Dependence Concepts

by
Marta Monika Wawrzyniak

Master Thesis in
Risk and Environmental Modelling
Written under the direction of
Dr Dorota Kurowicka

Delft Institute of Applied Mathematics
Delft University of Technology
Delft 2006

Acknowledgements

I would like to thank to my three supervisors whose support in a difficult time for me was overwhelming. Without them this would never be possible.

First, I would like to thank my daily supervisor Dr Dorota Kurowicka for her help, time, patience and for giving me her valuable comments. Many thanks to Prof.dr Jolanta Misiewicz, my supervisor at University in Poland, for taking the effort to referee this thesis and more importantly for continuous interest and support for this work. Last, but not least I would like to thank to supervisor responsible for my MSc program, Prof. dr Roger Cooke, for many contributions to this thesis in one way or another and endless support. I would also want to thank Prof. dr M. Dekking for his useful comments while refereing this thesis.

I am very grateful for having been given the opportunity to work in this highly academic environment at TUDelft. A special thankyou go again to prof. dr Jolanta Misiewicz and Prof. dr Roger Cooke, who in the year 2002 gave me this chance studying abroad and granted with financial support.

I dedicate this thesis to Robert P. Williams for his friendship and the unforgettable moments.

Marta Wawrzyniak,

February 2006.

Contents

Chapter 1

Introduction

1.1 Abstract

Without a doubt the dependence relations between random variables play a very important role in many fields of mathematics and is one of the most widely studied subjects in probability and statistics. A large variety of dependence concepts have been studied by a number of authors, offering proper definitions and useful properties with applications, just to mention an encyclopedic work of H. Joe *Multivariate model and dependence concepts* (1997) and other: Dall'Aglio, S. Kotz and G. Salinetti *Advances in Probability Distributions with Given Marginals*(1991), K.V.Mardia *Families of Bivariate Distributions*(1967),B. Shweizer, A. Sklar *Probabilistic Metric Spaces* (1983). In this thesis I will mainly use results from the following books: R.Nelsen (1999) [1], D.Kurowicka, R.Cooke (2005)[7] and J. Whittaker (1990)[6].

This thesis elaborates on the dependence concept and describes well known measures of dependence. We show how the dependence can be measured or expressed and in which distributions we can measure this dependency. For better understanding, we first concentrate on dependence concepts in bivariate case and then we generalize these concepts to higher dimensions.

We examine well known measures of dependence such as: product moment correlation, Spearman's rank correlation and Kendall's tau. Their properties, advantages, disadvantages and applications in describing the dependence between random variables are discussed. Then, a more sophisticated measure of dependence - measure of interaction (described by Whittaker) comes into play. We explore its properties and compare with other measures of dependence.

The most widely used measure of dependence is the product moment correlation (also called Pearson's linear correlation). This correlation measures the linear relationship between X and Y and can attain any value from $[-1, 1]$. Product moment correlation is easy to calculate and is a very attractive measure for the family of elliptical distributions (because for this distribution zero correlation implies independence). However, it has many disadvantages: it does not exist if the expectations and/or variances are not finite, its possible values depend on marginal distributions and it is not invariant under nonlinear strictly increasing transformations.

A more flexible measure of dependence is rank correlation (also called Spearman's rank correlation) which, in contrast to linear correlation, always exists, does not depend on marginal distributions and is invariant under monotonic strictly increasing transformations. Another widely used measure of dependence is Kendall's tau, which has a simple interpretation and can be easily calculated.

These correlations are used in the joint distribution function of (X, Y) to model the dependence between random variables X and Y. The most popular bivariate distribution is the normal distribution, which belongs to the large family of elliptical distributions. Elliptical distributions are very often used, particulary in risk and financial mathematics. A special case of bivariate distributions are copulas, which are defined on the unit square with uniformly distributed marginals. The relationship between a joint distribution and a copula allows us to study the dependency structure of (X, Y) separately from the marginal distributions. The most widely used copulas are: normal, archimedean and elliptical.

All of the measures (product moment, rank correlation, Kendall's tau) can be computed from data (either directly from the data or from the ranks of the data) but fail to measure more complicated dependence structure. In contrast, the measures of interaction described by Whittaker in [6] require information about joint distribution function and not the data. Then, the complex dependence structure between two variables can be estimated by identifying regions of positive, zero and negative values of interaction.

The concepts of dependence between two random variables is extended to the concepts of multivariate dependency. The conditional correlation, describes the relationship between two variables while conditioning on other variables. Conditional correlation is simply a product moment correlation of the conditional distributions, hence conditional correlation possess the same the disadvantages as the product moment correlation. The conditional correlation is equal to partial correlations in the family of elliptical distributions. The partial correlation can be easily calculated from recursive formulas and therefore the conditional correlation is often approximated by, or replaced by the partial correlation. The meaning of partial correlation for non-elliptical variables is less clear.

To represent multivariate dependence all of the measures of dependence between pairs of variables must be collected in a correlation matrix. This matrix has to be complete and positive definite. In order to avoid those problems with the correlation matrix, another way of representing multivariate distributions called vines, was introduced by Bredford, Cooke [10]. Using a copula-vine method we can construct multivariate distributions in a straightforward way by specifying marginal distributions and the dependence structure. This dependence structure can be represented by a vine. The main advantage of this approach is that, by assigning the conditional rank correlations between pairs of variables to a vine, we do not have worry about the positive definiteness.

Finally the mixed derivative measures of conditional interactions are defined as the extensions of bivariate interactions. We study and investigate possible relations between interactions and a copula-vine specification of the multivariate distributions.

1.2 Preliminaries

Let $(\Omega, \mathcal{A}, P)$ be a probability space, i.e., Ω is a non-empty set, $\mathcal{A}$ is a σ-algebra of subsets of Ω (collection of events), and P is a probability function $P : \mathcal{A} \to [0, 1]$ such that: $P(\Omega) = 1$, and if $\{A_n, n \geq 1\}$ is a sequence of sets of $\mathcal{A}$, where A_n and A_m are disjoint, then $P(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n)$. Let $\mathcal{B}(\mathbb{R})$ be the σ-algebra generated by the borelian sets in $\mathbb{R}$. A *random* *variable* X is a measurable function on probability space $X : \omega \to \mathbb{R}$. The *cumulative distribution function* of a random variable X is defined to be the function $F(x) = P(X \leq x)$ for $x \in \mathbb{R}$. The cumulative distribution function is non-decreasing, right-continuous, and $\lim_{x\to-\infty} F(x) = 0$, $\lim_{x\to\infty} F(x) = 1$. The *probability density function* $f(x)$ or simply density function of a continuous distribution is defined as the derivative of the (cumulative) distribution function $F(x)$, so:

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(x)dx.$$

For a random vector (X, Y) the joint distribution function is the function $H : \mathbb{R}^2 \to [0, 1]$ defined by: $H(x, y) = P(X \leq x, Y \leq y)$ for $(x, y) \in \mathbb{R}^2$. For n-dimensional random vector $(X_1, \cdots, X_n)$, the joint cumulative distribution function is defined by $F(x_1, \cdots, x_n) = P(X_1 \leq x_1, \cdots, X_n \leq x_n)$ for $(x_1, \cdots, x_n) \in \mathbb{R}^n$. The joint distribution function completely characterizes the behavior of $(X_1, \cdots, X_n)$: it defines the distribution of each of its components, called *marginal* distributions, and determines their relationships.

For better understanding of the independence concept, let us start with the independence of events. Let A and B be two events defined on the probability space $(\Omega, \mathcal{A}, P)$. Then A and B are independent if and only if $P(A \cap B) = P(A)P(B)$. If two events are independent, then the conditional probability of A given B is the same as the unconditional probability of A, that is:

$$A \perp B \Leftrightarrow P(A|B) = P(A).$$

Here the *conditional probability of A given B* is given by

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

if only $P(B) \neq 0$.

For our purposes it is more suitable to talk about random variables, whose values are described by a probability distribution function. From now on, the random variables are assumed to be continuous. We say that the two random variables X and Y are *independent*, denoted by $X \perp Y$, if and only if the joint probability density function of vector (X, Y) f_{XY}, is equal to the product of their marginal density functions:

$$f_{XY}(x, y) = f_X(x) f_Y(y)$$

for all values of x and y, where f_X and f_Y are marginal densities of X and Y.

The conditional density function of X given Y is defined as $\frac{f_{XY}}{f_Y}$, where f_Y is non-zero function. We can equivalently rewrite the definition of independence of random variables in terms of conditional formulation:

$$X \perp Y \Leftrightarrow f_{X|Y}(x; y) = f_X(x).$$

The two random variables X and Y are *conditionally independent* given Z if and only if there exist functions g and h such that

$$f_{XYZ}(x, y, z) = g(x, z) h(y, z) \tag{1.1}$$

for all x,y and z such that $f_Z(z) > 0$, this is called the factorization criterion.

We can extend the independence of two variables to the multivariate case. Then random variables $X_1, \cdots, X_n$ are independent if and only if:

$$f_{1\cdots n}(x_1, \cdots, x_n) = f_1(x_1) \cdots f_n(x_n),$$

where $f_{1\cdots n}$ is a joint probability density function of random vector $(X_1, \cdots, X_n)$, and f_i's are marginal densities, $i = 1, \cdots, n$.

The factorization criterion can be also extended to higher dimensions. Then, X and Y are conditionally independent given $\mathbf{Z} = (Z_1, \cdots, Z_k)$ if and only if there exist functions g and h such that

$$f_{XY\mathbf{Z}}(x, y, z_1, \cdots, z_k) = g(x, z_1, \cdots, z_k)h(y, z_1, \cdots, z_k)$$

for all x,y and $\mathbf{z} = (z_1, \cdots, z_n)$ such that $f_{\mathbf{z}}(z_1, \cdots, z_k) > 0$.

1.3 Thesis Overview

This thesis consists of two parts: the bivariate and the multivariate dependence concept. Bivariate dependence modelling is a subject of Chapter 2, while the multivariate aspects are studied in Chapter 3.

Chapter 2 is devoted to the dependence between two random variables X and Y. The well known measures of dependence such as: Pearson's linear correlation, Spearman's rank correlation and Kendall's tau (their definitions and properties) are described. Then we focus on the bivariate distribution functions of X and Y. We consider such distributions as elliptical distributions (normal and Cauchy distribution) and bivariate distributions with uniformly distributed marginals -copulas.

There exists a relationship between joint distribution and copulas. The joint distribution can be described as a product of the marginal distributions and the appropriate copula. This makes copulas a very special method to model the dependence in bivariate distributions. The most commonly used in applications are: the normal copula and the Archimedean copulas. We present a few types of copulas in this thesis: normal, Gumble, Clayton, Frank and elliptical copulas.

A significant part of this chapter is devoted to measures of interactions previously studied and described by Joe Whittaker in [6]. They are defined as mixed derivatives of the logarithm of the density. It is interesting that the whole information about interactions between random variables is contained in copula corresponding to the joint distribution between these random variables.

Chapter 3 discusses multidimensional concepts of dependence in a random vector $(X_1, \cdots, X_n)$.

First, as a key notions in multivariate modelling, the partial and conditional correlation (which correspond to linear correlation), are described.

Further, we talk about multivariate distributions and as an example the multivariate normal distribution is described. Multivariate distributions, like bivariate distribution, can be described in terms of copulas. We can study the dependence structure separately from marginal distributions. Further, we use vines as another way to define multivariate distributions. We describe the copula-vine method, which builds a joint density of random variables as a product of copulas and conditional copulas.

We finalize this thesis by considering conditional interactions for copula-vine distributions. Some conclusions and future research topics related to this thesis are placed in the last chapter of this thesis.

Chapter 2

Bivariate Dependence Concept

In this chapter we discuss some dependence notions which have been discussed by Kurowicka, Cooke [7], Nelsen [1],Whittaker [6].

When talking about bivariate dependence we need to discuss the following aspects:

- how to measure the dependence between two random variables,

- and in which bivariate distributions this dependence can be measured.

The answer to the first question produces measures of dependence (linear correlation, rank correlation, partial and conditional correlations), those measures are then used to study dependence concepts in bivariate distribution (copula approach).

2.1 Product Moment Correlation

Let X and Y be random variables. The covariance of X and Y is defined as $Cov(X,Y) = E[(X - E(X))(Y - E(Y))] = E(XY) - E(X)E(Y)$. We can standardize it by dividing by the square root of variances of each variable involved. This coefficient is often called linear correlation or Pearson's correlation coefficient.

The following definition of product moment correlation is adapted from Karl Pearson's *"Mathematical Contributions... III. Regression, Heredity, and Panmixia"* published in 1896.

Definition 1. *For any random variables X and Y with finite means and variances, the* product moment correlation *is defined as:*

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}. \tag{2.1}$$

Pearson's correlation coefficient is most widely known measure of dependence because it can be easily calculated. Its properties are listed below:

- product moment correlation measures the linear relationship between two variables.
- It ranges from -1 to $+1$;
- A correlation of $+1$ means that there is a perfect positive linear relationship between variables X and Y, hence $Y = aX + b$ almost surely for $a > 0$, $b \in \mathbf{R}$;

- A correlation of -1 means that there is a perfect negative linear relationship between variables X and Y, hence $Y = -aX + b$ almost surely for $a > 0$,$b \in \mathbf{R}$;

- Product moment correlation is invariant under strictly increasing linear transformations, i.e: $\rho(aX + b, Y) = \rho(X, Y)$ if $a > 0$ and $\rho(aX + b, Y) = -\rho(X, Y)$ if $a < 0$;

- If X and Y are independent, then $\rho(X, Y) = 0$.

In general, the converse of the last bullet does not hold. Zero correlation does not imply the independence of X and Y, as the following example shows.

Example 1. *Consider a standard normally distributed random variable X and a random variable $Y = X^2$, which is surely not independent of X. We have:*

$$Cov(X, Y) = E(XY) - E(X)E(Y) = E(X^3) = 0$$

because $E(X) = 0$ and $E(X^2) = 1$, therefore $\rho(X, Y) = 0$ as well.

We can observe that the equivalence between zero correlation and independent variables holds for the elliptical distributions[1].

Remark 1. *For two elliptical distributed random variables X and Y, the following is true:*

$$X \perp Y \Leftrightarrow \rho(X, Y) = 0.$$

[1]Elliptical distributions will be discussed in section 2.4.1.

So in case of elliptical distributed random variables, product moment correlation coefficient seems to be a good and effective measure of dependence. However, not all of the real problems have elliptical distributions. And for non-elliptical distributions product moment correlation may be very misleading.

Product moment correlation can be calculated as follows:

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \overline{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \overline{Y})^2}},$$

where $\overline{X}$ is the average of X, and $\overline{Y}$ is the average of Y.

Another well known measure of dependence is rank correlation called Spearman's rho.

2.2 Spearman's rho

The value of Spearman's rho, denoted by ρ_S is equivalent to the Pearson product moment correlation coefficient for the correlation between the ranked data. Spearman's rho was developed by Charles Spearman (1904).

The definition of this measure is as follows:

Definition 2. *For X and Y with cumulative distribution functions F_X and F_Y respectively, Spearman's rank correlation is defined as*

$$\rho_S(X,Y) = \rho(F_X(X), F_Y(Y)).$$

Another way of defining rank correlation is by introducing the so called population version:

$$\rho_S = 3\left(P[(X_1 - X_2)(Y_1 - Y_3) > 0] - P[(X_1 - X_2)(Y_1 - Y_3) < 0]\right) \tag{2.2}$$

where $(X_1, Y_1), (X_2, Y_2), (X_3, Y_3)$ are three independent identically distributed random vectors.

The most important properties of ρ_S are:

- Spearman's rho always exists,
- it is independent of marginal distributions,
- it is invariant under non-linear strictly increasing transformations,i.e: $\rho_S(X, Y) = \rho_S(G(X), Y)$ if $G : \mathbb{R} \to \mathbb{R}$ is strictly increasing function, and $\rho_S(X, Y) = -\rho_S(G(X), Y)$ if $G : \mathbb{R} \to \mathbb{R}$ is strictly decreasing function,
- if $\rho_S(X, Y) = 1$ then there exists a strictly increasing function $G : \mathbb{R} \to \mathbb{R}$ such that $X = G(Y)$

This measure is not perfect either, we can show an example that zero rank correlation is not equivalent with independent random variables.

