Copulas and Local Gaussian Correlation

Tommy Nordbø

Master's thesis in statistics

Mathematical statistics

Department of Mathematics
University of Bergen

May 2012

Contents

1 Introduction **5**

2 Copula **6**
2.1 Mathematical introduction 6
2.2 Sklars theorem 7
2.3 Properties and examples 8
2.3.1 Density 8
2.3.2 Increasing transformations 9
2.4 Survival copulas 9
2.5 Symmetry 10
2.5.1 Radial symmetry 10
2.5.2 Exchangeable symmetry 11

3 Global dependence measures **13**
3.1 Linear correlation 13
3.2 Perfect dependence 14
3.3 Kendalls tau and Spearmans rho 14
3.4 Quadrant dependence 16
3.5 Tail dependence 17

4 Local Gaussian Correlation **19**
4.1 Introduction 19
4.2 Existence of $\theta_b(x)$ 20
4.3 Non-linear transformations of Gaussian variables 21
4.4 Theoretical LGC 22
4.5 Some properties 24
4.5.1 Limits 24
4.5.2 LGC and tail dependence 24
4.6 Local correlation plot 24
4.7 Example: The Gaussian distribution 25
4.8 Symmetry 26
4.8.1 Radial symmetry 26
4.8.2 Reflection symmetry 27
4.8.3 Exchange symmetry 28

5 LGC and copulas **29**
5.1 Theoretical LGC for a copula 29
5.2 Some properties 31
5.2.1 The sign of the theoretical LGC 31
5.2.2 Independence 31
5.3 Comments on problems and limitations 31
5.3.1 Uniqueness 31
5.3.2 Restrictions on C 32
5.3.3 Complicated expression 32
5.3.4 Choice of margins 32

6 Ellitpical distributions and copulas **34**
6.1 Elliptical distributions 34
6.2 Elliptical copulas 37
6.2.1 Theoretical LGC 37
6.2.2 The Gaussian copula 39
6.2.3 t copula 41

7 Skew elliptical distributions and copulas **53**
7.1 Skewed normal distribution 53
7.2 SN copula 54
7.2.1 Theoretical LGC 55
7.2.2 Plots 56

8 Archimedean copulas **62**
8.1 Definitions and properties 62
8.1.1 Definition 62
8.1.2 Properties 63
8.1.3 Kendall's tau 63
8.1.4 Tail dependence 63
8.1.5 Theoretical LGC 64
8.2 Clayton 65
8.2.1 Plots 66
8.3 Gumbel copula 74
8.3.1 Plots 75
8.4 Frank copula 81
8.4.1 Plots 81
8.5 CG-copula 87
8.5.1 Plots 87

9 Estimation, selection and goodness of fit **93**
9.1 Estimation 93
9.1.1 Parametric estimation 93
9.1.2 Semi-parametric estimation 94
9.1.3 Empirical copula 95
9.1.4 Kernel methods 96
9.2 Detecting dependence 96
9.2.1 Chi plots 97
9.2.2 K-plot 98
9.2.3 LGC plot 99

9.3 Model diagnostics and Goodness-of-Fit 99
9.3.1 Graphical diagnostic 99
9.3.2 GoF 99
9.3.3 Probability integral transformation 100
9.3.4 Dimension reduction approaches 100
9.3.5 Other GoF 101
9.3.6 GoF based on the LGC 102

10 Conclusion **103**

11 Bibliography **104**

A Plots **108**

Chapter 1

Introduction

Dependence and copula theory are important and much studied subjects in statistics. In this thesis we will describe some of the work that has been in done in this field. We will also present a recently developed local dependence measure called local Gaussian correlation, and try to use this to characterize some of the best known copula models, together with som less knowns. In chapter 2 we define the copula concept, and present some results that will be of use to us in this thesis. Popular global dependence measures are presented in chapter 3, where we specially are interested in its connection with copulas. In chapter 4 we introduce the concept of local Gaussian correlation (LGC) and in chapter 5 we show how a theoretical version of this dependence measure can be developed for copula models. Elliptical distributions are used a lot in applications and in chapter 6 we look closer at this class of distributions, specially the Gaussian and the t distribution. Copula models is constructed from these distributions, and the theoretical LGC is derived and analysed. In chapter 7 we have a quick look at skewed versions of elliptical distributions, a copula is constructed from the skewed normal distribution and the theoretical LGC is calculated for this model. A method for calculating the theoretical LGC for Arhimedean copulas is presented in chapter 8, and then 4 different copulas from this class is analysed with the help of the LGC and other dependence measures. In chapter 9 we mention some of the existing methods for estimating, selecting and testing different copula models, and try to point out where the LGC can be used. Plots will be presented when appropriate throughout this thesis, and in the end there is an appendix with additional plots.

Chapter 2

Copula

The term copula was first introduced by Sklar (1959), but the interest in it did not really explode before recent years. One of the main reasons for the interest in copula theory is the many applications in finance. The copula function describe the dependence structure between stochastic variables, and it gives one the opportunity to separate the dependence structure and the marginal distributions. Nelsen (2006) is a classic and good introductory book to copulas. We are now going to define the concept of copulas and look at some properties. For the most part we will consider 2-dimensional copulas in this thesis. Still we will define the concept for the general n-dimensional case, and some of the result will also be stated in n-dimensions where there are no serious complication by doing so.

2.1 Mathematical introduction

In order to define the copula function in a proper way we need to define a couple of other terms first and state a lemma, which all is from Embrechts, Lindskog and McNeil (2001).

Definition 2.1. *Let $S_1, \ldots, S_n$ be nonempty subsets of $\bar{\mathbb{R}} = [-\infty, \infty]$ and let H be a real valued function of n-variables with domain $S_1 \times \cdots \times S_n$. For $a = [a_1, \ldots, a_n] \leq b = [b_1, \ldots, b_n]$, that is $a_k \leq b_k$ for all $k \epsilon \{1, 2, \ldots, n\}$, let $B = [a, b] = [a_1, b_1] \times \cdots \times [a_n, b_n]$ be an n-box whose vertices are in the domain of H.*
We then say that the H-volume is given by

$$V_H(B) = \sum sgn(c)H(c).$$

*The sum is taken over all vertices **c** of B, and the sgn-function is*

$$sgn(c) = \begin{cases} 1, & c_k = a_k \text{ for even } k\text{'s} \\ -1, & c_k = a_k \text{ for odd } k\text{'s} \end{cases}$$

For 2 variables we get $V_H(B) = H(b_1, b_2) - H(b_1, a_2) - H(a_1, b_2) + H(a_1, a_2)$.

Definition 2.2. *A real function H of n variables is n-increasing if $V_H(B) \geq 0$ for all n-boxes B whose vertices lie in the domain of H.*

Definition 2.3. *Let H be real function with domain $S_1 \times \cdots \times S_n$.*

- *Let each S_k has a smallest element a_k. We say H is grounded if $H(t) = 0$ for all t in the domain where $t_k = a_k$ for at least one k.*

- *If we also require that S_k is nonempty and has greatest element b_k we say that H has margins. The one dimensional margins to H are functions H_k with domain S_k such that $H_k(x) = H(b_1, \ldots, b_{k-1}, x, b_{k+1}, \ldots, b_n)$ for all x in S_k.*

Let $(t_1, \ldots, t_{k-1}, x, t_{k+1}, \ldots, t_n)$ and $(t_1, \ldots, t_{k-1}, y, t_{k+1}, \ldots, t_n)$ be in the domain of the real valued function H, and let $x \leq y$, then we say that H is increasing in every argument if $H(t_1, \ldots, t_{k-1}, x, t_{k+1}, \ldots, t_n) \leq H(t_1, \ldots, t_{k-1}, y, t_{k+1}, \ldots, t_n)$.

Lemma 2.4. *Let $S_1, \ldots, S_n$ be nonempty subsets of $\bar{\mathbb{R}}$ and let H be a grounded and n-increasing function with domain $S_1 \times \cdots \times S_n$, then H is increasing in every argument.*

Definition 2.5. *Let H be an n-dimensional function with domain $\bar{\mathbb{R}^n}$. We say that H is a distribution function if it is grounded, n-increasing and we have that $H(\infty, \ldots, \infty) = 1$.*

Now we are ready to define a copula.

Definition 2.6. *A n-dimensional copula is a function C with domain $[0, 1]^n$ such that*

1. *C is grounded and n-increasing*
2. *C has margins C_k, where $k = \{1, \ldots, n\}$, which satisfies $C_k(u) = u$ for all u in $[0, 1]$.*

In other words we can say that an n-copula is a distribution function of a stochastic vector in $\mathbb{R}^n$ with uniform $[0, 1]$ margins.

2.2 Sklars theorem

This theorem is very important, and used a lot in applications. In the text I will use $RanF$ as a short cut for the range of a function F.

Theorem 1. *Let H be a n-dimensional distribution function with margins $F_1, \ldots, F_n$. Then there exists a n-copula C such that for all $x \in \mathbb{R}^n$ we have*

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

If all $F_1, \ldots, F_n$ are continuous, then C will be unique. Otherwise C will be uniquely determined on $RanF_1, \ldots, RanF_n$. Conversely, if C is a n-copula and $F_1, \ldots, F_n$ are distribution functions, then H is a n-dimensional distribution function with margins $F_1, \ldots, F_n$.

For proof see Nelsen (2006) pages 17-24.

If F is an univariate distribution function we define $F^{-1}(t) = \inf\{x \in \mathbb{R} | F(x) \geq t\}$ for t in $[0,1]$. We then have a corollary following from Sklars theorem.

