:

$$\begin{aligned} C = W \quad &\Leftrightarrow \quad Y = T(X) \text{ a.s.}, \ T = F_2^{-1} \circ (1 - F_1) \text{ decreasing}, \\ C = M \quad &\Leftrightarrow \quad Y = T(X) \text{ a.s.}, \ T = F_2^{-1} \circ F_1 \text{ increasing}. \end{aligned}$$

3.3 Concordance

Let (x_i, y_i) and (x_j, y_j) be two observations from a random vector (X, Y) of continuous random variables. We say that (x_i, y_i) and (x_j, y_j) are concordant if $x_i < x_j$ and $y_i < y_j$, or if $x_i > x_j$ and $y_i > y_j$. Similarly we say that (x_i, y_i) and (x_j, y_j) are discordant if $x_i < x_j$ and $y_i > y_j$, or if $x_i > x_j$ and $y_i < y_j$. This can be formulated shorter: (x_i, y_i) and (x_j, y_j) are concordant if $(x_i - x_j)(y_i - y_j) > 0$, and discordant if $(x_i - x_j)(y_i - y_j) < 0$.

3.3.1 Kendall's tau

The sample version of the measure of association known as Kendall's tau is defined in terms of concordance as follows: Let $\{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$ denote a sample of n observations from a random vector (X, Y) of continuous random variables. Each of the $\binom{n}{2}$ distinct pairs (x_i, y_i) and (x_j, y_j) of observations in the sample is either concordant or discordant. Let c and d denote the number of concordant and discordant pairs respectively. Then Kendall's tau for the sample is

$$\frac{c-d}{c+d} = (c-d) / \binom{n}{2}.$$

The population version of Kendall's tau (simply denoted Kendall's tau) for continuous random variables is defined similarly. Let (X_1, Y_1) and (X_2, Y_2) be independent and identically distributed random vectors, each with joint distribution function H. Then

$$\tau = \tau_{X,Y} = \mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) > 0] - \mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) < 0] \tag{3.3.1}$$

Theorem 3.2. *Let (X_1, Y_1) and (X_2, Y_2) be independent vectors of continuous random variables with joint distribution functions H_1 and H_2, respectively, with common margins F (of X_1 and X_2) and G (of Y_1 and Y_2). Let C_1 and C_2 denote the copulas of (X_1, Y_1) and (X_2, Y_2), respectively, so that $H_1(x, y) = C_1(F(x), G(y))$ and $H_2(x, y) = C_2(F(x), G(y))$. Let Q denote the difference between the probability of concordance and discordance of (X_1, Y_1) and (X_2, Y_2), i.e., let*

$$Q = \mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) > 0] - \mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) < 0]. \tag{3.3.2}$$

Then

$$Q = Q(C_1, C_2) = 4 \iint_{\mathbf{I}^2} C_2(u, v)\, \mathrm{d}C_1(u, v) - 1. \tag{3.3.3}$$

Proof. Since the random variables are continuous, $\mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) < 0] = 1 - \mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) > 0]$ and hence

$$Q = 2\mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) > 0] - 1. \tag{3.3.4}$$

But $\mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) > 0] = \mathbb{P}[X_1 > X_2, Y_1 > Y_2] + \mathbb{P}[X_1 < X_2, Y_1 < Y_2]$, and these probabilities can be evaluated by integrating over the distribution of one of the vectors (X_1, Y_1) or (X_2, Y_2), say (X_1, Y_1). First we have

$$\begin{aligned}
\mathbb{P}[X_1 > X_2, Y_1 > Y_2] &= \mathbb{P}[X_2 < X_1, Y_2 < Y_1] \\
&= \iint_{\mathbb{R}^2} P[X_2 < x, Y_2 < y]\, \mathrm{d}C_1(F(x), G(y)) \\
&= \iint_{\mathbb{R}^2} C_2(F(x), G(y))\, \mathrm{d}C_1(F(x), G(y)),
\end{aligned}$$

so that employing the probability transforms $u = F(x)$ and $v = G(y)$ yields

$$\mathbb{P}[X_1 > X_2, Y_1 > Y_2] = \iint_{\mathbf{I}^2} C_2(u, v)\, \mathrm{d}C_1(u, v).$$

Similarly,

$$\begin{aligned}
&\mathbb{P}[X_1 < X_2, Y_1 < Y_2] = \\
= &\iint_{\mathbb{R}^2} \mathbb{P}[X_2 > x, Y_2 > y]\, \mathrm{d}C_1(F(x), G(y)) \\
= &\iint_{\mathbb{R}^2} [1 - F(x) - G(y) + C_2(F(x), G(y))]\, \mathrm{d}C_1(F(x), G(y)) \\
= &\iint_{\mathbf{I}^2} [1 - u - v + C_2(u, v)]\, \mathrm{d}C_1(u, v).
\end{aligned}$$

But since C_1 is the joint distribution function of a pair (U, V) of uniform $(0, 1)$ random variables, $\mathbb{E}(U) = \mathbb{E}(V) = 1/2$, and hence

$$\begin{aligned}
\mathbb{P}[X_1 < X_2, Y_1 < Y_2] &= 1 - \frac{1}{2} - \frac{1}{2} + \iint_{\mathbf{I}^2} C_2(u, v)\, \mathrm{d}C_1(u, v), \\
&= \iint_{\mathbf{I}^2} C_2(u, v)\, \mathrm{d}C_1(u, v).
\end{aligned}$$

Thus

$$\mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) > 0] = 2 \iint_{\mathbf{I}^2} C_2(u,v)\,\mathrm{d}C_1(u,v),$$

and the conclusion follows. □

Corollary 3.1. *Let C_1, C_2, and Q be as given in Theorem 3.2. Then*

1. *Q is symmetric in its arguments: $Q(C_1, C_2) = Q(C_2, C_1)$.*
2. *Q is nondecreasing in each argument: if $C_1 \prec C_1'$ and $C_2 \prec C_2'$ for all (u,v) in $\mathbf{I}^2$, then $Q(C_1, C_2) \le Q(C_1', C_2')$.*
3. *Copulas can be replaced by survival copulas in Q, i.e., $Q(C_1, C_2) = Q(\hat{C}_1, \hat{C}_2)$.*

Example 3.1. Because the support of M and W is just the first and second main diagonal of $\mathbf{I}^2$, the function Q can easily be evaluated for pairs of the basic copulas W, M and Π.

If g is integrable function whose domain is $\mathbf{I}^2$, then

$$\begin{aligned} \textstyle\iint_{\mathbf{I}^2} g(u,v)\,\mathrm{d}M(u,v) &= \textstyle\int_0^1 g(u,u)\,\mathrm{d}u, \\ \textstyle\iint_{\mathbf{I}^2} g(u,v)\,\mathrm{d}W(u,v) &= \textstyle\int_0^1 g(u,1-u)\,\mathrm{d}u, \\ \textstyle\iint_{\mathbf{I}^2} g(u,v)\,\mathrm{d}\Pi(u,v) &= \textstyle\iint_{\mathbf{I}^2} g(u,v)\,\mathrm{d}u\,\mathrm{d}v. \end{aligned}$$

From this it follows that

$$\begin{aligned} &Q(M, M) = 1, \\ &Q(M, \Pi) = 1/3, \\ &Q(M, W) = 0, \\ &Q(W, \Pi) = -1/3, \\ &Q(W, W) = -1, \\ &Q(\Pi, \Pi) = 0. \end{aligned}$$

The last equation follows from the fact $\mathrm{d}\Pi(u,v) = \mathrm{d}u\,\mathrm{d}v$. From Corollary 3.1 and the values of Q above it follows that for an arbitrary copula C

$$Q(C, M) \in [0,1], \quad Q(C, W) \in [-1, 0], \quad Q(C, \Pi) \in [-1/3, 1/3].$$

Theorem 3.3. *Let X and Y be continuous random variables whose copula is C. Then Kendall's tau for X and Y (denoted $\tau_{X,Y}$ or τ_C) is given by*

$$\tau_{X,Y} = \tau_C = Q(C, C) = 4 \iint_{\mathbf{I}^2} C(u,v)\,\mathrm{d}C(u,v) - 1. \tag{3.3.5}$$

Note that the integral in (3.3.5) is the expected value of the function $C(U,V)$ of uniform $(0,1)$ random variables with joint distribution function C, i.e.,

$$\tau_C = 4\mathbb{E}(C(U,V)) - 1 \tag{3.3.6}$$

3.3.2 Spearman's rho

Let (X_1, Y_1), (X_2, Y_2) and (X_3, Y_3) be three independent random vectors with common joint distribution function H, whose margins are F and G, and copula C. Then Spearman's rho is defined as

$$\rho = \rho_{X,Y} = 3(\mathbb{P}[(X_1 - X_2)(Y_1 - Y_3) > 0] - \mathbb{P}[(X_1 - X_2)(Y_1 - Y_3) < 0]) \tag{3.3.7}$$

Note that the joint distribution of (X_1, Y_1) is $H(x,y)$ and the joint distribution of (X_2, Y_3) is $F(x)G(y)$. Thus the copula of X_2 and Y_3 is Π. Using Theorem 3.2 and first part of Corollary 3.1 we get

Theorem 3.4. *Let X and Y be continuous random variables whose copula is C. Then Spearman's rho for X and Y (denoted $\rho_{X,Y}$ or ρ_C) is given by*

$$\begin{aligned}\rho_{X,Y} = \rho_C &= 3Q(C,\Pi), && (3.3.8)\\ &= 12\iint_{I^2} uv\,\mathrm{d}C(u,v) - 3, && (3.3.9)\\ &= 12\iint_{I^2} C(u,v)\,\mathrm{d}u\,\mathrm{d}v - 3. && (3.3.10)\end{aligned}$$

Note that the integral in (3.3.9) is just the expected value of the product of two uniform $(0,1)$ random variables U and V whose joint distribution is the copula C. Thus

$$\begin{aligned}\rho_{X,Y} = \rho_C &= 12\iint_{\mathbf{I}^2} uv\,\mathrm{d}C(u,v) - 3 = 12\mathbb{E}(U,V) - 3\\ &= \frac{\mathbb{E}(UV) - 1/4}{1/12} = \frac{\mathbb{E}(UV) - \mathbb{E}(U)\mathbb{E}(V)}{\sqrt{Var(U)}\sqrt{Var(V)}}.\end{aligned}$$

If x and y are observations from random variables X and Y with distribution functions F and G, respectively, then the "grades" (called ranks for the sample analogue) of x and y are given by $u = F(x)$ and $v = G(y)$. Note the the grades are observations from the uniform $(0,1)$ random variables $U = F(X)$ and $V = G(Y)$ whose joint distribution function is the copula C. Hence we have (let ρ_l denote the linear correlation coefficient)

$$\rho_{X,Y} = \rho_l(F(X), G(Y))$$

This also provides an easy way to approximatively calculate $\rho_{X,Y}$ given a sample from (X,Y) and the margins.

Definition 10. A numeric measure κ of association between two continuous random variables X and Y whose copula is C is a measure of concordance if it satisfies the following properties (we use $\kappa_{X,Y}$ or κ_C when convenient):

1. κ is defined for every pair X, Y of continuous random variables;
2. $-1 \leq \kappa_{X,Y} \leq 1$, $\kappa_{X,X} = 1$ and $\kappa_{X,-X} = -1$;
3. $\kappa_{X,Y} = \kappa_{Y,X}$;
4. if X and Y are independent, then $\kappa_{X,Y} = \kappa_\Pi = 0$;
5. $\kappa_{-X,Y} = \kappa_{X,-Y} = -\kappa_{X,Y}$;
6. if C_1 and C_2 are copulas such that $C_1 \prec C_2$, then $\kappa_{C_1} \leq \kappa_{C_2}$;
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 $\lim_{n\to\infty} \kappa_{C_n} = \kappa_C$.

Theorem 3.5. *Let κ be a measure of concordance for continuous random variables X and Y.*

1. *if Y is almost surely an increasing function of X, then $\kappa_{X,Y} = \kappa_M = 1$;*
2. *if Y is almost surely a decreasing function of X, then $\kappa_{X,Y} = \kappa_W = -1$;*
3. *if α and β are almost surely strictly increasing functions on RanX and RanY, respectively, then $\kappa_{\alpha(X),\beta(Y)} = \kappa_{X,Y}$.*

In the next theorem we will see that Kendall's tau and Spearman's rho are concordance measures according to the above definition.

Theorem 3.6. *If X and Y are continuous random variables whose copula is C, then Kendall's tau and Spearman's rho satisfy the properties in Definition 10 for a measure of concordance.*

Proof. For both Kendall's tau and Spearman's rho the first six properties in Definition 10 follows from properties of Q in Theorem 3.2, Corollary 3.1. and Example 2.1. For the seventh property the Lipschitz condition (2.1.4) implies that any family of copulas is equicontinuous, from which it follows that the convergence of $\{C_n\}$ to C is uniform. □

Example 3.2. It follows from the third part of Theorem 3.5 that Kendall's tau and Spearman's rho of the random variables X and Y is invariant under strictly increasing transforms of X and Y. Does this property also hold for linear correlation? If not then linear correlation can not be a measure of concordance.
Let X and Y be standard exponential random variables with copula C, where C is a member of the Farlie-Gumbel-Morgenstern family, i.e., C is given by

$$C(u,v) = C_\theta(u,v) = uv + \theta uv(1-u)(1-v),$$

For some θ in $[-1,1]$, and the joint distribution function H of X and Y is given by

$$H(x,y) = C(1-e^{-x}, 1-e^{-y}).$$

Let ρ_l denote the linear correlation coefficient. Then

$$\rho_l(X,Y) = \frac{\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}} = \mathbb{E}(XY) - 1,$$

where

$$\begin{aligned}
\mathbb{E}(XY) &= \int_0^\infty\int_0^\infty xy \,\mathrm{d}H(x,y) \\
&= \int_0^\infty\int_0^\infty xy\big(e^{-x-y} + \theta e^{-2x-2y} - \theta e^{-2x-y}(1-e^{-y}) \\
&\quad -\theta e^{-x-2y}(1-e^{-x}) + \theta(1-e^{-x})(1-e^{-y})e^{-x-y}\big)\,\mathrm{d}x\,\mathrm{d}y \\
&= 1 + \frac{\theta}{4}.
\end{aligned}$$

Hence $\rho_l(X,Y) = \theta/4$. But

$$\begin{aligned}
\rho_l(1-e^{-X}, 1-e^{-Y}) &= \rho_C \\
&= 12\iint_{\mathbf{I}^2} C(u,v)\,\mathrm{d}u\,\mathrm{d}v - 3 \\
&= 12\iint_{\mathbf{I}^2} (uv + \theta uv(1-u)(1-v))\,\mathrm{d}u\,\mathrm{d}v - 3 \\
&= 12(\frac{1}{4} + \frac{\theta}{36}) - 3 \\
&= \theta/3.
\end{aligned}$$

Hence $\rho_l(X,Y)$ is not invariant under strictly increasing transforms of X and Y from which follows that linear correlation is not a measure of concordance.

Although the properties of a measure of concordance are nice properties there are some additional properties that would make a measure of association satisfying those even more useful. From Theorem 3.5 we know that for a pair of random variables (X,Y) with copula C

$$\begin{aligned}
C &= M \implies \tau_C = \rho_C = 1, \\
C &= W \implies \tau_C = \rho_C = -1,
\end{aligned}$$

but does the converse hold? The following theorem states that Kendall's tau and Spearman's rho satisfy some additional properties.

Theorem 3.7. *Let X and Y be continuous random variables with copula C, and let δ denote Kendall's tau or Spearman's rho. Then the following are true:*

1. $\delta(X,Y) = \delta(Y,X)$
2. *If X and Y are independent, then* $\delta(X,Y) = 0$.
3. $-1 \leq \delta(X,Y) \leq 1$
4. *If T is strictly monotone on RanX, then* $$\delta(T(X),Y) = \begin{cases} \delta(X,Y), & T \textit{ increasing,} \\ -\delta(X,Y), & T \textit{ decreasing.} \end{cases}$$
5. $\delta(X,Y) = 1 \iff C = M$
6. $\delta(X,Y) = -1 \iff C = W$

Here

1., 2., and 3. follows directly from the fact that δ is a measure of concordance,

4. follows from the fact that δ can be expressed in term of the copula of $T(X)$ and Y which is given by $C_{T(X),Y}(u,v) = C_{X,Y}(u,v)$ when T is strictly increasing and $C_{T(X),Y}(u,v) = v - C_{X,Y}(1-u,v)$ when T is strictly decreasing.

5. and 6. See Embrechts, McNeil and Straumann (1999) [3].

From the definitions of Kendall's tau and Spearman's rho it follows that both are increasing functions of the value of the copula under consideration. Thus Kendall's tau and Spearman's rho are increasing with respect to the concordance ordering given in Definition 7.

For measures of association such as Kendall's tau or Spearman's rho all correlation values in the interval $[-1,1]$ can be obtained by a suitable choice of the copula, since these measures are expressed in terms of copulas. This is however not the case with linear correlation. This is shown in the following illustrative example by Embrechts, McNeil and Straumann (1999) [3].

Example 3.3. Let $X \sim \text{Lognormal}(0,1)$ and $Y \sim \text{Lognormal}(0,\sigma^2), \sigma > 0$. Then $\rho_{\min} = \rho(e^Z, e^{-\sigma Z})$ and $\rho_{\max} = \rho(e^Z, e^{\sigma Z})$, where $Z \sim \mathcal{N}(0,1)$. From this $\rho_{\min}$ and $\rho_{\max}$ can be calculated analytically, yielding:

$$\rho_{min} = \frac{e^{-\sigma} - 1}{\sqrt{e-1}\sqrt{e^{\sigma^2}-1}}, \quad \rho_{max} = \frac{e^{\sigma} - 1}{\sqrt{e-1}\sqrt{e^{\sigma^2}-1}},$$

from which follows that $\lim_{\sigma\to\infty} \rho_{\min} = \lim_{\sigma\to\infty} \rho_{\max} = 0$. Hence the linear correlation can be almost zero, even though X and Y are comonotonic or countermonotonic.

Kendall's tau and Spearman's rho are measures of association between two random variables. However the extension to higher dimensions is obvious, we simply write pairwise correlations in an $n \times n$-matrix in the same way as is done for linear correlation.

3.4 Tail Dependence

The concept of tail dependence relates to the amount of dependence in the upper-quadrant tail or lower-quadrant tail of a bivariate distribution. It is a concept that is relevant to dependence in extreme values. Furthermore, tail dependence between two 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 11. Let X and Y be continuous random variables with distribution functions F_1 and F_2. The coefficient of upper tail dependence of X and Y is

$$\lim_{u \to 1-} \mathbb{P}[Y > F_2^{-1}(u) | X > F_1^{-1}(u)] = \lambda_U$$

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.

An alternative and equivalent definition from which it is seen that the concept of tail dependence is indeed a copula property is the following.

Definition 12. If a bivariate copula C is such that

$$\lim_{u \to 1-} \overline{C}(u,u)/(1-u) = \lambda_U$$

exists, then C has upper tail dependence if $\lambda_U \in (0,1]$, and no upper tail dependence if $\lambda_U = 0$. $\overline{C}(u,u) = 1 - 2u + C(u,u)$.

Example 3.4. Consider the bivariate Gumbel family of copulas

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

for $\theta \geq 1$. Then

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

and here l'Hospitals rule can be applied yielding

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

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

For copulas without a simple closed form, such as the Gaussian family of copulas with bivariate copulas given by

$$C_\rho(u,v) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\{-\frac{s^2 - 2\rho st + t^2}{2(1-\rho^2)}\} \,\mathrm{d}s\,\mathrm{d}t,$$

where $-1 < \rho < 1$ and Φ is the univariate standard normal distribution function, an alternative formula for λ_U is more useful. Consider a pair of uniform $(0,1)$ random variables (U,V) with copula C. First note that

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

and

$$\mathbb{P}[V > v | U = u] = 1 - \frac{\partial}{\partial u} C(u,v),$$

and similarly when conditioning on V. Then

$$\begin{aligned}\lambda_U &= \lim_{u\to 1-} \overline{C}(u,u)/(1-u)\\ &= -\lim_{u\to 1-} \frac{\mathrm{d}\overline{C}(u,u)}{\mathrm{d}u}\\ &= -\lim_{u\to 1-}(-2+\frac{\partial}{\partial s}C(s,t)\big|_{s=t=u}+\frac{\partial}{\partial t}C(s,t)\big|_{s=t=u})\\ &= \lim_{u\to 1-}(\mathbb{P}[V>u|U=u]+\mathbb{P}[U>u|V=u]).\end{aligned}$$

Furthermore if C is a exchangeable copula, i.e. $C(u,v)=C(v,u)$, then the expression for λ_U simplifies into

$$\lambda_U = 2\lim_{u\to 1-}\mathbb{P}[V>u|U=u].$$

Example 3.5. Let X and Y have the bivariate standard normal distribution function with correlation parameter ρ. That is $(X,Y)\sim C(\Phi(x),\Phi(y))$, where C is a member of the Gaussian family given above. Since copulas in this family are exchangeable

$$\lambda_U = 2\lim_{u\to 1-}\mathbb{P}[V>u|U=u],$$

and because Φ is a distribution function with infinite right endpoint

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

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

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

from which it follows that $\lambda_U=0$ for $\rho<1$. Hence the Gaussian copula C does not have upper tail dependence.

Quite naturally, the concept of lower tail dependence can be defined in a similar way. If

$$\lim_{u\to 0+} C(u,u)/u = \lambda_L \qquad (3.4.1)$$

exists, C has lower tail dependence if $\lambda_L\in(0,1]$, and no lower tail dependence if $\lambda_L=0$.

For copulas without a simple closed form an alternative formula for λ_L is more useful. Consider a pair of random variables (U,V) with copula C. Then

$$\begin{aligned}\lambda_L &= \lim_{u\to 0+} C(u,u)/u\\ &= \lim_{u\to 0+}\frac{\mathrm{d}C(u,u)}{\mathrm{d}u}\\ &= \lim_{u\to 0+}(\frac{\partial}{\partial s}C(s,t)\big|_{s=t=u}+\frac{\partial}{\partial t}C(s,t)\big|_{s=t=u})\\ &= \lim_{u\to 0+}(\mathbb{P}[V<u|U=u]+\mathbb{P}[U<u|V=u]).\end{aligned}$$

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

$$\lambda_L = 2 \lim_{u \to 0+} \mathbb{P}[V < u | U = u].$$

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

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

and the joint survival function for two uniform $(0,1)$ variables whose joint distribution function is C is given by

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

Hence it follows that

$$\begin{aligned} \lim_{u \to 1-} \overline{C}(u,u)/(1-u) &= \lim_{u \to 1-} \hat{C}(1-u, 1-u)/(1-u) \\ &= \lim_{u \to 0+} \hat{C}(u,u)/u, \end{aligned}$$

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

$$\begin{aligned} \lim_{u \to 0+} C(u,u)/u &= \lim_{u \to 1-} C(1-u, 1-u)/(1-u) \\ &= \lim_{u \to 1-} (1 - 2u + \hat{C}(u,u))/(1-u) \\ &= \lim_{u \to 1-} \overline{\hat{C}}(u,u)/(1-u), \end{aligned}$$

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

4 Techniques for Construction of Multivariate Copulas

There are many ways of constructing copulas. One natural approach is to start with constructing families of 2-copulas with certain nice properties. These properties can be statistical properties or just nice mathematical properties. Once this is done the next natural step is trying to find a multivariate extension. A multivariate family of copulas is an extension of a bivariate family if: (1) all bivariate margins of the multivariate copula are in the given bivariate family; (2) all multivariate margins of order 3 to $n-1$ have the same multivariate forms. The problem of finding multivariate extensions of bivariate copula families is not trivial. It is quite common that bivariate copulas does not have a natural multivariate extension and if they have such the resulting dependence structure is often quite limited.

4.1 The Farlie-Gumbel-Morgenstern Family

A nice mathematical property of a copula family could be that the copula expression is given by low degree polynomials in the arguments u and v. Lets consider copulas of the form

$$C(u,v) = a(v)u^2 + b(v)u + c(v),$$

for some functions a, b, and c. The boundary conditions for copulas give

$$C(0,v) = 0 = c(v) \text{ and } C(1,v) = v = a(v) + b(v).$$

Let $a(v) = -\Psi(v)$, then $b(v) = v - a(v) = v + \Psi(v)$, and

$$C(u,v) = uv + \Psi(v)u(1-u) \tag{4.1.1}$$

where Ψ is chosen so that C is 2-increasing and $\Psi(0) = \Psi(1) = 0$ so that $C(u,0) = 0$ and $C(u,1) = u$.

Theorem 4.1. *Let Ψ be a function with domain $\boldsymbol{I}$ and let C be given by* (4.1.1) *for u, v in $\boldsymbol{I}$. Then C is a copula if and only if:*

1. $\Psi(0) = \Psi(1) = 0$;
2. *$\Psi(v)$ satisfies the Lipschitz condition*

$$\left|\Psi(v_2) - \Psi(v_1)\right| \leq \left|v_2 - v_1\right|$$

for all v_1, v_2 in $\boldsymbol{I}$. Furthermore, C is absolutely continuous.