Example 2 ([7]). *Let U and V be uniform on $(0, 1)$ random variables. And M, W are bivariate distributions such that mass is distributed uniformly on the main diagonal, i.e. $P(U = V) = 1$ or anti-diagonal i.e. $P(U + V = 1) = 1$ respectively. M and W are called the Fréchet-Hoeffding upper and lower bound respectively and are discussed in section 2.5. The graph (2.1) represent the main diagonal of unit square $u = v$ (left) and anti-diagonal $u = 1 - v$ (right).*

M and W describe positive and negative dependence between U and V respectively. Then if U and V have joint distribution M then $\rho_S(U, V) = 1$ and if they're joined with W then $\rho_S(U, V) = -1$.

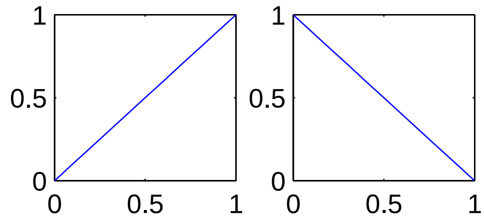

Figure 2.1: The support of Fréchet-Hoeffding bounds: upper M (left) and lower W (right)

Let us take mixture of the Fréchet-Hoeffding bounds, for which the mass is concentrated on the diagonal and anti-diagonal depending on parameter $\alpha \in [0,1]$*:*

$$C_\alpha(u,v) = (1-\alpha)W(u,v) + \alpha M(u,v)$$

for $(u,v) \in [0,1]^2$.

If we take $\alpha = \frac{1}{2}$*, then the variables* U *and* V *joined by* C_α *have rank correlation equal to zero. But this mixture is not independent.*

For Spearman's Rank Correlation Coefficient the calculations are carried out on the ranks of the data.

The values of the variable are put in order and numbered so that the lowest value is given rank 1, and the second lowest is given rank 2 etc. If two data values are the same for a variable, then they are given averaged ranks. This ranking needs to be done for both variables. Spearman rank is calculated by taking the product moment correlation of the ranks of the data. For given data points: $(x_1,y_1),(x_2,y_2),\cdots,(x_n,y_n)$ we assign ranks in the following manner:

r_i = rank of x_i, and s_i = rank of y_i, then:

$$\rho_S = \frac{\sum_{i=1}^{n}(r_i - \overline{r})(s_i - \overline{s})}{\sqrt{\sum_{i=1}^{n}(r_i - \overline{r})^2}\sqrt{\sum_{i=1}^{n}(s_i - \overline{s})^2}} = 1 - \frac{6}{n(n^2-1)\sum_{1=1}^{n}(r_i - s_i)^2}.$$

Since r and s are the ranks, then $\overline{r} = \overline{s} = \frac{n+1}{2}$.

In the 1940s Maurice Kendall developed another rank correlation, which is now called *Kendall's tau.*

2.3 Kendall's tau

One of the definitions of this measure is a population version of Kendall's tau:

$$\tau = P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0] \tag{2.3}$$

for two independent identically distributed random vectors $(X_1, Y_1), (X_2, Y_2)$.

We can see that Kendall's τ is symmetric, i.e: $\tau(X_1, X_2) = \tau(X_2, X_1)$, and normalized to the interval $[-1, 1]$. Also the following holds:

Proposition 1. *Let X_1 and X_2 be continuous random variables, then:*

$$X_1 \perp X_2 \Leftrightarrow \tau(X_1, X_2) = 0.$$

Kendall's tau and Spearman's rho have another important property which linear correlation does not have, they are copula-based measures and can be specified in terms of copulas but this will be discussed later on.

Kendall's tau can be estimated from an underlying data set by:

$$\tau(X_1, X_2) = \frac{\sum_{i<j} sign\lfloor(x_{1i} - x_{1j})(x_{2i} - x_{2j})\rfloor}{\binom{n}{2}}.$$

This summation is across all possible pairs of observations.

Spearman's rank correlation is a more widely used measure of rank correlation because it is much easier to compute than Kendall's tau. Both rank correlation coefficients are more useful in describing the dependence, however it is difficult to understand their exact meaning.

Product moment, rank correlation and Kendall's tau are not the only measures of dependence. There are of course many more. Just to name few: *Gini's* γ*, Blomqvist's* β*, Schweizer and Wolff's* σ described by Nelsen in [1] in chapter 5. They will not be discussed in this thesis.

Let us now concentrate on bivariate distributions.

2.4 Bivariate Distributions

The joint distribution of the random vector (X, Y) captures the dependence between random variables X and Y. The most visual way to specify a random vector is the probability density function (density). There are many bivariate distributions, the most popular and widely used among them is the normal distribution. The normal distribution is a special case of the larger class of elliptical distributions.

2.4.1 Elliptical Distributions

The 2-dimensional random vector $\mathbf{X}$ is said to be *elliptically distributed*, symbolically $\mathbf{X} \sim EC(\mu, \Sigma, \phi)$, if its characteristic function may be expressed in the form:

$$\psi(t) = E[e^{it^T\mathbf{X}}] = e^{it^T\mu}\phi(t^T\Sigma t),$$

with μ a 2-dimensional vector, Σ a positive 2×2 matrix, and ϕ a scalar function called *characteristic generator*.

The density of an elliptical random vector $\mathbf{X} \sim EC(\mu, \Sigma, \phi)$ has the form:

$$f(x) = \frac{c}{\sqrt{|\Sigma|}}\phi\left((x-\mu)^T\Sigma^{-1}(x-\mu)\right).$$

This density is constant on ellipses, that means when viewed from above, the contour lines of the distribution are ellipses. This is the reason for calling this family of distributions *elliptical*.

Elliptical random vector have the following properties, for more details refer to Fang et al.(1990):

- any linear combination of elliptically distributed variables is elliptical;
- Marginal distributions of elliptically distributed random vector are elliptical.
- Suppose $X \sim EC(\mu, \Sigma, \phi)$ possess k moments, if $k \geq 1$, then $E(X) = \mu$, and if $k \geq 2$, then $Cov(X) = -2\psi^T(0)\Sigma$;
- it can be easily verified that the normal and t-distribution are members of the class of elliptical distributions.

Elliptical distributions are the easiest distributions to work with but not always realistic.

We discuss two bivariate elliptical distributions: the normal which is the most famous member of this family, and Cauchy distribution which is a special case of the t-distribution.

2.4.2 Bivariate Normal Distribution

The normal distribution, also called Gaussian, is an elliptical distribution with characteristic generator:

$$\phi(t) = e^{-\frac{1}{2}t}.$$

Normally distributed random variable is given by two parameters: the mean μ and standard deviation σ, symbolically denoted by $X \sim N(\mu, \sigma)$

The probability density function of such distribution is:

$$f(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

The standard normal distribution is the normal distribution with a mean of zero and a standard deviation one:

$$f(x; 0, 1) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right).$$

The bivariate normal distribution is a joint distribution of two normal variables X and Y. The joint normal density of $(X, Y) \sim N([\mu_1, \mu_2], [\sigma_1, \sigma_2])$ is given by:

$$f(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left(-\frac{(\frac{x^2-\mu_1}{\sigma_1})^2 - 2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + (\frac{y^2-\mu_2}{\sigma_2})^2}{2(1-\rho^2)}\right) \tag{2.4}$$

And the density of the standard bivariate normal $(X, Y) \sim N([0, 0], [1, 1])$ has the following form:

$$f(x, y) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)}\right) \tag{2.5}$$

for $x, y \in (-\infty, \infty)$ and a parameter ρ (Pearson's product moment correlation) $\in [-1, 1]$.

The properties of normal distribution make this distribution to be so important and mostly used in many fields of mathematics. Some of them, relevant in this thesis are listed below, for more details and proofs I refer to Kurowicka, Cooke [7].

If (X, Y) has bivariate standard normal distribution with parameter ρ, then:

- the marginal distributions of X and Y are standard normal;
- $\rho(X, Y) = 0 \Leftrightarrow X \perp Y$, this in general, holds for any elliptical distributed random vector;
- the relationship between product moment and Spearman's rank correlations also known as a *Pearson's transformation*:

$$\rho(X, Y) = 2\sin(\frac{\pi}{6}\rho_S(X, Y));$$

- the relationship between product moment correlation and Kendall's tau:

$$\rho(X, Y) = \sin(\frac{\pi}{2}\tau(X, Y)).$$

2.4.3 The Cauchy Distribution

The Cauchy distribution belongs also to the the family of elliptical distributions and is characterized by the location parameter x_0 and the scale parameter $\gamma > 0$.

The Cauchy distribution is a special case of the student t distribution with one degree of freedom[2].

The probability density function of the univariate Cauchy distribution is defined as:

$$f(x; x_0, \gamma) = \frac{1}{\pi\gamma[1 + \left(\frac{x-x_0}{\gamma}\right)^2]},$$

and the cumulative distribution function is:

$$F(x; x_0, \gamma) = \frac{1}{\pi} \arctan\left(\frac{x - x_0}{\gamma}\right) + \frac{1}{2}.$$

The special case when $x_0 = 0$ and $\gamma = 1$ is called the standard Cauchy distribution. The bivariate standard Cauchy has the following probability density function:

$$f(x_1, x_2) = \frac{1}{\pi(1 + x_1^2 + x_2^2)^{3/2}}. \tag{2.6}$$

It is interesting that when U and V are two independent standard normal distributions, then the ratio $\frac{U}{V}$ has the standard Cauchy distribution.

The Cauchy distribution is a distribution for which expectation, variance or any higher moments are not defined.

[2]The density of the n dimensional t distribution with v degrees of freedom is defined by:

$$f(x_1, \cdots, x_n) = \frac{\Gamma(\frac{v+n}{2})}{\Gamma(\frac{v}{2})\sqrt{(v\pi)^n|\Sigma|}} \left(1 + \frac{1}{v}\mathbf{x}'\Sigma^{-1}\mathbf{x}\right)^{-\frac{v+n}{2}},$$

where $\frac{v}{v-2}\Sigma$ is the covariance matrix and is defined only if $v > 2$.

The bivariate distributions on unit square with uniform marginal distributions are called copulas. This class of distributions allows us to separate marginal distributions and the information about dependence in joint distribution

Copulas were characterized by Abe Sklar in 1959 but they were studied by other authors earlier. Recently, they became very popular and have been widely investigated, see e.g. Nelsen [1].

2.5 Bivariate Copulas

The copulas are usually defined on the unit square $\mathbf{I}^2$, where $\mathbf{I} = [0, 1]$, this interval can be transformed to, for instance: $[-1/2, 1/2]$ or $[-1, 1]$. According to Nelsen [1], the definition of a copula is:

Definition 3. *A 2-dimensional function C from $\mathbf{I}^2$ to $\mathbf{I}$ is called a* copula *if it satisfies the following properties:*

1. *For every $u, v \in \mathbf{I}$*

$$C(u, 0) = C(0, v) = 0$$

and

$$C(u, 1) = u, C(1, v) = v;$$

2. *For every $u_1, u_2, v_1, v_2 \in \mathbf{I}$ such that $u_1 \leq u_2$ and $v_1 \leq v_2$,*

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

An important result (Fréchet, 1951) states that any copula has a lower and an upper bound. Let C be a distribution function of a copula. Then for every $(u, v) \in \mathbf{I}^2$ the copula C must lie in the following interval:

$$W(u, v) = max(u + v - 1, 0) \leq C(u, v) \leq min(u, v) = M(u, v).$$

The bonds M and W are themselves copulas and are called *Fréchet-Hoeffding upper bond* and *Fréchet-Hoeffding lower bond* respectively. Another very important copula is the *product copula* $\prod(u, v) = uv$, also known as independent copula.

The next theorem plays undoubtedly the main role in the theory of copulas.

2.5.1 Sklar's theorem

Sklar's theorem describes the relationship between distribution function H and corresponding copula C.

Theorem 2 (Nelsen [1]). *Let X and Y be random variables with margins F_X and F_Y respectively and joint distribution function H. Then there exists a copula C such that for all $x, y \in \mathbf{R}^2$*

$$H(x, y) = C(F_X(x), F_Y(y)). \tag{2.7}$$

If F_X and F_Y are continuous, then C is unique.
Conversely, if C is a copula and F_X and F_Y are distribution functions, then the function H defined as above is a joint distribution function with margins F_X and F_Y.

Directly from this theorem comes the following definition:

Definition 4. *Random variables X and Y are joint by copula C if and only if their joint distribution F_{XY} can be written:*

$$F_{XY}(u, v) = C(F_X^{-1}(u), F_Y^{-1}(v)).$$

The above result provides a method of constructing copulas from joint distributions.

Since the copula corresponding to a joint distribution describes its dependence structure, it might be appropriate to use measures of dependence which are copula-based, so called measures of concordance. Spearman's rho and Kendall's tau are examples of such concordance measures. And they can be expressed in terms of copulas in the following way ([1]):

$$\tau(X, Y) = 4 \int\int_{[0,1]^2} C(u, v) dC(u, v) - 1, \tag{2.8}$$

$$\rho_S(X, Y) = 12 \int\int_{[0,1]^2} C(u, v) du dv - 3. \tag{2.9}$$

As with standard distribution functions, copulas have associated densities.

2.5.2 Copula as a density function

The density of a copula C is denoted by lowercase c, then the density of bivariate copula is just the mixed derivative of C, i.e:

$$c(u_1, u_2) = \frac{\partial^2 C(u_1, u_2)}{\partial u_1 \partial u_2}.$$

If X_1 and X_2 are random variables with densities f_1, f_2 and distribution functions F_1, F_2 respectively, then the joint density function of a pair of random variables (X_1, X_2) may be written as ([7]):

$$f_{12}(x_1, x_2) = c(F_1(x_1), F_2(x_2))f_1(x_1)f_2(x_2). \tag{2.10}$$

2.5.3 Parametric families of copulas

Parametric distributions are those bivariate distributions which are characterized by a vector of parameters, for instance the normal distribution is given by $(\mu, \sigma) \in \mathbf{R} \times [0, \infty)$. It is much more interesting to model bivariate distributions with the same dependence structure given by a vector of parameters.

Among all, the most significant is a family of normal copulas. The standard normal copula is described below.

Normal Copula

The normal copula allow us to create a family of bivariate normal distributions with a specified correlation coefficient $\rho \in [-1, 1]$.

The density of a bivariate standard normal copula is given by:

$$c_N(u_1, u_2) = \frac{1}{\sqrt{(1-\rho^2)}} \exp\left(-\frac{\zeta_1^2 - 2\rho\zeta_1\zeta_2 + \zeta_2^2}{2(1-\rho^2)}\right) \exp\left(\frac{1}{2}(\zeta_1^2 + \zeta_2^2)\right) \tag{2.11}$$

where $\zeta_1 = \Phi^{(-1)}(u_1),\ \zeta_2 = \Phi^{(-1)}(u_2)$.

Figure (2.2) represents the density of an example of a normal copula.

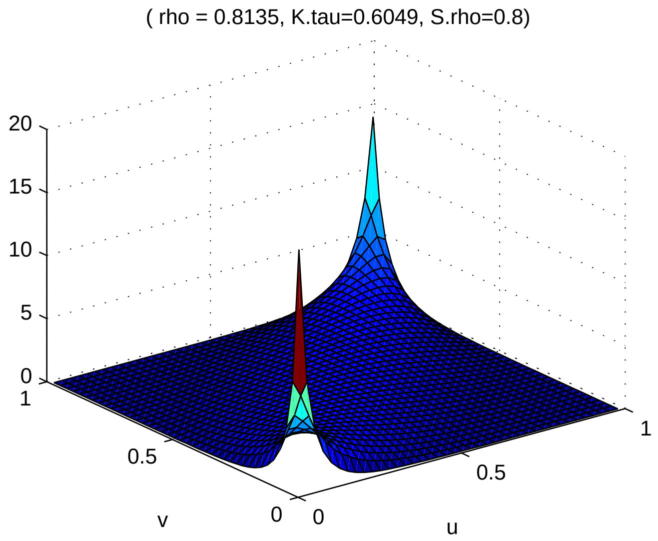

Figure 2.2: The density function of the normal copula $\theta = 0.8135$ and corresponding rank correlations: $\tau = 0.6049, \rho_S = 0.8$).