Corollary 2. *Let H be a n-dimensional distribution function with continuous margins $F_1, \ldots, F_n$ and copula C. Then for all u in $[0,1]^n$ we have that*

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

2.3 Properties and examples

An example of a trivial copula is the independence copula, or the product copula as it also is called, $\prod_d(u) = u_1 \cdots u_d$, where the components to $U = (U_1, \ldots, U_d)$ are independent and uniformly distributed on [0,1]. We will from now one denote the uniform distribution on the interval $[a, b]$ as U(a,b). The two functions $M_2(u) = min\{u_1, u_2\}$ and $W_2(u) = max\{u_1 + u_2 - 1, 0\}$ are also copulas. We can see this by noting that if U is U(0,1) we have

$$M_2(u) = P[U \leq u_1, U \leq u_2]$$

$$W_2(u) = P[U \leq u_1, 1 - U \leq u_2]$$

That is M_2 and W_2 are bivariate distribution functions for the vectors (U,U) and (U,1-U), and it follows that they are copulas. For $n \geq 3$ $M_n(u) = min\{u_1, \ldots, u_n\}$ will still be a copula, while $W_n(\mathbf{u}) = max\{u_1 + \cdots + u_n + 1 - n, 0\}$ never will be (Embrechts, Lindskog, McNeil (2001)). These two functions appear in a well known theorem which gives an upper and lower bound for every copula. We call it the Fréchet-Hoeffding Bounds.

Theorem 3. *Let C be any n-copula. Then for all u in $[0,1]^n$ we have that*

$$W_n(u) \leq C(u) \leq M_n(u)$$

Proof is given in Frèchet (1957).

If we have random variables $X_1, \ldots, X_n$ with joint distribution function H, we know that the random variables are independent if and only if we have $H(x_1, \ldots, x_n) = F_1(x_1) \cdots F_n(x_n)$. From Sklars theorem the following result follows.

Theorem 4. *Let $X_1, \ldots, X_n$ be a vector with continuous random variables with copula C. Then $X_1, \ldots, X_n$ are independent if and only if $C(u_1, \ldots, u_n) = u_1 \cdots u_n$. That is the product copula.*

2.3.1 Density

For a copula C we have that the mixed kth order derivative, $\frac{\partial^k C(u)}{\partial u_1 \ldots \partial u_k}$, exists for almost all u in $[0,1]^n$ (Embrechts, Lindskog, McNeil (2001)), and we have that

$$0 \leq \frac{\partial^k C(u)}{\partial u_1 \ldots \partial u_k} \leq 1. \tag{2.1}$$

The density for a n-copula C is in general given by

$$C(u) = \frac{\partial^n C(u)}{\partial u_1 \ldots \partial u_n}. \tag{2.2}$$

If we have a continuous n-dimensional distribution F, with density f, and continuous margins $F_1, \ldots, F_n$ with densities $f_1, \ldots, f_n$, then the density for a implicit copula can be written

$$c(u) = \frac{f(F_1^{-1}(u_1), \ldots, F_n^{-1}(u_n))}{f_1(F_1^{-1}(u_1)) \cdots f_n(F_n^{-1}(u_n))}. \tag{2.3}$$

Hence

$$c(F_1(x_1), \ldots, F_n(x_n)) = \frac{f(x_1, \ldots, x_n)}{f_1(x_1) \cdots f_n(x_n)}. \tag{2.4}$$

This means we can write a general n-dimensional density f as

$$f(x_1, \ldots, x_n) = c(F_1(x_1), \ldots, F_n(x_n)) f_1(x_1) \cdots f_n(x_n). \tag{2.5}$$

2.3.2 Increasing transformations

When working with a model of financial returns, one may want to change it to a model of the logarithm of these returns. We then have the useful property that this transformation will not affect the dependence structure given by the copula C. In general we have that copula functions are invariant under strictly increasing transformations.

Theorem 5. *Let $(X_1, \ldots, X_n)$ be a vector of continuous random variables with copula C. If $\alpha_1, \ldots, \alpha_n$ are strictly increasing transformations on $RanX_1, \ldots, RanX_n$, respectively, then $(\alpha_1(X_1), \ldots, \alpha_n(X_n))$ also has copula C.*

Proof is given in Embrechts, Lindskog and McNeil (2001).

2.4 Survival copulas

Some times in application we meet what we call the survival function, that is $\overline{F}(x) = P(X > x) = 1 - F(x)$, where F is the distribution function to the random variable X. For two random variables X, Y with joint distribution function H, we have the joint survival function $\overline{H}(x, y) = P(X > x, Y > y)$. For copulas we can define the function $\widehat{C}$: $I^2 \to I$ as

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

and we call it the survival copula. This is not to be confused with the survival function to the distribution function C with uniform margins. We have

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

The reason for the definition becomes clear when we look at

$$\begin{aligned}
\overline{H}(x, y) &= 1 - F(x) - G(y) + H(x, y) \\
&= \overline{F}(x) + \overline{G}(y) - 1 + C(F(x), G(y)) \\
&= \overline{F}(x) + \overline{G}(y) - 1 + C(1 - \overline{F}(x), 1 - \overline{G}(y)).
\end{aligned}$$

So we have

$$\overline{H}(x, y) = \widehat{C}(\overline{F}(x), \overline{G}(y)) \tag{2.6}$$

2.5 Symmetry

When we are going to characterize the different types of copula functions, saying something about the symmetry is essential. If we let X be a univariate random variable we say that it is symmetric about a if $F(a + x) = \overline{F}(a - x)$. In the bivariate case it is not that simple to know what is meant by symmetry. We are going to consider two different symmetry concepts, and show how it effects the copula. More details and also other kinds of symmetry can be found in Nelsen (2006).

2.5.1 Radial symmetry

We say that (X,Y) is radially symmetric about (a,b) if $(X - a, Y - b)$ and $(a-X, b-Y)$ has the same distribution. When working with continuous random variables it can be shown that (X,Y) is radially symmetric about (a,b) if and only if

$$H(a + x, b + y) = \overline{H}(a - x, b - y) \tag{2.7}$$

for all (x,y) in $\mathbb{R}^2$. Here H is the joint distribution function of (X,Y). The points $(a + x, b + y)$ and $(a - x, b - y)$ lie on rays emanating in opposite directions from (a, b), and that is where the term "radial" comes from. Figure 2.1 shows areas with equal probability when we have radial symmetry around (a, b).

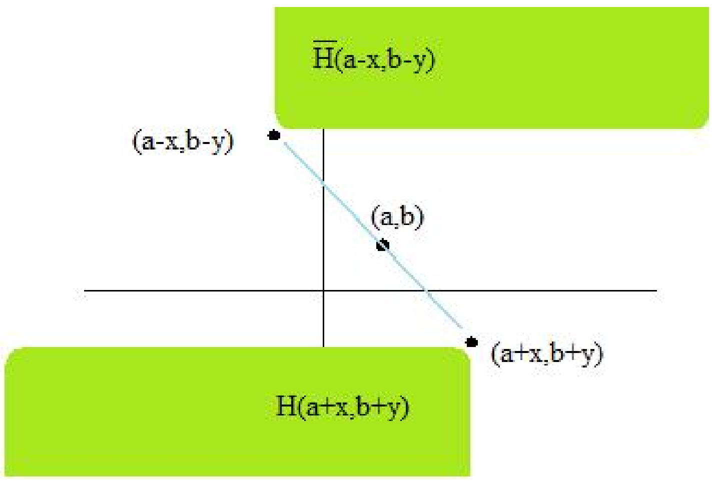

Figure 2.1: Shows the areas with the same probability when we have radial symmetric random variables.

The next theorem gives a link between the copula function and radial symmetry.

Theorem 6. *Let X and Y be continuous random variables with joint distribution function H, marginal distribution functions F and G respectively, and copula C.*

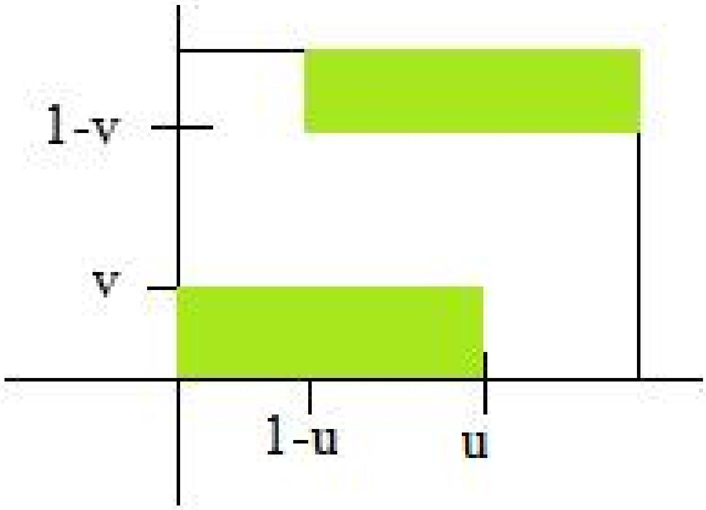

Figure 2.2: Shows the areas with the same probability when we have radial symmetric random variables.

Assume that X is symmetric about a and Y is symmetric about b. Then (X, Y) is radially symmetric about (a,b) if and only if $C = \widehat{C}$, that is if and only if

$$C(u, v) = u + v - 1 + C(1 - u, 1 - v)$$

for all (u,v) in I^2.

For proof see Nelsen (2006).

As shown in figure 2.2 this theorem has the geometrically interpretation that the rectangles $[0, u] \times [0, v]$ and $[1 - u, 1] \times [1 - v, 1]$ have equal C-volume.

2.5.2 Exchangeable symmetry

We say that X,Y is exchangeable if (X,Y) and (Y,X) are identically distributed. So if we have two random variables X and Y, with joint distribution function H, margins F and G, and copula C, we can write

$$C(u, v) = H(F^{-1}(u), G^{-1}(v)) = H(G^{-1}(v), F^{-1}(u)) = C(v, u). \tag{2.8}$$

Here we have clearly used Sklars theorem again. This shows that the exchange symmetry of random variables is inherited by their copula. We can state this more formally in a theorem.

Theorem 7. *Let X and Y be continuous random variables with joint distribution function H, margins F and G, respectively, and copula C. Then X and Y are exchangeable if and only if F=G and C(u,v)=C(v,u) for all (u,v) in I^2.*

Copulas with the property that C(u,v)=C(v,u) for all (u,v) in I^2 are often referred to only as being symmetric.

Chapter 3

Global dependence measures

We are going to consider some of the most used global dependence measures, and try to present the connections between them and the copula concept.