Proof. As noted above the boundary conditions $C(u,0) = 0$ and $C(u,1) = u$ are equivalent to $\Psi(0) = \Psi(1) = 0$. C is 2-increasing if and only if

$$V_C([u_1,u_2] \times [v_1,v_2]) = (u_2 - u_1)\{v_2 - v_1 + [\Psi(v_2) - \Psi(v_1)](1 - u_1 - u_2)\} \geq 0.$$

If $u_1 = u_2, v_1 = v_2$, or if $u_1 + u_2 = 1$, then $V_C([u_1,u_2] \times [v_1,v_2]) = 0$. So for $u_1 < u_2$ and $v_1 < v_2$, we have

$$\frac{\Psi(v_2) - \Psi(v_1)}{v_2 - v_1} \leq \frac{1}{u_2 + u_1 - 1} \quad \text{if } u_1 + u_2 > 1,$$

and

$$\frac{\Psi(v_2)-\Psi(v_1)}{v_2-v_1} \geq \frac{1}{u_2+u_1-1} \quad \text{if } u_1+u_2<1.$$

But $\inf\{1/(u_1+u_2-1)|0 \leq u_1 \leq u_2 \leq 1, u_1+u_2>1\} = 1$ and $\sup\{1/(u_1+u_2-1)|0 \leq u_1 \leq u_2 \leq 1, u_1+u_2<1\} = -1$, and hence C is 2-increasing if and only if

$$-1 \leq \frac{\Psi(v_2)-\Psi(v_1)}{v_2-v_1} \leq 1$$

for v_1, v_2 in $\mathbf{I}$ such that $v_1 < v_2$, from which the conclusion follows. From the Lipschitz condition it follows that $|\Psi'(v)| \leq 1$ almost everywhere on $\mathbf{I}$. Thus Ψ is absolutely continuous. The absolute continuity of C then follows. □

One choice of Ψ satisfying the above conditions is $\Psi(v) = \theta v(1-v)$ for θ in $[-1,1]$. This choice results in a copula family called the Farlie-Gumbel-Morgenstern (FGM) family.

The FGM family is positively ordered since $\theta_2 uv(1-u)(1-v) \geq \theta_1 uv(1-u)(1-v)$ for $\theta_2 \geq \theta_1$. Thus maximal and minimal rank correlations are obtained for $\theta = 1$ and $\theta = -1$ respectively. Simple calculations yields

$$\begin{aligned} \tau_C &= 4 \textstyle\iint_{\mathbf{I}^2} C(u,v)\,\mathrm{d}C(u,v) - 1 &&= 2\theta/9, \\ \rho_C &= 12 \textstyle\iint_{\mathbf{I}^2} C(u,v)\,\mathrm{d}u\,udv - 3 &&= \theta/3. \end{aligned}$$

Hence $|\tau_C| \leq 2/9$ and $|\rho_C| \leq 1/3$. Note that this limited range of dependence is one of the reasons why the FGM-family is not often used in practice.
The following calculation shows that the FGM-copulas do not have upper tail dependence.

$$\begin{aligned} \lim_{u\to 1-} \frac{\overline{C}(u,u)}{1-u} &= \lim_{u\to 1-} \frac{1-2u+u^2+\theta u^2(1-u)^2}{1-u} \\ &= \lim_{u\to 1-} \frac{(1-u)^2(1+\theta u^2)}{1-u} \\ &= \lim_{u\to 1-} (1-u)(1+\theta u^2) = 0 \end{aligned}$$

The FGM family of copulas provides a natural extension to higher dimensions, as seen from the following expression for the (2^n-n-1)-parameter n-copula C.

$$C(\mathbf{u}) = u_1u_2\ldots u_n\big[1+\sum_2^n \sum_{1\leq j_1<\ldots j_k\leq n} \theta_{j_1j_2\ldots j_k}(1-u_{j_1})(1-u_{j_2})\ldots(1-u_{j_k})\big]. \tag{4.1.2}$$

Each copula in this family is absolutely continuous with density

$$\frac{\partial^n C(\mathbf{u})}{\partial u_1 \partial u_2 \ldots \partial u_n} = 1+\sum_2^n \sum_{1\leq j_1<\ldots j_k\leq n} \theta_{j_1j_2\ldots j_k}(1-2u_{j_1})(1-2u_{j_2})\ldots(1-2u_{j_k}).$$

Since $C(\mathbf{u})$ is quadratic in each variable the density is linear in each variable. Hence the density will be nonnegative on $\mathbf{I}^n$ if and only if it is nonnegative at each of the 2^n vertices of $\mathbf{I}^n$. This gives the following 2^n parameter constraints:

$$1+\sum_2^n \sum_{1\leq j_1<\ldots j_k\leq n} \varepsilon_{j_1}\varepsilon_{j_2}\ldots\varepsilon_{j_k}\theta_{j_1j_2\ldots j_k} \geq 0, \quad \varepsilon_{j_1}, \varepsilon_{j_2}, \ldots, \varepsilon_{j_k} \in \{-1,1\}.$$

As a consequence each parameter, θ, must satisfy $|\theta| < 1$. Note that this is a proper multivariate extension since each k-margin, $3 \leq k < n$, is of the same form and the 2-margins are in the bivariate FGM family.

Consider the $n(n-1)/2$-parameter subfamily of the family of n-copulas described above resulting from setting $\theta_{j_1 j_2 \dots j_k} = 0$ when $k \geq 3$. This gives the following subfamily of the copula family given by (4.1.2):

$$C(\mathbf{u}) = u_1 u_2 \dots u_n \big[1 + \sum_{j<k} \theta_{jk}(1-u_j)(1-u_k)\big].$$

Given an arbitrary rank correlation matrix with obtainable rank correlations ($\left|\tau_{ij}\right| \leq 2/9$ or $\left|\rho_{ij}\right| \leq 1/3$ depending on what measure of association is used) this gives a one-to-one correspondence between the upper (lower) diagonal elements in the rank correlation matrix and the copula parameters.

Random variate generation from this copula family presents no problem. We simply apply the conditional method. With the same notation as used to describe the conditional method we get

$$\begin{aligned}
& \frac{\partial^{k-1} C_{k-1}(u_1, \dots, u_{k-1})}{\partial u_1 \dots \partial u_{k-1}} = \\
= \; & 1 + \theta_{12}(1-2u_1)(1-2u_2) + \dots + \theta_{1(k-1)}(1-2u_1)(1-2u_{k-1}) + \\
& + \theta_{23}(1-2u_2)(1-2u_3) + \dots + \theta_{2(k-1)}(1-2u_2)(1-2u_{k-1}) + \\
& + \dots + \\
& + \theta_{(k-2)(k-1)}(1-2u_{k-2})(1-2u_{k-1}) \\
= \; & c_k
\end{aligned}$$

and

$$\begin{aligned}
& \frac{\partial^{k-1} C_k(u_1, \dots, u_k)}{\partial u_1 \dots \partial u_k} = \\
= \; & c_k u_k + \big(\theta_{1k}(1-2u_1) + \dots + \theta_{(k-1)k}(1-2u_{k-1})\big) u_k - \\
& - \big(\theta_{1k}(1-2u_1) + \dots + \theta_{(k-1)k}(1-2u_{k-1})\big) u_k^2 \\
= \; & a_k u_k^2 + b_k u_k,
\end{aligned}$$

where $u_1, \dots, u_{k-1}$ (and hence also a_k, b_k and c_k) are determined from the previous steps in the algorithm.

Hence to simulate a value u_k from $C_k(u_k | u_1, \dots u_{k-1})$, where $u_1, \dots u_{k-1}$ are given by the previous steps:

- Simulate a value q from $U(0,1)$ independent of $u_1, \dots, u_{k-1}$,
- Solve for u_k

$$\begin{array}{rcl}
C_k(u_k | u_1, \dots u_{k-1}) = q & & \Longleftrightarrow \\
\frac{a_k u_k^2 + b_k u_k}{c_k} = q & & \Longrightarrow \\
u_k = -\frac{b_k}{2a_k} \pm \sqrt{(\frac{b_k}{2a_k})^2 + \frac{q c_k}{a_k}} & &
\end{array}$$

In the last equation the sign is chosen so that $u_k \in [0,1]$.

4.2 The Marshall-Olkin Family

Instead of constructing copulas from certain nice mathematical properties, statistical properties can provide the basis of construction.

Consider a two component system where the components are subjects to shocks, which are fatal to one or both components. Let X_1 and X_2 denote the lifetimes of the two components. Furthermore assume that the shocks form three independent Poisson processes with parameters $\lambda_1, \lambda_2, \lambda_{12} \geq 0$, where the index indicate whether the shocks kill 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 random variables with parameters λ_1, λ_2 and λ_{12}, respectively. Hence

$$\begin{aligned}\overline{H}(x_1, x_2) &= \mathbb{P}[X_1 > x_1, X_2 > x_2] \\ &= \mathbb{P}[Z_1 > x_1]\mathbb{P}[Z_2 > x_2]\mathbb{P}[Z_{12} > \max(x_1, x_2)].\end{aligned}$$

The univariate survival functions for X_1 and X_2 are $\overline{F}_1(x_1) = \exp(-(\lambda_1 + \lambda_{12})x_1)$ and $\overline{F}_2(x_2) = \exp(-(\lambda_2+\lambda_{12})x_2)$. Furthermore $\max(x_1, x_2) = x_1+x_2-\min(x_1, x_2)$ yielding

$$\begin{aligned}\overline{H}(x_1, x_2) &= \exp(-(\lambda_1 + \lambda_{12})x_1 - (\lambda_2 + \lambda_{12})x_2 + \lambda_{12}\min(x_1, x_2)) \\ &= \overline{F}_1(x_1)\overline{F}_2(x_2)\min(\exp(\lambda_{12}x_1), \exp(\lambda_{12}x_2)).\end{aligned}$$

Let $\alpha_1 = \lambda_{12}/(\lambda_1 + \lambda_{12})$ and $\alpha_2 = \lambda_{12}/(\lambda_2 + \lambda_{12})$. Then $\exp(\lambda_{12}x_1) = \overline{F}_1(x_1)^{-\alpha_1}$ and $\exp(\lambda_{12}x_2) = \overline{F}_2(x_2)^{-\alpha_2}$, and hence the survival copula is given by

$$\hat{C}(u_1, u_2) = u_1u_2\min(u_1^{-\alpha_1}, u_2^{-\alpha_2}) = \min(u_1^{1-\alpha_1}u_2, u_1u_2^{1-\alpha_2}).$$

The survival copulas for the Marshall-Olkin bivariate exponential distribution yields a copula family given by

$$C_{\alpha_1,\alpha_2}(u_1, u_2) = \min(u_1^{1-\alpha_1}u_2, u_1u_2^{1-\alpha_2}) = \begin{cases} u_1^{1-\alpha_1}u_2, & u_1^{\alpha_1} \geq u_2^{\alpha_2}, \\ u_1u_2^{1-\alpha_2}, & u_1^{\alpha_1} \leq u_2^{\alpha_2}. \end{cases}$$

This family is known as the Marshall-Olkin family.

The Marshall-Olkin copulas have both an absolutely continuous and a singular component. Since

$$\frac{\partial^2}{\partial u_1 \partial u_2}C_{\alpha_1,\alpha_2}(u_1, u_2) = \begin{cases} u_1^{-\alpha_1}, & u_1^{\alpha_1} > u_2^{\alpha_2}, \\ u_2^{-\alpha_2}, & u_1^{\alpha_1} < u_2^{\alpha_2}, \end{cases}$$

the mass of the singular component is concentrated on the curve $u_1^{\alpha_1} = u_2^{\alpha_2}$ in $\mathbf{I}^2$ as seen in figure 4.1.

Kendall's tau and Spearman's rho are quite easily evaluated for this copula family.
For Spearman's rho applying the result obtained in Theorem 3.4 yields:

$$\begin{aligned}\rho_{\alpha_1,\alpha_2} &= 12\iint_{\mathbf{I}^2} C_{\alpha_1,\alpha_2}(u, v)\,\mathrm{d}u\,\mathrm{d}v - 3 \\ &= 12\int_0^1 (\int_0^{u^{\alpha_1/\alpha_2}} u^{1-\alpha_1}v\,\mathrm{d}v + \int_{u^{\alpha_1/\alpha_2}}^1 uv^{1-\alpha_2}\,\mathrm{d}v)\,\mathrm{d}u - 3 \\ &= \ldots \\ &= \frac{3\alpha_1\alpha_2}{2\alpha_1 + 2\alpha_2 - \alpha_1\alpha_2}.\end{aligned}$$

To evaluate Kendall's tau we use the following theorem, which proof is found in Nelsen (1999) [13]:

Theorem 4.2. *Let C be a copula such that the product $\big(\partial C/\partial u\big)\big(\partial C/\partial v\big)$ is integrable on $\boldsymbol{I}^2$. Then*

$$\iint_{\boldsymbol{I}^2} C(u, v)\,\mathrm{d}C(u, v) = \frac{1}{2} - \iint_{\boldsymbol{I}^2} \frac{\partial}{\partial u}C(u, v)\frac{\partial}{\partial u}C(u, v)\,\mathrm{d}u\,\mathrm{d}v.$$

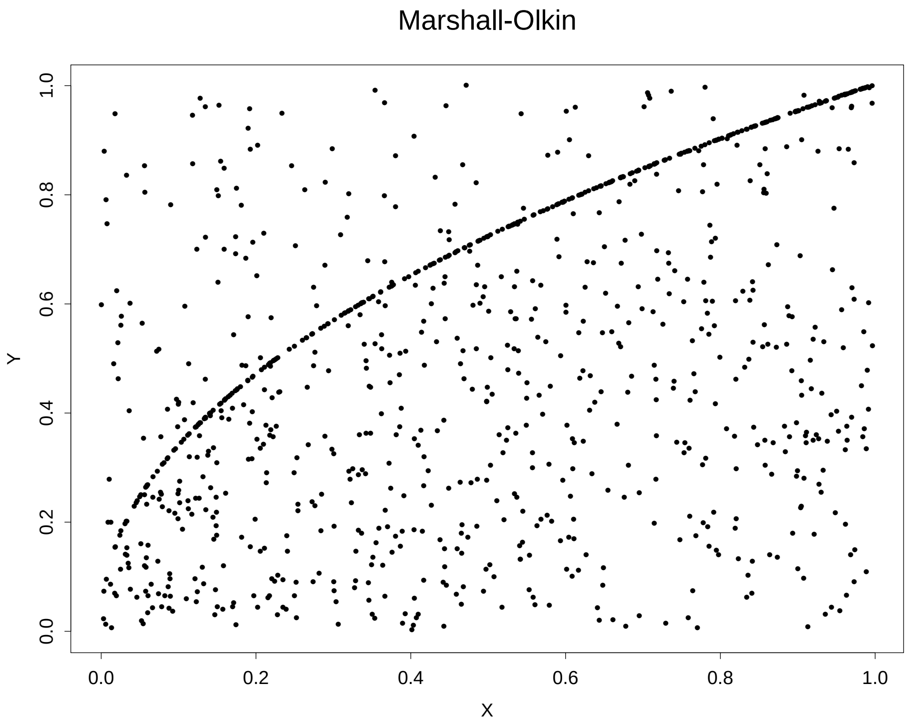

Figure 4.1: Samples from the Marshall-Olkin copula. $\lambda_1 = 1.1, \lambda_2 = 0.2$ and $\lambda_{12} = 0.6$

Using the result in Theorem 3.3 and the theorem above yields:

$$\begin{aligned}\tau_{\alpha_1,\alpha_2} &= 4\iint_{\mathbf{I}^2} C_{\alpha_1,\alpha_2}(u,v)\,\mathrm{d}C_{\alpha_1,\alpha_2}(u,v) - 1\\ &= 4\left(\frac{1}{2} - \iint_{\mathbf{I}^2} \frac{\partial}{\partial u}C_{\alpha_1,\alpha_2}(u,v)\frac{\partial}{\partial u}C_{\alpha_1,\alpha_2}(u,v)\,\mathrm{d}u\,\mathrm{d}v\right) - 1\\ &= \dots\\ &= \frac{\alpha_1\alpha_2}{\alpha_1+\alpha_2-\alpha_1\alpha_2}.\end{aligned}$$

Thus all values in the interval $[0,1]$ can be obtained for ρ_{α_1,α_2} and τ_{α_1,α_2}. The Marshall-Olkin copulas have upper tail dependence. Without loss of generality assume that $\alpha_1 > \alpha_2$.

$$\begin{aligned}\lim_{u\to 1-}\frac{\overline{C}(u,u)}{1-u} &= \lim_{u\to 1-}\frac{1-2u+u^2\min(u^{-\alpha_1},u^{-\alpha_2})}{1-u}\\ &= \lim_{u\to 1-}\frac{1-2u+u^2u^{-\alpha_2}}{1-u}\\ &= \lim_{u\to 1-}(2-2u^{1-\alpha_2}+\alpha_2 u^{1-\alpha_2}) = \alpha_2,\end{aligned}$$

and hence

$$\lambda_U = \min(\alpha_1,\alpha_2).$$

We now present the natural multivariate extension of the bivariate Marshall-Olkin family.

Let $X_1, \dots, X_n$ be an n-component system where the components are subject to $\sum_{k=1}^{n} \binom{n}{k}$ shocks. Each shock is in one-to-one correspondence with one of the $\sum_{k=1}^{n} \binom{n}{k}$ non-empty subsets of $\{X_1, \dots, X_n\}$, and each of those subsets are assigned to a shock according to the mapping Δ given by

$$\Delta(\{j_1 \dots j_k\}) = \{X_{j_1}, \dots, X_{j_k}\},$$

for $k \in \{1, \dots, n\}$ and $1 \leq j_1 < \dots j_k \leq n$.

Furthermore the shocks are assumed to form independent Poisson processes with parameters

$$\lambda_{j_1,\dots,j_k} \geq 0, k = 1, \dots, n, 1 \leq j_1 < \dots j_k \leq n.$$

The times of occurrence of these shocks are independent exponential random variables with the parameters above.

Example 4.1. Let $n = 4$. Then

$$\begin{aligned} X_1 &= \min(Z_1, Z_{12}, Z_{13}, Z_{14}, Z_{123}, Z_{124}, Z_{134}, Z_{1234}), \\ X_2 &= \min(Z_2, Z_{12}, Z_{23}, Z_{24}, Z_{123}, Z_{124}, Z_{234}, Z_{1234}), \\ X_3 &= \min(Z_3, Z_{13}, Z_{23}, Z_{34}, Z_{123}, Z_{134}, Z_{234}, Z_{1234}), \\ X_4 &= \min(Z_4, Z_{14}, Z_{24}, Z_{34}, Z_{124}, Z_{134}, Z_{234}, Z_{1234}). \end{aligned}$$

If for example $\lambda_{13} = 0$, then $Z_{13} = \infty$ almost surely.

Random variate generation from the extended Marshall-Olkin family is easy as will be shown by the algorithm below.

First of all order the shocks in some order (the order suggested by the above example seems however most natural). A parameter λ_k is then referred to as the parameter of the k:th shock according to this order. Define the $n \times \sum_{k=1}^{n} \binom{n}{k}$ matrix (a_{ij}) via:

$$a_{ij} = \begin{cases} 1 & \text{if shock } j \text{ kills component } i, \\ 0 & \text{otherwise.} \end{cases}$$

Example 4.2. Let $n = 4$. Then with the same order as in Example 4.1

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

The following algorithm generates random variates from the extended Marshall-Olkin copula:

Algorithm 2.

- Generate $l = \sum_{k=1}^{n} \binom{n}{k}$ independent values $v_1, \dots, v_l$ from $U(0, 1)$.
- Set $x_i = \min_{1 \leq k \leq l, a_{ik}=1, \lambda_k \neq 0} - \ln v_k / \lambda_k, \quad i = 1, \dots, n.$
- Set $\Lambda_i = \sum_{k=1}^{l} a_{ik} \lambda_k, \quad i = 1, \dots, n.$
- Set $u_i = \exp(-\Lambda_i x_i), \quad i = 1, \dots, n.$

$(x_1, \dots, x_n)$ is an n-variate from the extended Marshall-Olkin distribution and $(u_1, \dots, u_n)$ is an n-variate from the extended Marshall-Olkin copula. Furthermore Λ_i is the shock intensity “felt” by component i.

Kendall's tau and Spearman's rho are easily evaluated for the extended Marshall-Olkin copula since the expressions are on the same form as in the bivariate case. For $i, j \in \{1, \dots, n\}, i \neq j$:

$$\begin{aligned}\alpha_i &= \big(\sum_{k=1}^{l} a_{ik} a_{jk} \lambda_k\big) / \big(\sum_{k=1}^{l} a_{ik} \lambda_k\big),\\ \alpha_j &= \big(\sum_{k=1}^{l} a_{ik} a_{jk} \lambda_k\big) / \big(\sum_{k=1}^{l} a_{jk} \lambda_k\big).\end{aligned}$$

Then the Kendall's tau rank correlation matrix $\big(\tau_{ij}\big)$ is given by:

$$\tau_{ij} = \frac{\alpha_i \alpha_j}{\alpha_i + \alpha_j - \alpha_i \alpha_j},$$

and the Spearman's rho rank correlation matrix $\big(\rho_{ij}\big)$ is given by:

$$\rho_{ij} = \frac{3\alpha_i \alpha_j}{2\alpha_i + 2\alpha_j - \alpha_i \alpha_j}.$$

As seen above evaluating the rank correlation matrix given the full parameterization of the extended Marshall-Olkin copula is straight forward. However given a (Kandall's tau or Spearman's rho) rank correlation matrix we can not in general obtain a unique parameterization of the copula. By setting the shock intensities for subgroups with more then two elements to zero, we obtain the perhaps most natural parameterization of the copula in this situation. However this also means that the copula only has bivariate dependence.

5 Archimedean Copulas

In this chapter we discuss an important class of copulas called Archimedean copulas. This class of copulas is worth studying for a number of reasons: 1) The many parametric families of copulas belonging to this class 2) The great variety of different dependence structures 3) The many nice properties possessed by members of this class 4) The ease with which they can be constructed and simulated from. At the end of this chapter we present one possible multivariate extension of Archimedean copulas and a general algorithm for random variate generation for those families of n-copulas. For other multivariate extensions we refer to Joe (1997) [6].

5.1 Convex Sums

Let $\{C_i\}_{i=1}^m$ be a collections of n-copulas. Then every convex combination

$$\Upsilon(u_1,\dots,u_n)=\sum_{i=1}^m \lambda_i C_i(u_1,\dots,u_n),$$

is also a copula, since

$$\begin{aligned}
\Upsilon(u_1,\dots,0,\dots,u_n) &= \sum_{i=1}^m \lambda_i 0 = 0,\\
\Upsilon(1,\dots,u_k,\dots,1) &= \sum_{i=1}^m \lambda_i u_k = u_k,\\
V_\Upsilon(B) = \sum_{\mathbf{c}} \operatorname{sgn}(\mathbf{c})\Upsilon(\mathbf{c}) &= \sum_{\mathbf{c}} \operatorname{sgn}(\mathbf{c}) \sum_{i=1}^m \lambda_i C_i(\mathbf{c})\\
= \sum_{i=1}^m \lambda_i \sum_{\mathbf{c}} \operatorname{sgn}(\mathbf{c}) C_i(\mathbf{c}) &= \sum_{i=1}^m \lambda_i V_{C_i}(B) \geq 0.
\end{aligned}$$

This can be extended to infinite collections of copulas indexed by a continuous parameter θ. Consider the parameter θ as an observation of an continuous random variable Θ with distribution function Λ. If C' is given by

$$C'(u_1,\dots,u_n)=\int_{\mathbb{R}} C_\theta(u_1,\dots,u_n)\,\mathrm{d}\Lambda(\theta), \tag{5.1.1}$$

then C' is a copula, called the convex sum of $\{C_\theta\}$ with respect to Λ. We call Λ the mixing distribution of the family $\{C_\theta\}$.

Consider the representation given by equation (5.1.1) and for simplicity the bivariate case. It can be extended by replacing C_θ by more general bivariate functions. For example, set

$$H(u,v)=\int_0^\infty F^\theta(u)G^\theta(v)\,\mathrm{d}\Lambda(\theta), \tag{5.1.2}$$

e.g., let H be mixture of powers of distribution functions F and G. Furthermore assume $\Lambda(0)=0$. Let $\Psi(t)$ denote the Laplace transform of the mixing distribution Λ, i.e.,

$$\Psi(t)=\int_0^\infty e^{-\theta t}\,\mathrm{d}\Lambda(\theta).$$

Let F and G be distribution functions given by $F(u) = \exp(-\Psi^{-1}(u))$ and $G(v) = \exp(-\Psi^{-1}(v))$ for all u, v in $\mathbf{I}$. Then (5.1.2) becomes

$$\begin{aligned} H(u,v) &= \int_0^\infty \exp\left[-\theta\left(\Psi^{-1}(u)+\Psi^{-1}(v)\right)\right] \mathrm{d}\Lambda(\theta) \\ &= \Psi\left(\Psi^{-1}(u)+\Psi^{-1}(v)\right), \end{aligned}$$

which can be shown to be a bivariate distribution function. Furthermore, since $\Psi^{-1}(1) = 0$, its margins are uniform. Thus H is a copula. So when Ψ is the Laplace transform of a distribution function, then the function C given by

$$C(u,v) = \Psi\left(\Psi^{-1}(u)+\Psi^{-1}(v)\right) \tag{5.1.3}$$

is a copula. C given by (5.1.3) is a copula for a larger class of functions than Laplace transforms. These copulas are called Archimedean copulas and are the subject of this chapter.