From this figure, we can see that a normal copula is symmetric. In this example we observe strong positive dependence between the variables because the mass of the density is concentrated around the diagonal $u = v$ (when the mass is concentrated on diagonal $u = 1-v$ we talk about negative dependence).

Another very important class of copulas are the Archimedean copulas.

Archimedean Copulae

There are many families of Archimedean copulae, which are characterized by a generator function. They have nice properties and are very useful in many applications.

Let ϕ be a continuous, strictly decreasing convex function $\phi : (0,1] \to [0,\infty]$ with a positive second derivative such that $\phi(1) = 0$ and $\phi(u) + \phi(v) \leq \phi(0)$.

Definition 5. *Copula $C(u,v)$ is an* Archimedean Copula *with generator ϕ if:*

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

The density function c is then:

$$c(u,v) = -\frac{\phi''(C)\phi'(u)\phi'(v)}{(\phi'(C))^3}.$$

As we know, Kendall's tau and Spearman's rho can be defined in terms of copula and they are given by (2.8) and (2.9) respectively.

We can represent Kendall's tau in terms of the generator function (Nelsen, [1]), i.e:

$$\tau = 1 + 4 \int_0^1 \frac{\phi(t)}{\phi'(t)} dt. \tag{2.12}$$

The relation between Archimedean copulas and Spearman's rho is less known. To compute Spearman's rho we use the general definition of this coefficient, i.e:

$$\rho_S = 12 \int_0^1 \int_0^1 C(u,v)\, dudv - 3. \tag{2.13}$$

Nelsen in [1], Chapter 4. parameterized 22 families of Archimedean copulae, amongst them:

- Frank Copula,
- Gumbel Copula and
- Clayton Copula.

that will be discussed discussed here. These are the most popular Archimedean copulas.

Frank's copula The generating function of Frank's family is:

$$\phi(x) = -\log \frac{e^{-\theta x} - 1}{e^{-\theta} - 1},$$

Where $\theta \in (-\infty, \infty) \setminus \{0\}$. Parameter $\theta \to 0$ implies independence, $\theta \to \infty$ means perfect positive, and $\theta \to -\infty$ perfect negative dependence.

The Frank's copula is given by:

$$C(u, v; \theta)_F = -\frac{1}{\theta} \log \left(1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{(e^{-\theta} - 1)} \right). \tag{2.14}$$

The density of Frank's copula for random variables U, V and parameter θ is given by:

$$c_\theta^F(u, v) = \frac{\theta(1 - e^{-\theta}) e^{-\theta(u+v)}}{[1 - e^{-\theta} - (1 - e^{-\theta u})(1 - e^{-\theta v})]}. \tag{2.15}$$

For Frank copula the relationship between Kendall's tau ρ_τ, Spearman's rho ρ_S and parameter θ is the following ([1]):

$$\tau(\theta) = 1 - \frac{4}{\theta} \left(1 - \frac{1}{\theta} \int_0^\theta \frac{a}{e^a - 1} da \right),$$

$$\rho_S(\theta) = 1 - \frac{12}{\theta} \left(\frac{4}{\theta} \int_0^\theta \frac{a}{e^a - 1} da - \frac{2}{\theta^2} \int_0^\theta \frac{a^2}{e^a - 1} da \right).$$

The figure (2.3) presents the Frank's copula density. We can see that the mass is distributed in a similar way as with the normal copula.

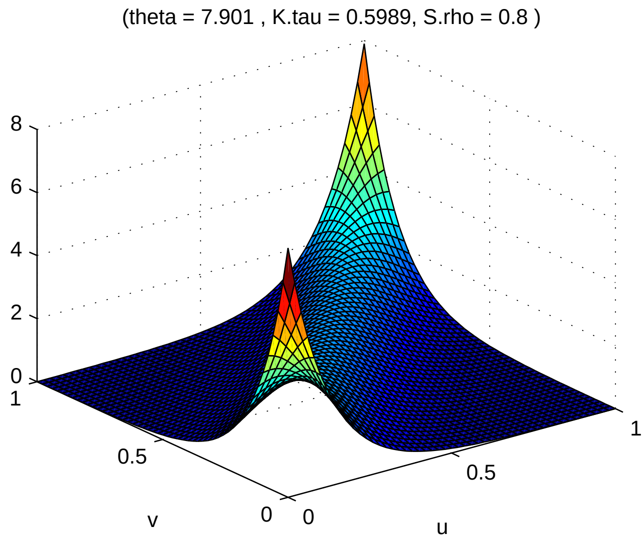

Figure 2.3: The Frank copula density function for $\theta = 7.901$ and the corresponding rank correlations: $\tau = 0.5989$, $\rho_S = 0.8$.

Gumbel's copula The Gumbel copula $C_\theta^G(u, v)$ is another member of the Archimedean family with generator

$$\phi(t) = (-\log t)^\theta.$$

The distribution function of this copula can be written as follows:

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

The parameter $\theta \geq 1$ controls the degree of dependence between variables joint by this copula. Parameter $\theta = 1$ implies an independent relationship and $\theta \to \infty$ means perfect positive dependence (the perfect negative dependence does not exist for this copula).

The density function of the Gumbel copula is given by ([18]) :

$$c_\theta^G(u,v) = \frac{(-\ln u)^{\theta-1}(-\ln v)^{\theta-1}}{uv} \exp\left(-[(-\ln u)^\theta + (-\ln v)^\theta]^{\frac{1}{\theta}}\right) \left([(-\ln u)^\theta + (-\ln v)^\theta]^{\left(\frac{1-\theta}{\theta}\right)^2} + (\theta-1)[(-\ln u)^\theta + (-\ln v)^\theta]^{\frac{1-2\theta}{\theta}}\right) \tag{2.16}$$

There is a simple relationship between the parameter θ and Kendall's tau τ.

$$\tau = 1 - \theta^{-1},$$

where $\tau \in [0,1]$.

Spearman's rho for this copula is given by (2.13) and can be calculated numerically in MATLAB;

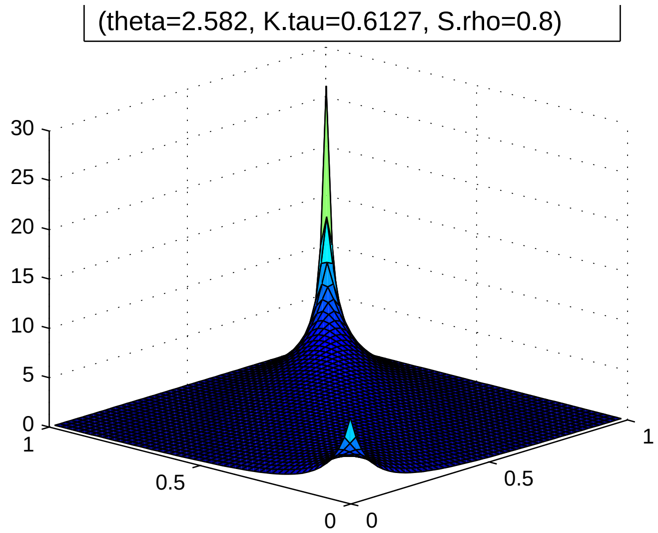

Figure 2.4: The density function Gumbel copula for parameter $\theta = 2.582$ and corresponding rank correlations: $\tau = 0.6127$, $\rho_S = 0.8$.

The Gumbel copula density is presented in Figure (2.4). From this picture we can see that the density surface is very peaky in a right-upper corner, that means the mass for Gumbel copula is concentrated in this corner. The higher the peak is the stronger the (positive) dependence is.

Clayton's copula The Clayton copula is an archimedean asymmetric copula with generator function

$$\phi(t) = \frac{t^{-\theta} - 1}{\theta},$$

where $\theta \in [-1, \infty) \setminus \{0\}$ is a parameter controlling the dependence.

The distribution of Clayton copula is equal to:

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

Perfect positive dependence is obtained if $\theta \to \infty$ and perfect negative dependence if $\theta \to -1$, while $\theta \to 0$ implies independence.

For the Clayton copula parameter θ is related with Kendall's tau in the following manner:

$$\tau = \frac{\theta}{\theta + 2}.$$

The numerical value of Spearman's rho given by (2.13) can be calculated in MATLAB.

The implicit formula of a density of Clayton copula is given by ([17])

$$c_\theta^{Cl}(u, v) = (1 + \theta)(uv)^{-1-\theta}(u^{-\theta} + v^{-\theta} - 1)^{-\frac{1}{\theta}-2}. \tag{2.17}$$

As we can see from the figure (2.5), the density of Clayton copula is peaky in the lower left corner of the unit square.

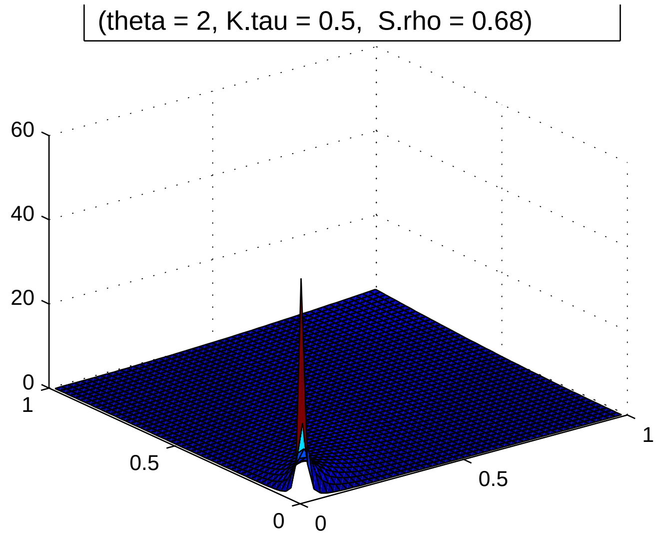

Figure 2.5: The Clayton copula density for parameter $\theta = 2.582$ and corresponding rank correlations: $\tau = 0.6127$, $\rho_S = 0.8$.

Elliptical Copula

The elliptical copula was constructed by projecting the uniform distribution on the ellipsoid in $\mathbb{R}^3$ to two dimensions (that is why this copula is called elliptical). This construction was proposed by Hardin (1982) and Misiewicz (1996).

A density function of the elliptical copula with correlation $\rho \in (-1, 1)$ is the following ([7]):

$$c_\rho^{El}(x,y) = \begin{cases} \frac{1}{\pi\sqrt{\frac{1}{4}-\frac{1}{4}\rho^2-x^2-y^2-2\rho xy}}, & \text{if } (x,y) \in B; \\ 0, & \text{if } (x,y) \notin B. \end{cases} \tag{2.18}$$

where

$$B = \{(x,y) : x^2 + (\frac{y-\rho x}{\sqrt{1-\rho^2}})^2 < \frac{1}{4}\}.$$

Note that the elliptical copula is absolutely continuous and realizes any correlation value in an interval $(-1, 1)$. You can see the graph of elliptical copula in Figure 2.6.

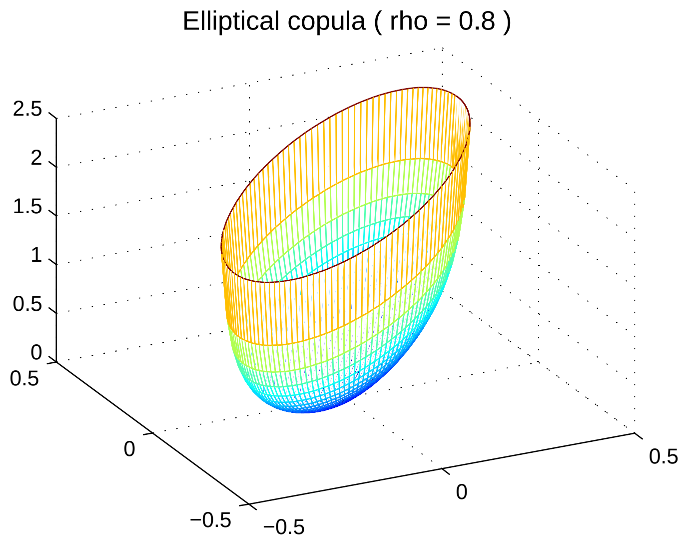

Figure 2.6: The elliptical copula density function with parameter $\rho = 0.8$.

There are elliptical copulas that correspond to elliptical distributions. The elliptical copula inherits some properties of the normal distribution, e.g. conditional correlations are constant and are equal to partial correlations (see Kurowicka, Cooke, [7] for more details).

For zero correlation the mass of the elliptical copula is concentrated on a disk (variables are not independent). So zero correlation is not a sufficient condition for independence.

Let us now concentrate on another, more sophisticated measure of dependence: the interactions.

2.6 Mixed Derivative Measures of Interaction

A measure of interaction is a function that measure dependence between two random variables. Interaction is an alternative to scalar dependence measures, such as correlation.

The interaction measure was first proposed by Holland and Wang in 1987 and then described by Whittaker in [6]. It is the mixed partial derivative of the logarithm of the density function.

Definition 6. *When the variables X_1 and X_2 are continuous the* mixed derivative measures of interaction *between X_1 and X_2 is:*

$$i_{12}(x_1, x_2) = D^2_{12} \log f_{12}(x_1, x_2) \tag{2.19}$$

where D_j denotes the ordinary partial derivative with respect to X_j, i.e. $D_j = \frac{\partial}{\partial x_j}$. The second mixed partial derivative with respect to j and k is: $D^2_{jk} = \frac{\partial^2}{\partial x_j \partial x_k}$.

From the facorisation criterion (1.1) we know that, if X_1 and X_2 are independent and if their joint density function $f_{12}(x_1, x_2)$ is (sufficiently) differentiable, then there exists functions g and h such that joint density factorises, i.e., $f_{12}(x_1, x_2) = g(x_1)h(x_2)$ for all x_1 and x_2. Hence it is easy to prove the following theorem (Whittaker, [6])

Theorem 3. *Let (X_1, X_2) be bivariate random vector, and suppose their joint* *density is differentiable then:*

$$X_1 \perp X_2 \Leftrightarrow i_{12}(x_1, x_2) = 0.$$

Proof. ($\Rightarrow$) If $X_1 \perp X_2$ then $f_{12}(x_1, x_2) = g(x_1)h(x_2)$. And the interaction of f_{12} is then:

$$\begin{aligned} i_{12}(x_1, x_2) &= D^2_{12} \log f_{12}(x_1, x_2) = D^2_{12}(\log g(x_1) + \log h(x_2)) \\ &= D_1[D_2 \log g(x_1) + D_2 \log h(x_2)] = D_1(D_2 \log h(x_2)) = 0. \end{aligned}$$

($\Leftarrow$) Let us denote $\bar{f}(x_1, x_2) = \log f_{12}(x_1, x_2)$.

If $i_{12}(x_1, x_2) = 0$ then $D^2_{12}\bar{f}(x_1, x_2) = 0$.

Integrating the above with respect to x_1 we get: $D^1_2 \bar{f}(x_1, x_2) = a(x_2)$ for some function a depending on x_2.

Integrating once more with respect to x_2 this time, we get: $\bar{f}(x_1, x_2) = A(x_2) + b(x_1)$, for some function $b(x_1)$ and $A(x_2) = \int a(x_2)dx_2$. Hence the joint density can be written as:

$$f_{12}(x_1, x_2) = e^{A(x_2)+b(x_1)} = g(x_1)h(x_2)$$

for some functions g and h, and this implies the independence of X_1 and X_2.

$\square$

The above theorem shows that in contrast to correlations, the interactions equal to zero are also sufficient for independence.

2.6.1 Bivariate distributions

We now calculate interaction measures of dependence for a few distributions (normal, Cauchy distributions, archimedean copulas).