3.1 Linear correlation

The linear correlation coefficient is used in many applications, it is for example used a lot as a dependence measure in financial theory. In Embrechts, McNeil and Straumanm (1999) we can find a pretty thorough investigation of the use of linear correlation as a dependence measure in finance and insurance. We are only going to mention a couple of important points from there.

Definition 8. *If we have two random variables X and Y with finite variances the linear correlation between X and Y is*

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

where $Cov(X, Y) = E[XY] - E[X]E[Y]$ are the covariance of X and Y. ρ is called the linear correlation coefficient because it measures the linear dependence between random variables. Actually knowledge about $\rho(X, Y)$ is equivalent to the coefficient β of the linear regression $Y = \beta X + \epsilon$. Here ϵ is a residual which is linearly uncorrelated of X. The correlation coefficient will be in the interval $[-1, 1]$, and in case of perfect linear dependence we have that $|\rho(X, Y)| = 1$. By perfect linear dependence we mean that $Y = aX + b$ almost surely, where a is a real number except zero and b is a real number. If A and B are $m \times n$ matrices, a and b $\in \mathbb{R}^m$, and X and Y are stochastic n-vectors, we have that

$$Cov(AX + a, BY + b) = ACov(X, Y)B^T$$

which again gives for $\alpha \in \mathbb{R}^n$

$$Var(\alpha^T X) = \alpha^T Cov(X)\alpha$$

where Cov(X) is defined as an $n \times n$ matrix where the ij-element of the matrix is given by $Cov(X_i, X_j)$. In this way we can decide the variance of a linear combination in the portfolio theory by considering the covariances of the components in pairs.

It follows from the definition of covariance that if we have independent stochastic variables the covariance, and thus the correlation, becomes zero. But if two stochastic variables are uncorrelated they are not necessarily independent. As an example we can look at $X \sim U(-1, 1)$, $Z \sim U(0, \frac{1}{10})$ and $Y = X^2 + Z$, where X and Z are independent. Here U is the continuous uniform distribution. If we look at the conditional distribution of Y given $X = x$, that is $Y = x^2 + Z$, it will have distribution $U(x^2, x^2 + \frac{1}{10})$. We see that X and Y obviously will have some sort of dependency. To calculate the covariance we use that $E[X] = E[X^3] = 0$ since $X \sim U(-1, 1)$ and $E[XZ] = E[X]E[Z]$ since X and Y are independent. We get $Cov(X, Y) = E[X(X^2 + Z)] - E[X]E[X^2 + Z] = E[X^3] + E[XZ] - 0 = 0 + E[X]E[Z] = 0$. Only when considering the multivariate normal distribution does $\rho = 0$ imply independence. Another possible problem with the correlation coefficient is that it does not exist for distributions where the variance is not finite, for example the bivariate student t distribution with degree of freedom less then or equal two.

3.2 Perfect dependence

We remember the Frèchet-Hoeffding bounds, in the bivariate case they where $M_2(u) = min\{u_1, u_2\}$ and $W_2(u) = max\{u_1 + u_2 - 1, 0\}$, and they are both copulas. We say that M_2 represents perfect positive dependence and W_2 perfect negative dependence. The following theorem formalize this. See Embechts, McNeil and Strauman (1999) for proof and further references.

Theorem 9. *Let (X, Y) have one of the copulas M_2 or W_2. Then there exists two monotone functions f,g: $\mathbb{R} \rightarrow \mathbb{R}$ and a real-valued random variable Z such that*

$$(X, Y) =_d (f(Z), g(Z)) \tag{3.1}$$

In the case of W_2 f will be increasing and g decreasing, and in the case of M_2 both will be increasing. The converse of the result is also true.

If X and Y has continuous margins, respectively F_X and F_Y, we have the following stronger result (Embechts, McNeil and Strauman (1999)).

$$C_{X,Y} = W_2 \Leftrightarrow Y = F_Y^{-1}(1 - F_X(X)) \tag{3.2}$$

$$C_{X,Y} = M_2 \Leftrightarrow Y = F_Y^{-1}(F_X(X)) \tag{3.3}$$

We say that X and Y is *comonotonic* if (X,Y) has copula M_2 and *countermonotonic* if (X,Y) has copula W_2.

3.3 Kendalls tau and Spearmans rho

If we have two observations (x, y) and $(\tilde{x}, \tilde{y})$ from a vector (X, Y) with continuous stochastic variables. Then we say that (x, y) and $(\tilde{x}, \tilde{y})$ are *concordant* if $(x - \tilde{x})(y - \tilde{y}) > 0$ and *discordant* if $(x - \tilde{x})(y - \tilde{y}) < 0$.

Definition 10. *If $(\tilde{X}, \tilde{Y})$ is an independent copy of (X, Y), then Kendalls tau for the stochastic vector (X, Y) is*

$$\tau(X,Y) = P[(X-\tilde{X})(Y-\tilde{Y}) > 0] - P[(X-\tilde{X})(Y-\tilde{Y}) < 0].$$

With other words Kendalls tau is the probability of concordance subtracted the probability of discordance. If we let X and Y be assets, we can from a financial point of view consider Kendalls tau to be a comparison between the probability of the two assets rising (or falling) together with the probability that one of the assets rise (fall) while the other fall (rise). This is important when trying to set up a portfolio with a diversification effect.

Definition 11. *If $(\tilde{X}, \tilde{Y})$, (X', Y') and (X, Y) are independent copies, then Spearmans rho for the stochastic vector (X, Y) is*

$$\rho_s(X,Y) = 3(P[(X-\tilde{X})(Y-Y') > 0] - P[(X-\tilde{X})(Y-Y') < 0]).$$

To see how we can describe these to dependence measures with the help of copula functions we need the following theorem.

Theorem 12. *Let (X, Y) and $(\tilde{X}, \tilde{Y})$ be independent vectors of continuous stochastic variables with joint distribution functions, H and $\tilde{H}$ respectively , and with common margins, that is marginal distribution function F for X and $\tilde{X}$, and G for Y and $\tilde{Y}$. Further on let C and $\tilde{C}$ be copulas for (X, Y) and $(\tilde{X}, \tilde{Y})$ respectively, such that $H(x, y) = C(F(x), G(y))$ and $\tilde{H}(x, y) = \tilde{C}(F(x), G(y))$. Let Q be*

$$Q = P[(X-\tilde{X})(Y-\tilde{Y}) > 0] - P[(X-\tilde{X})(Y-\tilde{Y}) < 0].$$

Then

$$Q = Q(C, \tilde{C}) = 4 \int\int_{[0,1]^2} \tilde{C}(u,v) dC(u,v) - 1.$$

Proof is found in Embrechts, Lindskog and McNeil (2001).

We see by the definition of Kendalls tau that it fits with this theorem. In this case we have $H = \tilde{H}$ and $C = \tilde{C}$, such that we get

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

Here U and V are U(0,1).
The theorem also gives an expression for Spearmans rho.

$$\rho_s(X,Y) = 3Q(C,\Pi) = 12 \int\int_{[0,1]^2} uv dC(u,v) - 3 = 12 \int\int_{[0,1]^2} C(u,v) du dv - 3$$

Where Π is the independence copula, $\Pi(u, v) = uv$.

By using this expression for ρ_s we can also find a relation with the linear correlation ρ. Let X and Y have marginal distribution functions F and G respectively, and let U=F(X) and V=G(Y). By remembering the definition of covariance, and that E(U)=1/2 and Var(U)=1/12, we get

$$\rho_s(X,Y) = 12\int\int_{[0,1]^2} uvdC(u,v) - 3 = 12E[UV] - 3 = \frac{E[UV]-\frac{1}{4}}{\frac{1}{12}}$$
$$= \frac{Cov[UV]-\frac{1}{4}+E[U]E[V]}{\frac{1}{12}} = \frac{Cov[UV]-\frac{1}{4}+\frac{1}{4}}{\frac{1}{12}}$$
$$\frac{Cov[UV]}{\sqrt{Var(U)}\sqrt{Var(V)}} = \rho(F(X),G(Y)).$$

This shows that ρ_s is nothing more than the linear correlation of the uniformly distributed random variables $U = F(X)$ and $V = G(Y)$.

Kendalls tau and Spearmans rho shares many properties, we list her some essential common ones. We are considering two continuous random variables X and Y with copula C, and we let k be either Kendalls tau or Spearmans rho. For proof and more properties we refer to Embrechts McNeil and Strauman (1999)

1. k is defined for any pair of continuous random variables.
2. k is symmetric, that is we have $k(X,Y) = k(Y,X)$.
3. k=0 for independent random variables.
4. We have that $k \in [-1,1]$, and k(X,Y)=1 if and only if C=M and k(X,Y)=-1 if and only if C=W.

3.4 Quadrant dependence

Definition 13. *We say that two random variables X and Y are positive quadrant dependent, and write PQD(X,Y), if for all (x,y) in $\mathbb{R}^2$ we have*

$$P(X \le x, Y \le y) \ge P(X \le x)P(Y \le y),$$

or equivalent

$$P(X \ge x, Y \ge y) \ge P(X \ge x)P(Y \le y).$$

If X and Y has joint distribution function H, continuous margins F and G, and copula C, we can say that X and Y are positive quadrant dependent if

$$H(x,y) \ge F(x)G(y) \text{ for all (x,y) in } \mathbb{R}^2$$

or

$$C(u,v) \ge uv \text{ for all (u,v) in } [0,1]^2$$

We can intuitively think that X and Y are PQD if the probability that they are simultaneously small is at least as big as if they would be independent. PQD is a copula property, and the graph of the copula of (X,Y) must lie on or above the graph of the independence copula Π if X and Y are PQD. In the same way as described for PQD we can define negative quadrant dependence (NQD) by switching the inequalities in the equations above. In general quadrant dependence is a global property, that is it must hold for all (x,y) in $\mathbb{R}^2$, but we can look at the inequalities given above in subsets of $\mathbb{R}^2$ and in that way get local PQD/NQD. See Nelsen (2006) for more and also connections with other dependence measures.

3.5 Tail dependence

Let us take a closer look at a local version of positive quadrant dependence. In finance and insurance we are often specially interested in what happens in the tails, that is in the lover left corner and upper right corner. So let us define tail dependence.