5.2 Definitions

Definition 13. Let φ be a continuous, strictly decreasing function from $\mathbf{I}$ to $[0,\infty]$ such that $\varphi(1) = 0$. The pseudo-inverse of φ is the function $\varphi^{[-1]}$ with Dom $\varphi^{[-1]} = [0,\infty]$ and Ran $\varphi^{[-1]} = \mathbf{I}$ given by

$$\varphi^{[-1]}(t) = \begin{cases} \varphi^{-1}(t), & 0 \le t \le \varphi(0), \\ 0, & \varphi(0) \le t \le \infty. \end{cases} \tag{5.2.1}$$

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

Lemma 5.1. *Let φ be a continuous, strictly decreasing function from $\mathbf{I}$ to $[0,\infty]$ such that $\varphi(1) = 0$, and let $\varphi^{[-1]}$ be the pseudo-inverse of φ defined by (5.2.1). Let C be the function from $\mathbf{I}^2$ to $\mathbf{I}$ given by*

$$C(u,v) = \varphi^{[-1]}(\varphi(u)+\varphi(v)). \tag{5.2.2}$$

Then C satisfies the boundary conditions for a copula.

Proof. $C(u,0) = \varphi^{[-1]}(\varphi(u)+\varphi(0)) = 0$, and $C(u,1) = \varphi^{[-1]}(\varphi(u)+\varphi(1)) = \varphi^{[-1]}(\varphi(u)) = u$. By symmetry, $C(0,v) = 0$ and $C(1,v) = v$. ☐

Lemma 5.2. *Let φ, $\varphi^{[-1]}$ and C satisfy the hypotheses of Lemma 5.1 Then the function C is 2-increasing if and only if whenever $u_1 \le u_2$,*

$$C(u_2,v) - C(u_1,v) \le u_2 - u_1. \tag{5.2.3}$$

Proof. Since (5.2.3) is equivalent to $V_C([u_1,u_2]\times[v,1]) \ge 0$, it holds whenever C is 2-increasing. Hence assume that C satisfies (3.1.3). Choose v_1, v_2 in $\mathbf{I}$ such that $v_1 \le v_2$ and note that $C(0,v_2) = 0 \le v_1 \le v_2 = C(1,v_2)$. But C is continuous (since φ and $\varphi^{[-1]}$ are), and thus there is a t in $\mathbf{I}$ such that $C(t, v_2) = v_1$, or $\varphi(v_2) + \varphi(t) = \varphi(v_1)$. Hence

$$\begin{aligned} C(u_2, v_1) - C(u_1, v_1) &= \varphi^{[-1]}(\varphi(u_2) + \varphi(v_1)) - \varphi^{[-1]}(\varphi(u_1) + \varphi(v_1)) \\ &= \varphi^{[-1]}(\varphi(u_2) + \varphi(v_2) + \varphi(t)) - \\ &\quad\ \varphi^{[-1]}(\varphi(u_1) + \varphi(v_2) + \varphi(t)) \\ &= C(C(u_2, v_2), t) - C(C(u_1, v_2), t) \\ &\leq C(u_2, v_2) - C(u_1, v_2), \end{aligned}$$

so that C is 2-increasing. □

Theorem 5.1. *Let φ be a continuous, strictly decreasing function from $\boldsymbol{I}$ to $[0, \infty]$ such that $\varphi(1) = 0$, and let $\varphi^{[-1]}$ be the pseudo-inverse of φ defined by* (5.2.1)*. The function C from $\boldsymbol{I}^2$ to $\boldsymbol{I}$ given by* (5.2.2) *is a copula if and only if φ is convex.*

Proof. We have already shown that C satisfies the boundary conditions for a copula, and as a consequence of the preceding lemma, we need only prove that (5.2.3) holds if and only if φ is convex. Note that φ is convex if and only if $\varphi^{[-1]}$ is convex. Observe that (5.2.3) is equivalent to

$$u_1 + \varphi^{[-1]}(\varphi(u_2) + \varphi(v)) \leq u_2 + \varphi^{[-1]}(\varphi(u_1) + \varphi(v))$$

for $u_1 \leq u_2$, so if we set $a = \varphi(u_1)$, $b = \varphi(u_2)$, and $c = \varphi(v)$, then (5.2.3) is equivalent to

$$\varphi^{[-1]}(a) + \varphi^{[-1]}(b + c) \leq \varphi^{[-1]}(b) + \varphi^{[-1]}(a + c), \qquad (5.2.4)$$

where $a \geq b$ and $c \geq 0$. Now suppose (5.2.3) holds, i.e., suppose that $\varphi^{[-1]}$ satisfies (5.2.4). Choose any s, t in $[0, \infty]$ such that $0 \leq s < t$. If we set $a = (s + t)/2$, $b = s$, and $c = (t - s)/2$ in (5.2.4), we have

$$\varphi^{[-1]}\left(\frac{s + t}{2}\right) \leq \frac{\varphi^{[-1]}(s) + \varphi^{[-1]}(t)}{2}. \qquad (5.2.5)$$

Thus $\varphi^{[-1]}$ is midconvex, and since $\varphi^{[-1]}$ is continuous it follows that $\varphi^{[-1]}$ is convex. In the other direction, assume $\varphi^{[-1]}$ is convex. Fix a, b and c in $\mathbf{I}$ such that $a \geq b$ and $c \geq 0$; and let $\gamma = (a - b)/(a - b + c)$. Now $a = (1 - \gamma)b + \gamma(a + c)$ and

$$\varphi^{[-1]}(a) \leq (1 - \gamma)\varphi^{[-1]}(b) + \gamma\varphi^{[-1]}(a + c).$$

and

$$\varphi^{[-1]}(b + c) \leq \gamma\varphi^{[-1]}(b) + (1 - \gamma)\varphi^{[-1]}(a + c).$$

Adding these inequalities yields (5.2.5), which completes the proof. □

Copulas of the form (5.2.2) 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.

Example 5.1. Let $\varphi(t) = (-\ln t)^{\theta}$, where $\theta \geq 1$. Clearly $\varphi(t)$ is continuous and $\varphi(1) = 0$. $\varphi'(t) = -\theta(-\ln t)^{\theta - 1}\frac{1}{t}$, so φ is a strictly decreasing function from $\mathbf{I}$ to $[0, \infty]$. $\varphi''(t) = \theta(\theta - 1)(-\ln t)^{\theta - 2}\frac{1}{t^2} + \theta(-\ln t)^{\theta - 1}\frac{1}{t^2} \geq 0$ on $\mathbf{I}$, so φ is convex. Moreover $\varphi(0) = \infty$, so φ is a strict generator. From (2) we get

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

Furthermore $C_1 = \Pi$ and $C_{\infty} = M$. This copula family is called the Gumbel family. As shown in Example 3.4 this copula family has upper tail dependence.

Example 5.2. Let $\varphi(t) = (t^{-\theta} - 1)/\theta$, where $\theta \in [-1, \infty)\backslash\{0\}$. This gives the Clayton family

$$C_\theta(u, v) = \max([u^{-\theta} + v^{-\theta} - 1]^{-1/\theta}, 0). \tag{5.2.6}$$

For $\theta > 0$ the copulas are strict and the copula expression simplifies to

$$C_\theta(u, v) = (u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}. \tag{5.2.7}$$

The Clayton family has lower tail dependence for $\theta \geq 0$ and is comprehensive, i.e., $C_{-1} = W$, $C_\infty = M$ and $\lim_{\theta \to 0} C_\theta = \Pi$. Since most of the following results are results for strict Archimedean copulas we will refer to (5.2.7) as the Clayton family and refer to (5.2.6) as the extension to negative dependence of the Clayton family.

Example 5.3. Let $\varphi(t) = -\ln \frac{e^{-\theta t} - 1}{e^{\theta} - 1}$, where $\theta \in \mathbb{R}\backslash\{0\}$. This gives the Frank family

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

The Frank copulas are strict copulas and are comprehensive, i.e., $C_{-\infty} = W$, $C_\infty = M$ and $\lim_{\theta \to 0} C_\theta = \Pi$. Members of the Frank family are the only Archimedean copulas which satisfy the equation $C(u, v) = \hat{C}(u, v)$ for so called radial symmetry.

Example 5.4. Let $\varphi(t) = 1 - t$ for t in $[0, 1]$. Then $\varphi^{[-1]}(t) = 1 - t$ for t in $[0, 1]$ and 0 for $t > 1$; i.e., $\varphi^{[-1]}(t) = \max(1 - t, 0)$. Hence $C(u, v) = \max(u + v - 1, 0) = W(u, v)$. Hence W is Archimedean.

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

Theorem 5.2. *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 $\boldsymbol{I}$;*
2. *C is associative; i.e., $C(C(u, v), w) = C(u, C(v, w))$ for all u, v, w in $\boldsymbol{I}$.*

Proof. The first part follows directly from (5.2.2).

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

and hence C is associative. □

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

Example 5.5. 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

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

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

5.3 Properties

For convenience, we let Ω denote the set of continuous strictly decreasing convex functions φ from $\mathbf{I}$ to $[0,1]$ with $\varphi(1)=0$.

Theorem 5.3. *Let C be an Archimedean copula generated by φ in Ω. Let $K_C(t)$ denote the C-measure of the set $\{(u,v)\in \boldsymbol{I}^2 | C(u,v)\le t\}$. Then for any t in $\boldsymbol{I}$,*

$$K_C(t) = t - \frac{\varphi(t)}{\varphi'(t^+)}. \tag{5.3.1}$$

Proof. Let t be in $(0,1)$, and set $w=\varphi(t)$. Let n be a fixed positive integer, and consider the partition of the interval $[t,1]$ induced by the partition $\{0, w/n, \dots, kw/n, \dots, w\}$ of $[0,w]$, i.e., the partition $\{t=t_0, t_1, \dots, t_k, \dots, t_n = 1\}$ where $t_{n-k}=\varphi^{[-1]}(kw/n)$, $k = 0,1,\dots,n$. Since $w<\varphi(0)$, it follows from (1) that

$$\begin{aligned} C(t_j,t_k) &= \varphi^{[-1]}(\varphi(t_j)+\varphi(t_k)) \\ &= \varphi^{[-1]}\left(\frac{n-j}{n}w + \frac{n-k}{n}w\right) = \varphi^{[-1]}\left(w + \frac{n-j-k}{n}w\right). \end{aligned}$$

In particular, $C(t_j,t_{n-j}) = \varphi^{[-1]}(w) = t$.
Let R_k denote the rectangle $[t_{k-1},t_k]\times[0,t_{n-k+1}]$, and set $S_n = \bigcup_{k=1}^{n} R_k$. From the convexity of $\varphi^{[-1]}$ it follows that

$$0 \le t_1 - t_0 \le t_2 - t_1 \le \dots \le t_n - t_{n-1} = 1 - t_{n-1},$$

and clearly $\lim_{n\to\infty}(1-t_n) = 1-\varphi^{[-1]}(0) = 0$. Hence $K_C(t)$ is given by the sum of the C-measure of $[0,t]\times$I and $\lim_{n\to\infty} V_C(S_n)$, i.e., $K_C(t) = t +\lim_{n\to\infty} V_C(S_n)$. For each k we have

$$\begin{aligned} V_C(R_k) &= C(t_k,t_{n-k+1}) - C(t_{k-1},t_{n-k+1}) - C(t_k,0) + C(t_{k-1},0) \\ &= C(t_k,t_{n-k+1}) - t = \varphi^{[-1]}(w-w/n) - \varphi^{[-1]}(w), \end{aligned}$$

and hence

$$\begin{aligned} V_C(S_n) &= \sum_{k=1}^{n} V_C(R_k) \\ &= -w[\frac{\varphi^{[-1]}(w) - \varphi^{[-1]}(w-w/n)}{w/n}] \end{aligned}$$

from which it follows that

$$lim_{n\to\infty} V_C(S_n) = -w{\varphi'}^{[-1]}(w^+) = -\varphi(t)/\varphi'(t^+)$$

□

Corollary 5.1. *Let U and V be uniform $(0,1)$ random variables whose joint distribution function is the Archimedean copula C generated by φ in Ω. Then the function K_C given by (5.3.1) is the distribution function of the random variable $C(U,V)$.*

In general evaluating Kendall's tau for a certain copula C requires evaluation of the double integral in (3.3.5). In many cases this might not even be possible analytically. For Archimedean copulas, the situation is simpler, because Kendall's tau can be evaluated directly from the generator, as shown in the following theorem.

Theorem 5.4. *Let X and Y be random variables with an Archimedean copula C generated by φ in Ω. Kendall's tau of X and Y is given by*

$$\tau_C = 1 + 4\int_0^1 \frac{\varphi(t)}{\varphi'(t)}\,\mathrm{d}t. \tag{5.3.2}$$

Proof. Let U and V be uniform $(0,1)$ random variables with joint distribution function C, and let K_C denote the distribution function of $C(U,V)$. Then from (3.3.6) we have

$$\begin{aligned}\tau_C &= 4\mathbb{E}(C(U,V))-1,\\ &= 4\int_1^0 t\,\mathrm{d}K_C(t)-1,\\ &= 4\big([tK_C(t)]_0^1-\int_0^1 K_C(t)\,\mathrm{d}t\big)-1,,\\ &= 3-\int_0^1 K_C(t)\,\mathrm{d}t.\end{aligned}$$

From Theorem 5.3 and Corollary 5.1 it follows that the distribution function K_C of $C(U,V)$ is

$$K_C(t)=t-\frac{\varphi(t)}{\varphi'(t^+)},$$

and hence

$$\tau_C=3-4\int_0^1\big(t-\frac{\varphi(t)}{\varphi'(t^+)}\big)\,\mathrm{d}t=1+4\int_0^1\frac{\varphi(t)}{\varphi'(t)}\,\mathrm{d}t,$$

where $\varphi'(t^+)$ is replaced by $\varphi'(t)$ in the denominator of the integral because concave functions are differentiable almost everywhere. ☐

Example 5.6. Consider the Gumbel family with generator $\varphi(t)=(-\ln t)^\theta$, for $\theta\geq 1$. Then

$$\frac{\varphi(t)}{\varphi'(t)}=\frac{t\ln t}{\theta}.$$

Using Theorem 5.4 we can calculate Kendall's tau for the Gumbel family.

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

Example 5.7. Consider the Clayton family with generator $\varphi(t)=(t^{-\theta}-1)/\theta$, for $\theta\in[-1,\infty)\backslash\{0\}$. Then

$$\frac{\varphi(t)}{\varphi'(t)}=\frac{t^{\theta+1}-t}{\theta}.$$

Using Theorem 5.4 we can calculate Kendall's tau for the Clayton family.

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

Example 5.8. Consider the Frank family presented in Example 5.3. It can be shown that Kendall's tau is

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

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

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

for any positive integer k. Furthermore Spearman's rho is

$$\rho_\theta = 1 - \frac{12}{\theta}\left(D_1(\theta) - D_2(\theta)\right).$$

The next theorem will provide the basis in 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, needed in the proof of this theorem.

From (5.2.2) it follows that

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

and hence

$$\begin{aligned} \frac{\partial^2}{\partial u \partial v}C(u,v) = -\frac{\varphi''(C(u,v))\frac{\partial}{\partial u}C(u,v)\frac{\partial}{\partial v}C(u,v)}{\varphi'(C(u,v))} &= \\ -\frac{\varphi''(C(u,v))\varphi'(u)\varphi'(v)}{[\varphi'(C(u,v))]}{}_3 & \end{aligned}$$

Thus, when C is absolutely continuous, its density is given by

$$\frac{\partial^2}{\partial u \partial v}C(u,v) = -\frac{\varphi''(C(u,v))\varphi'(u)\varphi'(v)^3}{[\varphi'(C(u,v))]}. \tag{5.3.3}$$

Theorem 5.5. *Under the hypotheses of Corollary 5.1, 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* $\boldsymbol{I}^2$. *Hence* S *and* T *are independent, and* S *is uniformly distributed on* $(0,1)$.

Proof. We presents a proof for the case when C is absolutely continuous. The joint density $h(s,t)$ of S and T is given by

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

in terms of s and t, where $\partial^2 C(u,v)/\partial u \partial v$ is given by (5.3.3) and $\partial(u,v)/\partial(s,t)$ denotes the Jacobian of the transformation $\varphi(u) = s\varphi(t)$, $\varphi(u) = (1-s)\varphi(t)$. But

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

and hence

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

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

and the conclusion follows. □

An application of Theorem 5.5 is the following algorithm for generating random variates (u,v) whose joint distribution is an Archimedean copula C with generator φ in Ω:

Algorithm 3.

- Generate two independent uniform $(0,1)$ variates s and q;
- Set $t = K_C^{(-1)}(q)$, where $K_C^{(-1)}$ denotes the quasi-inverse of the distribution function K_C;
- Set $u = \varphi^{[-1]}(s\varphi(t))$ and $v = \varphi^{[-1]}((1-s)\varphi(t))$;
- The desired pair is (u,v)

Note that the variates s and t correspond to the random variables S and T in Theorem 5.5 and from the proof it follows that this algorithm is correct.

5.3.1 Tail Dependence

For Archimedean copulas tail dependence can be expressed in terms of the generators.

Theorem 5.6. *Let C be a strict Archimedean bivariate copula. If ${\varphi^{-1}}'(0)$ is finite, then*

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

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

$$\lambda_U = 2 - 2\lim_{s\to 0}[{\varphi^{-1}}'(2s)/{\varphi^{-1}}'(s)].$$

Proof.

$$\begin{aligned}\lim_{u\to 1}\overline{C}(u,u)/(1-u) &= \lim_{u\to 1}[1 - 2u + \varphi^{-1}(2\varphi(u))]/(1-u)\\ &= 2 - 2\lim_{u\to 1}{\varphi^{-1}}'(2\varphi(u))/{\varphi^{-1}}'(\varphi(u))\\ &= 2 - 2\lim_{s\to 0}[{\varphi^{-1}}'(2s)/{\varphi^{-1}}'(s)].\end{aligned}$$

If ${\varphi^{-1}}'(0) \in (-\infty, 0)$, then the limit is zero and C does not have upper tail dependence. Since ${\varphi^{-1}}'(0) < 0$ the result follows. □

Example 5.9. The Gumbel copulas are a strict Archimedean copulas with generator $\varphi(t) = (-\ln t)^\theta$. Hence $\varphi^{-1}(s) = \exp(-s^{1/\theta})$ and ${\varphi^{-1}}'(s) = -s^{1/\theta - 1}\exp(-s^{1/\theta})/\theta$. By using Theorem 5.6 we get

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

Theorem 5.7. *Let φ be a strict generator. The coefficient of lower tail dependence for the copula $C(u,v) = \varphi^{-1}(\varphi(u) + \varphi(v))$ is equal to*

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

Proof. The proof is similar to that of Theorem 5.6. □

Example 5.10. Consider the Clayton family given by

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

for $\theta \geq 0$. 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 5.7 shows that the Clayton family has lower tail dependence.

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

Example 5.11. 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 $\varphi^{-1}(s) = -\frac{1}{\theta} \ln(1 - (1 - e^{-\theta})e^{-s})$ and $\varphi^{-1\prime}(s) = -\frac{1}{\theta}(1 - e^{-\theta})e^{-s})/(1 - (1 - e^{-\theta})e^{-s})$. Since

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

is finite the Frank family does not have upper tail dependence according to Theorem 5.6. Furthermore

$$\frac{\varphi^{-1\prime}(2s)}{\varphi^{-1\prime}(s)} = e^{-s} \frac{1 - (1 - e^{-\theta})e^{-s}}{1 - (1 - e^{-\theta})e^{-2s}}$$

and hence

$$\lim_{s\to\infty} \frac{\varphi^{-1\prime}(2s)}{\varphi^{-1\prime}(s)} = 0.$$

Thus the Frank family does not have lower tail dependence.

5.4 Order

Recall the concordance ordering of copulas suggested by the Fréchet-Hoeffding inequality. In this section we will give conditions in terms of the generators, that enables us to determine among other things which parametric families are positively ordered.

Definition 14. A function f defined on $[0,\infty)$ is subadditive if for all x, y in $[0,\infty)$,

$$f(x + y) \leq f(x) + f(y). \tag{5.4.1}$$

Theorem 5.8. *Let* C_1 *and* C_2 *be Archimedean copulas generated, respectively, by* φ_1 *and* φ_2 *in* Ω. *Then* $C_1 \prec C_2$ *if and only if* $\varphi_1 \circ \varphi_2^{[-1]}$ *is subadditive.*

Proof. Let $f = \varphi_1 \circ \varphi_2^{[-1]}$. Note that f is continuous, nondecreasing, and $f(0) = 0$. From Definition 7, $C_1 \prec C_2$ if and only if for all u, v in $\mathbf{I}$,

$$\varphi_1^{[-1]}(\varphi_1(u) + \varphi_1(v)) \leq \varphi_2^{[-1]}(\varphi_2(u) + \varphi_2(v)). \tag{5.4.2}$$

Let $x = \varphi_2(u)$ and $y = \varphi_2(v)$, then (5.4.2) is equivalent to

$$\varphi_1^{[-1]}(f(x) + f(y)) \leq \varphi_2^{[-1]}(x + y) \tag{5.4.3}$$

for all x, y in $[0, \varphi_2(0)]$. Moreover if $x > \varphi_2(0)$ or $y > \varphi_2(0)$, then each side of (5.4.3) is equal to 0.

Now suppose that $C_1 \prec C_2$. Applying φ_1 to both sides of (5.4.3) and noting that $\varphi_1 \circ \varphi_1^{[-1]}(w) \leq w$ for all $w \geq 0$ yields (5.4.1) for all x, y in $[0, \infty)$, hence f is subadditive. Conversely, if f satisfies (5.4.1), then applying $\varphi_1^{[-1]}$ to both sides and noting that $\varphi_1^{[-1]} \circ f = \varphi_2^{[-1]}$ yields (5.4.2), completing the proof. □

Lemma 5.3. *Let* f *be defined on* $[0, \infty)$. *If* f *is concave and* $f(0) = 0$, *then* f *is subadditive.*

Proof. Let x, y be in $[0, \infty)$. If $x + y = 0$, so that with $f(0) = 0$, (5.4.1) is trivial. So assume $x + y > 0$, so that

$$x = \frac{x}{x+y}(x+y) + \frac{y}{x+y}(0) \text{ and } y = \frac{x}{x+y}(0) + \frac{y}{x+y}(x+y).$$

If f is concave and $f(0) = 0$, then

$$f(x) \geq \frac{x}{x+y} f(x+y) + \frac{y}{x+y} f(0) = \frac{x}{x+y} f(x+y)$$

and

$$f(y) \geq \frac{x}{x+y} f(0) + \frac{y}{x+y} f(x+y) = \frac{y}{x+y} f(x+y).$$

from which (5.4.1) follows and f is subadditive. □

Corollary 5.2. *Under the hypotheses of Theorem 5.1, if* $\varphi_1 \circ \varphi_2^{[-1]}$ *is concave, then* $C_1 \prec C_2$.

Example 5.12. Let C_{θ_1} and C_{θ_2} be members of the Gumbel family with parameters θ_1 and θ_2, so that the generators of C_{θ_1} and C_{θ_2} are φ_{θ_1} and φ_{θ_2}, respectively, where $\varphi_{\theta_1}(t) = (-\ln t)^{\theta_1}$ and $\varphi_{\theta_2}(t) = (-\ln t)^{\theta_2}$. Then $\varphi_{\theta_1} \circ \varphi_{\theta_2}^{[-1]}(t) = \varphi_{\theta_1} \circ \varphi_{\theta_2}^{-1}(t) = (-\ln(\exp(-t^{1/\theta_2})))^{\theta_1} = t^{\theta 1/\theta_2}$. So if $\theta_1 \leq \theta_2$, then $\varphi_{\theta_1} \circ \varphi_{\theta_2}^{[-1]}$ is concave and $C_{\theta_1} \prec C_{\theta_2}$. Hence the Gumbel family is positively ordered.

5.5 Multivariate Archimedean Copulas

In this chapter we look at construction of and random variate generation from Archimedean n-copulas.

The expression for the product copula Π can be written in the form

$$\Pi(u, v) = uv = \exp(-[(-\ln u) + (-\ln v)]).$$

The extension of this idea to n dimensions, with $\mathbf{u} = (u_1, u_2, \ldots, u_n)$, results in the n-dimensional product copula Π^n in the form

$$\Pi^n(\mathbf{u}) = u_1 u_2 ... u_n = \exp(-[(-\ln u_1) + (-\ln u_2) + \ldots + (\ln u_n)]).$$

This leads naturally to the following generalization of (5.2.2):

$$C^n(\mathbf{u}) = \varphi^{[-1]}(\varphi(u_1) + \varphi(u_2) + \ldots + \varphi(u_n)). \tag{5.5.1}$$

The functions C^n are the serial iterates of the Archimedean 2-copulas generated by φ, that is, if we set $C^2(u_1, u_2) = C(u_1, u_2) = \varphi^{[-1]}(\varphi(u_1) + \varphi(u_2))$, then for $n \geq 3$, $C^n(u_1, u_2, \ldots, u_n) = C(C^{n-1}(u_1, u_2, \ldots, u_{n-1}), u_n)$. Note that this technique of composing copulas generally fails. But since Archimedean copulas are symmetric and associative it seems more likely that C^n, given certain additional properties of φ (and $\varphi^{[-1]}$) , is a copula for $n \geq 3$.

Definition 15. A function $g(t)$ is completely monotonic on the interval J if it is continuous there and has derivatives of all orders which alternate in sign, i.e., if it satisfies

$$(-1)^k \frac{d^k}{dt^k} g(t) \geq 0 \tag{5.5.2}$$

for all t in the interior of J and $k = 0, 1, 2, \ldots$.