Example 3 (The Bivariate Standard Normal Distribution). *Let $X = (X_1, X_2)$ have the standard Normal distribution $X \sim N([0,0],[1,1])$, then its joint density function is given by (2.4.1).*

Let Q be:

$$Q(x_1, x_2) = \frac{1}{(1-\rho^2)}(x_1^2 - 2\rho x_1 x_2 + x_2^2)$$

Then the logarithm of the density (2.4.1) is given by:

$$\log f_{12}(x_1, x_2; \rho) = -\log(2\pi) - \frac{1}{2}\log(1-\rho^2) - \frac{1}{2}Q(x_1, x_2)$$

The mixed derivative measure of interaction between X_1 and X_2 is then:

$$i_{12}(x_1, x_2) = -\frac{1}{2}D_{12}^2 Q(x_1, x_2) = \frac{\rho}{1-\rho^2}.$$

The normal distribution is a very special one, because the interaction for this distribution is constant. We can see that

$$i_{12}(x_1, x_2) = 0 \Leftrightarrow \rho = 0.$$

and the interaction $i_{12}(x_1, x_2)$ increases as ρ increases.

Remark 4 (The Bivariate Normal Distribution). *Let (X_1, X_2) be normally distributed random vector, $(X_1, X_2)\backslash N([\mu_1, \mu_2]; [\sigma_1, \sigma_2])$ with joint density function described by (2.4).*

Then its mixed derivative measure of interaction is:

$$i_{12}(x_1, x_2) = \frac{\rho}{(1-\rho^2)\sigma_1\sigma_2}.$$

Example 4 (Cauchy Distribution). *The bivariate Cauchy distribution was discussed in section 2.4.1 and its joint density was given by (2.6).*

And now taking the mixed derivative of the logarithm of $f_{12}(x_1, x_2)$ *we obtain the following function for the Cauchy interaction:*

$$i_{12}(x_1, x_2) = \frac{6x_1x_2}{(1 + x_1^2 + x_2^2)^2}$$

Its graph is presented in Figure 2.7.

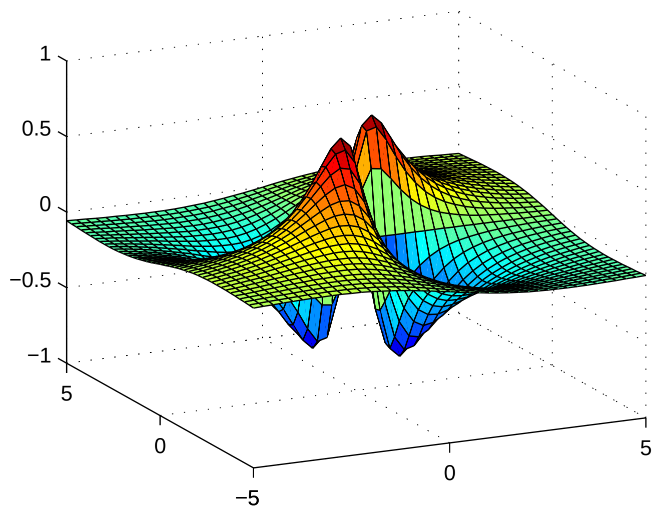

Figure 2.7: The interaction of the standard Cauchy distribution.

Observe that the interaction measure for bivariate Cauchy density changes sign, it is positive in the first and the third quadrant and negative in the second and fourth. The positive (negative) values of the interaction can be interpreted as positive (negative) dependence between variables X_1 *and* X_2*. It shows also that a constant number of measure of dependence like Pearson's correlation,* *rank correlation or Kendall's τ can not describe the sign-varying dependence structure.*

2.6.2 Copulas

Copulas were presented in section 2.5. From the definition of the density of a copula (2.10), the joint distribution can be represented as a product of marginal distributions and the copula:

$$f(x_1, x_2) = f_1(x_1)f_2(x_2)c(F_1(x_1), F_2(x_2)).$$

Observe that there is an equivalence between the interaction of this joint density and the interaction of the appropriate copula:

$$\log f(x_1, x_2) = \log f_1(x_1) + \log f_2(x_2) + \log c(F_1(x_1), F_2(x_2)),$$

and taking mixed derivatives, we obtain:

$$D_{12}^2 \log f(x_1, x_2) = D_{12}^2 \log c(F_1(x_1), F_2(x_2)),$$

which means that the interaction of the density function is equal to the interaction of the corresponding copula.

Let us determine the interactions for the copulas discussed in section 2.5.3.

Normal copula

The interaction for Normal copula of random variables with correlation coefficient ρ, is equal to:

$$i_{12}(u, v) = \frac{\rho}{1 - \rho^2},$$

which is always constant and depends on values of parameter $\rho \in [-1, 1]$.

The interaction increases for bigger ρ. The plot below (fig. 2.8) shows the interaction of normal copula as a function of parameter ρ.

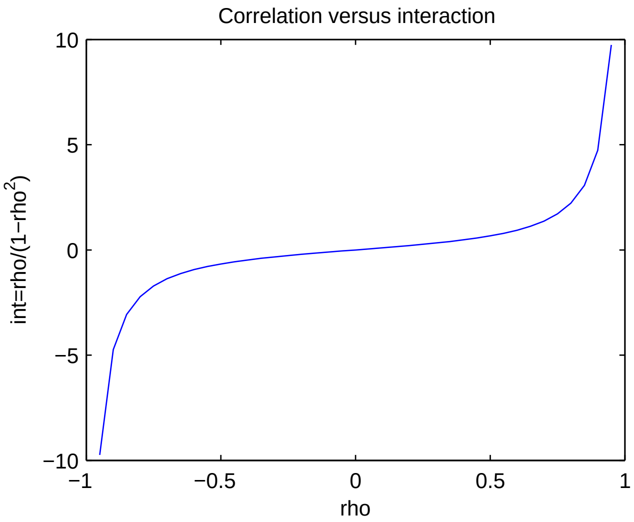

Figure 2.8: The interaction of normal copula, i_{12} as a function of ρ.

Archimedean copulas

The Bivariate Archimedean copulas were discussed in section 2.5.3, and three families of Archimedean copulas were described (Frank, Gumbel and Clayton copula). Let us calculate interactions for these families.

Frank's copula The findings for the interaction of Frank's copula are very interesting. The interaction for this copula is calculated in Maple, and takes the following form:

$$i^F(u,v) = \frac{2\theta^2(1-e^{-\theta})e^{-\theta(u+v)}}{[1-e^{-\theta}-(1-e^{-\theta u})(1-e^{-\theta v})]^2}. \tag{2.20}$$

This interaction is shown in Figure 2.9.

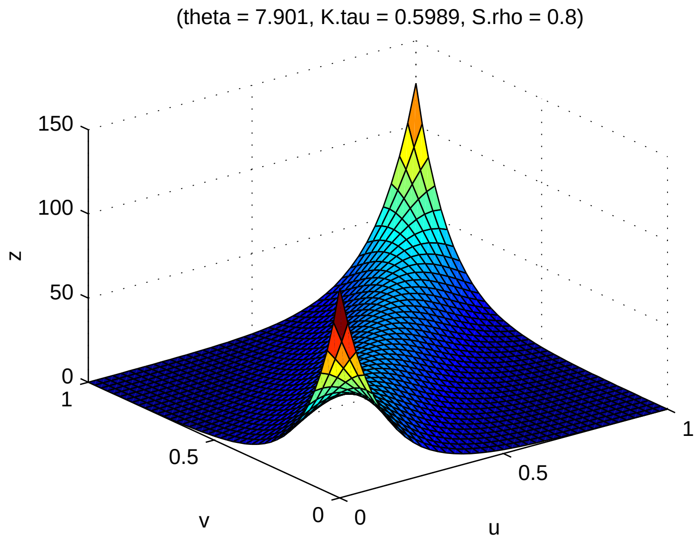

Figure 2.9: The interaction for Frank copula with parameter *theta* $= 7.901$ and rank correlations: $\tau = 0.5989$, $\rho_S = 0.8$.

If you compare this graph with the graph of Frank's copula, figure (2.3), presented in section 2.5.3, then you'll see the similarities. Their shapes are the same but the values of the interaction are much bigger than the values of the density.

Just to remind, the Frank's copula density has the following form:

$$c_\theta^F(u,v) = \frac{\theta(1-e^{-\theta})e^{-\theta(u+v)}}{[1-e^{-\theta}-(1-e^{-\theta u})(1-e^{-\theta v})]^2}. \tag{2.21}$$

Comparing (2.20) with (2.21), we see that:

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

and therefore:

$$\int_0^1\int_0^1 i^F(u,v)dudv = 2\theta\int_0^1\int_0^1 c_\theta^F(u,v)dudv = 2\theta.$$

Hence, we can normalize interactions and obtain the density of the Frank's copula. This is a very special property of this copula, which we have not confirmed for any other copulas.

Proposition 2. *Frank's copula is equal to normalized interaction of this copula, i.e:*

$$c_\theta^F(u,v) = \frac{i^F(u,v)}{\int_0^1\int_0^1 i^F(u,v)dudv}.$$

In the figure (2.10) one can see that the normalized interaction is in fact identic to the density of Frank's copula.

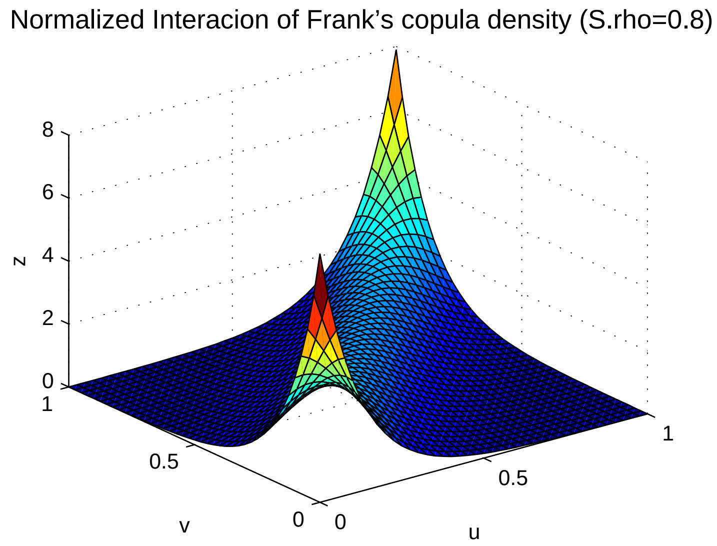

Figure 2.10: The Normalized interaction of Frank copula with parameter $theta = 7.901$ and rank correlations: $\tau = 0.5989$, $\rho_S = 0.8$.

Using the Taylor's expansion we can also show that the limit of $i^F(u,v;\theta)$ is zero when $\theta \to 0$.

We can approximate some function $f(x)$ by $T_2(x)$, where $T_2(x)$ is the quadratic approximation or a second Taylor polynomial for f based at b, such that:

$$T_2(x) = f(b) + f'(b)(x-b) + \frac{1}{2}f''(b)(x-b)^2.$$

Let us denote, the Frank's copula interaction as the fraction of two functions of θ:

$$\frac{h(\theta)}{g(\theta)} = \frac{2\theta^2(1-e^{-\theta})e^{-\theta(u+v)}}{[1-e^{-\theta}-(1-e^{-\theta u})(1-e^{-\theta v})]^2},$$

We can compute the first and the second derivative of h in Maple. We obtain: $h(0) = h'(0) = h''(0) = 0$, so the quadratic approximation $T_2(\theta)$ for the numerator h based at $\theta = 0$ is equal zero.

Hence, the approximation for the interaction based at $\theta = 0$ is also zero.

This result shows that the random variables are independent if $\theta \to 0$.

Frank's copula seems to be a very interesting case, the shape of the density function of Normal copula and the shape of Frank's copula density are so much alike, but their interactions so different. Interaction of normal copula is constant while normalized Frank's interaction is equal to the Frank's density.

Although considerable research has been devoted to the relation between interaction and Frank's copula density, it remains unclear what makes the interaction of Frank's copula so extraordinary.

It would be of interest to study properties of the Frank's copula that lead to this result.

We examine now two more examples of Archimedean copula to see if there is a similar behavior of the interactions.

Gumbel's copula The density of the Gumbel copula is given by (2.16). The interaction of this density can be computed in MATLAB. However, the general formula for this interaction is very long and is enclosed in appendix 4. The simplified form of the Gumbel copula interaction for parameter $\theta = 2$ or equivalently for Kendall's tau $\tau = 1 - \theta^{-1} = 0.5$, looks as follows:

$$
\begin{aligned}
i^G(u,v) = & D^2_{uv} \log(c^G_\theta(u,v)) \\
& = \{\log(u)^2\sqrt{\log(u)^2+\log(v)^2} + 6(\log(u)^2+\log(v)^2) \\
& + 12\sqrt{\log(u)^2+\log(v)^2} + \log(v)^2\sqrt{\log(u)^2+\log(v)^2} + 6\} \\
& \log(u)\log(v)(\log(u)^2+\log(v)^2)^2 uv((\log(u)^2+\log(v)^2)^{1/2}+1)^{-2};
\end{aligned}
$$

The interaction of Gumbel's copula is presented on the next figure (2.11).

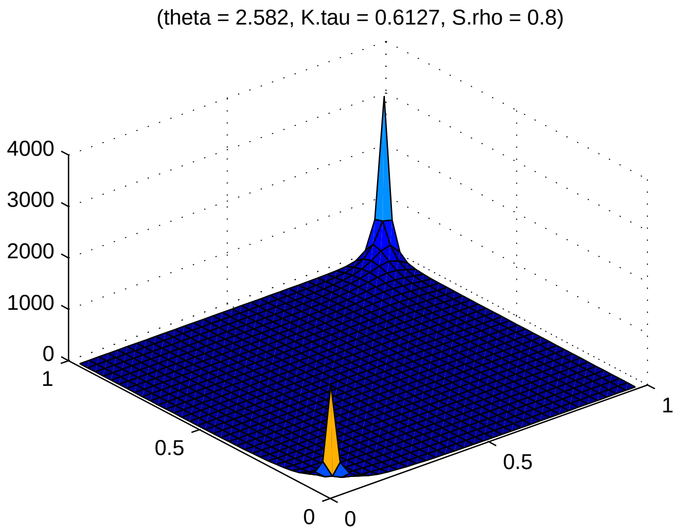

Figure 2.11: The interaction of Gumbel copula with parameter $\theta = 2.582$ and $\tau = 0.6127$, $\rho_S = 0.8$

Unfortunately the Gumbel interaction does not possess Frank's interaction property.

From the figure it appears that high values of interaction correspond to high values of the copula density, and the bigger interactions correspond with stronger dependence.

Clayton's copula The density of Clayton's copula is expressed by (2.17). The interaction of this density is computed in MATLAB and can be found in an appendix.

The interaction for $\theta = 1$ takes the simple form:

$$i^{Cl}(u,v) = \frac{-4(-v^2 + 4uv - uv^2 - u^2 - u^2v + 2u^2v^2)}{(-v - u + uv)^5}.$$

On the Figure 2.12 you can see the interaction for Clayton's copula with parameter $\theta = 2$ for which Kendall's tau is equal to: $\tau = \frac{\theta}{\theta+2} = 0.5$ and Spearman's rho $\rho_S = 0.68$.

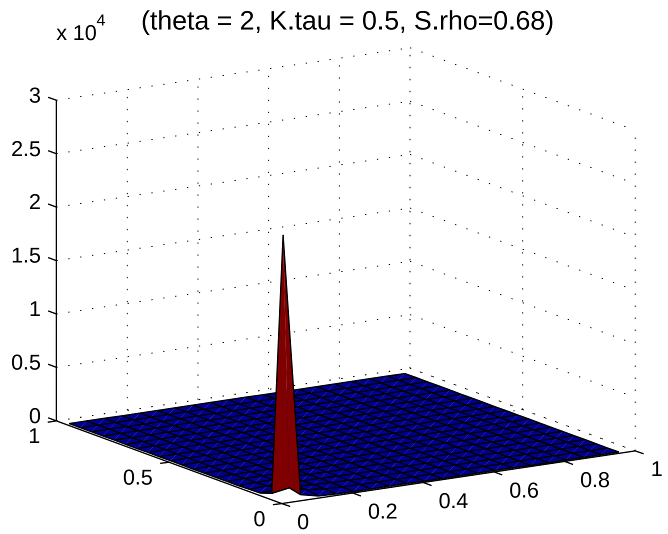

Figure 2.12: The interaction of Clayton copula with parameter $\theta = 2$ and $\tau = 0.5$, $\rho_S = 0.68$

In the figures of Gumbel's and Clayton's interactions we can see that the values in one of the corners rise up to infinity very quickly. Intuitively, this means that Clayton copula assigns more probability mass to the region in the upper left corner while Gumbel assigns more probability mass to the region in the right upper corner.