Definition 14. *Let (X,Y) be a vector of continuous random variables with marginal distribution functions F_X and F_Y. The coefficient for upper tail dependence is*

$$\lambda_u = \lim_{t \to 1^-} P[Y > F_Y^{-1}(t) | X > F_X^{-1}(t)], \tag{3.4}$$

and the coefficient for lower tail dependence is

$$\lambda_l = \lim_{t \to 0^+} P[Y \leq F_Y^{-1}(t) | X \leq F_X^{-1}(t)].$$

This is provided that the limits exists.

When $\lambda_u = 0$ we say that X and Y are *asymptotic independent* in the upper tail, and when $\lambda_u \in (0, 1]$ we say that they have upper tail dependence and large events tend to occur simultaneously. For λ_l it is similar.

Tail dependence is a copula property, something that is shown by the next theorem.

Theorem 15. *Let (X,Y), F_X, F_Y, λ_u and λ_l be as in the definition of tail dependence, and let C be the copula of X and Y. Then*

$$\lambda_u = \lim_{t \to 1^-} \frac{\overline{C}(t,t)}{1-t}$$

and

$$\lambda_l = \lim_{t \to 0^+} \frac{C(t,t)}{t}, \tag{3.5}$$

if the limits exists.

For proof see Nelsen (2006).

When we further on uses the notation λ it means that we are considering either λ_u or λ_l. Let us look at X and Y which can be considered independent for sufficient large values. That is where we have

$$\lim_{x,y \to \infty} \frac{F(x,y)}{F_X(x) F_Y(y)} = 1 \tag{3.6}$$

If this is the case it can be shown (Malevergne and Sornette 2006) that

$$\lim_{t \to 1^-} P(X > F_X^{-1}(t) | Y > F_Y^{-1}(t)) = \lim_{t \to 1^-} 1 - F_X(F_X^{-1}(t)) = \lim_{t \to 1^-} 1 - t = 0. \tag{3.7}$$

That is $\lambda = 0$ for independent variables, but it does not imply independence. Later we will look at an example in the Gaussian case. We can now follow in the direction of Coles, Heffernan and Tawn (1999) and define a new coefficient which will tell us more.

$$\overline{\lambda_u} = \lim_{t \to 1^-} \frac{2 \log P(X > F_X^{-1}(t))}{\log P(X > F_X^{-1}(t), Y > F_Y^{-1}(t))} - 1 = \lim_{t \to 1^-} \frac{2 \log(1-t)}{\log \overline{C}(t,t)} - 1. \quad (3.8)$$

We will call this the alternative tail coefficient, and generally denote it $\overline{\lambda}$. We have that $\overline{\lambda_u}$ will be in the interval [-1,1], where [-1,1) corresponds to asymptotic independence. In the case of asymptotic independence $\overline{\lambda_u}$ will be an increasing measure with respect to dependence strength. So when $\lambda = 0$ we will have the necessary condition $\overline{\lambda} = 0$ for true independence.

We remember that the survival copula is given by $\widehat{C}(u,v) = u+v-1+C(1-u, 1-v)$, while the survival function of two standard uniform distributed random variables U and V with distribution C is given by $\overline{C}(u,v) = 1-u-v+C(u,v) = \widehat{C}(1-u, 1-v)$. A sometimes useful property is to note that lower tail dependence coefficient of C is the upper tail dependence coefficient of $\widehat{C}$, and also the upper tail coefficient of C is the lower tail dependence coefficient of $\widehat{C}$. These two properties is easily shown so we only show the last one. (Embrechts, Lindskog and McNeil (2001))

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

Chapter 4

Local Gaussian Correlation

4.1 Introduction

In this thesis we will focus on an approach presented by Tjøstheim and Hufthammer (2012), which is called Local Gaussian Correlation, or just LGC for short. It is well known that in a multivariate normal distribution the dependence is completely determined by the correlation or covariance matrix. But in applications one often encounters other distributions than the normal distribution. The idea behind the LGC is to locally approximate the real density of a sample with the Gaussian distribution in every point. That is

$$\phi(u, v, \mu_1(x), \mu_2(x), \sigma_1(x), \sigma_2(x), \rho(x)) =$$
$$\frac{1}{2\pi\sigma_1(x)\sigma_2(x)\sqrt{1-\rho(x)^2}} \exp\{-\frac{1}{2(1-\rho(x)^2)}((\frac{u-\mu_1(x)}{\sigma_1(x)})^2 +$$
$$(\frac{v-\mu_2(x)}{\sigma_2(x)})^2 - 2\rho(x)(\frac{u-\mu_1(x)}{\sigma_1(x)})(\frac{v-\mu_2(x)}{\sigma_2(x)}))\},$$

and then use the correlation parameter from the Gaussian approximation as a measure of dependence locally. Then we can characterize dependence locally, and we can also use the other properties of the Gaussian distribution on a local scale. So if we now have a sample of n iid bivariate random variables $X^{(i)} = (X_1^{(i)}, X_2^{(i)})$ with real density $f(x)$, we want to approximate it with the bivariate Gaussian distribution $\phi(w, \theta(x)) = \phi(u, v, \theta(x))$ in the neighbourhood of every point x. Here $\theta(x)$ is the 5 dimensional parameter vector

$$\theta(x) = (\mu_1(x), \mu_2(x), \sigma_1(x), \sigma_2(x), \rho(x)). \tag{4.1}$$

For clarification we note that in some cases we will write the local parameter vector as $\theta(x) = (\mu(x), \Sigma(x))$, where $\mu(x) = (\mu_1(x), \mu_2(x))$ and $\Sigma(x) = (\sigma_{ij}(x))$ is the local covariance matrix. The local correlation coefficient is $\rho(x) = \frac{\sigma_{12}(x)}{\sigma_1(x)\sigma_2(x)}$. We will also denote the j'th element in the vector $\theta(x)$ as θ_j. To estimate $\theta(x)$ we will use the modified local log likelihood given in Tjøstheim and Hufthammer(2012)

$$L(X^{(1)}, ..., X^{(n)}, \theta(x)) = n^{-1}\sum_i K_b(X^{(i)}-x)log(\phi(X^{(i)}, \theta(x))) - \int K_b(w-x)\phi(w, \theta)dw \tag{4.2}$$

Here $K_b(X^{(i)}-x) = (b_1 b_2)^{-1} K(b_1^{-1}(X_1^{(i)}-x_1))K(b_2^{-1}(X_2^{(i)}-x_2))$ is a product kernel with bandwidth $b = [b_1, b_2]^T$ in the x_1 and x_2 direction. Setting $u_j = \frac{\partial \log \phi}{\partial \theta_j}$, we get the derivatives of the modified log likelihood

$$\frac{\partial L}{\partial \theta_j} = n^{-1} \sum_i K_b(X^{(i)}-x) u_j(X^{(i)}, \theta(x)) - \int K_b(w-x) u_j(w, \theta(x)) \phi(w, \theta(x)) dw$$

We can find an estimate $\widehat{\theta}(x) = \widehat{\theta}_{n,b}(x)$ for fixed values of n and b by solving the equations $\frac{\delta L}{\delta \theta_j} = 0$ for $j = 1, ..., 5$. By fixing b and letting $n \to \infty$ we get the equation

$$\int K_b(w-x) u_j(w, \theta_b(x)) \left[f(w) - \phi(w, \theta_b(x))\right] dw = 0 \tag{4.3}$$

A population value $\theta_b(x)$ can be defined as a solution to these equations. It can be shown that, if we assume there is a bandwidth b_0 such that there exists a unique solution $\theta_b(x)$ for every b with $0 < b < b_0$, $\widehat{\theta}_{n,b}(x)$ will converge to $\theta_b(x)$ (Tjøstheim and Hufthammer(2012)). By letting $b \to 0$ we also get a population vector $\theta(x)$, where $\theta_b(x) \to \theta(x)$. Letting b approach 0 give us the equation

$$u_j(x, \theta(x)) \left[f(x) - \phi(x, \theta(x))\right] + O(b^T b) = 0 \tag{4.4}$$

If we ignore the solution $u_j = 0$ we see that the local likelihood estimates requires $\phi(x, \theta(x))$ to be close to $f(x)$. The equation $f(x) - \phi(x, \theta(x))$ will in general have infinitely many solutions for the unknown θ, so we must look at it in the context of the local likelihood function. The limits here are calculated under some regularity conditions, see Tjøstheim and Hufthammer (2012) for details. It is also possible to make the same argument as above where we consider the observations to be from an ergodic time series $\{X_t\}$.

4.2 Existence of $\theta_b(x)$

If f has global Gaussian distribution with parameter vector $\theta = (\mu_1, \mu_2, \sigma_1, \sigma_2, \rho)$, the existence of a solution is easily established. That is, $\theta_b = \theta$ will satisfy equation (4.3). In the next step we follow Tjøstheim and Hufthammer (2012) and start by defining a piecewise linear function g_s as

$$X = g_s(Z) = \sum_{i=1}^{k} (a_i + A_i Z) 1(Z \in R_i). \tag{4.5}$$

Here $Z \sim N(0, I_2)$, where I_2 is the two dimensional identity matrix. R_i for $i = 1, \ldots, k$ is non overlapping regions of $\mathbb{R}^2$, such that $\mathbb{R}^2 = \bigcup_{i=1}^k R_i$. The a_is are corresponding vectors in $\mathbb{R}^2$ and the A_is are corresponding 2×2 non-singular matrices. Further on we define the region S_i to be $S_i = \{x : x = a_i + A_i z, z \in R_i\}$, and assumes that $S_i \bigcap S_j = \emptyset$ for $i \neq j$ and $\bigcup_{i=1}^k S_i = \mathbb{R}^2$. Now let the Kernel function K have a compact support, and let x be an interior point of S_i. Now b can be made small enough so that $w - x \in S_i$ if $w - x$ is in the support of K. Now if we set $\mu_i = a_i$ and $\Sigma_i = A_i A_i^T$, this restriction on b gives us the solution $\theta_b(x) := \theta_i = (\mu_i, \Sigma_i)$. We now have a local Gaussian approximation $\phi(x, \theta_b(x))$.