As a consequence, if $g(t)$ is completely monotonic on $[0, \infty)$ and $g(c) = 0$ for some $c > 0$, then g must be identically zero on $[0, \infty)$. So if the pseudo-inverse $\varphi^{[-1]}$ of an Archimedean generator φ is completely monotonic, it must be positive on $[0, \infty)$, i.e., φ is strict and $\varphi^{[-1]} = \varphi^{-1}$.

Theorem 5.9. *Let φ be a continuous strictly decreasing function from $\boldsymbol{I}$ to $[0, \infty]$ such that $\varphi(0) = \infty$ and $\varphi(1) = 0$, and let φ^{-1} denote the inverse of φ. If C^n is the function from $\boldsymbol{I}$ to $\boldsymbol{I}^n$ given by (5.5.1), then C^n is an n-copula for all $n \geq 2$ if and only if φ^{-1} is completely monotonic on $[0, \infty)$.*

This theorem can be partially extended to the case when φ is non-strict and $\varphi^{[-1]}$ is m-monotonic on $[0, \infty)$ for some $m \geq 2$, that is, the derivatives of $\varphi^{[-1]}$ alter sign up to and including the mth order on $[0, \infty)$. Then the function C^n given by (5.5.1) is an n-copula for $2 \leq n \leq m$.

Corollary 5.3. *If the inverse φ^{-1} of a strict generator φ of an Archimedean copula C is completely monotonic, then $C \succ \Pi$.*

Proof. We begin by proving that for a strict Archimedean copula C generated by φ in Ω if $-\ln \varphi^{-1}$ is concave on $(0, \infty)$, then $C \succ \Pi$.
Assume that $-\ln \varphi^{-1}$ is concave on $(0, \infty)$. Since $-\ln \varphi^{-1}(0) = -\ln 1 = 0$, $-\ln \varphi^{-1}$ is subadditive according to Lemma 5.3. Hence
$-\ln \varphi^{-1}(\varphi(u) + \varphi(v)) \leq -\ln \varphi^{-1}(\varphi(u)) - \ln \varphi^{-1}(\varphi(v)) \Leftrightarrow$
$-\ln C(u, v) \leq -\ln uv \Leftrightarrow C(u, v) \geq uv$.
This means that $C \succ \Pi$.
The concavity of $-\ln \varphi^{-1}$ on $(0, \infty)$ is equivalent to requiring that
$(\varphi^{-1\prime})^2 - \varphi^{-1}\varphi^{-1\prime\prime} \leq 0$, on $(0, \infty)$, and this inequality holds for all completely monotonic functions. □

While it is simple to generate n-copulas of the form given by (5.5.1) they suffer from a very limited dependence structure since all k-margins are identical. One would like to have a multivariate extension of the Archimedean 2-copula given by

(5.2.2) with a wide range of dependence structure. One example of such an extension is given next.

If generators φ_i are chosen so that certain conditions are satisfied, then multivariate copulas can be obtained such that each bivariate margin has the form of (5.2.2) for some i. However, the number of distinct generators φ_i among the $n(n-1)/2$ bivariate margins is only $n-1$, so that the resulting dependence structure is one of partial exchangeability. However a general algorithm for random variate generation from this class can be obtained which among other things provides an algorithm for random variate generation from the multivariate extreme value copula resulting from the multivariate extension of the Gumbel copula.

Since the general multivariate result is notationally complex, we indicate the pattern and conditions from the trivariate and 4-variate extensions of (5.2.2). The trivariate generalization of (5.2.2) is

$$\varphi_1^{-1}(\varphi_1 \circ \varphi_2^{-1}(\varphi_2(u_1)+\varphi_2(u_2))+\varphi_1(u_3)), \tag{5.5.3}$$

where φ_1 and φ_2 are generators of strict Archimedean copulas. The 4-variate generalization of (5.2.2) is

$$\varphi_1^{-1}(\varphi_1 \circ \varphi_2^{-1}(\varphi_2 \circ \varphi_3^{-1}(\varphi_3(u_1)+\varphi_3(u_2))+\varphi_2(u_3))+\varphi_1(u_4)), \tag{5.5.4}$$

where φ_1, φ_2 and φ_3 are generators of strict Archimedean copulas.

Clearly the generators have to satisfy the necessary conditions for the n-copula given by (5.5.1) in order to make (5.5.3) and (5.5.4) valid copula expressions, since (5.5.1) is a special case of the more general multivariate results indicated by (5.5.3) and (5.5.4). What other conditions are needed to make these proper copulas? To answer that question we now introduce function classes $\mathcal{L}_n$ and $\mathcal{L}_n^*$.

Let ϕ be a strictly decreasing differentiable function.

$$\mathcal{L}_n = \{\phi : [0,\infty) \to [0,1] \mid \phi(0)=1,\ \phi(\infty)=0,$$
$$(-1)^j\phi^{(j)} \geq 0,\ j=1,\dots,n\},$$

$n = 1,2,\dots,\infty$, with $\mathcal{L}_\infty$ being the class of Laplace transforms.

$$\mathcal{L}_n^* = \{\omega : [0,\infty) \to [0,\infty) \mid \omega(0)=0,\ \omega(\infty)=\infty,$$
$$(-1)^{j-1}\omega^{(j)} \geq 0, j=1,\dots,n\},$$

$n = 1,2,\dots,\infty$. Note that $\varphi^{-1} \in \mathcal{L}_1$ when φ is the generator of a strict Archimedean copula. The functions in $\mathcal{L}_n^*$ are usually compositions of the form $\psi^{-1} \circ \phi$ with ψ, $\phi \in \mathcal{L}_1$.

Note also that with this notation, the necessary and sufficient conditions for (5.5.1) to be a proper copula is that $\varphi^{-1} \in \mathcal{L}_n$ and that if (5.5.1) is a copula for all n, then φ^{-1} must be completely monotonic and hence be a Laplace transform.

It turns out that if φ_1^{-1} and φ_2^{-1} are completely monotonic (Laplace transforms) and $\varphi_1 \circ \varphi_2^{-1} \in \mathcal{L}_\infty^*$, then (5.5.3) is a proper copula. Note that (5.5.3) has $(1,2)$ bivariate margin of the form (5.2.2) with generator φ_2 and $(1,3)$ and $(2,3)$ bivariate margins of the form (5.2.2) with generator φ_1. Also (5.5.1) is the special case of (5.5.3) when $\varphi_1 = \varphi_2$. As a result of Corollary 5.2 the trivariate copula in (5.5.3) has a $(1,2)$ bivariate margin copula which is larger than the $(1,3)$ and $(2,3)$ bivariate margin copulas (which are identical).

As one would expect there are similar conditions for the 4-variate case. If φ_1^{-1}, φ_2^{-1} and φ_3^{-1} are completely monotonic (Laplace transforms) and $\varphi_1 \circ \varphi_2^{-1}$ and

$\varphi_2 \circ \varphi_3^{-1}$ are in $\mathcal{L}_\infty^*$, then (5.5.4) is a proper copula. Note that all trivariate margins of (5.5.4) have the form (5.5.3) and all bivariate margins have the form (5.2.2). Clearly the idea of (5.5.3) and (5.5.4) generalizes to higher dimensions.

Example 5.13. Let $\varphi_{\theta_1}(t) = (-\ln t)^{\theta_1}$ and $\varphi_{\theta_2}(t) = (-\ln t)^{\theta_2}$, where $\theta_1, \theta_2 \geq 1$. For $\theta_1 \leq \theta_2$ we get the trivariate extension of the Gumbel family. The copula expression (5.5.3) becomes

$$C(\mathbf{u}; \theta_1, \theta_2) = \exp\{-([(-\ln u_1)^{\theta_2} + (-\ln u_2)^{\theta_2}]^{\theta_1/\theta_2} + (-\ln u_3)^{\theta_1})^{1/\theta_1}\}.$$

We can now present an algorithm for random variate generation from the multivariate extension of (5.2.2) indicated by (5.5.3) and (5.5.4). Because the recursive nature of the multivariate extension of (5.2.2) indicated by (5.5.3) and (5.5.4) the algorithm is basically an extension of the algorithm for random variate generation from an Archimedean copula of the form (5.2.2) already presented.

Algorithm 4.

```
• Generate a uniform $(0,1)$ variate $q$.

• Set $t_1 = K_{C_{\theta_1}}^{-1}(q)$.
• Generate a uniform $(0,1)$ variate $s_1$ independent of $q$.
• Set $a_1 = \varphi_{\theta_1}^{-1}(s_1 \varphi_{\theta_1}(t_1))$ and $u_n = \varphi_{\theta_1}^{-1}((1-s_1)\varphi_{\theta_1}(t_1))$

  ⋮

• Set $t_i = K_{C_{\theta_i}}^{-1}(a_{i-1})$.
• Generate a uniform $(0,1)$ variate $s_k$ independent of $q$ and $s_1, \ldots, s_{i-1}$.
• Set $a_i = \varphi_{\theta_i}^{-1}(s_i \varphi_{\theta_i}(t_k))$ and $u_{n-i+1} = \varphi_{\theta_i}^{-1}((1-s_i)\varphi_{\theta_i}(t_i))$

  ⋮

• Set $t_{n-1} = K_{C_{\theta_{n-1}}}^{-1}(a_{n-2})$.
• Generate a uniform $(0,1)$ variate $s_{n-1}$ independent of $q, s_1, \ldots, s_{n-2}$.
• Set $u_1 = \varphi_{\theta_{n-1}}^{-1}(s_{n-1}\varphi_{\theta_{n-1}}(t_{n-1}))$ and $u_2 = \varphi_{\theta_{n-1}}^{-1}((1-s_{n-1})\varphi_{\theta_{n-1}}(t_{n-1}))$

• $(u_1, u_2, \ldots, u_n)$ is the desired random variate from
  $C^n(u_1, u_2, \ldots, u_n; \theta_1, \ldots, \theta_{n-1})$,
```

where $i \in \{2, \ldots, n-2\}$ and

$$\begin{aligned} C^k(u_1, u_2, \ldots, u_k; \theta_1, \ldots, \theta_{k-1}) &= C(C^{k-1}(u_1, u_2, \ldots, u_{k-1}; \theta_2, \ldots, \theta_{k-1}), u_k; \theta_1) \\ C(\cdot, \cdot; \theta_j) &= C_{\theta_j}(\cdot, \cdot) = \varphi_{\theta_j}^{-1}(\varphi_{\theta_j}(\cdot) + \varphi_{\theta_j}(\cdot)), \end{aligned}$$

for $k = 3, 4, \ldots, n$ and $j = 1, 2, \ldots, k-1$.
Let $U_1, U_2, \ldots, U_n$ be the uniform $(0,1)$ random variables given by the Algorithm 4. Then

$$K_{C_{\theta_1}}(C_{\theta_1}(K_{C_{\theta_2}}(C_{\theta_2}(\ldots K_{C_{\theta_{n-1}}}(C_{\theta_{n-1}}(U_1, U_2)), \ldots, U_{n-1})), U_n)) \tag{5.5.5}$$

is a uniform $(0,1)$ random variable.

The correctness of the algorithm is stated in the following theorem.

Theorem 5.10. *Let the random variables* $U_1, \dots, U_n$ *be generated according to Algorithm 4. Then*

(U_1, U_2) *have copula* $C_{\theta_{n-1}}$,
$(U_1, U_3), (U_2, U_3)$ *have copula* $C_{\theta_{n-2}}$,

$\vdots$

$(U_1, U_n), (U_2, U_n), \dots, (U_{n-1}, U_n)$ *have copula* C_{θ_1}.

Hence $(U_1, U_2, \dots, U_n)$ *have copula* $C^n(u_1, u_2, \dots, u_n; \theta_1, \dots, \theta_{n-1})$.

Proof. From the algorithm it follows that:
⋆ (U_1, U_2) have the copula $C_{\theta_{n-1}}$.
⋆ $(K_{C_{\theta_{n-1}}}(C_{\theta_{n-1}}(U_1, U_2)), U_3)$ have copula $C_{\theta_{n-2}}$, and since $K_{C_{\theta_{n-1}}}$ is strictly increasing on **I**, $(C_{\theta_{n-1}}(U_1, U_2), U_3)$ have copula $C_{\theta_{n-2}}$.

$$\begin{aligned} U_1 &= \varphi^{-1}_{\theta_{n-1}}(S_{n-1}\varphi_{\theta_{n-1}}(C_{\theta_{n-1}}(U_1, U_2)) \\ &= f_{n-1}(C_{\theta_{n-1}}(U_1, U_2)), \\ U_2 &= \varphi^{-1}_{\theta_{n-1}}((1 - S_{n-1})\varphi_{\theta_{n-1}}(C_{\theta_{n-1}}(U_1, U_2)) \\ &= g_{n-1}(C_{\theta_{n-1}}(U_1, U_2)), \end{aligned}$$

where f_{n-1} and g_{n-1} are strictly increasing on **I**. Hence (U_1, U_3) and (U_2, U_3) have copula $C_{\theta_{n-2}}$.
⋆ $(K_{C_{\theta_{n-2}}}(C_{\theta_{n-2}}(A_{n-2}, U_3)), U_4)$ have copula $C_{\theta_{n-3}}$,
where $A_{n-2} = K_{C_{\theta_{n-1}}}(C_{\theta_{n-1}}(U_1, U_2))$. Since $K_{C_{\theta_{n-2}}}$ is strictly increasing on **I**, $(C_{\theta_{n-2}}(K_{C_{\theta_{n-1}}}(C_{\theta_{n-1}}(U_1, U_2)), U_3), U_4)$ have copula $C_{\theta_{n-3}}$.

$$\begin{aligned} A_{n-2} &= \varphi^{-1}_{\theta_{n-2}}(S_{n-2}\varphi_{\theta_{n-2}}(C_{\theta_{n-2}}(A_{n-2}, U_3)) \\ &= f_{n-2}(C_{\theta_{n-2}}(A_{n-2}, U_3)), \\ U_3 &= \varphi^{-1}_{\theta_{n-2}}((1 - S_{n-2})\varphi_{\theta_{n-2}}(C_{\theta_{n-2}}(A_{n-2}, U_3)) \\ &= g_{n-2}(C_{\theta_{n-2}}(A_{n-2}, U_3)), \end{aligned}$$

where f_{n-2} and g_{n-2} are strictly increasing on **I**. Hence (A_{n-2}, U_4) and (U_3, U_4) have copula $C_{\theta_{n-3}}$. Furthermore

$$\begin{aligned} U_1 &= f_{n-1}(C_{\theta_{n-1}}(U_1, U_2)) \\ &= f_{n-1}(K^{-1}_{C_{\theta_{n-1}}}(K_{C_{\theta_{n-1}}}(C_{\theta_{n-1}}(U_1, U_2)))) \\ &= f_{n-1}(K^{-1}_{C_{\theta_{n-1}}}(A_{n-2}), \\ U_2 &= g_{n-1}(C_{\theta_{n-1}}(U_1, U_2)) \\ &= g_{n-1}(K^{-1}_{C_{\theta_{n-1}}}(K_{C_{\theta_{n-1}}}(C_{\theta_{n-1}}(U_1, U_2)))) \\ &= g_{n-1}(K^{-1}_{C_{\theta_{n-1}}}(A_{n-2}), \end{aligned}$$

and hence (U_1, U_4) and (U_2, U_4) have copula $C_{\theta_{n-3}}$.

By continuing this way, it follows that
(U_1, U_2) have copula $C_{\theta_{n-1}}$,
$(U_1, U_3), (U_2, U_3)$ have copula $C_{\theta_{n-2}}$,

$\vdots$

$(U_1, U_n), (U_2, U_n), \dots, (U_{n-1}, U_n)$ have copula C_{θ_1}.

Hence $(U_1, U_2, \dots, U_n)$ have copula $C^n(u_1, u_2, \dots, u_n; \theta_1, \dots, \theta_{n-1})$. □

Example 5.14. Suppose we want to generate random variates from a 4-dimensional distribution with standard exponential margins $F_i(x) = 1 - e^{-x}$, and a Gumbel copula $C(u_1, u_2, u_3, u_4; \theta_1, \theta_2, \theta_3)$ with $\theta_1 = 3/2$, $\theta_2 = 2$ and $\theta_3 = 5/2$, where the Gumbel copula C is a 4-variate extension given by (5.5.4). To generate a random variate from C we simply apply Algorithm 4 with

$$\begin{aligned}
&\varphi_{\theta_i}(t) = (-\ln t)^{\theta_i}, \\
&\varphi_{\theta_i}^{-1}(t) = \exp(-t^{1/\theta_i}), \\
&K_{C_{\theta_i}}(t) = t - \frac{\varphi_{\theta_i}(t)}{\varphi'_{\theta_i}(t)} = t - \frac{t \ln t}{\theta_i}.
\end{aligned}$$

$K_{C_{\theta_i}}^{-1}(t)$ is evaluated using numeric rootfinding. This gives us the random variate (u_1, u_2, u_3, u_4) from C and hence the desired random variate from the prespecified distribution is $(F_1^{-1}(u_1), \dots, F_4^{-1}(u_4)) = (-\ln(1-u_1), \dots, -\ln(1-u_4))$. Furthermore, the Kendall's tau rank correlation matrix for this distribution is given by

$$\tau = \begin{pmatrix} 1 & 1-1/\theta_3 & 1-1/\theta_2 & 1-1/\theta_1 \\ 1-1/\theta_3 & 1 & 1-1/\theta_2 & 1-1/\theta_1 \\ 1-1/\theta_2 & 1-1/\theta_2 & 1 & 1-1/\theta_1 \\ 1-1/\theta_1 & 1-1/\theta_1 & 1-1/\theta_1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 3/5 & 1/2 & 1/3 \\ 3/5 & 1 & 1/2 & 1/3 \\ 1/2 & 1/2 & 1 & 1/3 \\ 1/3 & 1/3 & 1/3 & 1 \end{pmatrix}.$$

Note that the results presented in this chapter (5.5) holds for *strict* Archimedean copulas. With some additional constraints most of the results can be generalized to hold also for non-strict Archimedean copulas. However for practical purposes it is sufficient to only consider strict Archimedean copulas. This basically means (there are exceptions such as the Frank family) that we consider copula families with only positive dependence. Furthermore, risk models are often designed to model positive dependence, since in some sense it is the "dangerous" dependence.

6 Elliptical Copulas

A spherical distribution distribution is an extension of the standard multinormal distribution $\mathcal{N}(\mathbf{0}, \mathbf{I}_n)$ and an elliptical distribution is an extension of $\mathcal{N}(\mu, \Sigma)$. Recall that $\mathcal{N}(\mu, \Sigma)$ can be defined from the standard multivariate normal distribution $\mathcal{N}(\mathbf{0}, \mathbf{I}_n)$ via

$$\mathbf{X} =_d \mu + A\mathbf{Y}, \tag{6.0.6}$$

where $\mathbf{X} \sim \mathcal{N}(\mu, \Sigma), \mathbf{Y} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_n)$ and $\Sigma = AA^T$. With this in mind it seems natural to start with spherical distributions and then define elliptical distributions from the spherical in the way indicated by (6.0.6).

Definition 16. An $n \times 1$ random vector $\mathbf{X}$ is said to have a spherically symmetric distribution (or simply spherical distribution) if for every $\Gamma \in \mathcal{O}(n)$,

$$\Gamma\mathbf{X} =_d \mathbf{X}, \tag{6.0.7}$$

where $\mathcal{O}(n)$ denotes the set of $n \times n$ orthogonal matrices.

Note that an orthogonal matrix represents a rotation transformation. Hence spherical distributions are geometrically interpreted as those invariant under rotations.

Theorem 6.1. *An n-vector $\mathbf{X}$ has a spherical distributions if and only if its characteristic function $\Psi(t)$ satisfies one of the following equivalent conditions:*

1. $\Psi(\Gamma^T\mathbf{t}) = \Psi(\mathbf{t})$ *for any* $\Gamma \in \mathcal{O}(n)$*;*
2. *There exist a function $\phi(\cdot)$ of a scalar variable such that* $\Psi(\mathbf{t}) = \phi(\mathbf{t}^T\mathbf{t})$.

For the proof, see Fang, Kotz and Ng (1987) [5].

From the theorem above we see that for spherical distributions the characteristic function $\Psi(\mathbf{t}) = \mathbb{E}(\exp(i\mathbf{t}^T\mathbf{X})) = \phi(\mathbf{t}^T\mathbf{t})$, and for this reason we will use the notation $\mathbf{X} \sim S_n(\phi)$ when $\mathbf{X}$ is spherically distributed. The function ϕ given above is called the characteristic generator of the spherical distribution.

Example 6.1. Let $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_n)$. Since the components $X_i \sim \mathcal{N}(0,1), i = 1, \ldots, n$ are independent and the characteristic function of X_i is $\exp(-t_i^2/2)$, the characteristic function of $\mathbf{X}$ is

$$\exp\{-\frac{1}{2}(t_1^2 + \ldots + t_n^2)\} = \exp\{-\frac{1}{2}\mathbf{t}^T\mathbf{t}\}.$$

From Theorem 6.1 it then follows that $\mathbf{X} \sim S_n(\phi)$ where $\phi(s) = \exp(-s/2)$.

The spherical distributions can be defined in an alternative and equivalent way. This definition also provides the basis for many random variate generation techniques.

Definition 17. An $n \times 1$ random vector $\mathbf{X}$ is said to have a spherically symmetric distribution (or simply spherical distribution) if it has the stochastic representation

$$\mathbf{X} =_d R\mathbf{U} \tag{6.0.8}$$

for some positive random variable R independent of the random vector $\mathbf{U}$ uniformly distributed on the unit hypersphere $S_{n-1} = \{\mathbf{z} \in \mathbb{R}^n | \mathbf{z}^T\mathbf{z} = 1\}$.

Spherical distributions can thus be interpretated as mixtures of uniform distributions on hyperspheres with different radius.

Example 6.2. For the standard multivariate normal distribution $R \sim \sqrt{\chi_n^2}$. For the standard multivariate t-distribution with ν degrees of freedom R satisfies $R^2/n \sim F(n,\nu)$, where $F(n,\nu)$ denotes the F-distribution with n and ν degrees of freedom.

A random vector $\mathbf{X} \sim S_n(\phi)$ does not necessarily have a density. However it seems reasonable for our purpose to only consider spherical (and later elliptical) distributions that possess densities. In this case it follows from Theorem 6.1 that the densities must be of the form $g(\mathbf{x}^T\mathbf{x})$ for some nonnegative function g of one scalar variable. Hence the contours of equal density form spheres in $\mathbb{R}^n$.

Definition 18. An $n \times 1$ random vector $\mathbf{X}$ is said to have an elliptically symmetric distribution (or simply elliptical distribution) with parameters μ and Σ if

$$\mathbf{X} =_d \mu + A\mathbf{Y}, \quad \mathbf{Y} \sim S_k(\phi) \tag{6.0.9}$$

where $A : n \times k$ and $AA^T = \Sigma$ with $\operatorname{rank}(\Sigma) = k$.

It follows that the characteristic function can be written as

$$\begin{aligned}\Psi(\mathbf{t}) &= \mathbb{E}(\exp(i\mathbf{t}^T\mathbf{X})) = \mathbb{E}(\exp(i\mathbf{t}^T(A\mathbf{Y}+\mu))) \\ &= \exp(i\mathbf{t}^T\mu)\mathbb{E}(\exp(i(A^T\mathbf{t})^T\mathbf{Y})) = \exp(i\mathbf{t}^T\mu)\phi((A^T\mathbf{t})^T(A^T\mathbf{t})) \\ &= \exp(i\mathbf{t}^T\mu)\phi(\mathbf{t}^T\Sigma\mathbf{t}).\end{aligned}$$

We write $\mathbf{X} \sim EC_n(\mu, \Sigma, \phi)$.

If $\mathbf{Y}$ has density $g(\mathbf{y}^T\mathbf{y})$ and Σ is positive definite, then $A\mathbf{Y} + \mu$ has density

$$\frac{1}{\sqrt{\det(\Sigma)}} g((\mathbf{x}-\mu)^T\Sigma^{-1}(\mathbf{x}-\mu)),$$

so the contours of equal density form ellipsoids in $\mathbb{R}^n$.

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

Given the distribution of $\mathbf{X}$, the representation $EC_n(\mu, \Sigma, \phi)$ is not unique. It uniquely determines μ but Σ and ϕ are only determined up to a positive constant, see Fang, Kotz and Ng (1987) [5]. We would like to have a representation such that $\operatorname{Cov}(\mathbf{X}) = \Sigma$. Is there such a representation?

$$\begin{aligned}\operatorname{Cov}(\mathbf{X}) &= \operatorname{Cov}(\mu + A\mathbf{Y}) = AA^T\operatorname{Cov}(\mathbf{Y}) \\ &= AA^T\mathbb{E}(R^2)\operatorname{Cov}(\mathbf{U}),\end{aligned}$$

provided $\mathbb{E}(R^2) < \infty$.