Concluding, the above examples show that the Frank copula is a special distribution, when normalize, the mixed derivative measure of interaction is equal to the joint density.

This surely is a starting point for further investigation. It would be of interest to learn more about Frank's copula property, and search for other families of copulas with similar properties.

Elliptical Copula

For the elliptical copula with the density (2.18) the interaction is as follows:

$$i^{El}(u,v;\rho) = \frac{2uv + \rho u^2 + \rho v^2 + 0.25\rho(1-\rho^2)}{[0.25 - 0.25\rho^2 - u^2 - v^2 - 2uv\rho]^2}$$

for all u, v such that the point (u, v) belongs to: $\{u^2 + (\frac{v-\rho u}{\sqrt{1-\rho^2}})^2 < \frac{1}{4}\}$.

For parameter value $\rho = 0$ the interaction is equal to $i^{El}(u,v;0) = \frac{2uv}{(\frac{1}{4}-(u^2+v^2))^2}$.

2.7 Summary

This chapter was concerned with the bivariate dependence concepts. We summarize shortly our findings. The dependency between two random variables is perfectly characterized by their joint distribution. The dependence between random variables can be measured by correlation coefficients (linear correlation, rank correlation, Kendall's tau). Well known bivariate distribution are elliptical distributions, and especially the normal distribution. For this distribution there exists a relationship between the product moment, rank correlation and Kendall's tau. One can study the marginals separately from the dependency structure by means of copulas. The most popular copulas are: normal and family of Archimedean copulas, for which Kendall's tau plays very important role. We can also discuss the dependence by simply observing the joint density or functions of the joint density, in particular so called interactions. Zero interaction correspond to independence. In contrast to correlations, the converse is also true. It is easy to show that interaction of joint distribution equals theinteraction of corresponding copula. For normal copula (or equivalently normal distribution) the interaction is constant while the interaction of Frank's copula is equal to normalized density of this copula.

The next chapter is an extension of those dependence concepts to n dimensions.

Chapter 3

Multidimensional Dependence Concept

This chapter is dedicated to the multivariate dependence concepts. Previously introduced dependence concepts between two random variables are extended to the dependency in n dimensional random vector $(X_1, \cdots, X_n)$, where two random variables are conditioned on all the other variables.

I start this chapter with introducing the measures of conditional dependence: a conditional correlation and a partial correlation coefficient, which in a sense, correspond to the product moment correlation. Partial and conditional correlations are equal for the elliptical distributions, but in general they are not.

The multivariate joint distribution of a random vector contains whole information about this vector, hence also contain the information about the dependence between random variables. As an example, the multivariate normal distribution is described. All multivariate distributions with continuous margins have their corresponding multivariate copula, which is an extension of the bivariate copula. Multivariate normal and Archimedean copulas are presented.

To represent multivariate dependence we need to collect all the bivariate measures of dependence between two variables into a positive definite matrix (for instance joint multivariate normal distribution requires covariance matrix). This however imposes difficulties. Kurowicka, Cooke [7] show an example, that positive definite rank correlation matrix transformed to correlation matrix is no longer positive definite. Moreover, for large matrices it is very unlikely to get a positive definite matrix. Therefore, other methods for specifying dependence must be used.

The copula-vine method uses the conditional dependence to construct multidimensional distributions from the marginal distributions and from the dependence structure between variables. We can represent the dependence structure via regular vine. In case of joint normal distribution, for which partial and conditional correlation are equal, we can specify the vine with conditional rank correlation or either conditional or partial correlation. Bredford, Cooke [10] show, that the partial correlations on a regular vine are algebraically independent and there is one-to-one correspondence with correlation matrices. This approach is considered to be alternative way of specifying multivariate distribution.

In chapter 2 we discussed measures of interactions, now we define the conditional interactions for multidimensional random vector. Conditional interactions are based onthe multivariate joint distributions. They are mixed distributions of logarithm of conditional distribution with respect to conditional variables. We can write joint density using the copula-vine method and then take the conditional interactions of such density function. Later, we investigate interactions for vine distributions.

Lets start with measures of conditional independence. The notions introduced here, were described in [7].

3.1 Conditional Correlation

While product moment correlation describes the linear relationship between two random variables, the conditional correlation describes the relationship between two variables while conditioning on other variables.

Let us take the partition (X_i, X_j, X_a) of the n-dimensional random vector $X = (X_1, X_2, \cdots, X_n)$, where $a = \{1, 2, \cdots, n\} \setminus \{i, j\}$, so the vector X_a is a vector composed of all the other variables except X_i and X_j.

Definition 7 (Conditional Correlation). *The* conditional correlation of X_i and X_j given X_a *denoted by* $\rho_{X_iX_j|X_a}$ *(or simply* $\rho_{ij|a}$*) is a product moment correlation of* X_i *and* X_j *conditioned on* X_a *with respect to the conditional distribution between* X_i *and* X_j *conditioned on* X_a*.*

$$\rho_{X_iX_j|X_a} = \rho(X_i|X_a, X_j|X_a) = \frac{E(X_iX_j|X_a) - E(X_i|X_a)E(X_j|X_a)}{\sigma(X_j|X_a)\sigma(X_j|X_a)} \tag{3.1}$$

The conditional correlation is 'an extension' of ordinary linear correlation and has similar properties :

- it ranges from -1 to $+1$,
- if X_i and X_j are independent given X_a, denoted by $X_i \perp X_j | X_a$, then $\rho_{X_iX_j|X_a} = 0$,

- if $\rho_{X_iX_j|X_a} = 0$ then $X_i \perp X_j | X_a$ if and only if X is elliptically distributed random vector.

3.2 Partial Correlation

A partial correlation describes the relationship between two variables, whilst the other variables are kept constant.

The partial correlation $\rho_{12;3,\cdots,n}$ represents the correlation between the orthogonal projections of X_1 and X_2 on the plane orthogonal to the space spanned by $X_3, \cdots, X_n$

The partial correlation is defined in the following way:

Definition 8. *Let X_i be random variables with $E(X_i) = 0$ and standard deviations $\sigma_i = 1$ for $i = 1, \cdots, n$ and let the numbers $b_{12;3,\cdots,n}, \cdots, b_{1n;2,\cdots,n-1}$ minimize the following expected value:*

$$E\left((X_1 - b_{12;3,\cdots,n}X_2 - \cdots - b_{1n;2,\cdots,n-1}X_n)^2\right)$$

Then partial correlation is defined as ([7]):

$$\rho_{12;3,\cdots,n} = sgn(b_{12;3,\cdots,n})\sqrt{b_{12;3,\cdots,n}b_{21;3,\cdots,n}}. \tag{3.2}$$

Partial correlations can be computed from correlations with the following recursive formula ([7]):

$$\rho_{12;3,\cdots,n} = \frac{\rho_{12;3,\cdots,n-1} - \rho_{1n;3,\cdots,n-1}\rho_{2n;3,\cdots,n-1}}{\sqrt{1-\rho^2_{1n;3,\cdots,n-1}}\sqrt{1-\rho^2_{2n;3,\cdots,n-1}}} \tag{3.3}$$

The partial correlation has similar properties to those of conditional correlation, moreover there exists relationship between partial and conditional correlation. Because partial correlation is easier to calculate, sometimes it is more convenient to replace conditional with the partial correlation.

Baba, Shibata and Sibuya in [14] have discussed the relationship between the partial correlation and the conditional correlation, showing that they are equivalent for elliptical distributions. They suggest, that the linearity of conditional expectation is a key property for this equivalence.

In general, outside the family of elliptical distributions, the partial and conditional correlations are not equal.

Kurowicka, Cooke in [7] show an example, where the zero conditional correlation does not imply zero partial correlation. They took X, Y, Z such that: X was uniformly distributed on $[0,1]$, $Y \perp Z | X$, hence $\rho_{YZ|X} = 0$. And $Y|X$, $Z|X$ were uniformly distributed on $[0, X^k]$, $k > 0$. For those kind of random variables the difference: $|\rho_{YZ|X} - \rho_{YZ;X}|$ converge to $\frac{3}{4}$ as $k \to \infty$.

So for conditionally independent Y, Z given X, their partial correlation is not zero.

The conditional and partial correlations measure the dependence but this is a distribution of random vector $X = (X_1, \cdots, X_n)$ that contain all information about the dependence between those variables.

3.3 Multidimensional distributions

If we have an n-dimensional random vector $(X_1, \cdots, X_n)$, then its joint distribution is called multivariate or n-dimensional distribution. To represent multivariate dependence all of the measures of dependence between pairs of variables (correlation, rank correlation or Kendall's tau) must be collected into $n \times n$ dependence matrix. This matrix must be complete and positive definite.

The multivariate normal distribution discussed below, is a specific and very important distribution because it possess many desired properties in probability theory and statistics.

3.3.1 Multivariate normal distribution

A random vector $X = (X_1, \cdots, X_n)$ follows a multivariate normal distribution, symbolically denoted by $X \sim N(\mu, \Sigma)$, if there is a vector $\mu = (\mu_1, \cdots, \mu_n)$ and a symmetric, positive definite covariance matrix Σ ($n \times n$ matrix), such that X has density:

$$f(x) = \frac{1}{\sqrt{(2\pi)^2|\Sigma|}} \exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right),$$

where $x = (x_1, \cdots, x_n)$, $|\Sigma|$ is the determinant of Σ, μ is a vector of expected values of X ($\mu_i = E(X_i)$), while components of Σ are the covariances ($\Sigma_{ij} = Cov(X_i, X_j)$).

Of course it is true that: if a random vector has a multivariate normal distribution then, any two or more of its components that are uncorrelated, are independent (because zero correlation implies independence for elliptical distributions). But it is not true, that two separate random variables that are normally distributed and uncorrelated, are independent (two random variables that are normally distributed may fail to be jointly normally distributed).

Example 5. *Let* $\mathbf{X}$ *be a bivariate random vector with components* X_1 *and* X_2*. Let* X_1 *and* Z *be independent standard normal random variables, and define:*

$$X_2 = sign(X_1)|Z|$$

where sign function returns 1 *if* $X_1 \geq 0$ *and returns* -1 *if* $X_1 < 0$*. In this case, both* X_1 *and* X_2 *are standard normal, but the vector* $\mathbf{X}$ *is not joint normal. The distribution of* $X_1 + X_2$ *has a substantial probability of being equal to* 0*, whereas the normal distribution, as a continuous distribution has no discrete part. Consequently* X *and* Y *are not jointly normally distributed, even though they are separately normally distributed.*

Another very important property of normal distribution is the equivalence of partial and conditional correlation, i.e:

$$\rho_(X_i X_j | X_a) = \rho_{X_i X_j; X_a}.$$

For the proof refer to Kurowicka, Cooke [7].

Just like bivariate distribution, we can represent any continuous multidimensional distribution as a product of marginals and a corresponding multidimensional copula. The copula, being a dependence structure, tells us about the relations between random variables $X_1, \cdots, X_n$.

The notions of multivariate copulas described here, were introduced by Nelsen in [1].

3.4 Multivariate Copulas

The concept of copula a bivariate distribution on the unit square with uniform marginals, can be extended to the multivariate case. We can define multivariate copula as a multidimensional distribution on an unit hypercube with uniform marginal distributions.

Let $X_1, \cdots, X_n$ be random variables. The role of copulas as dependence functions justifies the Sklar's theorem. Sklar's theorem was previously presented for bivariate distributions in section 2.5.1, here is its multivariate version.

Theorem 5 (Sklar's theorem). *Let H denote n dimensional distribution function with marginal distributions $F_1, \cdots, F_n$. Then there exists the n dimensional copula C such that for all $(x_1, \cdots, x_n)$:*

$$H(x_1, \cdots, x_n) = C(F(x_1), \cdots, F(x_n))$$

If all marginals are continuous then the copula is unique. The converse of the above statement is also true.

Proposition 3. *Let $F_1^{-1}, \cdots, F_n^{-1}$ denote the inverses of marginal distributions, then for every $(u_1, \cdots, u_n)$ there exists unique copula C such that:*

$$C(u_1, \cdots, u_n) = H(F_1^{-1}(u_1), \cdots, F_1^{-1}(u_n)).$$

From this proposition we know that given any marginals and a copula we can construct a joint distribution. The copula density contains all information about dependence in the random vector.

By applying Sklar's theorem we can derive the multivariate copula density:

$$c[F_1(x_1), \cdots, F_n(x_n)] = \frac{f(x_1, \cdots, x_n)}{\prod_{i=1}^{n} f_i(x_i)}.$$

And hence, the multivariate joint density can be written as:

$$f(x_1, \cdots, x_n) = f_1(x_1) \cdots f_n(x_n) c[F_1(x_1), \cdots F_n(x_n)]$$

The most popular copula in multidimensional modelling is the normal copula, also known as Gaussian.

3.4.1 Multivariate Normal Copula

If $\mathbf{\Phi_\Sigma}$ is the multivariate normal cumulative distribution function with correlation matrix Σ then the distribution function of normal copula is given by:

$$C_N(u_1, \cdots, u_N) = \mathbf{\Phi_\Sigma}^n(\Phi^{(-1)}(u_1), \cdots, \Phi^{(-1)}(u_n)).$$

The expression for the copula density c_N is as follows:

$$c_N(u_1, \cdots, u_N) = \frac{1}{\sqrt{|\Sigma|}} \exp\left(-\frac{1}{2}\zeta^T(\Sigma^{-1} - \mathbf{I})\zeta\right)$$

where $\zeta = (\Phi^{(-1)}(u_1), \cdots, \Phi^{(-1)}(u_n))$, $\mathbf{I}$ is the $n \times n$ identity matrix and $\Phi^{(-1)}$ is the inverse of the standard univariate normal distribution.

Using c_N as a dependence function, the joint multivariate normal density is given by:

$$f(x_1, \cdots, x_n) = f_1(x_1) \cdots f_n(x_n) exp\left(-\frac{1}{2}\zeta^T(\Sigma^{-1} - \mathbf{I})\zeta\right)$$

where $f_i(x_i)$ is a marginal density function of X_i, $i = 1, \cdots, n$.

Another important class of multivariate copulas is a family of Archimedean copulae. The bivariate Archimedean copulas described in chapter 2 can be generalized to higher dimensions, for more details on Archimedean copulas refer to Nelsen [1].

3.4.2 Archimedean multivariate copulae

The extension of the 2-dimensional Archimedean copulas defined in the definition 5, results in writing the n-dimensional Archimedean copula with $u = (u_1, \cdots, u_n)$, in the following form:

$$C^n(u) = \phi^{-1}(\phi(u_1) + \cdots + \phi(u_n))$$

The function ϕ is defined like in the definition 5, and the functions C^n are the *serial iterates* of the bivariate Archimedean copula generated by ϕ, i.e.:

$$C^n(u_1, \cdots, u_n) = C(C^{n-1}(u_1, \cdots, u_{n-1}), u_n).$$

For instance, if we set

$$C^2(u, v) = C(u, v) = \phi^{-1}(\phi(u) + \phi(v)) = \tilde{u}.$$

Then

$$\begin{aligned} C^3(u, v, z) =& C(C^2(u, v), z) = C(\tilde{u}, z) = \phi^{-1}[\phi(\tilde{u}) + \phi(z)] \\ &= \phi^{-1}[\phi(\phi^{-1}(\phi(u) + \phi(v))) + \phi(z)] \\ &= \phi^{-1}[\phi(u) + \phi(v) + \phi(z)]. \end{aligned}$$

Frank's copula Let $\phi(t) = -\log\left(\frac{e^{-\theta t}-1}{e^{-\theta}-1}\right)$ and $\theta > 0$, this function generates the bivariate Frank's family.

We can generalize the Frank's family of 2-dimensional copulas to a family of n-dimensional copulas for any $n \geq 2$:

$$C^n(u) = -\frac{1}{\theta} \log \left(1 + \frac{\prod_{i=1}^{n} (e^{-\theta u_i} - 1)}{(e^{-\theta} - 1)^{n-1}} \right).$$

Notice that there is only one parameter θ that can be specified. Hence this distribution specifies exchangeable model. There are ways to specify that way more complicated structures, but it is not trivial to find out which parameter choices specify consistent model [Joe,(1997)].