4.3 Non-linear transformations of Gaussian variables

By increasing k and letting the regions R_i be smaller we can use a sequence of functions like in equation (4.5) to approximate more general non linear continuous functions. We are now going to see how we can use non linear transformations of a bivariate normal distribution to find what we will call the theoretical LGC. This is also found in Tjøstheim and Hufthammer (2012). Now let $Z \sim N(0, I_2)$ and g be a one-to-one vector function $g : \Re^2 \to \Re^2$ with inverse $h = g^{-1}$. The Jacobi matrix is denoted

$$\frac{\partial g}{\partial z} = \begin{bmatrix} \frac{\partial g_1}{\partial z_1} & \frac{\partial g_1}{\partial z_2} \\ \frac{\partial g_2}{\partial z_1} & \frac{\partial g_2}{\partial z_2} \end{bmatrix}$$

We assume that g has continuous second order derivatives, so we can make a Taylor expansion around $z = (z_1, z_2)$ and get

$$X_i = g_i(Z) = g_i(z) + \sum_{j=1}^{2} \frac{\partial g_i}{\partial z_j}(z)(Z_j - z_j) + \frac{1}{2}\sum_{j=1}^{2}\sum_{k=1}^{2} \frac{\partial^2 g_i}{\partial z_j \partial z_k}(\xi)(Z_j - z_j)(Z_k - z_k) \tag{4.6}$$

for $i = 1, 2$. The mean value theorem produces the intermediate value ξ. We let b_z and b_x be locality defined bandwidths for the Z and X variable respectively. For the local likelihood method described earlier it is important to notice that we try to fit the best Gaussian approximation to a density f_X in a neighbourhood around the point of interest x. $X = g(Z)$ will therefore be considered in the neighbourhoods $N(z) = \{z^{'} = |z^{'} - z| \leq b_z\}$ and $N(x) = \{x^{'} = |x^{'} - x| \leq b_x\}$. When these neighbourhoods gets sufficiently small, that is when $b_z \to 0$ and $b_x \to 0$, the idea is that we can neglect the last term in the Taylor expansion in probability. This will give us

$$P(X \in N(x)) \sim P(U(z) \in N(x))$$

where

$$U(z) = g(z) + \frac{\partial g}{\partial z}(z)(Z - z).$$

The distribution of X is now approximated in the neighbourhood N(x) by U(z), which is Gaussian because it is an affine transformation of a Gaussian variable. In the limit U(z) give a Gaussian approximation of X at x. By computing the expectation and covariance of U(z) and then substitute $z = h(x)$, we can find a local mean and covariance at x. We can now try to define our local parameters as

$$\mu(x) = g(z) - \frac{\partial g}{\partial z}(z)z = x - (\frac{\partial h}{\partial x}(x))^{-1}h(x) \tag{4.7}$$

and

$$\Sigma(x) = \frac{\partial g}{\partial z}(z)(\frac{\partial g}{\partial z}(z))^T = (\frac{\partial h}{\partial x}(x))^{-1}((\frac{\partial h}{\partial x}(x))^{-1})^T \tag{4.8}$$

By standard transformation theory we can now show that the local parameters defined in (4.7) and (4.8) gives a representation $f_X(x) = \phi(x, \mu(x), \Sigma(x))$.

The first question now will be which distributions can be represented by such a g function together with a bivariate standard normal variable Z. To answer this we therefore state the following lemma

Lemma 16. *Let Y have a density $f_Y(y)$ on $\Re^2$ with cumulative distribution function $F_Y(y) = \int_{-\infty}^{y_1}\int_{-\infty}^{y_1} f_Y(w_1, w_2)dw_1 dw_2$. Then there exists a one-to-one function g such that $Y = g(Z)$ where $Z \sim N(0, I_2)$ with I_2 being the 2-dimensional identity matrix.*

Proof. We have $X = (X_1, X_2)$ and $Z = (Z_1, Z_2)$, where Z_1 and Z_2 will be independent. We also have that $f_X(x) = f_{X_1}(x_1)f_{X_2|X_1}(x_2|x_1)$ Let us denote the cumulative distribution function of the standard normal density by Φ. Now $U_1 = F_{X_1}^{-1}(X_1)$ will be uniform, and there will also exist a standard normal variable such that $U_1 = \Phi(X_1)$. In the same way there exists a uniform variable U_2, independent of U_1, such that $U_2 = F_{X_2|X_1}(X_2|X_1)$, and there exists a $Z_2 \sim N(0, 1)$ independent of Z_1 such that $U_2 = \Phi(Z_2)$. From this we now have

$$\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} F_{X_1}^{-1}(\Phi(Z_1)) \\ F_{X_2|X_1}^{-1}(\Phi(Z_2)|F_{X_1}^{-1}(\Phi(Z_1)) \end{bmatrix} = g(Z) \tag{4.9}$$

Details can be found in Tjøstheim and Hufthammer (2012) and Rosenblatt (1952). We see that the representation

$$\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} F_{X_1|X_2}^{-1}(\Phi(Z_1)|F_{X_2}^{-1}(\Phi(Z_2)) \\ F_{X_2}^{-1}(\Phi(Z_2)) \end{bmatrix} = g^{'}(Z^{'}) \tag{4.10}$$

also is a possibility. Here $g \neq g^{'}$ and $Z \neq Z^{'}$ in general. This unfortunately means that given a density f(x) the representation can be generated in several ways and therefore is non-unique.

4.4 Theoretical LGC

We are now going to follow the approach presented in Berentsen et al. (2012). Firstly we will illustrate the non-uniqueness by using the transformations from equation (4.9) and (4.10), which we call Rosenblatt 1 (R1) and Rosenblatt 2 (R2) respectively. For the random variable X with margins F_1 and F_2, the g function given by R1 is

$$\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} F_1^{-1}(\Phi(Z_1)) \\ F_{2|1}^{-1}(\Phi(Z_2)|F_1^{-1}(\Phi(Z_1)) \end{bmatrix} = g(Z),$$

with the inverse h given by

$$\begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} = \begin{bmatrix} \Phi^{-1}(F_1(X_1)) \\ \Phi^{-1}(F_{2|1}(X_2|X_1) \end{bmatrix} = h(X).$$

This gives us

$$(\frac{\partial h}{\partial x})^{-1} = (\frac{\partial h_1}{\partial x_1}\frac{\partial h_2}{\partial x_2})^{-1} \begin{bmatrix} \frac{\partial h_2}{\partial x_2} & 0 \\ \frac{-\partial h_2}{\partial x_1} & \frac{\partial h_1}{\partial x_1} \end{bmatrix},$$

which again gives us

$$\Sigma(x) = (\frac{\partial h_1}{\partial x_1}\frac{\partial h_2}{\partial x_2})^{-2}\begin{bmatrix} (\frac{\partial h_2}{\partial x_2})^2 & -\frac{\partial h_2}{\partial x_1}\frac{\partial h_2}{\partial x_2} \\ -\frac{\partial h_2}{\partial x_1}\frac{\partial h_2}{\partial x_2} & (\frac{\partial h_1}{\partial x_1})^2 + (\frac{\partial h_2}{\partial x_1})^2 \end{bmatrix}.$$

The local correlation would now be

$$\rho_{R1}(x) = \frac{\Sigma_{12}(x)}{\Sigma_{11}(x)\Sigma_{22}(x)} = \frac{-\frac{\partial h_2}{\partial x_1}}{\sqrt{(\frac{\partial h_1}{\partial x_1})^2 + (\frac{\partial h_2}{\partial x_1})^2}} \tag{4.11}$$

The g function given by R2 is

$$\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} F_{1|2}^{-1}(\Phi(Z_1^{'})|F_2^{-1}(\Phi(Z_2^{'})) \\ F_2^{-1}(\Phi(Z_2^{'})) \end{bmatrix} = g^{'}(Z^{'})$$

with invers $h^{'}$

$$\begin{bmatrix} Z_1^{'} \\ Z_2^{'} \end{bmatrix} = \begin{bmatrix} \Phi^{-1}(F_{1|2}(X_1|X_2)) \\ \Phi^{-1}(F_2(X_2)) \end{bmatrix} = h^{'}(X). \tag{4.12}$$

In a similar matter as R1 we get that

$$(\frac{\partial h^{'}}{\partial x})^{-1} = (\frac{\partial h_1^{'}}{\partial x_1}\frac{\partial h_2^{'}}{\partial x_2})^{-1}\begin{bmatrix} \frac{\partial h_2^{'}}{\partial x_2} & \frac{-\partial h_2^{'}}{\partial x_1} \\ 0 & \frac{\partial h_1^{'}}{\partial x_1} \end{bmatrix},$$

which again gives us

$$\Sigma(x) = (\frac{\partial h_1^{'}}{\partial x_1}\frac{\partial h_2^{'}}{\partial x_2})^{-2}\begin{bmatrix} (\frac{\partial h_2^{'}}{\partial x_2})^2 + (\frac{\partial h_1^{'}}{\partial x_2})^2 & -\frac{\partial h_1^{'}}{\partial x_1}\frac{\partial h_1^{'}}{\partial x_2} \\ -\frac{\partial h_1^{'}}{\partial x_1}\frac{\partial h_1^{'}}{\partial x_2} & (\frac{\partial h_1^{'}}{\partial x_1})^2 \end{bmatrix}.$$

The local correlation would in this case be

$$\rho_{R2}(x) = \frac{\Sigma_{12}(x)}{\Sigma_{11}(x)\Sigma_{22}(x)} = \frac{-\frac{\partial h_1^{'}}{\partial x_2}}{\sqrt{(\frac{\partial h_1^{'}}{\partial x_2})^2 + (\frac{\partial h_2^{'}}{\partial x_2})^2}}. \tag{4.13}$$

We can notice that h_1 is independent of x_2, while $h_2^{'}$ is independent of x_1. So the first local variance with respect to R1 would be $\sigma_1^2 = (\frac{\partial h_1}{\partial x_1})^{-2}$, that is only dependent on x_1. The second local variance with respect to R2 would be $\sigma_2^2 = (\frac{\partial h_2^{'}}{\partial x_2})^{-2}$, which is only dependent on x_2. This all means that if R1 and R2 was to define the same local parameters, the local variances σ_1^2 and σ_2^2 would only depend on x_1 and x_2 respectively. This is not the case in general for the likelihood approach.