Let $\mathbf{Y} \sim \mathcal{N}_n(\mathbf{0}, \mathbf{I}_n)$. Then we have $\mathbf{Y} =_d \|\mathbf{Y}\|\mathbf{U}$, where $\|\mathbf{Y}\|$ is independent of $\mathbf{U}$. Furthermore $\|\mathbf{Y}\|^2 \sim \chi_n^2$, so $\mathbb{E}(\|\mathbf{Y}\|^2) = n$. Since $\operatorname{Cov}(\mathbf{Y}) = \mathbf{I}_n$ we get the general result:

If $\mathbf{U}$ is uniformly distributed on the unit hypersphere in $\mathbb{R}^n$, then $\operatorname{Cov}(\mathbf{U}) = \mathbf{I}_n/n$. Thus $\operatorname{Cov}(\mathbf{X}) = AA^T\mathbb{E}(R^2)\mathbf{I}_n/n$. By choosing the characteristic generator $\phi'(s) = \phi(s/c)$, where $c = \mathbb{E}(R^2)/n$, we get $\operatorname{Cov}(X) = \Sigma$. Hence an elliptical distribution is fully described by its mean, it covariance matrix and its characteristic generator, which can be chosen so that $\operatorname{Cov}(\mathbf{X}) = \Sigma$.

Theorem 6.2. *If $\mathbf{X} \sim EC_m(\mu, \Sigma, \phi)$, B is an $m \times n$ matrix of rank m $(m \le n)$ and $\mathbf{b}$ is an $m \times 1$ vector, then*

$$\mathbf{b} + B\mathbf{X} \sim EC_m(\mathbf{b} + B\mu, B\Sigma B^T, \phi).$$

This result provides a way to relate elliptical and spherical distributions. Let $\mathbf{Y} \sim EC_n(\mathbf{0}, \mathbf{I}, \phi)$ and suppose that $\Sigma = AA^T$ for some $n \times n$ matrix A. We have

$$\mathbf{X} = \mu + A\mathbf{Y} \sim EC_n(\mu, \Sigma, \phi). \tag{6.0.10}$$

Thus for random variate generation from elliptical distributions, it suffices to generate variates from spherical distributions and then apply (6.0.10).

All marginal distributions of elliptical distributions are elliptical and have the same generator. Let $\mathbf{X} \sim EC_n(\mu, \Sigma, \phi)$ and let $\mathbf{X} = \binom{\mathbf{X}_1}{\mathbf{X}_2}$, where $\mathbf{X}_1 \in \mathbb{R}^p$ and $\mathbf{X}_2 \in \mathbb{R}^q$ $(p+q=n)$. Furthermore let $\mathbb{E}(\mathbf{X}_1) = \mu_\mathbf{1}$, $\mathbb{E}(\mathbf{X}_2) = \mu_\mathbf{2}$ and $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$. Then

$$\mathbf{X_1} \sim EC_p(\mu_\mathbf{1}, \Sigma_\mathbf{11}, \phi), \mathbf{X_2} \sim EC_q(\mu_\mathbf{2}, \Sigma_\mathbf{22}, \phi).$$

We assume that Σ is strictly positive definite. The conditional distribution of $\mathbf{X_1}$ given $\mathbf{X_2}$ is also elliptical, but in general with a different generator ϕ':

$$\mathbf{X_1}|\mathbf{X_2} \sim EC_p(\mu_\mathbf{1.2}, \Sigma_\mathbf{11.2}, \phi'),$$

where $\mu_\mathbf{1.2} = \mu_\mathbf{1} + \Sigma_\mathbf{12}{\Sigma_\mathbf{22}}^{-1}(\mathbf{X_2} - \mu_\mathbf{2})$ and $\Sigma_\mathbf{11.2} = \Sigma_\mathbf{11} - \Sigma_\mathbf{12}{\Sigma_\mathbf{22}}^{-1}\Sigma_\mathbf{21}$.

Because all marginal distributions of an elliptical distribution have the same characteristic generator, it follows that an elliptical distribution is uniquely determined by its mean, covariance matrix and knowledge of its type. Thus the copula of an multivariate elliptical distribution is uniquely determined by its correlation matrix and knowledge of its type.

6.1 The Gaussian Copula

The copula of the n-variate normal distribution with linear correlation matrix ρ_l is

$$C^{Ga}_{\rho_l}(\mathbf{u}) = \Phi^n_{\rho_l}(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_n)), \tag{6.1.1}$$

where $\Phi^n_{\rho_l}$ denotes the joint distribution function of the n-variate standard normal distribution function with linear correlation matrix ρ_l and Φ^{-1} denotes the inverse of the distribution function of the univariate standard normal distribution. Copulas on the form (6.1.1) are called Gaussian copulas.
In the bivariate case the copula expression can be written

$$C^{Ga}_{\rho_l}(u,v) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi(1-\rho_l^2)^{1/2}} \left\{ -\frac{s^2 - 2\rho_l st + t^2}{2(1-\rho_l^2)} \right\} \mathrm{d}s\,\mathrm{d}t.$$

Furthermore, if $U_1, \dots, U_n$ have joint distribution function $C^{Ga}_{\rho^*}$, where ρ^* is a linear correlation matrix then $U_1, \dots, U_n$ have a Spearman's rank correlation matrix (ρ_{ij}) given by

$$\rho_{ij} = \frac{6}{\pi} \arcsin \frac{\rho^*_{ij}}{2},$$

and a Kendall's rank correlation matrix (τ_{ij}) given by

$$\tau_{ij} = \frac{2}{\pi} \arcsin \rho^*_{ij}.$$

Example 3.5 shows that Gaussian copulas do not have upper tail dependence. With similar calculations it can easily be shown that Gaussian copulas do not have lower tail dependence.

We now address the question of random variate generation from Gaussian copulas.
Suppose $\Sigma = AA^T$ for some $n \times n$ matrix A. Then if $Z_1, \dots, Z_n \sim \mathcal{N}(0,1)$

$$\mathbf{X} = \mu + A\mathbf{Z} \sim \mathcal{N}_n(\mu, \Sigma),$$

since $\mathbb{E}((\mathbf{X}-\mu)(\mathbf{X}-\mu)^T) = A\mathbb{E}(\mathbf{Z}\mathbf{Z}^T)A^T = AA^T = \Sigma$. One natural way of finding A given Σ is by using the Cholesky decomposition of Σ. The Cholesky decomposition of Σ is the unique lower-triangular matrix L with $LL^T = \Sigma$. Furthermore the Cholesky decomposition is implemented in most mathematical software. This provides an easy algorithm for random variate generation from the Gaussian n-copula, $C^{Ga}_{\rho_l}$, parametrizised with the linear correlation matrix ρ_l.

Algorithm 5.

- Find the Cholesky decomposition A of ρ_l.
- Generate n independent variates $z_1, \dots, z_n$ from $\mathcal{N}(0,1)$.
- Set $\mathbf{x} = A\mathbf{z}$.
- Set $u_i = \Phi(x_i), i = 1, \dots, n$.
- $(u_1, \dots, u_n) \sim C^{Ga}_{\rho_l}$.

As usual Φ denotes the univariate standard normal distribution function.

6.2 The t-copula

The copula of the n-variate t-distribution with ν degrees of freedom and linear correlation matrix ρ_l is

$$C^t_{\nu,\rho_l}(\mathbf{u}) = \Theta^n_{\nu,\rho_l}(t_\nu^{-1}(u_1), \dots, t_\nu^{-1}(u_n)),$$

where Θ^n_{ν,ρ_l} denotes the joint distribution function of the n-variate standard t-distribution function with ν degrees of freedom and linear correlation matrix ρ_l and t_ν^{-1} denotes the inverse of the distribution function of the univariate standard t-distribution with ν degrees of freedom.
In the bivariate case the copula expression can be written

$$C^t_{\nu,\rho_l}(u,v) = \int_{-\infty}^{t_\nu^{-1}(u)} \int_{-\infty}^{t_\nu^{-1}(v)} \frac{1}{2\pi(1-\rho_l^2)^{1/2}} \left\{1 + \frac{s^2 - 2\rho_l st + t^2}{\nu(1-\rho_l^2)}\right\}^{-(\nu+2)/2} \mathrm{d}s\,\mathrm{d}t,$$

where ρ_l is the linear correlation coefficient.

The density function of the multivariate t-distribution is with ν degrees of freedom is

$$f(\mathbf{x}) = \frac{\Gamma[(\nu+p)/2]}{(\pi\nu)^{p/2}\Gamma(\nu/2)} |\Sigma|^{-1/2} [1 + \nu^{-1}(\mathbf{x}-\mu)^T \Sigma^{-1} (\mathbf{x}-\mu)]^{-(\nu+p)/2},$$

with $\nu > 0$. This distribution can be obtained by the transformation

$$\mathbf{X} = \frac{\sqrt{\nu}}{\sqrt{S}}\mathbf{Z} + \mu, \tag{6.2.1}$$

where $\mathbf{Z} \sim \mathcal{N}_p(\mathbf{0}, \Sigma)$ and independent of S, which is χ^2_ν.

If (X_1, X_2) has a standard bivariate t-distribution with ν degrees of freedom and linear correlation ρ_l, then $X_2|X_1 = x$ is t-distributed with $v+1$ degrees of freedom and

$$\mathbb{E}(X_2|X_1 = x) = \rho_l x, \quad Var(X_2|X_1 = x) = \left(\frac{\nu + x^2}{\nu + 1}\right)(1-\rho_l^2).$$

This can be used to show that the t-copula has upper (and lower) tail dependence.

$$\begin{aligned}\lambda_U &= 2\lim_{x\to\infty}\mathbb{P}(X_2 > x|X_1 = x)\\ &= 2\lim_{x\to\infty}\overline{t}_{\nu+1}\left(\left(\frac{\nu+1}{\nu+x^2}\right)^{1/2}\frac{x-\rho_l x}{\sqrt{1-\rho_l^2}}\right)\\ &= 2\lim_{x\to\infty}\overline{t}_{\nu+1}\left(\left(\frac{\nu+1}{\nu/x^2+1}\right)^{1/2}\frac{\sqrt{1-\rho_l}}{\sqrt{1+\rho_l}}\right)\\ &= 2\overline{t}_{\nu+1}\left(\frac{\sqrt{\nu+1}\sqrt{1-\rho_l}}{\sqrt{1+\rho_l}}\right)\end{aligned}$$

From this it is also seen that the upper tail dependence is increasing in ρ_l and decreasing in ν, as one would think. Furthermore since

$$\lim_{\nu\to\infty} t_{\nu,\rho_l} = \mathcal{N}_{\rho_l},$$

the upper tail dependence tends to zero as the number of degrees of freedom tends to infinity for $\rho_l < 1$.

Coefficient of upper tail dependence for the bivariate t-copula.

$\nu\backslash\rho_l$	-0.5	0	0.5	0.9	1
2	0.06	0.18	0.39	0.72	1
4	0.01	0.08	0.25	0.63	1
10	0.00	0.01	0.08	0.46	1
∞	0	0	0	0	1

The last row represents the Gaussian copula.

Equation (6.2.1) provides an easy algorithm for random variate generation from the multivariate t-copula, C^t_{ν,ρ_l}.

Algorithm 6.

- Find the Cholesky decomposition A of ρ_l.
- Generate n independent variates $z_1, \ldots, z_n$ from $\mathcal{N}(0,1)$.
- Generate a variate s from χ^2_ν independent of $z_1, \ldots, z_n$.
- Set $\mathbf{y} = A\mathbf{z}$.
- Set $\mathbf{x} = \frac{\sqrt{\nu}}{\sqrt{s}}\mathbf{y}$.
- Set $u_i = t_\nu(x_i), i = 1, \ldots, n$.
- $(u_1, \ldots, u_n) \sim C^t_{\nu,\rho_l}$.

Here, t_ν denotes the univariate standard t-distribution function with ν degrees of freedom.

Figure 6.1 and 6.2 show bivariate distributions with Gaussian and t-copulas. All four distributions have linear correlation 0.8. For the bivariate distributions with t-copulas the plots show that extreme values have a tendency to occur together. This is due to the tail dependence of the t-copulas. It can also be seen that this is not the case for the bivariate distributions with Gaussian copulas.

The algorithms presented for the Gaussian and t-copulas are fast and easy to implement. We want to emphasize the potential usefulness of t-copulas as an alternative to Gaussian copulas. Both Gaussian and t-copulas are easily parametrizised by the linear correlation matrix, but only t-copulas give dependence structures with tail dependence.

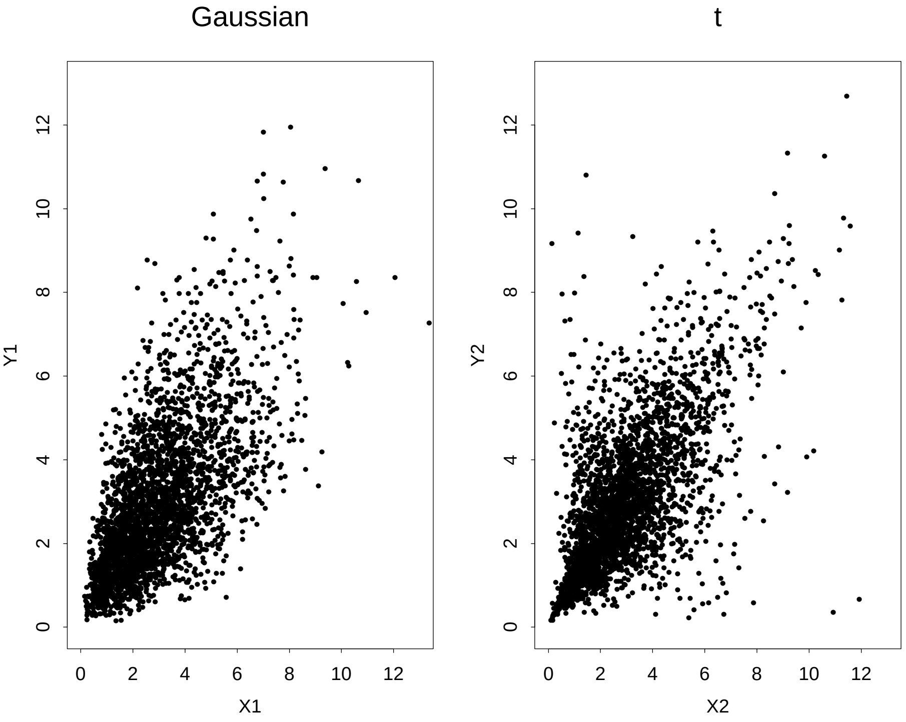

Figure 6.1: Samples from two distributions with Gamma $(3, 1)$ margins, Linear correlation 0.8 and different dependence structures. (X1,Y1) has a Gaussian copula and (X2,Y2) has a t_2-copula.

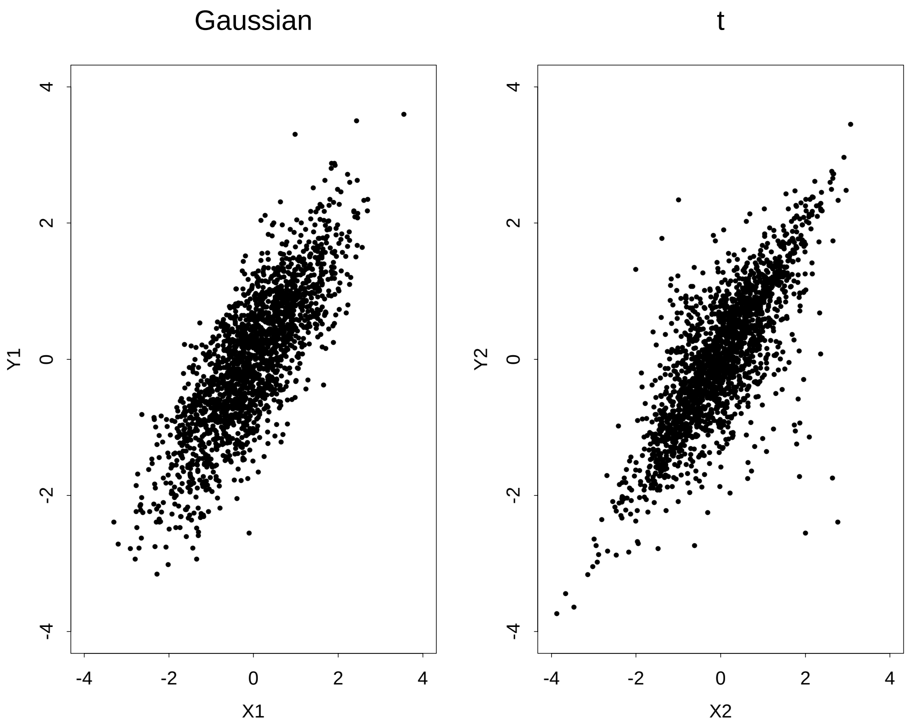

Figure 6.2: Samples from two distributions with standard normal margins, Linear correlation 0.8 and different dependence structures. (X1,Y1) has a Gaussian copula and (X2,Y2) has a t_2-copula.

7 Extreme Value Copulas

7.1 Univariate Extreme Value Theory

Let $X_1, X_2, \ldots$ be iid random variables with continuous distribution function F. Let $S_n = X_1 + \ldots + X_n, M_n = \max(X_1, \ldots, X_n)$ and $L_n = \min(X_1, \ldots, X_n)$. Furthermore let x_F, given by

$$x_F = \sup_x \{x \in \mathbb{R} : F(x) < 1\} \leq \infty$$

denote the right endpoint of F. It can be shown that $M_n \to x_F$ as $n \to \infty$. If X_i has finite mean and variance μ and σ^2 respectively, we know from the central limit theorem that

$$\mathbb{P}[(S_n - n\mu)/\sqrt{n\sigma^2} \leq x] \to \Phi(x) \text{ as } n \to \infty,$$

where Φ is the distribution function of the standard normal distribution. What about normalized maxima and minima? Consider possible limiting distributions for $(M_n - a_n)/b_n$ and $(L_n - c_n)/d_n$ as $n \to \infty$ for suitably chosen sequences $\{a_n\}, \{b_n\}, \{c_n\}, \{d_n\}$. Since

$$\min(X_1, \ldots, X_n) = -\max(-X_1, \ldots, -X_n)$$

it is sufficient to study the case of maxima. The sequences $\{a_n\}$ and $\{b_n\}$ we are interested in are such that

$$\mathbb{P}[(M_n - a_n)/b_n \leq x] = \mathbb{P}[M_n \leq a_n + b_n x] = F^n(a_n + b_n x) \to H(x),$$

as $n \to \infty$ for some non-degenerate distribution function H, i.e. $\{H(x) | x \in \mathbb{R}\} \neq \{0, 1\}$. The three possible types for H are called univariate (max) extreme value distributions and are given by:

- Fréchet

$$\phi_\alpha(x) = \begin{cases} 0, & x \leq 0 \\ \exp(-x^{-\alpha}), & x > 0 \end{cases} \quad \alpha > 0.$$

- Gumbel

$$\Lambda(x) = \exp(-e^{-x}), \quad x \in \mathbb{R}$$

- Weibull

$$\Psi_\alpha(x) = \begin{cases} \exp(-(-x)^\alpha), & x \leq 0 \\ 1, & x > 0 \end{cases} \quad \alpha > 0.$$

We say that two random variables U and V are of the same type if and only if

$$U =_d aV + b \text{ for } b > 0, a \in \mathbb{R}.$$

This means that

$$F_U(x) = F_V((x - b)/a)$$

and hence random variables of the same type have the same distribution function up to changes of scale and location. Thus in the non-degenerate case the only possible limits of $(M_n - a_n)/b_n$ are the location-scale families based on the above three distribution functions. In this case we say that F is in the maximum domain of attraction of H,

$$F \in MDA(H).$$

Example 7.1. Exponential. $F(x) = 1 - e^{-x}, x > 0$. Let $a_n = \ln n, b_n = 1$; then

$$F^n(a_n + b_n z) = (1 - e^{-\ln n - z})^n = (1 - \frac{1}{n} e^{-z})^n \to \exp(-e^{-z}),$$

$-\infty < z < \infty$. Hence $F \in MDA(\Lambda)$.

Pareto. $F(x) = 1 - x^{-1/\gamma}, x > 1, \gamma > 0$. Let $a_n = 0, b_n = n^\gamma$; then

$$F^n(a_n + b_n z) = (1 - \frac{1}{n} z^{-1/\gamma})^n \to \exp(-z^{-1/\gamma}),$$

$z > 0$. Hence $F \in MDA(\phi_{1/\gamma})$.

Beta. $F(x) = 1 - (1 - x)^{-1/\gamma}, 0 < x < 1, \gamma < 0$. Note that the right end point of F is finite. Let $a_n = 1, b_n = n^\gamma$; then

$$F^n(a_n + b_n z) = (1 - \frac{1}{n}(-z)^{-1/\gamma})^n \to \exp(-(-z)^{-1/\gamma}),$$

$z < 0$. Hence $F \in MDA(\Psi_{-1/\gamma})$.

As indicated by the above example, in the non-degenerate case the type of the limiting distribution (H) depends on how heavy tailed the distribution under consideration is.
After location/scale changes, the three types can be combined into the generalized extreme value (GEV) family

$$H(x;\gamma) = \begin{cases} \exp(-(1+\gamma x)^{-1/\gamma}), & \gamma \neq 0, \\ \exp(-e^{-x}), & \gamma = 0 \end{cases}$$

where $1 + \gamma x > 0$. It follows that

$$\begin{aligned} &H((x-1)/\gamma;\gamma) = \phi_{1/\gamma}(x), \quad \gamma > 0, \\ &H(x;0) = \Lambda(x), \quad \gamma = 0, \\ &H(-(x+1)/\gamma;\gamma) = \Psi_{-1/\gamma}(x), \quad \gamma < 0. \end{aligned}$$

Furthermore the three-parameter GEV model is given by

$$H_{\gamma,\mu,\sigma}(x) := H((x-\mu)/\sigma;\gamma).$$

Univariate extreme value theory provides many useful applications in various fields. However since we in this paper focus on modelling dependence we refer to Embrechts, Klüppelberg and Mikosch (1997) [4] and Joe (1997) [6] for further results on univariate extreme value theory.

7.2 Multivariate Extreme Value Theory

Let $(X_{i1}, \ldots, X_{im}), i = 1, 2, \ldots$ be iid m-variate random vectors with distribution function F. Let $M_{ij} = \max_{1 \le i \le n} X_{ij}, j = 1, \ldots, m$ be the componentwise maxima. Multivariate extreme value (MEV) distributions come from limits of random vectors $((M_{1n} - a_{1n})/b_{1n}, \ldots, (M_{mn} - a_{mn})/b_{mn})$. If a limiting distribution exists, then each univariate margin of this distribution is an univariate extreme value distribution (member of the GEV family). Furthermore, it can be written in the form

$$C(H(z_1;\gamma_1), \ldots, H(z_m;\gamma_m)), \tag{7.2.1}$$

where $H(z_j;\gamma_j)$ are GEV distributions and C is an m-copula. It is not surprising that the univariate margins are in the GEV family. Our interest lies rather in the

copula C given by (7.2.1).
Let $(M_{1n}, \dots, M_{mn})$ be a vector of componentwise maxima of iid random vectors $(X_{i1}, \dots, X_{im})$ from F, and suppose

$$\begin{aligned} G(\mathbf{z}) &= \lim_{n\to\infty} F^n(a_{1n} + b_{1n}z_1, \dots, a_{mn} + b_{mn}z_m) \\ &= \lim_{n\to\infty} \mathbb{P}[M_{1n} \leq a_{1n} + b_{1n}z_1, \dots, M_{mn} \leq a_{mn} + b_{mn}z_m)] \\ &= C(H(z_1;\gamma_1), \dots, H(z_m;\gamma_m)). \end{aligned}$$

Let r_j be a strictly increasing transform of X_{ij}, and let the transformed variables and maxima be denoted X^*_{ij} and M^*_{in} respectively. Furthermore suppose that $(M^*_{jn} - a^*_{jn})/b^*_{jn}$ converges in distribution as $n \to \infty$ for all j. Let

$$\begin{aligned} G(\mathbf{z}) &= \lim_{n\to\infty} \mathbb{P}[M^*_{1n} \leq a^*_{1n} + b^*_{1n}z^*_1, \dots, M^*_{mn} \leq a^*_{mn} + b^*_{mn}z^*_m)] \\ &= C^*(H(z^*_1;\gamma^*_1), \dots, H(z^*_m;\gamma^*_m)). \end{aligned}$$

It can be shown, see Joe (1997) [6], that $C = C^*$. This basically means that we can obtain MEV distributions by taking margins and constants a_{jn}, b_{jn} that lead to easier calculations. In Joe (1997) [6] it is also shown that for an m-variate MEV distribution with copula C

$$C(u_1^t, \dots, u_m^t) = C^t(u_1, \dots, u_m) \tag{7.2.2}$$

for all $t > 0$. We call copulas satisfying this condition extreme value copulas.

Example 7.2. Consider the bivariate Gumbel family of copulas given by

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

for $\theta \geq 1$. It is easy to show that these copulas satisfy (7.2.2) for extreme value copulas. Furthermore, the multivariate extension of this family discussed in chapter 5.5 also satisfies (7.2.2).

The next theorem states that for a bivariate copula with upper tail dependence, the extreme value limit has the same coefficient of upper tail dependence. Clearly the result also holds for bivariate margins of a multivariate copula.

Theorem 7.1. *Let C be a bivariate copula and let $F(x_1, x_2) = C(1-e^{-x_1}, 1-e^{-x_2})$. Suppose $\lim_{u\to 1-} \overline{C}(u,u)/(1-u) = \lambda_U$, where $\lambda_U \in (0,1]$ and $\lim_{n\to\infty} F^n(x_1 + \ln n, x_2 + \ln n) = H(x_1, x_2)$ with margins $\exp(-e^{-x_j}), j = 1,2$. Let $C^*(u_1,u_2) = H(-\ln(-\ln u_1), -\ln(-\ln u_2))$. Then $\lim_{u\to 1-} \overline{C}^*(u,u)/(1-u) = \lambda_U$.*

For the proof, see Joe (1997) [6].