The closed form of multivariate Frank's copula density is not known, however we can compute this density for small n in Maple. The three dimensional Frank copula density is:

$$c^F(u_1, u_2, u_3) = \frac{3\theta^2(1-e^{-\theta})^2 e^{-\theta(u_1+u_2+u_3)}[(1-e^{-\theta})^2 + (1-e^{-\theta u_1})(1-e^{-\theta u_2})(1-e^{-\theta u_3})]}{[(1-e^{-\theta})^2 - (1-e^{-\theta u_1})(1-e^{-\theta u_2})(1-e^{-\theta u_3})]^3}. \tag{3.4}$$

Clayton copula

We can generalize the Clayton family of bivariate copulas to a family of n-dimensional copulas for parameter $\theta > 0$, the generator $\phi(t) = t^{\theta} - 1$ and any $n \geq 2$:

$$C^n(u) = (u_1^{-\theta} + \cdots + u_n^{-\theta} - n + 1)^{-\frac{1}{\theta}}.$$

This copula has main advantage in the set of the multivariate Archimedean copulas, its density as shown in [19] is easy to compute and is given by:

$$c^n(u) = \left(1 - n + \sum_{i=1}^{n} u_i^{-\theta}\right)^{-n-\frac{1}{\theta}} \prod_{j=1}^{n} \left(u_j^{-\theta-1}\{(j-1)\theta+1\}\right). \tag{3.5}$$

Gumbel copula

We can generalize the Gumbel family of bivariate copulas to a family of n-dimensional copulas for $\theta \geq 1$, generator function $\phi(t) = (-\log t)^{\theta}$, and any $n \geq 2$:

$$C^n(u) = exp\left(-[(-\log u_1)^{\theta} + \cdots + (-\log u_n)^{\theta} - n + 1]^{\frac{1}{\theta}}\right).$$

The three dimensional density of Gumbel copula can be computed in MATLAB, however it appears as a very long expression and is not included in this thesis. In Appendix, reader can find MATLAB code which generates the density of three dimensional Gumbel copula as well as its interaction function.

Representing multivariate dependence in the ways described above has several disadvantages.

To represent multivariate dependence a correlation matrix must be specified, however there are several restrictions to this approach. The correlation matrix must be complete and positive definite. So when we have missing data we can not work with partially specified correlation structure (Kurowicka, Cooke [7]).

Moreover, to specify multivariate copula, a set of parameters has to be provided. It is in general not obvious which values of a parameter (or parameters) specify consistent model and which dependence structures can be obtained with given copula (Joe, 1997).

These drawbacks are reasons to seek for other methods of representing multidimensional distributions.

The method known as a *copula-vine method* uses conditional dependence to construct multidimensional distributions from marginal distributions (marginals can be obtained from data or experts) and the dependency structure between random variables represented by a vine (quantification of a vine is done by experts).

3.5 Vines

A *vine* is a graphical modek introduced by Bedford, Cooke [9, 10] (2001) and then studied by Kurowicka, Cooke [7, 8]. A *copula-vine method* allows us to specify joint distribution from bivariate and conditional bivariate pieces.

A *vine* on N variables is a nested set of trees [1], where the edges of tree j are the nodes of tree $j+1$; $j = 1, \cdots, N-2$, and each tree has the maximum number of edges. A regular vine on N variables is a vine in which two edges in tree j are joined by an edge in tree $j+1$ only if these edges share a common node, $j = 1, \cdots, N-2$.

Definition 9 (Vine). *V is a **vine** on n elements if*

1. $V = (T_1, \cdots, T_{n-1})$

[1] $T = (N, E)$ is a **tree** with nodes $N = 1, 2, \cdots, n$ and edges E where E is a subset of unordered pairs of N with no cycle. That is, there does not exist a sequence $a_1, \cdots, a_k$ where $k > 0$ of elements of N such that

$$a_1, a_2 \in E, \cdots\ a_{k-1}, a_k \in E, a_k, a_1 \in E.$$

The **degree** of node $a_i \in E$ is the number of edges attached to a_i.

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

   V is a ***regular vine*** *on n elements if additionally:*

3. ***proximity*** *For $i = 2, \cdots, n-1$, if $a = a_1, a_2$ and $b = b_1, b_2$ are nodes of T_i connected by an edge in T_i, then exactly one of the a_i equals one of the b_i.*

We recall two special kinds of vines, the D-vine and C-vine .

Definition 10. *A regular vine is called a:*

D-vine *if each node in T_1 has degree at most 2, see*

C-vine *if each tree T_i has a unique node of degree $n-i$. The node with maximal degree in T_1 is the root.*

For each edge of a vine we define *constraint, conditioned* and *conditioning sets* of this edge as follows ([7]):

Definition 11 (Conditioning, conditioned and constraint sets).

1. *For $j \in E_i, i \leq n-1$ the subset $U_j(k)$ of $E_{i-k} = N_{i-k+1}$, defined by*
   $U_j(k) = \{e | \exists e_{i-(k-1)} \in e_{i-(k-2)} \in \cdots \in j, e \in e_{i-(k-1)}\}$,
   is called the ***k-fold union***, *$U_j^* = U_j(i)$ is the* ***complete union*** *of j, that is, the subset of $\{1, \cdots, n\}$ consisting of m-descendants of j.*
   If $a \in N_1$ then $U_a^* = \emptyset$.
   $U_j(1) = \{j_1, j_2\} = j$.
   By definition we write $U_j(0) = \{j\}$.

2. *The **constraint set** associated with $e \in E_i$ is U_e^*.*

3. *For $i = 1, \cdots, n-1, e \in E_i, e = \{j, k\}$, the **conditioning set** associated with e is*

$$D_e = U_j^* \cap U_k^*$$

*and the **conditioned set** associated with e is*

$$\{C_{e,j}, C_{e,k}\} = U_j^* \triangle U_k^* = \{U_j^* \setminus D_e, U_k^* \setminus D_e\}$$

the order of node e is number of elements of D_e

Note that for $e \in E_1$, the conditioning set is empty. For $e \in E_i, i \leq n-1, e = \{j, k\}$ we have $U_e^* = U_j^* \cup U_k^*$.

3.5.1 Regular vine specification

Each edge in a regular vine is associated with a constant conditional rank correlation, such that conditioning variables are equal to conditioning set and conditioned variables to conditioned set, this rank correlation is denoted by $r_{C_{e,j},C_{e,k}|D_e}$.

In chapter 2., we have shown that the rank correlation is a measure of dependence between two random variables joined by the copula. So the vine builds the joint distribution by specifying copulas and conditional copulas according to the structure of a vine.

The rank correlation specification on regular vine and copula determines the whole joint distribution. The procedure of sampling such a distribution can be written for any regular vine, see [7] for more details on sampling a vine.

The figure (3.1) represents the rank correlation specification on a D-vine on four nodes.

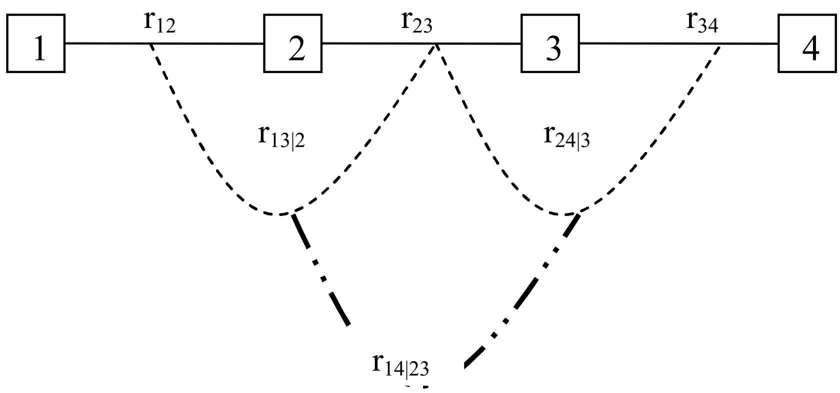

Figure 3.1: D-vine

3.5.2 The density function

Let assume that for each edge $e \in T_m$ where $m = 1, \cdots, n-1$ and each possible value of the variables in the conditioning set D_e a copula is specified. The corresponding copula to an edge ij given D_e is denoted as $C_{ij|D_e}$ and its density as $c_{ij|D_e}$ while the density of this copula is written as c_d. Also, marginal distributions F_j are specified for each $j \in N$.

The density of a vine is defined as follows:

Definition 12. *Let $V = (T_1, \cdots, T_n)$ be a regular vine on n elements. Given F_i and $C_{ij}|D_e$ defined as above, there is a unique vine dependent distribution with density given by*

$$f_{1\cdots n} = f_1 \cdots f_n \prod_{m=1}^{n-1} \prod_{e \in E_m} c_{ij|D_e} \left(F_{i|D_e}, F_{j|D_e}\right) \tag{3.6}$$

where e is an edge labelled $ij|D_e$.

Example 6. *This example shows how this can be applied to obtain the density* *of a vine dependent distribution for the regular vine in the figure 3.5.1.*

$$
\begin{aligned}
f_{1234} &= f_{14|23} f_{23} \\
&= c_{14|23}\left(F_{1|23}, F_{4|23}\right) f_{1|23} f_{4|23} f_{23} \\
&= c_{14|23}\left(F_{1|23}, F_{4|23}\right) \frac{f_{123} f_{234}}{f_{23}} \\
&= c_{14|23}\left(F_{1|23}, F_{4|23}\right) \frac{f_{13|2} f_2 f_{24|3} f_3}{f_{23}} \\
&= c_{14|23}\left(F_{1|23}, F_{4|23}\right) c_{13|2}\left(F_{1|2}, F_{3|2}\right) c_{24|3}\left(F_{2|3}, F_{4|3}\right) \frac{f_2 f_3 f_{1|2} f_{3|2} f_{2|3} f_{4|3}}{f_{23}} \\
&= c_{14|23}\left(F_{1|23}, F_{4|23}\right) c_{13|2}\left(F_{1|2}, F_{3|2}\right) c_{24|3}\left(F_{2|3}, F_{4|3}\right) \frac{f_{12} f_{23} f_{34}}{f_2 f_3} \\
&= c_{14|23}\left(F_{1|23}, F_{4|23}\right) c_{13|2}\left(F_{1|2}, F_{3|2}\right) c_{24|3}\left(F_{2|3}, F_{4|3}\right) \\
&\quad c_{12}\left(F_1, F_2\right) c_{23}\left(F_2, F_3\right) c_{34}\left(F_3, F_4\right) f_1 f_2 f_3 f_4 .
\end{aligned}
$$

3.5.3 Partial correlation specification

Here, the edges of a regular vine are associated with the partial correlations in the following manner: with values chosen arbitrarily in the interval $(-1, 1)$ in the following way: to every $e \in E_i$ with either conditioned and conditioning variables $\{j, k\}$ and D_e respectively, we associate partial correlation value: $\rho_{\mathbf{j},\mathbf{k};\mathbf{D_e}}$, where $i = 1, \cdots, n-1$. It is very convenient to do calculations with partial correlations, such vine is called **partial correlation vine**.

The following theorem says that each such partial correlation vine specification uniquely determines the correlation matrix, see Bedford, Cooke [10].

Theorem 6. *For any regular vine on n elements there is a one to one correspondence between the set $n \times n$ full rank correlation matrices and the set of partial correlation specification for the vine.*

All assignments of the numbers from $(-1, 1)$ to the edges of a partial correlation regular vine are consistent, in the sense that there is a joint distribution realizing these partial correlations, and all correlation matrices can be obtained this way.

3.5.4 Normal Vines

The normal vine is a special regular vine with rank and conditional rank correlation assigned to its edges. The copula used is a normal copula. This is equivalent with transforming all marginal distributions to standard normal and taking all conditional distributions to be normal distributions. Hence, the normal vine simply allows us to specify joint normal distribution with given correlation matrix, hence we can avoid problems encountered with positiveness and completeness of the dependency matrix.

This procedure may be described as follows. Suppose random variables $X_1, \cdots, X_n$ correspond to nodes of regular vine V with specified (conditional) rank correlations $r_{ij|D_e}$ for any edge $e \in V$. We can create a partial correlation vine V' by assigning partial correlation $\rho_{ij;D_e}$ to every edge of V, where:

$$\rho_{ij;D_e} = 2\sin(\frac{\pi}{6} r_{ij|D_e}).$$

Let R denote the correlation matrix determined by the partial correlations $\rho_{ij;D_e}$ of V'. Then we sample a joint normal distribution $(Y_1, \cdots, Y_n)$ with standard normal margins and correlation R.

In chapter 2 we have discussed mixed derivative measures of interactions, we have looked at its properties when applied to bivariate joint distributions.

Now, as an extensions of bivariate interactions, we investigate mixed derivatives of conditional interactions. These are related to multivariate distributions.

3.6 Mixed Derivative Measures of Conditional Interaction

Consider the partition (X_i, X_j, X_a) of the $X = (X_1, X_2, \cdots, X_n)$, where $a = \{1, 2, \cdots, n\} \setminus \{i, j\}$, so the vector X_a is a vector composed of all the other variables except X_i and X_j.

Whittaker in [6] defines mixed derivative measure of conditional interaction as following:

Definition 13. *When the variables X_i and X_j are continuous the* mixed derivative measures of conditional interaction *between X_i and X_j conditioning on X_a is:*

$$i_{ij|a}(x_i, x_j; x_a) = D^2_{ij} \log f_{ij|a}(x_i, x_j; x_a).$$

It is easy to show that the mixed derivatives of the joint and of the conditional density functions are the same; that is

$$D^2_{ij} \log f_{ij\cdots n}(x_i, \cdots, x_n) = D^2_{ij} \log f_{ij|a}(x_i, x_j; x_a),$$

because $\log f_{ij|a} = \log f_{ij\cdots n} - \log f_a$.

Analogous to bivariate case, the necessary and sufficient condition for conditional independence, $X_i \perp X_j | X_a$, is that

$$i_{ij|a}(x_i, x_j; x_a) = 0.$$

For better understanding next calculations are performed on 3 dimensional random vector. Later the discussion will be extended to n dimensions.

Theorem 7. *Let (X_1, X_2, X_3) be a 3-dimensional random vector. Suppose that their joint density function is differentiable. Then:*

$$X_1 \perp X_2 | X_3 \Leftrightarrow i_{12|3} = 0.$$

Proof. ($\Rightarrow$) Suppose $X_1 \perp X_2 | X_3$, then the joint density function f of (X_1, X_2, X_3) may be written as:

$$f(x_1, x_2, x_3) = g(x_1, x_3) h(x_2, x_3) m(x_3).$$

From the definition of the interaction measure:

$$\begin{aligned} i_{12|3} &= D^2_{12} \log f(x_1, x_2, x_3) = \\ &D_1[D_2 \log g(x_1, x_3) + D_2 \log h(x_2, x_3) + D_2 \log m(x_3)] = \\ &D_1 D_2 \log h(x_2, x_3) = 0 \end{aligned}$$

($\Leftarrow$) Now assume that $i_{12|3} = 0$ and let $\bar{f}(x_1, x_2, x_3) = \log f(x_1, x_2, x_3)$. Then

$$i_{12|3} = D^2_{12} \bar{f}(x_1, x_2, x_3) = 0.$$

Integrating first over x_1 we get:

$$D_2 \bar{f}(x_1, x_2, x_3) = a(x_2, x_3),$$

for some function $a(x_2, x_3)$. Integrating now with respect to x_2 we have

$$\bar{f}(x_1, x_2, x_3) = A(x_2, x_3) + b(x_1, x_3),$$

for some function $b(x_1, x_3)$, where $D_2 A(x_2, x_3) = a(x_2, x_3)$. Finally $f(x_1, x_2, x_3) = e^{A(x_2,x_3)+b(x_1,x_3)} = g(x_1, x_3)h(x_2, x_3)$ which implies that

$$X_1 \perp X_2 | X_3.$$

$\square$

This theorem shows the equivalence between conditional independence and zero conditional interaction. It was shown in chapter 2, that zero correlation do not imply the independence (except for elliptical distribution), while interactions do. This is main reason to consider interaction as a better measure than correlation.