Though $\rho_{R1}(x)$ and $\rho_{R2}(x)$ in general do not coincide, in some situations and for some subsets of $\mathbb{R}^2$ they do. Later we are going to show that $\rho_{R1}(s,s) = \rho_{R2}(s,s)$, that is they coincide on the diagonal, when X_1 and X_2 are exchangeable. More general along the curve given by $F_1(x_1) = F_2(x_2)$ we will also have the same local correlation. Actually, as pointed out in Berentsen et al. (2012), when the two Rosenblatt transformations coincides it means that we have uniqueness. So let us now assume we can find a unique $\theta(x) = (\mu(x), \Sigma(x))$ from the Rosenblatt transformations. If we now remember equation (4.5), we chose a stepwise linear representation with $a_i = \mu(x)$ and $A_i = \Sigma^{\frac{1}{2}}$, and make it so that $x \in S_i$. Now $\theta_b(x) = \theta_i = (a_i, A_i A_i^T) = (\mu(x), \Sigma(x))$ with a b small enough. We now have a θ_b which solves the equations given by 4.3, and which obviously converges to θ when $b \to 0$. We conclude that in some cases and in some points we have a unique LGC given by the R1 transformation, which coincides with the likelihood approach described earlier. We will call this the theoretical LGC. It must also be pointed out that in the cases where the Rosenblatt transformations do not coincides, we can still find an estimate which will converge towards a unique solution $\theta_b(x)$ given by equation (4.3).

4.5 Some properties

We are quickly going to mention a couple of properties of the LGC from Tjøstheim and Hufthammer (2012).

4.5.1 Limits

The LGC will have the same limits as the ordinary correlation, that is we have $-1 \leq \rho_h(x, y) \leq 1$ and $-1 \leq \widehat{\rho_{h,n}} \leq 1$. This is easily seen by noting that both equation 4.2 and 4.3 contains the expression $\sqrt{1 - \rho^2}$. We also mention that if X and Y are independent, then $\rho_h(x, y) = 0$

4.5.2 LGC and tail dependence

For a bivariate normal distribution $X \sim N(\mu, \Sigma)$ the lower tail dependence is given by (see section on Gaussian copula later in the thesis for justification)

$$\lambda_l = 2 \lim_{s \to -\infty} P(\frac{X_2 - \mu_2}{\sigma_2} \leq s | \frac{X_1 - \mu_1}{1} = s) = 2 \lim_{s \to -\infty} \Phi(s \frac{\sqrt{1 - \rho}}{\sqrt{1 + \rho}}).$$

So if we now have the Gaussian approximation $V_x = (V_{x,1}, V_{x,2}) \sim N(\mu(x), \Sigma(x))$ at the point $x = (s, s)$. We then have

$$\lambda_l = 2 \lim_{s \to -\infty} P(\frac{V_{x,2} - \mu_2(x)}{\sigma_2(x)} \leq s | \frac{V_{x,1} - \mu_1(x)}{\sigma_1(x)} = s) = 2 \lim_{s \to -\infty} \Phi(s \sqrt{\frac{1 - \rho(s, s)}{1 + \rho(s, s)}}).$$

This means that there will be no lower tail dependence if $\rho(s, s) < 1$ for all $s < 0$. On the other if we have lower tail dependence we must have $\rho(s, s) \to 1$ as $s \to -\infty$.

4.6 Local correlation plot

To illustrate the local dependence we will use the local Gaussian correlation plot shown in Tjøstheim and Hufthammer (2012) and Berentsen et al. (2012). That is a levelplot with the two variables (X, Y) as horizontal and vertical scales. The plot is divided into equally sized cells, where a colour indicates the correlation. Cyan indicates negative LGC, magenta indicates positive LGC, while white indicates zero LGC. The LGC value in each cell is also printed with two decimals precision. These leveplots contains a lot of information, and can sometimes be confusing. A possibility is to only plot the diagonal or other lines. This will often give a clearer picture of the dependence, and in many cases it is the LGC on the diagonal we are interested in. Using only the diagonal also makes it easier to compare different distributions, since we can have many diagonals in one plot. In the theoretical LGC case we can only plot it where it is defined, usually on the diagonal. If nothing else is mentioned we will use the Gaussian kernel when estimating LGC. The choice of kernel will usually not influence the result that much, while the bandwidth choice will have a big impact on the LGC estimate. The bandwidths will in general be chosen after some experiments with different values, where we will try to find a balance between low variance and closeness to the usual kernel estimate with default bandwidth. We will rather oversmooth than undersmooth and try not to get to much noise in our LGC plots. Still to heavy smoothing will make the local Gaussian approximation deteriorate. For a discussion on optimal bandwidth choice and suggestions on algorithms see Tjøstheim and Hufthammer (2012) and Berentsen et al. (2012). The estimated plots will be compared to the theortical LGC where this is possible. Using scatter plots of the observations is also a good way to assess the reliability of our estimates, since they will show us where observations are scarce. We will see some clear boundary effects on our plot. When we estimates the tails, we use points closer to the middle of the plot in the estimation process. This will in general give a underestimation if the dependence is increasing towards the tails, and a overestimation if the dependence is decreasing towards the tails.

4.7 Example: The Gaussian distribution

A trivial example is the bivariate Gaussian distribution, where we will have constant parameters, that is $\theta_b(z) = \theta$ for all b. The estimate will naturally vary with h and z, and in general improve with increasing values of b, since the local likelihood equation in that case will tend to the global likelihood equation. As an illustration we have made a Local Gaussian Correlation map (see figure 4.1) from 5000 observations from a bivariate Gaussian distribution with global correlation 0.5. we have used bandwidths $b_1 = b_2 = 1$. The plot shows, as expected, that the local correlation is approximately 0.5, with more deviating values at the boundaries. This is because of less data in those regions. The local Gaussian approximation could off course be improved by different methods, for example by increasing the bandwidth.

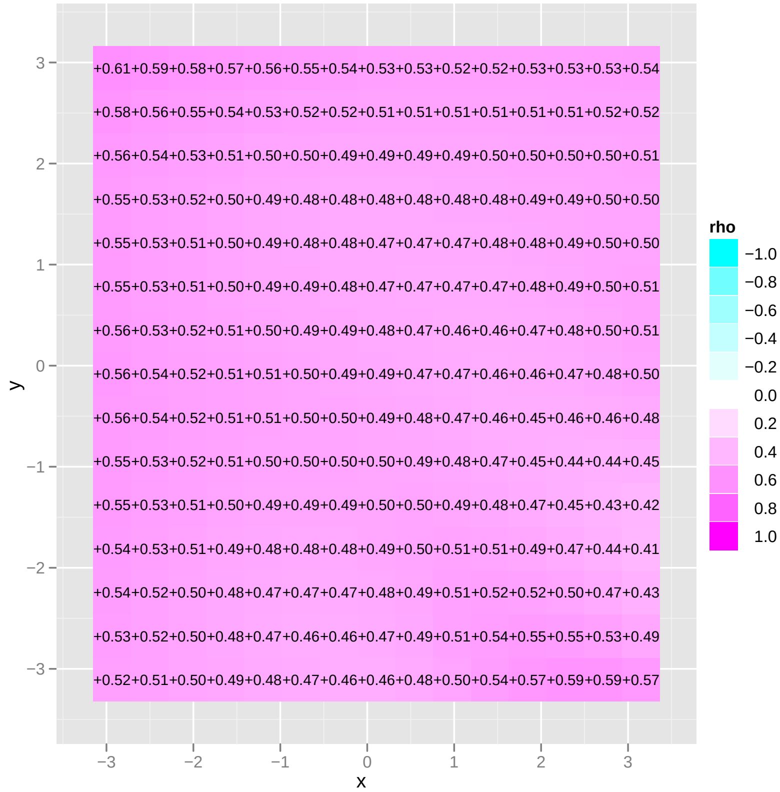

Figure 4.1: LGC of bivariate normal distribution with global correlation $\rho = 0.5$.

4.8 Symmetry

Let us have a look on how different forms of symmetry will effect the LGC. We will follow the presentation in Tjøstheim and Hufthammer (2012) and use the fact that these symmetries can be described by linear transformations. Let us look at a random variable $X = (X_1, X_2)$ with density f. It is assumed that $\mu = E(X) = 0$, since we otherwise can center the density at μ and discuss the symmetry around μ. If $y = Ax$ we have that $\Sigma_Y(y) = A\Sigma_X(x)A^T$ and $\mu_Y(y) = A\mu_X(x)$.

4.8.1 Radial symmetry

As we know radial symmetry around 0 means that $X = -X$, which means that we can write

$$-X = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} X = X. \tag{4.14}$$

This again lead us to

$$\Sigma(-x) = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \Sigma(x) \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} = \Sigma(x) \tag{4.15}$$

This shows that $\Sigma(x)$ and then also $\rho(x)$ has radial symmetry. The elliptic distributions, a class of distributions which we will discuss in more depth later, is known to have radial symmetry. So as an example let us look at an LGC-plot of a sample from an elliptic distribution, in this case a bivariate t-distribution with 4 degrees of freedom and $\rho = 0$. From the plot (figure 4.2) we can clearly see the radial symmetry around zero.