Example 7.3. Consider the bivariate copula

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

for $1 \leq \theta < \infty$. This is a strict Archimedean copula with

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

It follows from Theorem 5.6 that

$$\begin{aligned}\lambda_U &= 2-2\lim_{s\to 0}\left(\frac{{\varphi^{-1}}'(2s)}{{\varphi^{-1}}'(s)}\right)\\ &= 2-2\lim_{s\to 0}\left(\frac{(1-e^{-2s})^{1/\theta-1}e^{-2s}}{(1-e^{-2s})^{1/\theta-1}e^{-2s}}\right)\\ &= \{e^x = 1+x+O(x^2)\}\\ &= 2-2\lim_{s\to 0}\left(\frac{(-2s+O(s^2))^{1/\theta-1}}{(-s+O(s^2))^{1/\theta-1}}\right)\\ &= 2-22^{1/\theta-1} = 2-2^{1/\theta}.\end{aligned}$$

The upper extreme value limit of C_θ is the Gumbel family with coefficient of upper tail dependence $2-2^{1/\theta}$ as shown in Example 3.4.

Three one-parameter bivariate families of extreme value copulas are

$$C_\theta(u,v) = \exp(-[(-\ln u)^\theta + (-\ln v)^\theta]^{1/\theta}), \tag{7.2.3}$$

$$C_\theta(u,v) = uv\exp((-\ln u)^{-\theta} + (-\ln v)^{-\theta})^{-1/\theta}), \tag{7.2.4}$$

$$C_\theta(u,v) = \exp(\ln u\Phi(\frac{1}{\theta}+\frac{\theta}{2}\ln(\frac{\ln u}{\ln v})) + \ln v\Phi(\frac{1}{\theta}+\frac{\theta}{2}\ln(\frac{\ln v}{\ln u}))). \tag{7.2.5}$$

for $\theta \geq 1$, $\theta \geq 0$ and $\theta \geq 0$ respectively. (7.2.3) is the Gumbel family, (7.2.4) is the Galambo family and (7.2.5) is the Hüsler and Reiss family. All three have upper tail dependence and $C_\theta = \Pi$ for $\theta = 1, 0$ and 0 for (7.2.3), (7.2.4) and (7.2.5) respectively. Furthermore $C_\infty = M$ for all three. Multivariate extensions of these bivariate extreme value copulas are discussed in Joe (1997) [6]. One of these extensions is the multivariate extension of the Gumbel family presented in chapter 5.5.

Consider a multivariate distribution function F. Finding the multivariate limiting distribution is often much more difficult than in the univariate case since it means finding sequences $\{a_{jn}\}$, $\{b_{jn}\}$ for $j = 1,\ldots,m$ such that $((M_{1n}-a_{1n})/b_{1n},\ldots,(M_{mn}-a_{mn})/b_{mn})$ converges in distribution. Results concerning domains of attraction of multivariate extreme value distributions can be found in Marshal and Olkin (1983) [11]. One example is the following theorem.

Theorem 7.2. *Let G be a k-dimensional (max) extreme value distribution such that G_i is of type Λ for $i = 1,\ldots,k$. Let $\phi_i(t) = \overline{F_i}^{-1}(\overline{F_1}(t))$, $i = 2,\ldots,k$, $r_i(t) = F_i'(t)/\overline{F_i}(t)$ and $x_1^0 = \sup\{x : F(x) < 1\}$. Then $F \in MDA(G)$ if*

$$\lim_{t\to x_1^0}\frac{1-F(\frac{x_1}{r_1(t)}+t, \frac{x_2}{r_2(\phi_2(t))}+\phi_2(t),\ldots,\frac{x_k}{r_k(\phi_k(t))}+\phi_k(t))}{1-F_1(t)} = -\ln G(\mathbf{x}) \tag{7.2.6}$$

for all $\mathbf{x}$ such that $G(\mathbf{x}) > 0$.

An application of this result is shown in the following example.

Example 7.4. Consider the bivariate dependence model

$$X_1 = \min(Z_1, Z_{12}), X_2 = \min(Z_2, Z_{12}),$$

where Z_1, Z_2 and Z_{12} are independent exponentially distributed random variables with parameters λ_1, λ_2 and λ_{12} respectively. Let X_1 and X_2 have joint distribution function F, where F has margins F_1 and F_2.

$$\begin{aligned}\overline{F}(x_1,x_2) &= \mathbb{P}[X_1 > x_1, X_2 > x_2]\\ &= \mathbb{P}[Z_1 > x_1]\mathbb{P}[Z_2 > x_2]\mathbb{P}[Z_{12} > \max(x_1,x_2)]\end{aligned}$$

and hence $F_1(x) = 1 - \exp(-(\lambda_1 + \lambda_{12})x)$ and $F_2(x) = 1 - \exp(-(\lambda_2 + \lambda_{12})x)$. It follows that $F_1, F_2 \in MDA(\Lambda)$. With the notation used in Theorem 7.2 we have

$$\begin{aligned}
&r_1(t) = F_1'(t)/\overline{F_1}(t) = (\lambda_1 + \lambda_{12}) \exp(-(\lambda_1 + \lambda_{12})t)/\exp(-(\lambda_1 + \lambda_{12})t) \\
&\quad = \lambda_1 + \lambda_{12}, \\
&r_2(t) = \ldots = \lambda_2 + \lambda_{12}, \\
&\phi_2(t) = \overline{F_2}^{-1}(\overline{F_1}(t)) = -\tfrac{1}{\lambda_2+\lambda_{12}} \ln e^{-(\lambda_1+\lambda_{12})t} = \tfrac{\lambda_1+\lambda_{12}}{\lambda_2+\lambda_{12}} t, \\
&x_1^0 = \sup\{x : F_1(x) < 1\} = \infty.
\end{aligned}$$

Furthermore

$$\begin{aligned}
&F(x_1, x_2) = \overline{F}(x_1, x_2) + F_1(x_1) + F_2(x_2) - 1, \\
&1 - F(x_1, x_2) = \overline{F_1}(x_1) + \overline{F_2}(x_2) - \overline{F}(x_1, x_2) \\
&\quad = \{\overline{F}(x_1, x_2) = \overline{F_1}(x_1)\overline{F_1}(x_1) \exp(\lambda_{12} \min(x_1, x_2))\} \\
&\quad = \overline{F_1}(x_1) + \overline{F_2}(x_2) - \overline{F_1}(x_1)\overline{F_2}(x_2) \exp(\lambda_{12} \min(x_1, x_2)).
\end{aligned}$$

Hence

$$\frac{1 - F\left(\frac{x_1+(\lambda_1+\lambda_{12})t}{\lambda_1+\lambda_{12}}, \frac{x_2+(\lambda_1+\lambda_{12})t}{\lambda_2+\lambda_{12}}\right)}{1 - F_1(x_1)} = e^{-x_1} + e^{-x_2} - e^{-x_1-x_2} h(t),$$

where

$$h(t) = \exp(\min(-\lambda_1 t + \frac{\lambda_{12} x_1}{\lambda_1 + \lambda_{12}}, -\lambda_1 t - \lambda_{12} t + \frac{\lambda_{12} x_2}{\lambda_2 + \lambda_{12}} + \lambda_{12} \frac{\lambda_1 + \lambda_{12} t}{\lambda_2 + \lambda_{12}} t)).$$

Hence the right side of equation (7.2.6) is

$$e^{-x_1} + e^{-x_2} - e^{-x_1-x_2} \lim_{t\to\infty} h(t).$$

Consider the case when $\max(\lambda_1, \lambda_2) > 0$. Then $\lim_{t\to\infty} h(t) = 0$. From Theorem 7.2 it follows that

$$\begin{aligned}
-\ln G(x_1, x_2) &= e^{-x_1} + e^{-x_2} \quad \Longrightarrow \\
G(x_1, x_2) &= \exp(-e^{-x_1} - e^{-x_2}) = \Pi(e^{-e^{-x_1}}, e^{-e^{-x_2}}).
\end{aligned}$$

Consider the case when $\lambda_1 = \lambda_2 = 0, \lambda_{12} > 0$. Then $\lim_{t\to\infty} h(t) = \exp(\min(x_1, x_2))$. From Theorem 7.2 it follows that

$$\begin{aligned}
-\ln G(x_1, x_2) &= e^{-x_1} + e^{-x_2} - e^{-x_1-x_2+\min(x_1,x_2)} \\
&= e^{-x_1} + e^{-x_2} - e^{-\max(x_1,x_2)} \\
&= e^{-x_1} + e^{-x_2} - e^{\min(-x_1,-x_2)} \\
&= e^{-x_1} + e^{-x_2} - \min(e^{-x_1}, e^{-x_2}) \\
&= \max(e^{-x_1}, e^{-x_2}) \quad \Longrightarrow \\
G(x_1, x_2) &= \exp(-\max(e^{-x_1}, e^{-x_2})) \\
&= \exp(\min(-e^{-x_1}, -e^{-x_2})) \\
&= \min(e^{-e^{-x_1}}, e^{-e^{-x_2}}) \\
&= M(e^{-e^{-x_1}}, e^{-e^{-x_2}}).
\end{aligned}$$

Thus $F \in MDA(G)$ where G is given by

$$G(x_1, x_2) = \begin{cases} \Pi(e^{-e^{-x_1}}, e^{-e^{-x_2}}) & \max(\lambda_1, \lambda_2) > 0 \\ M(e^{-e^{-x_1}}, e^{-e^{-x_2}}) & \lambda_1 = \lambda_2 = 0, \lambda_{12} > 0 \end{cases}$$

This shows that bivariate vectors with a Marshall-Olkin copula gives perfect positive dependence or independence in the limit for componentwise maxima, depending on the parameterization of the copula.

8 Mixture of Extremal Distributions

Theorem 8.1. *Let X, Y and Z be random variables with joint distribution function H and continuous margins F_1, F_2 and F_3 respectively.*

1. *If (X, Y) and (Y, Z) are comonotonic then (X, Z) is also comonotonic and $H(x, y, z) = \min\{F_1(x), F_2(y), F_3(z)\}$.*
2. *If (X, Y) is comonotonic and (Y, Z) is countermonotonic then (X, Z) is countermonotonic and $H(x, y, z) = \max\{0, \min\{F_1(x), F_2(y)\} + F_3(z) - 1\}$.*
3. *If (X, Y) and (Y, Z) are countermonotonic then (X, Z) is comonotonic and $H(x, y, z) = \max\{0, \min\{F_1(x), F_3(z)\} + F_2(y) - 1\}$.*

For the proof, see Embrechts, McNeil and Straumann (1999) [3].

Let $\rho_{\max}$ and $\rho_{\min}$ denote the maximum and minimum attainable correlations between two random variables with given univariate distribution functions.

Theorem 8.2. *Let $F_1, \dots, F_n$, $\quad n \geq 3$, be continuous distribution functions and let ρ be a proper correlation matrix satisfying the following conditions for all $i \neq j$, $i \neq k$ and $j \neq k$:*

1. *$\rho_{ij} \in \{\rho_{\min}(F_i, F_j), \rho_{\max}(F_i, F_j)\}$,*
2. *If $\rho_{ij} = \rho_{\max}(F_i, F_j)$ and $\rho_{ik} = \rho_{\max}(F_i, F_k)$ then $\rho_{jk} = \rho_{\max}(F_j, F_k)$,*
3. *If $\rho_{ij} = \rho_{\max}(F_i, F_j)$ and $\rho_{ik} = \rho_{\min}(F_i, F_k)$ then $\rho_{jk} = \rho_{\min}(F_j, F_k)$,*
4. *If $\rho_{ij} = \rho_{\min}(F_i, F_j)$ and $\rho_{ik} = \rho_{\min}(F_i, F_k)$ then $\rho_{jk} = \rho_{\max}(F_j, F_k)$.*

Then there exists a unique distribution with margins $F_1, \dots, F_n$ and correlation matrix ρ. This distribution is known as an extremal distribution. In $\mathbb{R}^n$ there are 2^{n-1} possible extremal distributions.

Proof. Without loss of generality suppose

$$\rho_{1j} = \begin{cases} \rho_{\max}(F_1, F_j), & 2 \leq j \leq m \leq n, \\ \rho_{\min}(F_1, F_j), & m < j \leq n, \end{cases}$$

for some $2 \leq m \leq n$. The pairwise relationship of any two margins is determined by their pairwise relationship to the first margin. The margins for which ρ_{1j} takes a maximal value form an equivalence class, as do the margins for which ρ_{1j} takes a minimal value. Hence $(X_1, \dots, X_m)$ must be pairwise comonotonic and $(X_{m+1}, \dots, X_n)$ must be pairwise comonotonic, but (X_k, X_l) where $2 \leq k \leq m$ and $m < l \leq n$ must be countermonotonic. Thus for $m < n$ the random vector

$$(F_1^{-1}(U), \dots, F_m^{-1}(U), F_{m+1}^{-1}(U), \dots, F_n^{-1}(U))$$

has the required joint distribution, where $U \sim U(0, 1)$. Furthermore, for $m < n$

$$F(x_1, \dots, x_n) = \max\{0, \min_{1 \leq i \leq m} F_i(x_i) + \min_{m < i \leq n} F_i(x_i) - 1\},$$

and for $m = n$

$$F(x_1, \dots, x_n) = \min_{1 \leq i \leq n} F_i(x_i).$$

This also shows the uniqueness of distributions with pairwise extremal correlations. Since X_1 and X_i for $2 \leq i \leq n$ are either comonotonic or countermonotonic there are 2^{n-1} possible extremal distributions. □

If ρ denotes a measure of association such as Kendall's tau or Spearman's rho then all correlation values in the interval $[-1, 1]$ can be obtained by a suitable choice of the copula. This is however not the case when ρ denotes linear correlation as shown in Example 3.3. With this in mind we will in this chapter use rank correlations as measure of association. By only allowing rank correlations we also avoid having to make certain assumptions about the margins and can present most results only in terms of copulas. For the more general case we refer to Tiit (1996) [17].

When rank correlation is chosen as measure of association the above theorem can be restated in the following form.

Corollary 8.1. *Let* $F_1, \dots, F_n$, $n \geq 3$, *be continuous distribution functions and let* ρ *be a proper rank correlation matrix satisfying the following conditions for all* $i \neq j$, $i \neq k$ *and* $j \neq k$*:*

1. $\rho_{ij} \in \{-1, 1\}$,
2. *If* $\rho_{ij} = 1$ *and* $\rho_{ik} = 1$ *then* $\rho_{jk} = 1$,
3. *If* $\rho_{ij} = 1$ *and* $\rho_{ik} = -1$ *then* $\rho_{jk} = -1$,
4. *If* $\rho_{ij} = -1$ *and* $\rho_{ik} = -1$ *then* $\rho_{jk} = 1$.

Then there exists a unique distribution with margins $F_1, \dots, F_n$ *and rank correlation matrix* ρ*. This distribution is known as an extremal distribution. In* $\mathbb{R}^n$ *there are* 2^{n-1} *possible extremal distributions.*

Since the pairwise rank correlations of the continuous random variables $X_1, \dots, X_n$ are invariant under strictly increasing transforms of $X_1, \dots, X_n$ we can without loss of generality consider the uniform $(0, 1)$ random variables $U_1 \dots U_n$ instead. To get back to the situation presented in the theorem above we need only apply the strictly increasing transforms $F_1^{-1}, \dots, F_n^{-1}$.

Definition 19. Let $J^0 = \{1, \dots, n\}$ be a given index set, and let $J = \{i_1, \dots, i_q\}$ be an arbitrary subset of J^0, satisfying the condition

$$1 \in J.$$

Let $J^c = J^0 \backslash J = \{j_1, \dots, j_s\}$, where $s = n - q$. Then the pair (J, J^c) defines a partition of the index set J^0.

Theorem 8.3. *Let* (J, J^c) *be a given non-trivial partition of the index set* $\{1, \dots, n\}$*. Then the function* C^{J,J^c} *from* $\boldsymbol{I}^n$ *to* $\boldsymbol{I}$*, defined via*

$$C^{J,J^c}(u_1, \dots, u_n) = \max(0, (\min_{i \in J} u_i + \min_{j \in J^c} u_j - 1)),$$

is an n*-copula. If* $J = \{1, \dots, n\}$*, then*

$$C^{J,J^c}(u_1, \dots, u_n) = \min(u_1, \dots, u_n).$$

C^{J,J^c} *is said to be extremal.*

Proof. That C^{J,J^c} satisfies the boundary conditions for a copula is obvious. Furthermore, for all n-boxes B with vertices $(a_1, \dots, a_n), (b_1, \dots, b_n) \in \mathbf{I}^n$ with $a_k \leq b_k$ we have

$$\begin{aligned} V_{C^{J,J^c}}(B) &= \triangle_{a_n}^{b_n} \triangle_{a_{n-1}}^{b_{n-1}} \dots \triangle_{a_1}^{b_1} C^{J,J^c}(\mathbf{t}), \\ &= V_W(B'), \end{aligned}$$

where B' is the rectangle $[\min_{i \in J} a_i, \min_{j \in J^c} a_j] \times [\min_{i \in J} b_i, \min_{j \in J^c} b_j]$. But since $B' \subset \mathbf{I}^2$ and W is a copula $V_W(B') \geq 0$. Thus C^{J,J^c} is an n-copula. □

The copula C^{J,J^c} is the copula of random variables $X_1,\dots,X_n$ such that those in $\{X_i|i\in J\}$ are pairwise comonotonic and those in $\{X_i|i\in J^c\}$ are pairwise comonotonic but two random variables taken from different groups are countermonotonic. If $J=\{1,\dots,q\}$ then C^{J,J^c} is the copula of the random vector

$$(F_1^{-1}(U),\dots,F_q^{-1}(U),F_{q+1}^{-1}(1-U),\dots,F_n^{-1}(1-U)).$$

To clarify the results obtained so far we begin with presenting the 2^{n-1} random vectors in $\mathbf{I}^n$ having extremal distributions and their corresponding extremal rank correlation matrices for $n=3$.

The distinct random vectors (U_1,U_2,U_3) having extremal distributions, U is a uniform $(0,1)$ random variable:

$$(U,U,U),(U,U,1-U),(U,1-U,U),(U,1-U,1-U).$$

Their corresponding extremal matrices in the above order:

$$\begin{pmatrix}1&1&1\\1&1&1\\1&1&1\end{pmatrix},\begin{pmatrix}1&1&-1\\1&1&-1\\-1&-1&1\end{pmatrix},\begin{pmatrix}1&-1&1\\-1&1&-1\\1&-1&1\end{pmatrix},\begin{pmatrix}1&-1&-1\\-1&1&1\\-1&1&1\end{pmatrix}.$$

The extension to higher dimensions is obvious from the above random vectors and their corresponding rank correlation matrices. Some properties of extremal copulas are the following:

1. An extremal copula is singular.
2. The support of each extremal copula is one of the diagonals of $\mathbf{I}^n$.
3. Any k-marginal C^A of the extremal copula C^{J,J^c}, defined by the index set $A=\{a_1,\dots,a_k\}$, is the the copula M^k if $A\subset J$ or $A\subset J^c$.
4. The rank correlation matrix ρ^{J,J^c} of the extremal copula C^{J,J^c} is defined in the following way:

$$\rho_{ij}^{J,J^c}=\begin{cases}1 & \text{if } i\in J \text{ and } j\in J \text{ or } i\in J^c \text{ and } j\in J^c,\\ -1 & \text{if } i\in J \text{ and } j\in J^c \text{ or } i\in J^c \text{ and } j\in J.\end{cases}$$

A correlation matrix consisting of maximal and minimal correlations is said to be an extremal correlation matrix.

The extremal copulas and the extremal rank correlation matrices depend only on the partition (J,J^c). The connections between them is as stated in the following theorem.

Theorem 8.4. *The set of extremal copulas and the set of extremal rank correlation matrices are in one-to-one correspondence.*

Proof. From 3. and 4. above follows that the rank correlation matrix of an extremal copula is extremal. To establish the opposite correspondence, we show that a given rank correlation matrix ρ^* uniquely defines a partition (J,J^c) of the index set J^0. Given ρ^* it is easy to recover the partition (J,J^c) in the following way:

$$i\in J,\ \text{if } \rho^*_{1i}=1, i\in J^c,\ \text{if } \rho^*_{1i}=-1, i=1,\dots,n.$$

It is obvious that the partition given by this procedure is unique. □

Let $G_j^n, j = 1, \dots, 2^{n-1}$ denote the extremal distributions with margins $F_1, \dots, F_n$ and rank correlation matrices ρ_j^n. Convex combinations

$$G_\lambda^n = \sum_{j=1}^{2^{n-1}} \lambda_j G_j^n, \lambda_j \geq 0, \sum_{j=1}^{2^{n-1}} \lambda_j = 1, \tag{8.0.7}$$

have the same marginals and rank correlation matrix given by $\rho_\lambda^n = \sum_{j=1}^{2^{n-1}} \lambda_j \rho_j^n$. The subscript λ is to indicate that the chosen coefficients in the convex combination is relevant.
The problem we will study in this chapter is the following:

> Given marginal distributions $F_1, \dots, F_n$ and an $n \times n$ rank correlation matrix ρ, are there coefficients λ_j such that G_λ^n given above has those margins and rank correlation matrix?

Since all choices of coefficients λ_j in the convex combination will result in G_λ^n having the prescribed margins, the problem only lies in whether there exists λ_j:s such that ρ has a convex decomposition in the class of extremal rank correlation matrices $\rho_j, j = 1, \dots, 2^{n-1}$. Thus the problem can be restated:

> Given an arbitrary $n \times n$ rank correlation matrix ρ, is there an n-vector λ such that

$$\rho = \sum_{j=1}^{2^{n-1}} \lambda_j \rho_j^n, \lambda_j \geq 0, \sum_{j=1}^{2^{n-1}} \lambda_j = 1 \quad ?$$

This can be formulated in an alternative way:

> Given an arbitrary $n \times n$ rank correlation matrix ρ, is there a non-negative solution to the equation system

$$\sum_{j=1}^{2^{n-1}} \lambda_j \rho_j^n - \rho = 0, \sum_{j=1}^{2^{n-1}} \lambda_j - 1 = 0 \quad ? \tag{8.0.8}$$

However, there exist rank correlation matrices that are not decomposable in the class of extremal rank correlation matrices. And if such a decomposition exists, it is in general not unique. Examples of such matrices are easy to find.

Example 8.1. Consider the linear correlation matrix ρ given by

$$\rho = \begin{pmatrix} 1 & 0.3 & 0.2 & 0.5 \\ 0.3 & 1 & 0.4 & 0.7 \\ 0.2 & 0.4 & 1 & 0.8 \\ 0.5 & 0.7 & 0.8 & 1 \end{pmatrix}.$$

It is easily verified that ρ is positive-definite, and hence a proper linear correlation matrix. By applying the transformation

$$f : \rho_{ij} \mapsto \frac{6}{\pi} \arcsin \frac{\rho_{ij}}{2},$$

to all elements in ρ, $f(\rho)$ is a proper Spearman's rank correlation matrix for some Gaussian copula. We now look for a nonnegative solution to the equation system

(8.0.8); that is, we look for a convex-extremal decomposition of $f(\rho)$. Standard techniques for solving linear equation systems then gives

$$\begin{aligned}
\lambda_1 + \lambda_8 &= 0.713\ldots, \\
\lambda_2 - \lambda_8 &= -0.247\ldots, \\
\lambda_3 - \lambda_8 &= -0.100\ldots, \\
\lambda_4 + \lambda_8 &= 0.278\ldots, \\
\lambda_5 - \lambda_8 &= -0.098\ldots, \\
\lambda_6 + \lambda_8 &= 0.228\ldots, \\
\lambda_7 + \lambda_8 &= 0.226\ldots.
\end{aligned}$$

It is obvious that there is no $\lambda_8 \in [0,1]$ such that $\lambda_i \in [0,1]$ for $i = 1, \ldots, 7$. Thus $f(\rho)$ is not decomposable in the class of extremal rank correlation matrices.

Hence the question above is somewhat ill-posed. The problem is rather finding sufficient conditions on a rank correlation matrix ρ so that it has a convex decomposition.

This equation system consists of 2^{n-1} variables and $n(n-1)/2+1$ equations. For $n = 3$ it is quite easy to find sufficient conditions for the equation system above to have a nonnegative solution in terms of the three upper (lower) diagonal elements in ρ. For higher dimensions it looks somewhat harder.

If a convex-extremal decomposition of a rank correlation matrix exists, then there exists a decomposition not having more than $n(n-1)/2+1$ terms.

Even if a certain rank correlation matrix has a convex decomposition in the class of extremal rank correlation matrices it is not obvious how to find it for higher dimensions. When $n = 30$ solving the equation system (8.0.8) means solving an equation system with 436 equations and 536870912 variables.

Assume now that we have found a decomposition of the given rank correlation matrix ρ in terms of the coefficients λ_j in the convex combination. Then G_λ^n given by (8.0.7) is a mixture of extremal distributions having the given marginals $F_1, \ldots, F_n$ and rank correlation matrix ρ. Random variate generation from the distribution G_λ^n presents no problem.

The order of the extremal distributions G_j^n is as follows (presented for $n = 3$). $(F_1^{-1}(U), F_2^{-1}(U), F_3^{-1}(U)) \sim G_1^3$, $(F_1^{-1}(U), F_2^{-1}(U), F_3^{-1}(1-U)) \sim G_2^3$, $(F_1^{-1}(U), F_2^{-1}(1-U), F_3^{-1}(U)) \sim G_3^3$, $(F_1^{-1}(U), F_2^{-1}(1-U), F_3^{-1}(1-U)) \sim G_4^3$. For the general case note the analogue of the representation of $j-1$ as a binary number ($U \sim 0$ and $1-U \sim 1$).