Lets calculate the conditional interactions for three dimensional normal distribution The following calculations were presented by Whittaker in [6], therefore further on in this text I may refer to it, as a Whittaker approach.

3.6.1 The Three Dimensional Standard Normal Distribution

The standardized 3-dimensional normal probability density function of three dimensional random vector (X_1, X_2, X_3) can be written as:

$$f_{123}(x_1, x_2, x_3) = \frac{1}{\sqrt{(2\pi)^3 |\Sigma|}} \exp\{-\frac{1}{2} Q(x_1, x_2, x_3)\}, \qquad (3.7)$$

where Σ is covariance matrix, $|\Sigma|$ is the determinant of Σ, i.e: $|\Sigma| = det(\Sigma)$, and Σ^{-1} is its inverse matrix.

$$\Sigma = \begin{pmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{21} & 1 & \rho_{23} \\ \rho_{31} & \rho_{32} & 1 \end{pmatrix}$$

The matrix Σ is symmetric with determinant equal to:

$|\Sigma| = 1 + 2\rho_{12}\rho_{13}\rho_{23} - \rho_{12}^2 - \rho_{13}^2 - \rho_{23}^2$,

The inverse covariance matrix is given by:

$$\Sigma^{-1} = \frac{1}{|\Sigma|} \begin{pmatrix} 1-\rho_{23}^2 & \rho_{12}-\rho_{13}\rho_{23} & \rho_{12}\rho_{23}-\rho_{13} \\ \rho_{12}-\rho_{13}\rho_{23} & 1-\rho_{13}^2 & \rho_{23}-\rho_{12}\rho_{13} \\ \rho_{12}\rho_{23}-\rho_{13} & \rho_{23}-\rho_{12}\rho_{13} & 1-\rho_{12}^2 \end{pmatrix}$$

The quadratic form Q in (3.7) equals to:

$$\begin{aligned} Q(x_1, x_2, x_3) &= (x_1, x_2, x_3)'\Sigma^{-1}(x_1, x_2, x_3) \\ &= a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + 2a_{23}x_2x_3 \end{aligned}$$

where a_{ij} are elements of inverse covariance matrix Σ^{-1}.

Observe that in the expansion of $Q(x_1, x_2, x_3)$ there are no terms higher than quadratic. This is very special for normal distribution, that we can write the logarithm of its density (3.7) as:

$$\log f_{123}(x_1, x_2, x_3) = constant - \frac{1}{2}Q(x_1, x_2, x_3). \tag{3.8}$$

In order to get the mixed derivative measure of conditional interaction, $i_{23|1}$, we have to take the mixed partial derivative of (3.8) with respect to X_2 and

X_3 given X_3. In formula for $Q(x_1, x_2, x_3)$, there is only one $x_2 x_3$ term with coefficient: $2a_{23}$. That is:

$$\begin{aligned} i_{23|1}(x_1, x_2, x_3) &= D^2_{23} \log f_{123}(x_1, x_2, x_3) \\ &= -a_{23} \\ &= -\frac{1}{|\Sigma|}(\rho_{23} - \rho_{12}\rho_{13}). \end{aligned}$$

Similarly we derive expressions for $i_{13|2}$ and $i_{12|3}$, i.e.

$i_{13|2}(x_1, x_2, x_3) = -a_{13}$

$i_{12|3}(x_1, x_2, x_3) = -a_{12}$

You can see that these interactions are constant. This result is unusual, no other distribution was found for which the interaction takes constant values. The measures of interactions are usually functions, and hardly ever a constant.

Observe the relationship between the element of inverse covariance matrix a_{23}, and the partial correlation $\rho_{23;1}$, namely:

$$a_{23} = 0 \Leftrightarrow \rho_{23;1} = 0,$$

this comes directly from the definition of partial correlation:

$$\rho_{23;1} = \frac{\rho_{23} - \rho_{12}\rho_{13}}{\sqrt{(1-\rho^2_{12})(1-\rho^2_{13})}},$$

where

$$\rho_{23;1} = 0 \Leftrightarrow (\rho_{23} - \rho_{12}\rho_{13}) = 0.$$

Naturally, we would get the same result calculating interactions for the normal copula because we can represent joint normal distribution as a product of marginals and a normal copula, where marginals do not influence the dependence structure.

Whittaker considers the case when (X_1, X_2, X_3) is normally distributed random vector, characterized by complete 3×3 Pearson's correlation matrix Σ:

$$\Sigma = \begin{pmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{21} & 1 & \rho_{23} \\ \rho_{31} & \rho_{32} & 1 \end{pmatrix}.$$

But we can also represent the joint normal distribution via regular vine. It is enough to consider the partial correlation specification on a vine. Then the partial correlations determine a unique, complete correlation matrix (by theorem 6).

Assume, that we have given the partial correlations specification on a D-vine of three variables, see Figure 3.2.

From given partial correlations we can calculate correlation matrix Σ, where element ρ_{13} is obtained from the formula below:

$$\rho_{13} = \rho_{13;2}\sqrt{(1-\rho_{12}^2)(1-\rho_{23}^2)} + \rho_{12}\rho_{23}. \tag{3.9}$$

Having correlation matrix Σ, determined from partial correlations, the interactions are computed in the similar way as in Whittaker's approach.

The interaction $i_{13|2}$ is written in terms of given partial correlations ρ_{12},ρ_{23} and $\rho_{13;2}$:

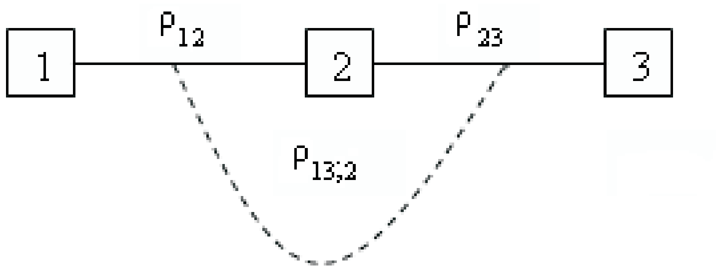

Figure 3.2: D-vine specified by the correlation : ρ_{12},ρ_{23} and $\rho_{13;2}$.

$$i_{13|2} = \frac{1}{|\Sigma|}(\rho_{12}\rho_{23} - \rho_{13}) = -\frac{1}{|\Sigma|}(\rho_{13;2}\sqrt{(1-\rho_{12}^2)(1-\rho_{23}^2)}).$$

Therefore $i_{13|2} = 0 \Leftrightarrow \rho_{13;2} = 0$.

Let us look at the other interaction:

$$i_{23|1} = \frac{1}{|\Sigma|}(\rho_{23} - \rho_{12}\rho_{13}) = \frac{1}{|\Sigma|}(\rho_{23} - \rho_{12}\rho_{13;2}\sqrt{(1-\rho_{12}^2)(1-\rho_{23}^2)} - \rho_{12}^2\rho_{23}). \tag{3.10}$$

We know (from Whittaker's approach) that $i_{23|1} = 0 \Leftrightarrow \rho_{23;1} = 0$. Because $i_{23|1}$ is a function of ρ_{12},ρ_{23} and $\rho_{13;2}$, we can find the relationship between these partial correlations and $\rho_{23;1}$ by solving the equation: $i_{23|1} = 0$.

This solution is obtained via MAPLE:

$$i_{23|1} = 0 \Leftrightarrow \{\rho_{12} = 0, \rho_{23} = 0\}, \{\rho_{13;2} = \frac{\rho_{23}(1-\rho_{12}^2)}{\rho_{12}\sqrt{(1-\rho_{12}^2)(1-\rho_{23}^2)}}\}.$$

The relationship between given partial correlations specified on a vine (Figure 3.2) and the partial correlation $\rho_{23;1}$ is:

$\rho_{23;1}$ is equal to zero (or equivalently $X_2 \perp X_3 | X_1$) if either both ρ_{12} and ρ_{23} are 0, or the partial correlation $\rho_{13;2}$ is of the form: $\rho_{13;2} = \frac{\rho_{23}(1-\rho_{12}^2)}{\rho_{12}\sqrt{(1-\rho_{12}^2)(1-\rho_{23}^2)}}$.

3.6.2 Three dimensional Archimedean Copulae

The multivariate Archimedean copulas were discussed in section 3.4.2. The interactions of three dimensional Archimedean copulas can be calculated in MATLAB. They appear in a very long expressions and are not included in this thesis. However, in appendix you can find a MATLAB code to generate those interactions for Frank, Clayton and Gumbel copulas.

These three Archimedean copulas are characterized by one parameter only, hence we can consider the conditional interaction as a function of parameter θ.

$$i_{12|3}(\theta) = D_{12} \log c(u, v, z; \theta),$$

where c is the density of either Frank, Clayton or Gumbel copula.

The solution of $i_{12|3}(\theta) = 0$ for Clayton's copula is computed in MATLAB, and is equal to 0, i.e: $\theta = 0$.

For Frank's and Gumbel's copula the explicit solution could not be found, however the limits of their interactions, $\lim_{\theta \to 0} i_{12|3}(\theta)$ are zero.

3.6.3 D-Vine

Earlier, in section , we have discussed conditional interactions of three dimensional normal distribution specified by partial correlations on a d-vine on three nodes (figure(3.2)). Now, we generalize this approach and consider a d-vine, where nodes are joined by a copula and a conditional copula.

Let X_i be random variable with F_i-its cumulative distribution function and f_i-its density function, where $i = 1, 2, 3$ and c_{ij} a joint copula of (X_i, X_j). Lets consider d-vine with three nodes. Then the joint density function is defined as:

$$\begin{aligned} f_{123}(x_1, x_2, x_3) \quad &= f_1(x_1)f_2(x_2)f_3(x_3)c_{12}(F_1(x_1), F_2(x_2))c_{23}(F_2(x_2), F_3(x_3)) \\ &\quad c_{13|2}(F_{1|2}(x_1; x_2), F_{3|2}(x_3; x_2)). \end{aligned} \tag{3.11}$$

Then its logarithm is:

$$\begin{aligned} \log f_{123}(x_1, x_2, x_3) = \log f_1(x_1) + \log f_2(x_2) + \log f_3(x_3) \\ + \log c_{12}(F_1(x_1), F_2(x_2)) + \log c_{23}(F_2(x_2), F_3(x_3)) \\ + \log c_{13|2}(F_{1|2}(x_1; x_2), F_{3|2}(x_3; x_2)). \end{aligned}$$

And the interactions can be written as:

$$\begin{aligned} i_{13|2} &= D^2_{13} \log c_{13|2}(F_{1|2}(x_1; x_2), F_{3|2}(x_3; x_2)). \\ i_{12|3} &= D^2_{12} \log \left(c_{12}(F_1(x_1), F_2(x_2))c_{13|2}(F_{1|2}(x_1; x_2), F_{3|2}(x_3; x_2))\right); \\ i_{23|1} &= D^2_{12} \log \left(c_{23}(F_1(x_1), F_2(x_2))c_{13|2}(F_{1|2}(x_1; x_2), F_{3|2}(x_3; x_2))\right). \end{aligned}$$

The following proposition is a simple application of theorem (7)

Proposition 4. *If (X_1, X_2, X_3) is a 3-dimensional random vector with the density given by (3.11), then*

1. $X_1 \perp X_3 | X_2$ *if and only if*

$$c_{13|2}(F_{1|2}(x_1; x_2), F_{3|2}(x_3; x_2)) = e^{A(x_2,x_3)+B(x_1,x_2)};$$

2. $X_1 \perp X_2 | X_3$ *if and only if*

$$c_{12}(F_1(x_1), F_2(x_2))c_{13|2}(F_{1|2}(x_1; x_2), F_{3|2}(x_3; x_2)) = e^{A(x_1,x_3)+B(x_2,x_3)};$$

3. $X_2 \perp X_3 | X_1$ *if and only if*

$$c_{23}(F_1(x_1), F_2(x_2))c_{13|2}(F_{1|2}(x_1; x_2), F_{3|2}(x_3; x_2)) = e^{A(x_1,x_2)+B(x_1,x_3)}.$$

for some functions $A(x_2, x_3), B(x_1, x_2)$.

Hence, the interaction is equal 0 if appropriate copula or the product of two copulas can be represented as a function of separated variables.

The following example illustrates how this proposition applies to a normal copula.

Example 7 (Normal copula). *Let* (X_1, X_2, X_3) *be 3-dimensional normal random vector with (standard) normal marginal distributions* X_1, X_2, X_3.

A product of copulas as described in the point 2 of proposition 4 can be derived from (3.11), hence:

$$c_{12}c_{13|2} = \frac{f_{123}(x_1, x_2, x_3)}{f_1(x_1)f_2(x_2)f_3(x_3)c_{23}(F_2(x_2), F_3(x_3))}; \tag{3.12}$$

where $c_{12}c_{13|2}$ *is an abbreviation of* $c_{12}(F_1(x_1), F_2(x_2))c_{13|2}(F_{1|2}(x_1; x_2), F_{3|2}(x_3; x_2))$.

Lets remind the formula for the joint three dimensional standard normal density:

$$f_{123}(x_1, x_2, x_3) = \frac{1}{\sqrt{(2\pi)^3|\Sigma|}} \exp\{-\frac{1}{2}Q(x_1, x_2, x_3)\}, \tag{3.13}$$

The product of marginal distributions can be written as:

$$f_1(x_1)f_2(x_2)f_3(x_3) = \frac{1}{\sqrt{(2\pi)^3}} \exp\{-\frac{1}{2}(x_1^2 + x_2^2 + x_3^2)\}, \tag{3.14}$$

and the three dimensional normal copula as:

$$c_{23}(F_3(x_3), F_3(x_3)) = \frac{1}{\sqrt{1-\rho_{23}^2}} \exp\{-\frac{1}{2(1-\rho_{23}^2)}(x_2^2 - 2\rho_{23}x_2x_3 + x_3^2)\} \exp\{\frac{1}{2}(x_2^2 + x_3^2)\} \tag{3.15}$$

Inserting (3.13), (3.14) and (3.15) into the formula for $c_{12}c_{13|2}$ *(3.12) we derive the following formula:*

$$c_{12}c_{13|2} = \frac{\sqrt{1-\rho_{23}^2}}{|\Sigma|} \exp\{-\frac{1}{2}(a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + 2a_{23}x_2x_3)\}$$
$$\exp\{\frac{1}{2}(x_1^2)\} \exp\{\frac{1}{2(1-\rho_{23}^2)}(x_2^2 - 2\rho_{23}x_2x_3 + x_3^2)\}.$$

Grouping elements of the above expression in such way to contain products of functions of separated variables, we get

$$c_{12}c_{13|2} = \frac{\sqrt{1-\rho_{23}^2}}{|\Sigma|} \exp\{-\frac{1}{2}(a_{11}x_1^2 + a_{33}x_3^2 + 2a_{13}x_1x_3 - x_1^2)\}$$
$$\exp\{-\frac{1}{2}(a_{22}x_2^2 + 2a_{23}x_2x_3 - \frac{1}{(1-\rho_{23}^2)}(x_2^2 - 2\rho_{23}x_2x_3 + x_3^2))\} \exp\{-a_{12}x_1x_2\}$$

Hence, we have shown that, $c_{12}c_{13|2}$ *can be rewritten as a product of the following functions:*

$$c_{12}c_{13|2} = \exp\{A(x_1, x_3)\} \exp\{B(x_2, x_3)\} \exp\{-a_{12}x_1x_2\}.$$

This expression satisfies the point 2 of proposition 3.11 if and only if $a_{12} = 0$*. This is equivalent to:* $X_1 \perp X_2 | X_3 \Leftrightarrow \rho_{12;3} = 0$*, where* a_{12} *is an element of inverse covariance matrix* Σ*.*

So far we have considered conditional interactions of three dimensional random vector. We can now generalize our discussion to high dimensions. The multivariate normal distribution plays very important role in high dimensional modelling and is discussed in this section.

3.6.4 Multivariate standard normal distribution

The multivariate normal distribution was described section 3.3.