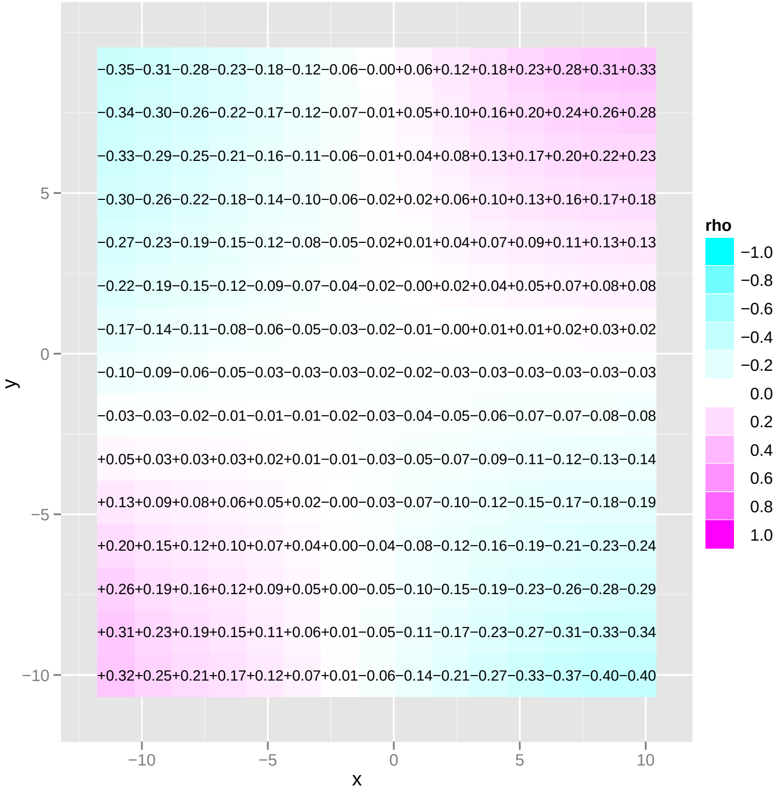

Figure 4.2: Local Gaussian correlation plot of bivariate t-distribution with 4 degrees of freedom and $\rho = 0$. Based on 5000 observations and bandwidth b=4.

4.8.2 Reflection symmetry

We say we have reflection symmetry if $f(-x_1, x_2) = f(x_1, x_2)$ and/or $f(x_1, -x_2) = f(x_1, x_2)$, that is reflection around one of the coordinate axis. With the matrices

$$A_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \text{ and } A_2 = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \tag{4.16}$$

we can describe the reflection symmetry by

$$\begin{bmatrix} x_1 \\ -x_2 \end{bmatrix} = A_1 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \text{ and } \begin{bmatrix} -x_1 \\ x_2 \end{bmatrix} = A_2 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}. \tag{4.17}$$

This gives us

$$\Sigma(x_1, -x_2) = A_1\Sigma(x)A_1^T = \begin{bmatrix} \sigma_1(x) & -\sigma_{12}(x) \\ -\sigma_{12}(x) & \sigma_2(x) \end{bmatrix}. \tag{4.18}$$

That is we have

$$\rho(x_1, -x_2) = -\rho(x_1, x_2) \text{ and } \rho(-x_1, x_2) = -\rho(x_1, x_2). \tag{4.19}$$

As a consequence of this we will have that $\rho(x)$ is zero along the coordinate axes. The t-distribution has reflection symmetry, and we can see the effect on the LGC-plot on figure 4.2.

4.8.3 Exchange symmetry

We know the random variables X_1 and X_2 are exchangeable if (X_1, X_2) and (X_2, X_1) are identically distributed. The matrix

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{4.20}$$

can be used to describe exchange symmetry by

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = A \begin{bmatrix} x_2 \\ x_1 \end{bmatrix}. \tag{4.21}$$

We have

$$\Sigma(x_1, x_2) = A\Sigma(x_2, x_1)A^T = \Sigma(x_2, x_1), \tag{4.22}$$

which again implies that $\rho(x_1, x_2) = \rho(x_2, x_1)$. So we can conclude that exchange symmetry of $f(x)$ implies exchange symmetry of $\rho(x)$. Again figure (4.2) illustrates this.

Chapter 5

LGC and copulas

If we have a continuous random variable (X, Y) with joint distribution function $F_{X,Y}$ and margins $F_X(x)$ and $F_Y(y)$ we know from Sklars theorem that we can write this as $F_{X,Y}(x,y) = C(F_X(x), F_Y(y))$, where C is the copula of (X, Y). On the other hand we can also couple two arbitrary distribution functions $F_X(x)$ and $F_Y(y)$ with a chosen copula an get a bivariate distribution function $F_{X,Y}(x,y) = C(F_X(x), F_Y(y))$. In this case F is often called a meta-distribution. We are now going to consider a general algorithm for generating observations from copula models and meta-distributions. This method is described for example in Embrechts, Lindskog and McNeil (2001). First we need to remark that for a copula C(u,v) we have the following expression for the conditional distribution for V given U=u:

$$C_u(u,v) = P(V \leq v | U = u) = \frac{\partial}{\partial u} C(u,v). \tag{5.1}$$

If we now have two independent standard uniformly distributed random variables U and T, then $(U, C_u^{-1}(T))$ has distribution C. The method is very general, and can be used for all copula models where C_u is invertible, and also for generating observations from a general n-copula. We get the observations (x, y) from $C(F_X(x), F_Y(y))$ by the following steps:

1. Generate two independent standard uniform variables u and t.
2. Set $v = C_u^{-1}(t)$.
3. Set $x = F_X^{-1}(u)$ and $y = F_y^{-1}(v)$.

5.1 Theoretical LGC for a copula

We can now define a function from $\mathbf{R}^2$ to $\mathbf{R}^2$ by

$$X = F_X^{-1}(\Phi(Z_1)),\ Y = F_Y^{-1}(C_{\Phi(Z_1)}^{-1}(\Phi(Z_2))).$$

and we immediately recognise this as the g function given by the R1 transformation discussed earlier. The inverse function h will be

$$Z_1 = \Phi^{-1}(F_X(X)),\ Z_2 = \Phi^{-1}(C_{F_X(X)}(F_Y(Y))). \tag{5.2}$$

We follow the approach in Berentsen et al. (2012) and introduce the notation

$$
\begin{aligned}
C_1(u,v) &= \frac{\partial}{\partial u} C(u,v) \\
C_{11}(u,v) &= \frac{\partial^2}{\partial u^2} C(u,v).
\end{aligned}
\tag{5.3}
$$

The h function from (5.2) can be written as

$$
h_1(x,y) = \Phi^{-1}(F_X(x)), \;\; h_2(x,y) = \Phi^{-1}(C_1(F_X(x), F_Y(y))).
$$

The partial derivatives of h is found by regular derivative rules, and we get

$$
\begin{aligned}
\frac{\partial h_1}{\partial x} &= \frac{f_X(x)}{\phi(\Phi^{-1}(F_X(x)))} \\
\frac{\partial h_2}{\partial x} &= \frac{C_{11}(F_X(x), F_Y(y)) f_X(x)}{\phi(\Phi^{-1}(C_1(F_X(x), F_Y(y))))}.
\end{aligned}
\tag{5.4}
$$

We now put these expression into the expression for the LGC under the R1 transformation (4.11), and find that

$$
\rho_{R1}(x,y) = -\frac{C_{11}(F_X(x), F_Y(y)) \phi(\Phi^{-1}(F_X(x)))}{\sqrt{\phi^2(\Phi^{-1}(C_1(F_X(x), F_Y(y)))) + C_{11}^2(F_X(x), F_Y(y)) \phi^2(\Phi^{-1}(F_X(x)))}}.
\tag{5.5}
$$

The same can be done for the R2 transformation. In this case we get the h' function

$$
h_1'(x,y) = \Phi^{-1}(C_2(F_X(x), F_Y(y))), \;\; h_2'(x,y) = \Phi^{-1}(F_Y(y)).
$$

The partial derivatives are

$$
\begin{aligned}
\frac{\partial h_1'}{\partial y} &= \frac{C_{22}(F_X(x), F_Y(y)) f_Y(y)}{\phi(\Phi^{-1}(C_2(F_X(x), F_Y(y))))} \\
\frac{\partial h_2'}{\partial y} &= \frac{f_Y(y)}{\phi(\Phi^{-1}(F_Y(y)))}.
\end{aligned}
\tag{5.6}
$$

Putting these into expression for the LGC under the R2 transformation (4.13) gives us

$$
\rho_{R2}(x,y) = -\frac{C_{22}(F_X(x), F_Y(y)) \phi(\Phi^{-1}(F_Y(y)))}{\sqrt{\phi^2(\Phi^{-1}(C_2(F_X(x), F_Y(y)))) + C_{22}^2(F_X(x), F_Y(y)) \phi^2(\Phi^{-1}(F_Y(y)))}}.
\tag{5.7}
$$

In general it is obvious that (5.5) and (5.7) does not necessarily give the same value. Let us now consider the diagonal, that is $x = y$, and look at the case when the two margins are equal, that is $F_X = F_Y = F$. Now we see that $\rho_{R1} = \rho_{R2}$ if $C_1 = C_2$ and $C_{11} = C_{22}$. We remember that for symmetric copula models (that is when they are exchangeable) we have $C(u,v) = C(v,u)$. This will imply that $C_1(u,v) = C_2(v,u)$ and $C_{11}(u,v) = C_{22}(v,u)$. If we now look at the diagonal we have that $C_1 = C_2$ and $C_{11} = C_{22}$. This means that for symmetric copulas with the same margins chosen, we have $\rho_{R1} = \rho_{R2} = \rho$ on the diagonal. We can also see that for margins that are not equal, (5.5) and (5.7) will coincide along the line given by the equation $F_X(x) = F_Y(y)$ when we have a symmetric copula. This means that we in this cases will consider (5.5) to be the theoretical LGC for a given copula model and margins.

5.2 Some properties

We give some properties presented in Berentsen et al. (2012).

5.2.1 The sign of the theoretical LGC

From (5.5) we see that the only element which is not necessarily positive is $-C_{11}(F_X(x), F_Y(y))$, which means that it will decide the sign of the theoretical LGC. This also have an intuitive interpretation. If we consider two stochastic variables X and Y which is positively dependent in the neighbourhood of the point (x_1, y_1). Positive dependence will in this setting mean that the function defined by $m(x) = P(Y \leq y_1 | X = x) = C_1(F_X(x), F_Y(y_1))$ is decreasing with regard to x around (x_1, y_1). This means that $m^{'}(x) = C_{11}(F_X(x), F_Y(y)) < 0$ and $\rho(x,y) > 0$ in the neighbourhood of (x_1, y_1).