Algorithm 7.

- Generate independent uniform $(0,1)$ variates q and u.
- Choose the index j such that

$$j = \min\{i \geq 1 | \sum_{k=1}^{i} \lambda_k \geq q\}.$$

- The desired random variate from G_λ^n is:
  $(F_1^{-1}(a_1^j(u)), \ldots, F_n^{-1}(a_n^j(u)))$, where

$$(a_1^j(x), \ldots, a_n^j(x)) = (x, \ldots, x) - (b_1, \ldots, b_n)(2x-1).$$

  $(b_1, \ldots, b_n)$ is the binary representation of $j-1$, where b_n is LSB and b_1 is MSB.

To clarify the last step:

$$\begin{aligned}
(a_1^1(x), a_2^1(x), a_3^1(x)) &= (x, x, x), \\
(a_1^2(x), a_2^2(x), a_3^2(x)) &= (x, x, 1-x), \\
(a_1^3(x), a_2^3(x), a_3^3(x)) &= (x, 1-x, x), \\
(a_1^4(x), a_2^4(x), a_3^4(x)) &= (x, 1-x, 1-x).
\end{aligned}$$

The algorithm can be explained in the following way: Pick one of the 2^{n-1} extremal copulas, where the kth extremal copula is picked with probability λ_k. Generate a random variate from this copula.
To clarify the results obtained so far consider the following example.

Example 8.2. Consider the problem of simulating from an extremal copula with rank correlation matrix, ρ, given by

$$\begin{pmatrix} 1 & 0.3 & 0.2 \\ 0.3 & 1 & 0.4 \\ 0.2 & 0.4 & 1 \end{pmatrix}.$$

First of all we need to find a convex decomposition of ρ in the class of extremal rank correlation matrices. That is, we need to find an nonnegative solution to the equation system

$$\sum_{j=1}^{4} \lambda_j \rho_j^3 - \rho = 0,$$

$$\sum_{j=1}^{4} \lambda_j - 1 = 0,$$

where

$$\rho_1^3 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \rho_2^3 = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix},$$

$$\rho_3^3 = \begin{pmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix}, \rho_4^3 = \begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & 1 \\ -1 & 1 & 1 \end{pmatrix}.$$

This equation system can be written

$$\begin{pmatrix} 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{pmatrix} = \begin{pmatrix} 0.3 \\ 0.2 \\ 0.4 \\ 1 \end{pmatrix}.$$

Solving the equation system yields

$$\lambda_1 = 0.425, \lambda_2 = 0.225, \lambda_3 = 0.175, \lambda_4 = 0.175.$$

From Theorem 8.3 it is clear that the resulting copula, C, is given by

$$\begin{aligned}
C(u_1, u_2, u_3) = \; & 0.425 \min(u_1, u_2, u_3) + \\
& 0.225 \max(\min(u_1, u_2) + u_3 - 1, 0) + \\
& 0.175 \max(\min(u_1, u_3) + u_2 - 1, 0) + \\
& 0.175 \max(u_1 + \min(u_2, u_3) - 1, 0).
\end{aligned}$$

To simulate from this copula we simply apply Algorithm 7. To clarify we show an example of how this algorithm produces one random variate.

- Generate independent uniform $(0,1)$ variates q and u.
  Assume that this gives $q = 0.6$ and $u = 0.2$.
- Choose the index j such that

  $$j = \min\{i \geq 1 | \sum_{k=1}^{i} \lambda_k \geq q\}.$$

  Since $\lambda_1 < q$ and $\lambda_1 + \lambda_2 > q$ we get $j = 2$.
- The desired random variate from C is:
  $(a_1^2(u), (a_2^2(u), (a_3^2(u))$, where

  $$\begin{aligned}(a_1^2(u), a_2^2(u), (a_3^2(u)) &= (u, u, u) - (0, 0, 1)(2u - 1) \\ &= (u, u, 1 - u) = (0.2, 0.2, 0.8).\end{aligned}$$

  Note the similarity with the binary representation of $j - 1 = 1$ ($u \sim 0$, $1 - u \sim 1$).

When many random variates are produced according to the algorithm, the rank correlation matrix of the data is very close to ρ.

Example 8.3. Consider the following problem: We are given margins $F_1, \dots, F_4$ and rank correlation matrix ρ given by

$$\begin{pmatrix} 1 & 0.1 & -0.1 & -1 \\ 0.1 & 1 & 0.7 & -0.1 \\ -0.1 & 0.7 & 1 & 0.1 \\ -1 & -0.1 & 0.1 & 1 \end{pmatrix}.$$

This is a positive-semi-definite matrix (however not positive-definite since $\det(\rho) = 0$) and hence a proper rank correlation matrix. We want to find a distribution with this rank correlation matrix, which belongs to the family of mixtures of extremal distributions obtained from convex combinations of extremal distributions with margins $F_1, \dots, F_4$. In other words we want to find a convex decomposition of ρ in the eight extremal rank correlation matrices. Simply solving the resulting equation system

$$\begin{pmatrix} 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & 0.1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 & -0.1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 & 0.7 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & -0.1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 0.1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix},$$

given by (8.0.8) and only allowing nonnegative solutions yields

$$\lambda_1 = 0, \lambda_2 = 0.425, \lambda_3 = 0, \lambda_4 = 0.125, \lambda_5 = 0, \lambda_6 = 0.25, \lambda_7 = 0, \lambda_8 = 0425,$$

as the only solution (note that in general the decomposition is not unique). From Theorem 8.3 it is clear that the resulting copula, C, is given by

$$\begin{aligned} C(u_1, u_2, u_3, u_4) \quad = \quad & 0.425 \max(\min(u_1, u_2, u_3) + u_4 - 1, 0) + \\ & 0.125 \max(\min(u_1, u_2) + \min(u_3, u_4) - 1, 0) + \\ & 0.25 \max(\min(u_1, u_3) + \min(u_2, u_4) - 1, 0) + \\ & 0.425 \max(u_1 + \min(u_2, u_3, u_4) - 1, 0). \end{aligned}$$

It then follows that the desired distribution is given by

$$F(x_1, x_2, x_3, x_4) = C(F_1(x_1), F_2(x_2), F_3(x_3), F_4(x_4)).$$

Random variate generation from the this distribution, a mixture of extremal distributions, is then straight forward according to the above algorithm.

Extremal distributions are not only of theoretical interest as shown in the following example.

Example 8.4. Consider prices on put and call options, with different exercise prices and a fixed time horizon T, on some underlying stock. At time $t < T$ it seems reasonable to model pairs of prices on one put and one call as countermonotonic and pairs of prices on only puts or only calls as comonotonic if the exercise prices are not too obscure and if $T - t$ is not too small.

9 Modelling Extremal Events in Practice

9.1 Pricing Risky Insurance Contracts

Consider a portfolio consisting of n risks $X_1, \dots, X_n$ which are positive. Let the risks represent potential losses in dependent lines of business for an insurance company. Suppose that losses not greater than $k_1, \dots, k_n$ can be accepted, and that the insurance company wants to buy protection against the situation where l or more of the losses simultaneously exceed their respective thresholds $k_1, \dots, k_n$. Suppose that a reinsurance company offers to sell a contract which makes a payout only in the case that at least l of the n risks exceeds their thresholds. To begin with we will look at the probability that a payout is made. Later we will look at the expected size of the payout under the assumption that losses which exceed their thresholds are paid in full.
Assume historical data are available allowing estimation of

- marginal distributions;
- pairwise *rank* correlations.

Let

$$N = \left|\left\{i \in \{1, \dots, n\} | X_i > k_i\right\}\right|$$

be the number of losses exceeding their thresholds and let

$$L_l = \mathbf{1}_{\{N \geq l\}} \sum_{i=1}^{n} \left(X_i \mathbf{1}_{\{X_i > k_i\}}\right)$$

be the loss to the reinsurer.

Suppose that the contract the insurance company is interested in is the one for which the losses are paid in full when all losses exceed their thresholds. If the insurance company holds this contract it has full protection against simultaneous big losses in all lines of business no matter how big the total loss is. The probability of payout is then given by

$$\mathbb{P}[N = n] = \overline{H}(k_1, \dots, k_n).$$

Unfortunately a good estimation of the joint distribution, H, is not realistic due to lack of data and theoretical difficulties. However, data from the previous year(s) enable estimation of the Kendall's tau rank correlation matrix via the sample version. Furthermore, estimation of univariate marginal distributions is well understood. Hence we can assume that rank correlations and margins are given. Since

$$\overline{H}(x_1, \dots, x_n) = \hat{C}(\overline{F_1}(x_1), \dots, \overline{F_n}(x_n)),$$

the payout probability can not be calculated without choosing a suitable copula representing the dependence structure among the n risks. The standard approach might be to choose a Gaussian copula, given by

$$C^{Ga}_{\rho_l}(\mathbf{u}) = \Phi^n_{\rho_l}(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_n)),$$

where $\Phi^n_{\rho_l}$ denotes the joint distribution function of the n-variate standard normal distribution function with linear correlation matrix ρ_l and Φ^{-1} denotes the inverse of the distribution function of the univariate standard normal distribution. To parametrizise the Gaussian copula from the estimated Kendall's tau rank correlation matrix the relation

$$\rho_l = \sin(\pi\tau/2)$$

is used, where τ stands for Kendall's tau.

However, this choice of copula might prove dangerous. Since the value of the contract depends on simultaneous exceedences, a copula such as the Gaussian would result in a price too low if the true dependence structure has the property of upper tail dependence and the thresholds were high enough. The seller of the contract would then tend to undervalue the contract. A safer approach from the sellers point of view would be to use a Gumbel copula. From the expression for the bivariate Gumbel survival copula it follows that for the bivariate case the probability of payout is given by

$$\begin{aligned}\mathbb{P}[N=2] &= \hat{C}(\overline{F_1}(k_1), \overline{F_2}(k_2)) \\ &= \overline{F_1}(k_1) + \overline{F_2}(k_2) - 1 + \exp\left(-[(-\ln F_1(k_1))^{\theta} + (-\ln F_2(k_2))^{\theta}]^{1/\theta}\right).\end{aligned}$$

For higher dimensions we use the multivariate extension of the Gumbel family presented in chapter 5.5. For the trivariate case we have

$$\begin{aligned}\mathbb{P}[N=3] &= \hat{C}(\overline{F_1}(k_1), \overline{F_2}(k_2)\overline{F_3}(k_3)) \\ &= \overline{F_1}(k_1) + \overline{F_2}(k_2) + \overline{F_3}(k_3) - 2 \\ &\quad + C_{\theta_2}(F_1(k_1), F_2(k_2)) + C_{\theta_1}(F_1(k_1), F_3(k_3)) + C_{\theta_1}(F_2(k_2), F_3(k_3)) \\ &\quad - C_{\theta_1,\theta_2}(F_1(k_1), F_2(k_2), F_3(k_3)),\end{aligned}$$

where

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

and

$$C_{\theta_1,\theta_2}(u_1, u_2, u_3) = \exp\{-\left([(-\ln u_1)^{\theta_2} + (-\ln u_2)^{\theta_2}]^{\theta_1/\theta_2} + (-\ln u_3)^{\theta_1}\right)^{1/\theta_1}\}.$$

To parametrizise the Gumbel copulas the relation

$$\theta_i = \frac{1}{1-\tau_{1,n-i+1}}, \quad i = 1, \ldots, n-1$$

is used.

To compare the effect of the Gumbel dependence structure with the Gaussian, let $X_i \sim$Lognormal$(0,1)$ for all i, $k_i = k$ for all i and $\tau(X_i, X_j) = 0.5$ for all $i \neq j$. This gives

$$\mathbb{P}[N=n] = 1 + (-1)^1 \binom{n}{1} C_1(F(k)) + \cdots + (-1)^n \binom{n}{n} C_n(F(k), \ldots, F(k)),$$

where $C_1(u) = u$ and $C_m(u_1, \ldots, u_m)$ is the m-margin of $C_n(u_1, \ldots, u_n)$ for $m \in \{2, \ldots, n-1\}$.

In the Gumbel case

$$C_m(F(k), \ldots, F(k)) = \exp\{-[(-\ln F(k))^{\theta} + \cdots + (-\ln F(k))^{\theta}]^{1/\theta}\} = F(k)^{m^{1/\theta}}$$

and in the Gaussian case

$$C_m(F(k), \ldots, F(k)) = \Phi^m_{\rho_l}(\Phi^{-1}(F(k)), \ldots, \Phi^{-1}(F(k))),$$

where (to avoid complicated notation) $\rho_l = \sin(\frac{\pi}{2}\tau)$ denotes the single parameter of the equicorrelated Gaussian copula. $\Phi^m_{\rho_l}(\Phi^{-1}(F(k)), \ldots, \Phi^{-1}(F(k)))$ can be calculated by numerical integration using the fact that (see Johnson and Kotz p.48 [8])

$$\Phi^m_{\rho_l}(a, \ldots, a) = \int_{-\infty}^{\infty} \phi(x) \big[\Phi\left(\frac{a - \sqrt{\rho_l}x}{\sqrt{1-\rho_l}}\right)\big]^m \, dx,$$

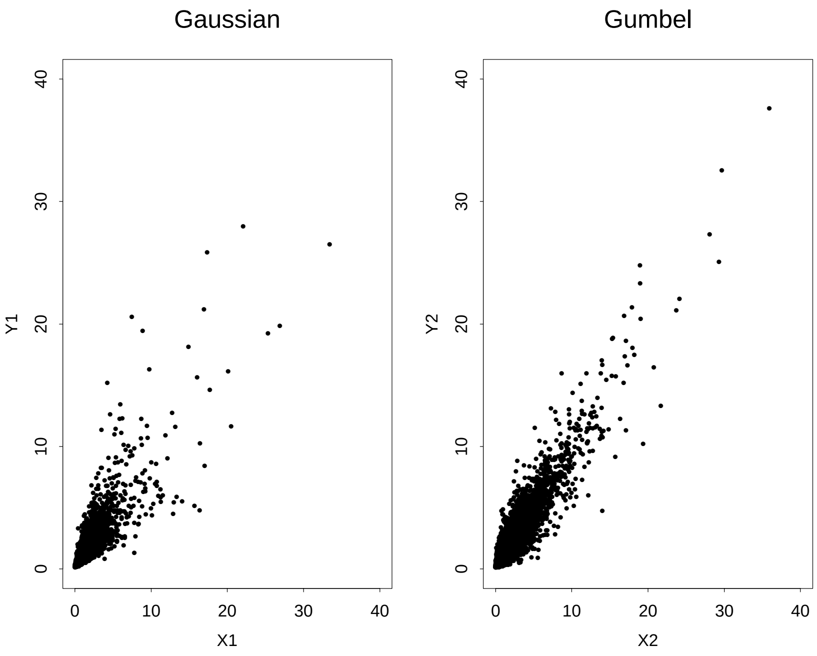

Figure 9.1: Samples from two distributions with standard lognormal margins, Kendall's tau 0.5 and different dependence structures.

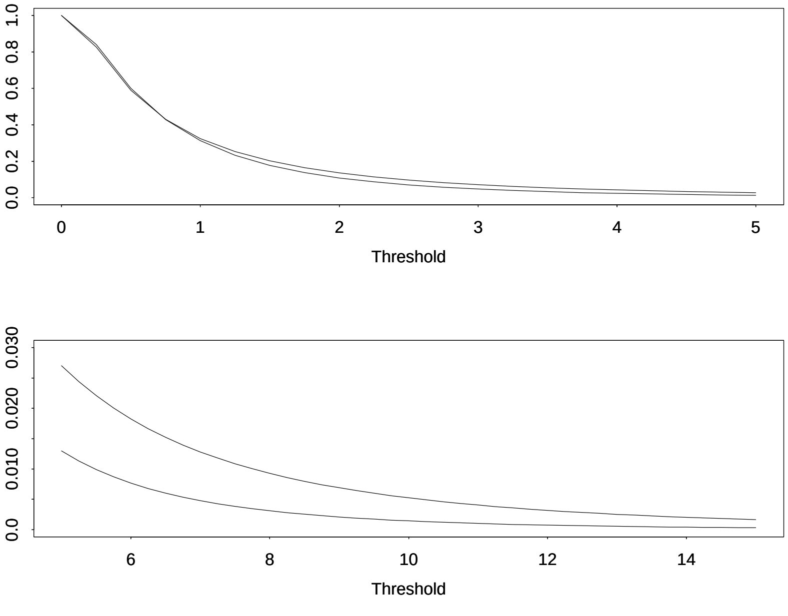

Figure 9.2: Probability of payout for $n = 3$ when the dependence structure is given by a Gumbel copula (upper curve) and Gaussian copula (lower curve).

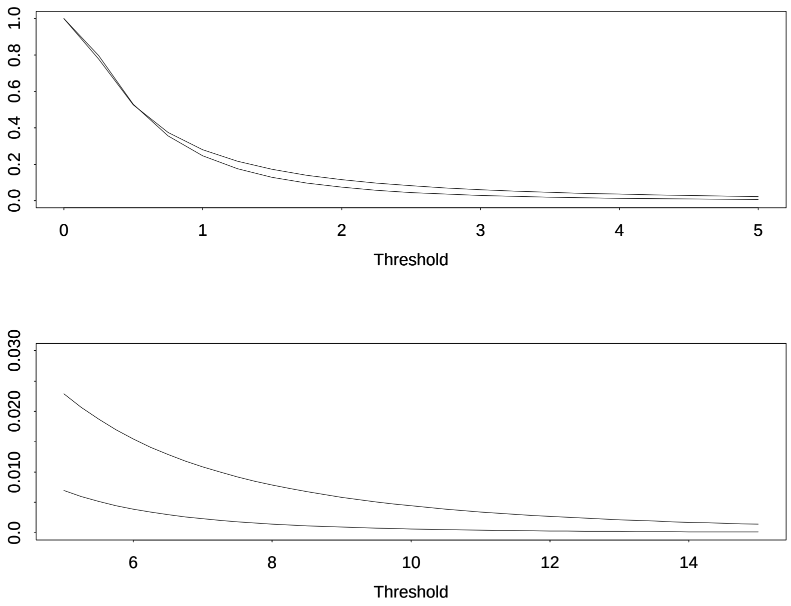

Figure 9.3: Probability of payout for $n = 5$ when the dependence structure is given by a Gumbel copula (upper curve) and Gaussian copula (lower curve).

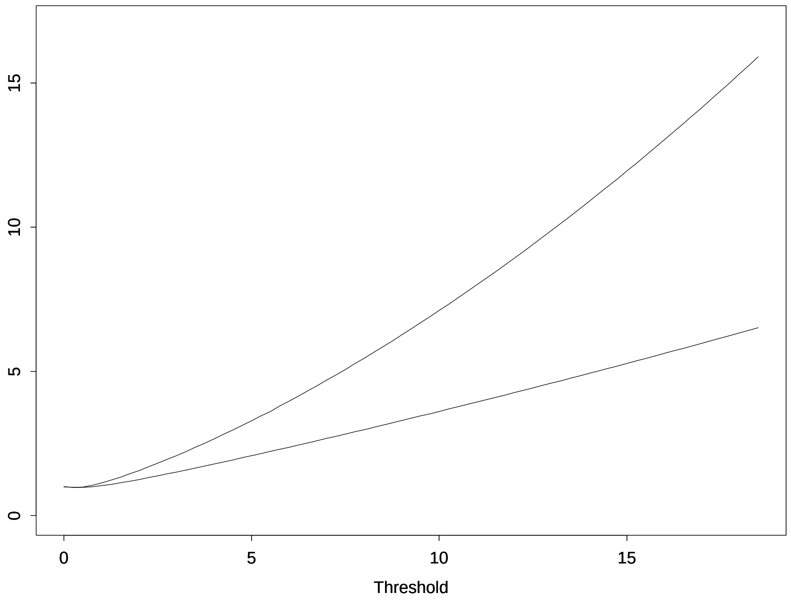

Figure 9.4: Ratios $\mathbf{P}^{\text{Gumbel}}[N = n]/\mathbf{P}^{\text{Gaussian}}[N = n]$ for $n = 3$ (lower curve) and $n = 5$ (upper curve).

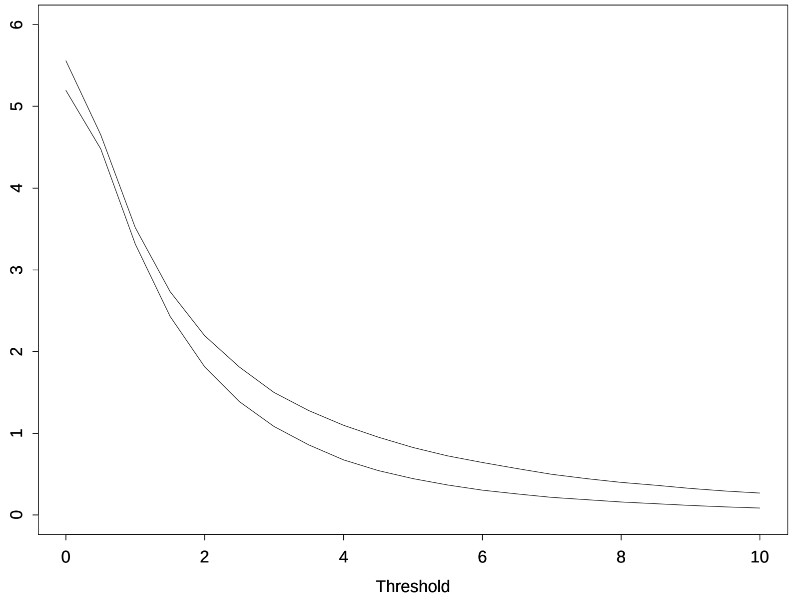

Figure 9.5: $\mathbf{E}^{\text{Gumbel}}(L_n)$ (upper curve) and $\mathbf{E}^{\text{Gaussian}}(L_n)$ (lower curve) for $n = 3$. Simulation results.

where ϕ denotes the univariate standard normal density function. Note that for $k = \text{VaR}_{0.99}(X_i) \approx 10.25$ the calculated payout probability calculated under Gaussian assumptions in the trivariate and 5-variate cases is a factor

$$\frac{\mathbf{P}^{\text{Gumbel}}[N=3]}{\mathbf{P}^{\text{Gaussian}}[N=3]} \approx 3.7 \text{ or } \frac{\mathbf{P}^{\text{Gumbel}}[N=5]}{\mathbf{P}^{\text{Gaussian}}[N=5]} \approx 7.3$$

too low if the true dependence structure is given by a Gumbel copula. So even if the estimates of margins and rank correlations are the best possible the model risk from careless dependence assumptions might be very big even for small n.
It is clear that for calculating the payout probability $\mathbf{P}[N = n]$ it is highly relevant to determine if the underlying dependence structure has the property of tail dependence. However, this is an asymptotic property and it is quite likely that due to the limited amount of data it is far from obvious whether the underlying dependence structure has tail dependence.

Suppose for example that the risks represent a portfolio of potential losses in nearby geographical areas due to damage of property. If a natural catastophy would occur, resulting in huge losses for a certain area it is very likely that the insurance company also faces huge losses for nearby areas. In such a case it seems reasonable to assume that the underlying dependence structure has the property of tail dependence even if the limited amount of historical data suggests the opposite. To price this contract we need to calculate the expected loss to the reinsurer, $\mathbb{E}(L_n)$. Figure 9.5 shows the expected loss to the reinsurer for the Gumbel and Gaussian dependence structure in the trivariate case. The results where obtained using Monte Carlo simulation (100000 runs).

Suppose all three losses exceed their 99% Value-at-Risk. Assuming a Gaussian dependence structure when the true dependence structure is given by a Gumbel copula will result in the loss to the reinsurer being approximately 4 times or more the loss expected. For a bigger n the situation will be a lot worse.

The preceding results show that when pricing contracts which depend on simultaneous exceedences of high thresholds, knowledge of pairwise correlations and marginal distributions is not enough.

The results could easily be generalized to the case with different margins, different pairwise correlations and $l < n$. However, it would then not be as easy to observe the effect of different copula choices.

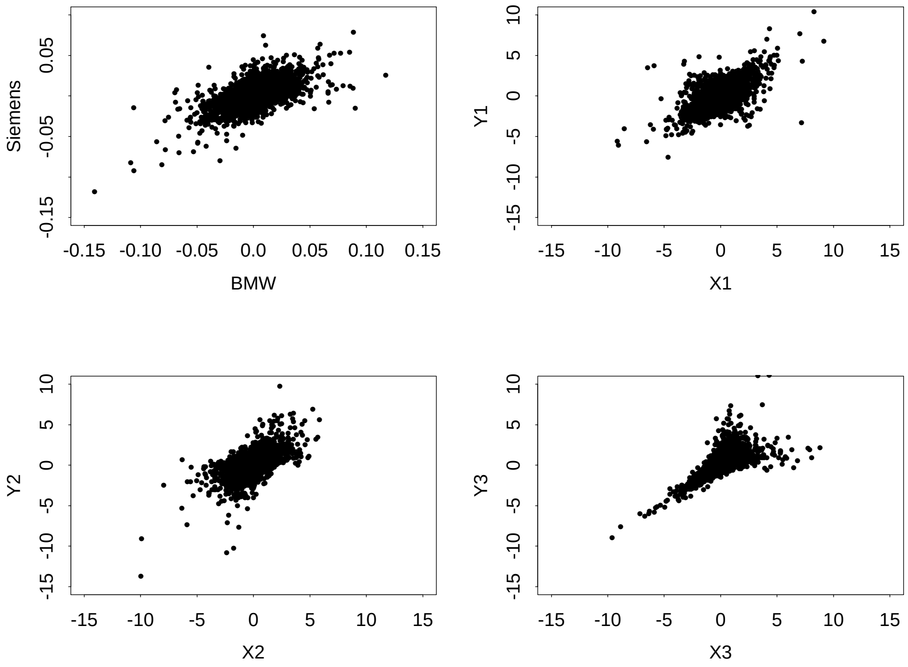

Figure 9.6: The first figure shows daily stock returns for BMW-Siemens. (X1,Y1) have a t_2-copula, (X2,Y2) have a Gaussian copula and (X3,Y3) have a Clayton copula. All margins are t_4-distributed and all pairs have linear correlation 0.64.