It was said that the normal density is a special one because it can be written in a log linear form,i.e:

$$\log f(\mathbf{x}) = const - \log(|\Sigma|) - \frac{1}{2}\mathbf{x}^T\Sigma^{-1}\mathbf{x}.$$

for $\mathbf{x} = (x_1, \cdots, x_n) \in R^n$, where Σ is positive definite covariance matrix, and $|\Sigma|$ is the determinant of Σ.

It is a quadratic form in terms of coordinates of $\mathbf{x}$ and can be expanded to:

$$\begin{aligned}\log f(x) &= const - \log(|\Sigma|) - \frac{1}{2}\left(\sum_{i}^{n} a_{ii}x_i^2 + \sum_{ij}^{n} 2a_{ij}x_ix_j\right)\\ &= const - \log(|\Sigma|) - \frac{1}{2}\sum_{i}^{n} a_{ii}x_i^2 - \sum_{ij}^{n} a_{ij}x_ix_j,\end{aligned}$$

where $i < j$ and a_{ij}'s are the elements of the inverse covariance matrix Σ.

Then the interaction between x_i and x_j (the i-th and j-th derivative of $\log f(x)$) is entirely determined by a_{ij}, and the following holds:

$$a_{ij} = 0 \Leftrightarrow X_i \perp X_j | X_a.$$

Whittaker shows in [6], that the coefficients a_{ij} can be interpreted as partial variances and partial correlations:

$$\begin{aligned} a_{ii} = &\quad \tfrac{1}{var(X_i|X_a)}, \text{ and} \\ a_{ij} = &\quad -\rho(X_i, X_j|X_a)\sqrt{a_{ii}a_{jj}} = -\rho(X_i, X_j|X_a)\sqrt{var(X_i|X_a)}\sqrt{var(X_j|X_a)}. \end{aligned}$$

So, the criterion for pairwise conditional independence is that $a_{ij} = 0$ and consequently the $\rho(X_i, X_j|X_a) = 0$.

Example 8 (The 4-dimensional Standard Normal Distribution). *The $Q(\mathbf{x})$ for 4-dimensional normally distributed random vector can be written as:*

$$Q(x) = (x_1, x_2, x_3, x_4) \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

and this is equal to the quadratic form:

$Q(x) = x_1^2 a_{11} + x_2^2 a_{22} + x_3^2 a_{33} + x_4^2 a_{44} + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + 2a_{14}x_1x_4 + 2a_{23}x_2x_3 + 2a_{34}x_3x_4$.

The logarithm of joint density f is:

$$\begin{aligned} \log f(\mathbf{x}) = const - \frac{1}{2}Q(x) = const - \frac{1}{2}(x_1^2 a_{11} + x_2^2 a_{22} + x_3^2 a_{33} + x_4^2 a_{44}) \\ - (a_{12}x_1x_2 + a_{13}x_1x_3 + a_{14}x_1x_4 + a_{23}x_2x_3 + a_{34}x_3x_4) \end{aligned}$$

and the mixed interactions correspond, of course to the elements a_{ij}, for instance:

$$i_{13|24} = -a_{13}$$

$$X_1 \perp X_3 | X_2 \Leftrightarrow i_{13|2} = 0$$

Chapter 4

Summary and conclusions

The thesis study different aspects of how to measure the dependence. The key terms throughout this work were:

- Product moment correlation ;
- Rank correlations (Spearman's rho and Kendall's tau);
- Partial and conditional correlation;
- Copulas;
- Vines;
- Mixed derivative measures of (conditional) interactions.

Product moment correlation is by far the most used measure to test dependence, it is easy to calculate but it has several disadvantages, just to name few:

- product moment correlation is only defined when the expectations and variances of random variables are finite.
- In general zero correlation does not imply independence (this become applicable for elliptical distributions);
- It depends on the marginal distributions.

The product moment correlation outside the world of elliptical distributions may induce misleading conclusions and therefore can not reveal all the desired information about the dependency hidden in a joint distribution. More flexible measures are rank correlations. These are: Spearman's ρ_S and Kendall's τ. As opposed to the product moment correlation coefficient, they can be expressed in terms of copula C.

Both, Kendall's tau and Spearman's rho are very useful. They possess properties which are not shared by linear correlation.

- They may be considered as measures of the degree of monotonic dependence, whereas product moment correlation measures the degree of linear dependence only;
- They are invariant under monotone transformations, while the linear correlation is not;
- $\rho_S(X, Y)$ and $\tau(X, Y)$ depend only on the copula of (X, Y);
- if $X \perp Y$ then $\rho_S(X, Y) = \tau(X, Y) = 0$.

The dependence between random variables is characterized by their joint distribution. We can talk about joint distributions in terms of copula, this is very convenient way when studying the dependence because the copula separate the marginal distributions from dependence structure. From a practical point of view, the advantages of the copula approach are:

- copula connects a given joint distribution function to its marginal distributions,
- we can link together any two marginal distributions and any copula and we'll get a valid joint distribution function,
- the marginal distributions for the components of a multivariate distribution can be selected freely, and then linked with a suitable copula. So the dependence structure can be modelled independently of the marginal distributions.

Often the dependence structure between X and Y is very complicated and describing it by scalar-based measures is not suitable. In this case we take an advantage of mixed derivative measure of interactions. Interactions which are functions calculated from the density, can reflect more complicated dependence structures.

Copulas are used to represent either bivariate and multivariate distributions. Alternative way of representing multivariate distributions is by specifying the dependence structure using vines, the graphical model for conditional dependence. The Vine model allows to give input conditional (rank) correlations. The advantage of this approach is that there are no joint restrictions on the correlations by contrast the correlation matrix for joint normal distribution must be positive definite in case of Pearson product moment correlations.

Finally, the conditional interactions measure the dependence for multivariate joint distributions. It is shown that the mixed derivative measure of conditional interaction when applied to multivariate normal distribution is equal to the appropriate element of the inverse covariance matrix. The elements of such matrix are interpreted as partial variances and partial correlations.

The aim of this thesis was to study the relationships between interactions and other well known measures of dependence. Throughout the research we have found many interesting properties of this measure of dependence, like:

- Interactions depend on the density function and not on the data;
- Interaction is a function that reflects complexity of the dependence;
- Interactions of the joint density and the corresponding copula are equal;
- The zero interaction for normal distribution corresponds to a zero partial correlation;
- The normalized Frank's interaction and the Frank's copula density are the same;
- They are sufficient and necessary condition for independence, i.e.: $X \perp Y \Leftrightarrow i(X,Y) = 0$.

Observe, that the last bullet does not hold either for product moment or rank correlations.

The interaction is a very attractive tool to model the dependence (as a function interaction contains much more information about the dependency than a scalar based measures). From the other side it is an extraordinary measure (for normal distributions it is constant and corresponds to the product moment correlation and for Frank's copula, when normalized, is equal to this copula density). It would be of interest to investigate and learn more about interactions and their properties. One can study the interactions of discreet random variables (the derivatives are replaced by differences in this case).

In applications, interactions can be used in multivariate vine distributions to measure the conditional dependence. Because the lack of the time unable me to study this approach, it is a starting point for future research.

Matlab codes for Archimedean copulas

1.Three dimensional Frank Copula

The interaction of 3-dimensional Frank copula (given by (3.4)) can be calculated using either MAPLE or MATLAB. The result produced by MATLAB is in more condensed form than the formula obtained in MAPLE, therefore here I enclose the expression I have obtained in MATLAB.

These commands produce the interaction `inter`, of three dimensional Frank copula:

```
syms u v z t
CDF=1/t*log(1-[(1-exp(-t*u))*(1-exp(-t*v))*...
    ...(1-exp(-t*z))]/(1-exp(-t))^2)
PDF=diff(diff(diff(CDF,u),v),z)
fun=log(PDF);
inter=diff(diff(fun,u),v)
```

Where `CDF` is a cumulative distribution function of Frank copula $C^F(u,v,z)$, the `PDF` is the density function $c^F(u,v,z)$ given by (3.4) and `inter` is the interaction of this density.

The following set of commands produces the values of Spearman's rho (`spearman`) and Kendall's tau (`kendall`) for Frank's copula with given parameter θ.

```
warning off all
spearman = 1 + 12/t*(2/t^2*quad(inline('a.^2./(exp(a)-1)','a'),0,t)...
          ...-1/t*quad(inline('a./(exp(a)-1)','a'),0,t))
kendall=1-(4/t)*(1-1/t*quad(inline('a./(exp(a)-1)','a'),0,t))
warning on
```

Where `t` denotes the parameter θ.

2.Bivariate Clayton Copula

The interaction of the density of Clayton copula is given by (2.17) and can be calculated using either MAPLE or MATLAB. The result produced by MATLAB is in more condensed form than the formula obtained in MAPLE, therefore here I enclose the expression I have obtained in MATLAB.

The following set of commands produces the interaction `inter` of bivariate Clayton copula.

```
syms u v t
fun=log((1+t)*(u.*v).^(-t-1).*(u.^(-t)+v.^(-t)-1).^(-2-1/t));
z=diff(diff(fun,u),v);
inter = simplify(z)
```

inter=

$-t(14v^{7t}tu^{2t} + 42u^{6t}v^{3t}t - 280v^{7t}u^{5t}t - 420u^{6t}v^{6t}t - 84u^{7t}tv^{3t} - 42v^{8t}u^{6t}t + 210u^{7t}v^{6t}t - 14u^{8t}v^{2t}t + 42u^{3t}v^{6t}t + 14v^{2t}tu^{7t} - 70u^{8t}tv^{4t} + 42u^{8t}tv^{3t} + 105v^{7t}u^{6t} + 21v^{8t}u^{3t} + 105u^{4t}v^{7t} + 7v^{8t}u^{7t} - 105v^{6t}u^{4t} - u^{8t}v^{8t} - 35u^{4t}v^{8t} + 35u^{5t}v^{4t} - 21v^{6t}u^{8t} + 210u^{6t}v^{5t} - 42v^{7t}u^{7t} - 140v^{5t}u^{5t} + 35u^{5t}v^{8t} + u^{8t}v^{t} + 7u^{8t}v^{7t} - 42u^{3t}v^{7t} + 210u^{5t}v^{6t} + 35u^{4t}v^{5t} + 35u^{8t}v^{5t} + 105u^{7t}v^{4t} + 7v^{7t}u^{2t} + 21u^{6t}v^{3t} - 140v^{7t}u^{5t} - 210u^{6t}v^{6t} - 42u^{7t}v^{3t} - 21v^{8t}u^{6t} + 105u^{7t}v^{6t} - 7u^{8t}v^{2t} + 21u^{3t}v^{6t} + 7v^{2t}u^{7t} - 35u^{8t}v^{4t} + 21u^{8t}v^{3t} - 105u^{6t}v^{4t} - 140u^{7t}v^{5t} - 7v^{8t}u^{2t} + u^{t}v^{8t} - 210u^{6t}v^{4t}t + 42v^{8t}tu^{3t} + 2u^{t}v^{8t}t + 210u^{4t}v^{7t}t + 14v^{8t}tu^{7t} - 210v^{6t}tu^{4t} - 2u^{8t}v^{8t}t - 70u^{4t}v^{8t}t + 70u^{5t}v^{4t}t - 42v^{6t}u^{8t}t + 420u^{6t}v^{5t}t - 84v^{7t}tu^{7t} - 280v^{5t}u^{5t}t + 210v^{7t}tu^{6t} + 70u^{5t}v^{8t}t + 2u^{8t}tv^{t} + 14u^{8t}v^{7t}t - 84u^{3t}v^{7t}t + 420u^{5t}v^{6t}t + 70u^{4t}v^{5t}t + 70u^{8t}v^{5t}t + 210u^{7t}v^{4t}t - 280u^{7t}v^{5t}t - 14v^{8t}tu^{2t})/u/v/(-v^{t} - u^{t} + u^{t}v^{t})^{4}/(-v^{2t} - 2u^{t}v^{t} + 2u^{t}v^{2t} - u^{2t} + 2u^{2t}v^{t} - u^{2t}v^{2t})/(v^{3t} + 3u^{t}v^{2t} - 3u^{t}v^{3t} + 3u^{2t}v^{t} - 6u^{2t}v^{2t} + 3u^{2t}v^{3t} + u^{3t} - 3u^{3t}v^{t} + 3u^{3t}v^{2t} - u^{3t}v^{3t})$

3. Three dimensional Clayton Copula

To generate the interaction of three dimensional Clayton's copula, we use the following set of commands:

```
   syms u v z t
CDF=(u^(-t)+v^(-t)+z^(-t)-2)^(-1/t);
PDF=diff(diff(diff(CDF,u),v),z);
fun=log(PDF);
inter=diff(diff(fun,u),v)
```

Where command `inter` produces this interaction.

4. Two dimensional Gumbel Copula

The following MATLAB code is used to calculate the interaction of the density of two dimensional Gumbel copula (2.16).

```
syms u v t;
z=log([exp(-[(-log(u)).^t+(-log(v)).^t].^(1/t)).*...
  ...[(-log(u)).*(-log(v))].^(t-1)]./...
  ...[u.*v.*((-log(u)).^t+(-log(v)).^t).^(2-1/t)].*...
  ...[((-log(u)).^t+(-log(v)).^t).^(1/t)+t-1]);
d=diff(diff(z,u),v)
```

where `z` is the logarithm of the density (2.16), and `d` is the mixed derivative of `z`, hence the command `d` produces the formula for the interaction of the Gumbel density.

5. Two dimensional Gumbel Copula

The next set of commands produces the 3-dimensional density (`PDF`) and 3-dimensional interaction of this density (`inter`) for three dimensional Gumbel copula given by the distribution function (`CDF`).

```
syms u v z t
CDF=exp(-[(-log(u))^t+(-log(v))^t+(-log(z))^t]^(1/t));
PDF=diff(diff(diff(CDF,u),v),z)
fun=log(PDF);
inter=diff(diff(fun,u),v)
```

Bibliography

[1] R.B. Nelsen *An Introduction to Copulas*, Springer-Verlag New York, Inc., 1999.

[2] R. T. Clemen, T. Reilly, *Correlations and copulas for decision and risk analysis*, 1997;

[3] M. Dorey, P. Joubert, *Modelling Copulas: An Overview*, The Staple Inn Actuarial Society;

[4] S. Demarta, A. J. McNeil,*The t Copula and Related Copulas*, Zurich, 2004;

[5] N. L. Johnson, S. Kotz, N. Balakrishan *Continuous univariate distributions, volume 1*, Wiley Series in Probability and Mathematical Statistics, 1994;

[6] J. Whittaker, *Graphical models in Applied Multivariate Statistics*, John Wiley & Sons, 1990;

[7] D. Kurowicka, R.M. Cooke, *Uncertainty Analysis with High Dimensional Dependence Modelling*, TUDelft, 2005;

[8] D. Kurowicka, R.M. Cooke, *Completion problem with partial correlation vines*, TUDelft, 2005;

[9] T. Bedford, R. Cooke, *Probability density decomposition for conditionally dependent random variables modeled by vines?*, Annals of Mathematics and Artifical Intelligence, 2001;

[10] T. Bedford, R. Cooke, *Vines - a new graphical model for dependent random variables*, 2001;

[11] T. S. Ferguson, *A Class of Symmetric Bivariate Uniform Distributions*, 1994;

[12] J. Bojarski, *A new class of band copulas - distributions with uniform marginals*, Juournal of Mathematical Sciences, Vol. 111, No. 3, 2002;

[13] D. Fantazzini, *Copula's Conditional Dependence Measures for Portfolio Management and Value at Risk*, University of Konstanz;

[14] K. Baba, R. Shibata, M. Sibuya *Partial Correlation and Conditional Correlation as MEasures of Conditional Independence*, Australian Statistical Publishing Association Inc., 2004;

[15] H. M. Kat, *The Dangers of Using Correlation to Measure Dependence*, City University, London, 2002;

[16] S. Nadarajah, K. Mitov, S. Kotz *Local dependence functions for extreme value distributions*, Journal of Applied Statistics, 2003;

[17] P.G. Sankaran, R.P. Gupta *Characterizations using local dependence function* Communications in statistics, 2004;

[18] K. Aas *Modelling the dependstructure of financial assets: A survey of four copulas*, Norwegian Computing Centre, 2004;

[19] E. Cuvelier, M. Noirhomme-Fraiture *Clayton copula and mixture decomposition*, University Notre-Dame de La Paix;