9.2 The Perfect Storm

The upper left plot in figure 9.6 shows daily stock returns during a 4000 day period for BMW and Siemens stocks. It appears that a big fall in the stock price for e.g. the BMW stock also means a simultaneous big fall in the stock price for the Siemens stock. The understanding of how to model such phenomena is highly relevant. Extreme, synchronized rises and falls in financial markets occur infrequently – but they do occur. The problem with most models used is that they do not assign a high enough chance of occurrence to the scenario in which many things go wrong at the same time – the "perfect storm" scenario.

In this chapter we will measure the "risk" of a linear portfolio of hypothetical equities. We will do this by considering risk measures which describe the tail of the distribution of the linear sum.

To model a risky scenario such as the one shown in figure 9.6 we clearly need copulas with lower tail dependence, since the dependence structure of the observed data seems to have that property. The plots apart from the upper left plot show samples from bivariate distributions with t_4-margins but different dependence structures. We have chosen t_4-margins because heavy tailed margins seems to fit the data and we want the margins to have finite third moments since this is in line with what is often observed. Furthermore, all four plots show data with linear correlation 0.64, which is the empirical estimate for the data.

One relevant question is the following: Given that a loss exceeds a high threshold, by how much can that threshold be exceeded? To try and answer this question, we will consider measuring the tail risk for a linear portfolio of stocks using a risk measure called expected shortfall. We begin with the concept of mean excess over thresholds which is closely related to expected shortfall. The study of mean excesses over thresholds is a classical idea used in insurance and financial risk management.

Let u be the threshold and define the excess distribution above the threshold for a random variable X with distribution function F to have the distribution function

$$\begin{aligned} F_u(x) &= \mathbb{P}[X-u \le x | X > u] \\ &= \frac{F(x+u)-F(u)}{1-F(u)}, \end{aligned}$$

for $0 \le x < x_F - u$ where $x_F \le \infty$ is the right endpoint of F.
Let the mean excess of the random variable X be given by

$$e_X(u) = \mathbb{E}(X-u|X>u).$$

We sometimes write $e(u)$ when it is clear which random variable we mean. The mean excess of X, $e_X(u)$, can also be expressed as

$$\begin{aligned} e_X(u) = \int_{x=0}^{x_F-u} x\,\mathrm{d}F_u(x) &= \int_{x=u}^{x_F} \frac{(x-u)\,\mathrm{d}F(x)}{1-F(u)} \\ \int_{x=u}^{x_F} \frac{\overline{F}(x)\,\mathrm{d}x}{\overline{F}(u)}. \end{aligned}$$

For the standard normal distribution

$$e(u) = \frac{\phi(u)}{\overline{\Phi}(u)} - u,$$

where ϕ and $\overline{\Phi}$ are the density and survival function of standard normal respectively. It can be shown that

$$\frac{\phi(u)}{\overline{\Phi}(u)} = u(1+u^{-2}+o(u^{-2})),$$

for large u, and it then follows that

$$\lim_{u\to\infty} \frac{u+e(u)}{u} = 1.$$

For the t-distribution with ν degrees of freedom it can be shown that

$$\frac{u+e(u)}{u} = \frac{\sqrt{\nu}(1+u^2/\nu)^{-(\nu-1)/2}}{(\nu-1)B(1/2,\nu/2)u\overline{t}_\nu(u)},$$

where $B(p,q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$ is the Beta function and $\overline{t}_\nu$ is the survival function of univariate standard t-distribution with ν degrees of freedom. It follows that

$$\lim_{u\to\infty} \frac{u+e(u)}{u} = \frac{\nu}{\nu-1} > 1.$$

A measure of risk based on this idea is the expected shortfall. Let X be a random variable with distribution function F. Then the expected shortfall of X at the probability level q is given by

$$S_q = \mathbb{E}(X|X>x_q),$$

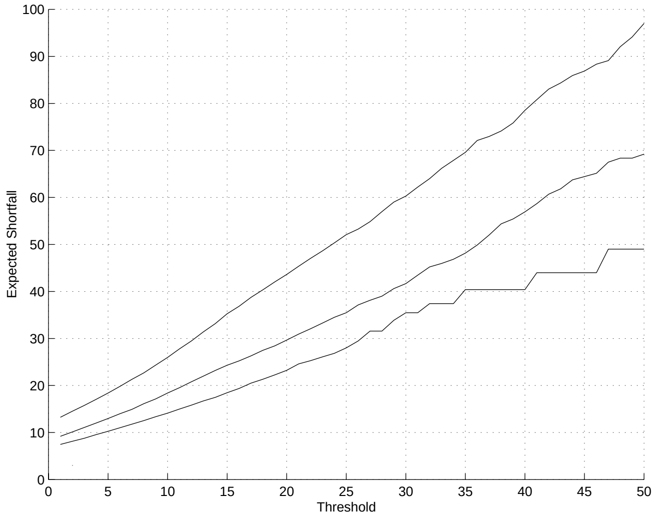

Figure 9.7: Expected losses over thresholds – simulation results.

where $x_q = F^{(-1)}(q)$. Equivalently $x_q = \mathrm{VaR}_q(X)$ which is the well known Value-at-Risk of X at probability level q. The expected shortfall of X can be expressed in terms of the mean excess of X as

$$\frac{S_q}{x_q} = \frac{x_q + e(x_q)}{x_q}.$$

It follows that if X is a normally distributed random variable, then $S_q/x_q \to 1$ as $q \to 1$. If X has a t-distribution with ν degrees of freedom, then $S_q/x_q \to \nu/(\nu-1)$ as $q \to 1$. If X represents a loss and we assume that the loss exceeds some high threshold u, then if X is normally distributed the expected loss exceeds u by only some small percentage. On the other hand if X has a t-distribution with 2 degrees of freedom the expected loss exceeds u by over 100%.

Consider a portfolio of n stocks. Suppose that the stock prices are $S_i(t)$, for $i = 1, \ldots, n$ and $t \in \mathbb{N}$. Let $X_i(t) = (S_i(t+1) - S_i(t))/S_i(t)$ be the stock returns. Suppose that the portfolio on day t gives the stochastic return

$$\sum_{k=1}^{n} \alpha_k X_k(t),$$

where α_k denotes the weight of stock k and $\sum_{k=1}^{n} \alpha_k = 1$. Assume for simplicity that $\alpha_k = 1/n$ for all k. Assume stationarity and drop the dependence on t. Then the portfolio gives the stochastic return

$$\frac{1}{n} \sum_{k=1}^{n} X_k.$$

Thus it seems reasonable to study expected shortfalls for the random variable $\sum_{k=1}^{n} X_k$, where n represents the number of stocks in the portfolio. Figure 9.6 suggests an n-copula with (lower) tail dependence. Furthermore the marginal distributions seem to be heavier tailed than the normal distribution.

Figure 9.7 shows simulation results for expected shortfalls for the random variable $\sum_{k=1}^{n} X_k$ in the case where $n = 10$ and all pairwise linear correlations are 0.7. The upper curve results from simulation from a 10-dimensional t_2-copula with

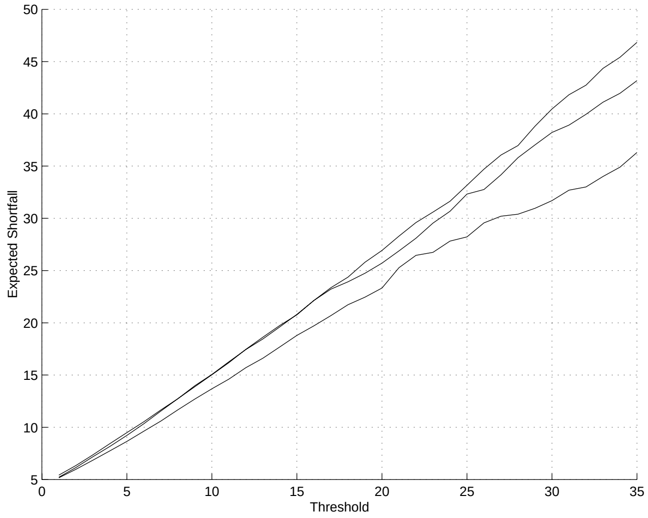

Figure 9.8: Expected losses over thresholds – simulation results.

standard t_2-margins, the lower curve results from a 10-dimensional Gaussian copula with standard normal margins and the middle curve from a 10-dimensional t_2-copula with standard t_4-margins. As seen in figure 9.7 the plot is erratic for large thresholds, when the averaging is over very few excesses. For the Gaussian copula (the lower curve) this is expected since it doesn't have lower tail dependence. Note that if $X_1, \dots, X_n$ are t_2-distributed random variables with mean zero, then so is $\sum_{k=1}^{n} X_k$. If $X_1, \dots, X_n$ are normally distributed with mean zero then so is $\sum_{k=1}^{n} X_k$. With this and the above results in mind, the simulation results shown in figure 9.7 are not surprising. Note also that there are many other multivariate copulas which can be used to model this scenario.

Figure 9.8 and 9.10 show simulation results of expected shortfall and Value-at-Risk for the random variable $\sum_{i=1}^{5} X_i$, where the X_i:s all have standard t_4-distributions and the pairwise Spearman's rank correlation is 0.7 for all pairs X_i, X_j, $i \neq j$. In both figures the curves from upper to lower represent the cases where $(X_1, \dots, X_5)$ have a Gumbel copula, a t-copula and a Gaussian copula respectively. Gumbel copulas have upper but no lower tail dependence. Therefor we here choose modelling losses as positive. Figure 9.9 shows expected shortfall for the random variable $\sum_{i=1}^{5} X_i$ for the two cases where the dependence structures among the risks are given by Gaussian and t_2-copulas. The upper curve shows the expected shortfall for the t_2 case and the lower curve for the Gaussian case.

Note that since the margins and rank correlations are the same in all cases, the differences are attributable to copulas alone.

Note also that it would be just as easy to study the the random variable $\sum_{k=1}^{n} \alpha_k X_k$ for different weights $\alpha_k, k = 1, \dots, n$, for different distributions for the X_ks and different pairwise rank correlations. However, the differences between the results for different choices of copulas would be less apparent.

The conclusions that can be drawn from this example is that when considering risk measures which describe the tail of the distribution the choice of the copula representing the dependence structure among the risks is very important. Furthermore, unlike expected shortfall, VaR does not quite reveal the extra risk inherent in a dependence structure with tail dependence.

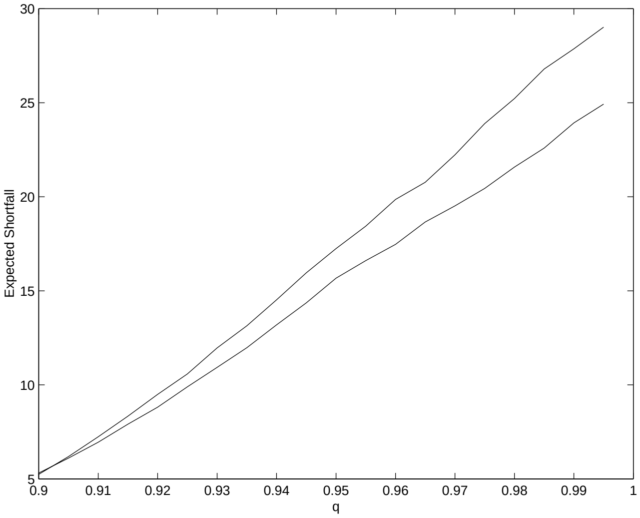

Figure 9.9: Expected shortfall to quantile ratio – simulation results.

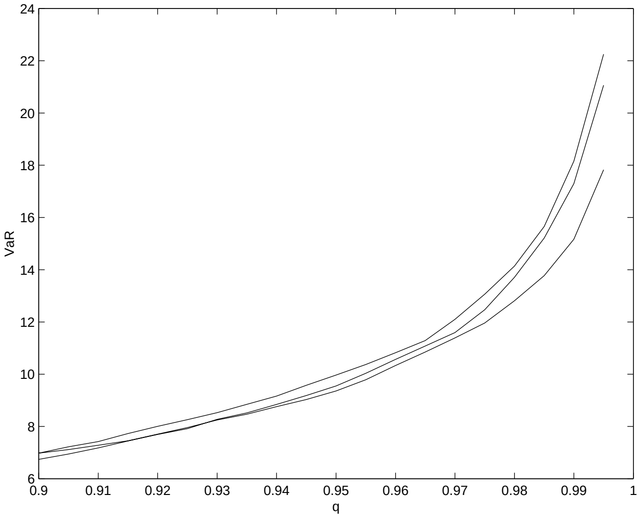

Figure 9.10: Value-at-Risk – simulation results.

10 Conclusions

In this paper we have shown why knowledge of copulas is essential for understanding dependence among random variables. We have presented copula based measures of association and dependence concepts such as tail dependence which provide a natural way of studying dependence. Furthermore, these measures and dependence concepts have the very nice property of invariance under strictly increasing transforms of the underlying random variables. This also means that the problem of finding multivariate models which are consistent with the prespecified margins is much simplified.

We have presented a large number of copula families and suggested natural multivariate extensions with interesting properties and dependence structures. Since much of the literature in this area is basically concerned with the bivariate case, we have focused on multivariate extensions of bivariate copula families and extensions of bivariate results to general dimensions. A main aim of this paper has been constructing efficient algorithms for random variate generation for the presented n-dimensional copula families. This is interesting from a theoretical point of view but perhaps more for practical purposes, since it provides a basis for general simulation tools for modelling dependence in practice. Applications for risk management in insurance and finance are discussed, where we have used copula based measures of association which do not suffer from some of the drawbacks of linear correlation.

It should be noted that although the study of copulas and their applications in statistics is a rather modern phenomenon, a basic knowledge of probability theory and statistics is enough to understand all results.

11 Appendix

To emphasize the usefulness of copula ideas when modelling dependence in practice, this chapter contains sample programs relating to the algorithms presented in the previous chapters and a few examples of how they are used. The programs are written in Splus.

11.1 Implementations

A few possible implementations of the algorithms discussed in the previous chapter are shown below.

```
# Generates random variates from a multivariate standard normal
# distribution with linear correlation matrix x using Cholesky
# decomposition.
# Returns a matrix with n rows and siz columns. Every row is a
# random variate.
rmn<-function(n, x = matrix(c(1, 0, 0, 1), 2, 2))
{
  V <- NULL
  U <- chol(x)
  siz <- dim(x)[1]

  for(i in 1:n)
    {
      Z <- rnorm(siz)
      res <- t(U) %*% Z
      V <- cbind(V,res)
    }
  t(V)
}

# Generates random variates from a multivariate t
# distribution with linear correlation matrix x.
# Returns a matrix with n rows and siz columns. Every row is a
# random variate.
rmt<-function(n, df, x = matrix(c(1, 0, 0, 1), 2, 2))
{
  dimen <- dim(x)[1]
  chi <- 2 * rgamma(n, shape = df/2)
  m1 <- rmn(n, x)
  m2 <- matrix(rep(sqrt(df)/sqrt(chi), dimen), ncol = dimen)
  return(m1 * m2)
}

# Generates random variates from the Gaussian n-copula
mvsncopula<-function(n, isKTau, rcor)
{
  dimen <- dim(rcor)[1]
  lcor <- matrix(0,nrow=dimen,ncol=dimen)
  row <- col <- 1
  while(row <= dimen)
    {
      while(col <= dimen)
```

```
        {
          if(rcor[row,col] == 1)
            lcor[row,col] <- 1
          if(!isKTau)
            lcor[row,col] <- 2*sin(pi*rcor[row,col]/6)
          if(isKTau)
            lcor[row,col] <- sin(pi*rcor[row,col]/2)

          col <- col + 1
        }
      row <- row + 1
      col <- 1
    }
  X <- rmn(n, lcor)

  Z <- NULL
  for(i in (1:n))
    Z <- rbind(Z, pnorm(X[i,], 0, 1))

  Z
}

# Generates random variates from the t n-copula
mvstcopula<-function(n, df, lcor)
{
  X <- rmt(n, df, lcor)

  Z <- NULL
  for(i in (1:n))
    Z <- rbind(Z, pt(X[i,], df))

  Z
}

# Distribution function for the random variable C(U,V), where C
# is the Gumbel copula and U and V are uniform (0,1) random var.
KCg<-function(x, teta, subt)
{
  x*(1 - log(x)/teta) - subt
}

# Generates random variates from the m-dimensional (m - 1)-parameter
# recursive extension of the Gumbel copula.
# The theta values in theta must be in non-decreasing order.
recGumbel<-function(n, theta, dimen)
{
  Z <- runif(n)

  U <- NULL
  cnt <- 1
  while(cnt < dimen)
    {
      Q <- runif(n)
```

```
      W <- NULL
      for(i in (1:n))
        {
          lval <- 0.001
          while(KCg(lval, theta[cnt], Z[i]) > 0)
            lval <- lval/10
          res <- uniroot(KCg,lower=lval,upper=1,teta=theta[cnt],
                         subt=Z[i])
          W <- rbind(W, unlist(res[1]))
        }

      U <- cbind(W^((1-Q)^(1/theta[cnt])),U)
      Z <- W^(Q^(1/theta[cnt]))

      cnt <- cnt + 1
    }
  U <- cbind(Z,U)

  U
}


# Quantile functions for various univariate distributions.

snormal<-function(sampledata)
{
  qnorm(sampledata)
}

suniform<-function(sampledata)
{
  sampledata
}

exponential<-function(sampledata, lambdaval)
{
  qexp(sampledata, lambdaval)
}

frechet<-function(sampledata)
{
  - 1/log(sampledata)
}

tdistr<-function(sampledata, df)
{
  qt(sampledata, df)
}

gdistr<-function(sampledata, param)
{
  qgamma(sampledata, param)
}
```

```
lognormal<-function(sampledata)
{
  qlnorm(sampledata)
}

# Transforms In samples to Rn samples according to the given
# marginals.
simulRn<-function(samples = stop("no samples arg"),
                  margins = stop("no margins arg"),
                  params = stop("no params arg"),
                  n)
{
  X <- NULL

  for(i in 1:n)
    {
      res<-switch(margins[i],
                  SNormal     = snormal(samples[,i]),
                  SUniform    = suniform(samples[,i]),
                  Exponential = exponential(samples[,i], params[i]),
                  Frechet     = frechet(samples[,i]),
                  Tdistr      = tdistr(samples[,i], params[i]),
                  Gdistr      = gdistr(samples[,i], params[i]),
                  Lognormal   = lognormal(samples[,i])
                  )
      X <- cbind(X, res)
    }
  X
}
```

11.2 Examples

An example of how to use the programs is shown below. We begin with simulating 2000 random variates from a Gaussian 5-copula with Spearman's rank correlation matrix *rcorm*. After that we check pairwise Spearman's rank correlations for the data and apply t-margins with different degrees of freedom. Then we proceed with checking that the rank correlations are the same and look at the linear correlation matrix for the transformed data. This shows that the linear correlations are not invariant under strictly increasing transformations of the margins. When we apply both t-margins and standard lognormal margins this is even more obvious. Finally, we simulate 2000 random variates from a trivariate extension of the bivariate Gumbel copula. The Spearman's rank correlations for the data is seen to close to the expected rank correlations $\rho_{12} = 0.7$ and $\rho_{13} = \rho_{23} = 0.5$.

```
> rcorm
     [,1] [,2] [,3] [,4] [,5]
[1,]  1.0  0.7  0.7  0.7  0.7
[2,]  0.7  1.0  0.7  0.7  0.7
[3,]  0.7  0.7  1.0  0.7  0.7
[4,]  0.7  0.7  0.7  1.0  0.7
[5,]  0.7  0.7  0.7  0.7  1.0
>
> copvals<-mvsncopula(2000,F,rcorm)
> cor(rank(copvals[,1]),rank(copvals[,2]))
```

```
[1] 0.6886402
> cor(rank(copvals[,1]),rank(copvals[,3]))
[1] 0.6976064
> cor(rank(copvals[,1]),rank(copvals[,4]))
[1] 0.6984134
> cor(rank(copvals[,1]),rank(copvals[,5]))
[1] 0.7181849
>
> rnvals<-simulRn(copvals,c("Tdistr","Tdistr","Tdistr","Tdistr",
"Tdistr"),c(3,4,5,2,6),5)
>
> cor(rank(rnvals[,1]),rank(rnvals[,2]))
[1] 0.6886402
> cor(rank(rnvals[,1]),rank(rnvals[,3]))
[1] 0.6976064
> cor(rank(rnvals[,1]),rank(rnvals[,4]))
[1] 0.6984134
> cor(rank(rnvals[,1]),rank(rnvals[,5]))
[1] 0.7181849
> cor(rnvals)
          res       res       res       res       res
res 1.0000000 0.6410928 0.6762915 0.5833469 0.6941502
res 0.6410928 1.0000001 0.6857604 0.6093035 0.6740947
res 0.6762915 0.6857604 1.0000000 0.6182865 0.7004315
res 0.5833469 0.6093035 0.6182865 1.0000001 0.6079597
res 0.6941502 0.6740947 0.7004315 0.6079597 1.0000001
>
> rnvals<-simulRn(copvals,c("Tdistr","Tdistr","Lognormal",
"Lognormal","Lognormal"),c(3,3,0,0,0),5)
>
> cor(rnvals)
          res       res       res       res       res
res 1.0000000 0.6198149 0.5630721 0.5432518 0.5884659
res 0.6198149 1.0000000 0.5224883 0.5557166 0.5431907
res 0.5630721 0.5224883 1.0000000 0.6004927 0.5706847
res 0.5432518 0.5557166 0.6004927 1.0000001 0.6274458
res 0.5884659 0.5431907 0.5706847 0.6274458 1.0000000
>
> copvals<-recGumbel(2000,c(1.54,2.07),3)
>
> cor(rank(copvals[,1]),rank(copvals[,2]))
[1] 0.7183862
> cor(rank(copvals[,1]),rank(copvals[,3]))
[1] 0.4746474
> cor(rank(copvals[,2]),rank(copvals[,3]))
[1] 0.4705604
>
```

References

[1] Dhaene J. and Denuit M. (1999) The safest dependence structure among risks, *Insurance: Mathematics and Economics 25*, 11-21.

[2] Embrechts P., McNeil A.J. and Straumann D. (1999) Correlation: Pitfalls and Alternatives, *RISK* **12**(5), 69-71.

[3] Embrechts P., McNeil A.J. and Straumann D. (1999) Correlation and Dependence in Risk Management: Properties and Pitfalls, *Preprint ETH Zürich*, available from http://www.math.ethz.ch/ ~embrechts.

[4] Embrechts P., Klüppelberg C. and Mikosch T. (1997) *Modelling Extremal Events for Insurance and Finance*, Springer, Berlin.

[5] Fang K.-T., Kotz S. and Ng K.-W. (1987) *Symmetric Multivariate and Related Distributions*, Chapman & Hall, London.

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

[7] Johnson M.E. (1987) *Multivariate Statistical Simulations*, John Wiley & Sons, New York.

[8] Johnson N.L. and Kotz S. (1972) *Distributions in Statistics: Continuous Multivariate Distributions*, John Wiley & Sons, New York.

[9] Lee L. (1979) Multivariate distributions having Weibull properties, *Journal of Multivariate Analysis*, **9**, 267-277.

[10] Li D.X. (1999) The valuation of basket credit derivatives, *CreditMetrix Monitor*, **4**, 34-50.

[11] Marshall A.W. and Olkin I. (1983) Domains of attraction of multivariate extreme value distributions, *The Annals of Probability*, **11**1, 168-177.

[12] Müller A. and Bäuerle N. (1998) Modelling and comparing dependencies in multivariate risk portfolios, *ASTIN Bulletin*, **28**1, 59-76.

[13] Nelsen R.B. (1999) *An Introduction to Copulas*, Springer, New York.

[14] Ripley B.D. (1987) *Stochastic Simulation*, John Wiley & Sons, New York.

[15] Schweizer B. and Sklar A. (1983) *Probabilistic Metric Spaces*, North-Holland, New York.

[16] Sklar A. (1996) Random variables, distribution functions, and copulas – a personal look backward and forward, *Distributions with Fixed Marginals and Related Topics*, ed. by L. Rüschendorff, B. Schweizer, and M.D. Taylor, 1-14, Hayward, CA. Institute of Mathematical Statistics

[17] Tiit, E. (1996) Mixtures of multivariate quasi-extremal distributions having given marginals, *Distributions with Fixed Marginals and Related Topics*, ed. by L. Rüschendorff, B. Schweizer, and M.D. Taylor, 337-357, Hayward, CA. Institute of Mathematical Statistics