5.2.2 Independence

If the two random variables in question are independent we will have the independence copula, that is $C(u,v) = uv$, which will imply that $\rho(x,y) = 0$ for all (x,y) because $C_{11}(u,v)$ clearly is zero in this case. If we now have $\rho(x,y) = 0$ this implies that $C_{11}(F_X(x), F_Y(y)) = 0$, which again lead us to conclude that $C_1(F_X(x), F_Y(y)) = P(Y \leq y | X = x)$ do not vary with respect to x. Naturally X and Y have to be independent. This result is not valid in general, only at areas where the thoretical LGC in (5.5) is defined.

5.3 Comments on problems and limitations

5.3.1 Uniqueness

The restriction on where the theoretical LGC is defined is of course a limitation to the theory. But even though it would be interesting to be able to calculate the theoretical LGC where ever we want, for the use of analysing the dependence structure of copula models the line given by $F_X(x) = F_Y(y)$ is often enough to get a good impression. We will in this text for the most part calculate (5.5) on the diagonal where we choose $F_X = F_Y$. This will give us a LGC line in the area where there usual is most observations, and also in the tails. If we want to use the LGC to measure the goodness of a fitted copula to a set of observations, for example by measuring the distance between the theoretical LGC and the estimated LGC, it is of course very limited to only be able to measure the fit on a grid consisting of points given by $F_X(x) = F_Y(y)$.

5.3.2 Restrictions on C

As mentioned this approach are restricted in the choice of copula and marginal distributions. Firstly we have only considered the continuous case, and secondly we are restricting us to symmetric copulas. Also we are assuming that the marginal distribution functions and $C_1(u,v)$ are invertible. That $C_1(u,v)$ is going to be invertible with respect to v is not the case for all copula models, but as we will see it is true for many of the most used and interesting models. But in general this is something that has to be checked before trying to calculate the LGC in this way from a general copula function. The derivative of $C_1(u,v)$ with respect to v is the density $c(u,v)$ of the copula, so one way to check that $C_1(u,v)$ is invertible will be to analyse the density function and look after possible points where $c(u,v) = 0$. As a final thing we observe that the final expression for the LGC contains C_{11} which has to exists and be calculated. These restrictions on the choice of copula models is not to limited, and poses no problems for the most common ones.

5.3.3 Complicated expression

The expression in equation (5.5) is not very nice, and in general it is difficult to get easy interpretable analytic functions $\rho(x,y)$ for real copula models. Some computer capacity is usually needed to calculate the theoretical LGC, but it is a lot faster than using the local likelihood approach. We have used the R programming language for calculating the LGC. In many cases (5.5) can be written with the help of already implemented functions in R, which makes it pretty straight forward to create easy algorithms for $\rho(x,y)$. One way to simplify is to choose standard normal distribution for the margins, that is $F_X(x) = \Phi(x)$ and $F_Y(y) = \Phi(y)$. This will simplify the calculations. Now we have $\phi_x = \phi(\Phi^{-1}(\Phi(x))) = \phi(x)$, which give us

$$\rho(x,y) = -\frac{C_{11}(\Phi(x),\Phi(y))\phi(x)}{\sqrt{(C_1(\Phi(x),\Phi(y)))^2 + (C_{11}(\Phi(x),\Phi(y)))^2\phi(x)^2}}. \tag{5.8}$$

5.3.4 Choice of margins

This procedure leads to an expression for the LGC of meta distributions, that is it depends on our choice of marginal distribution. It could sometimes be more interesting to look at copula model itself, that is the uniformly distributed random variables (U,V) with distribution C. This is because while modelling we would like to keep the fitting of the margins and the choice of dependence structure (the copula) separated. The algorithm described in the start of this section is used to generate observation (u,v) from a copula C also, in which we omit the third step. So it is natural now to try and find a LGC $\rho(u,v)$ by using the function g given by

$$U = \Phi(Z_1),\ V = C_u^{-1}(\Phi(Z_2)),$$

with the inverse function h given by

$$Z_1 = \Phi^{-1}(U),\ Z_2 = \Phi^{-1}(C_u(u,v)).$$

Now we can do more or less the same calculation as above, and we get the same expression only now without the marginal distribution, that is

$$\rho(u, v) = -\frac{C_{11}(u, v)\phi(\Phi^{-1}(C_1(u, v)))}{\sqrt{\phi(\Phi^{-1}(u))^2 + C_{11}(u, v)^2\phi(\Phi^{-1}(C_1(u, v)))^2}}.$$

But we have now overlooked some essential things, for example that we now have a g function which goes from $\mathbf{R}^2$ to $[0, 1]^2$ and that gives us problems. Actually the theory with the use of the non-linear transformation $X = g(Z)$ breaks down in the uniform case. The problem with the uniform distribution will also sometimes be noticed when we estimate the LGC with the local likelihood approach, especially at the boundaries. One question that arises now will be how the choice of margins affect the LGC. In theory the LGC should describe the whole dependence structure for the given data locally. But what is understood to be the whole dependence structure is debatable. In a lot of copula related articles it is often pointed out that the copula captures the whole dependence structure, and with this point of view different margins in the same copula model should not affect the dependence, nor preferably a dependence measure like the LGC. Some authors have criticized this separation of dependence and margins, and claims that the copula does not give a full picture of the dependence structure (see for example the critic from Mikosch (2006) and the response from Genest and Remillard (2006)). In this case the LGC might give a more complete insight into the dependence structure than copula based dependence measures. It is also possible that the LGC may contain other information than the dependence structure, in which case it could be affected by the margins. It is also off course an estimate, so the different margins may change the estimation process. In any case it might be reasonable to assume that the choice of standard normal margins will affect the procedure the least, and might come closest to showing the dependence structure of the copula.

Chapter 6

Ellitpical distributions and copulas

6.1 Elliptical distributions

The elliptical distributions is a wide class of multivariate distributions which have some of the same characteristics as the multivariate normal distribution. These distributions are easy to simulate from, and is used in many applications. The two most well-known elliptical distributions is the multivariate normal and the multivariate student t. From Sklars theorem it is easy to use this class of distributions to create useful Copula functions. Fang, Kotz and Ng (1987) gives a good introduction to elliptical distributions, and most definitions and theorems in this section is from their book. We start by defining spherical distributions, and then proceeds to the elliptical distributions. The spherical distributions are based on a famous property of the standard normal distribution, that is the invariance of orthogonal transformation.

Definition 17. *We say that a $n \times 1$ random vector X is spherical distributed if for every $n \times n$ orthogonal matrix A we have*

$$AX =_d X.$$

With the help of this definition we define elliptical distributions in the same way as we make general normal distribution from the standard normal distribution.

Definition 18. *We say that a $n \times 1$ random vector X is elliptical distributed with parameters μ ($n \times 1$ vector) and Σ ($n \times n$ matrix) if*

$$X =_d \mu + A^T Y,$$

where Y is spherical distributed and A is a $k \times n$ matrix such that $A^T A = \Sigma$ with rank(Σ)=k.

Elliptical distributions can be defined in other alternative ways. For example will the characteristic function $\psi(t)$ of an elliptical distribution X be of the form

$$\psi(t) = e^{it^T \mu} \phi(t^T \Sigma t)$$

for some scalar function ϕ. It is also true that X have the stochastic representation

$$X =_d \mu + rA^T U^{(k)} \tag{6.1}$$

where r and $U^{(k)}$ is independent random variables, $r > 0$ and $U^{(k)}$ is uniformly distributed on $\{z \in \mathbf{R}^k | z^t z = 1\}$. A and k are the same as in the definition. ϕ is called the characteristic generator, and it is often depending on a shape parameter, for example the degree of freedom in the multivariate t-distribution. We must point out that in the case of elliptical distributions ϕ is not the same as the density function of the standard normal distribution. The standard multivariate normal distribution has characteristic generator $\phi(u) = \exp(-\frac{u}{2})$. In general an elliptical distribution does not need to have a density. A necessary condition for the existence of a density is that rank(Σ)=n. We will mostly be interested in cases where the density exist, and then it must be of the form

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

where g is a non negative function and called the density generator. It is normal to denote an elliptical distribution as $E_n(\mu, \Sigma, \phi)$ or with the help of the density generator as $E_n(\mu, \Sigma, g)$. An elliptical distribution is fully described by μ, Σ and ϕ, but the representation $E_n(\mu, \Sigma, \phi)$ is not unique. (Embrechts, Lindskog and McNeil 2001) Σ and ϕ are only determined up to a positive constant, and if cov(X) exist one can choose ϕ so that cov(X)=Σ. We can write this more formally. That is if we have

$$X \sim E_n(\mu, \Sigma, \phi) \text{ and } X \sim E_n(\mu^*, \Sigma^*, \phi^*) \tag{6.2}$$

we get that

$$\mu^* = \mu,\ \Sigma^* = c\Sigma,\ \phi^*(\cdot) = \phi(\frac{\cdot}{c}) \tag{6.3}$$

where c is a positive constant. We get cov(X)=Σ by choosing $c = \frac{E(r^2)}{n}$.

Elliptical distributions have many desirable properties. Marginal distributions, linear combinations, and conditional distributions of elliptical variables will also bee elliptical. To write some of these important properties in a informal way we will list a few theorems, which proofs can be found in Fang, Kotz and Ng (1987).

Theorem 19. *Let $X \sim E_n(\mu, \Sigma, \phi)$ with rank(Σ)=k. Also let B be an $n \times m$ matrix and v be an $m \times 1$ vector. Then*

$$v + B^T X \sim E_n(v + B^T \mu, B^T \Sigma B, \phi).$$

We now partition X, μ and Σ into

$$X = \begin{bmatrix} X^{(1)} \\ X^{(2)} \end{bmatrix}, \mu = \begin{bmatrix} \mu^{(1)} \\ \mu^{(2)} \end{bmatrix}, \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}.$$

Here $X^{(1)}$ and $\mu^{(1)}$ is $m \times 1$ vectors and Σ_{11} is a $m \times m$ matrix.

Theorem 20. *Let $X \sim E_n(\mu, \Sigma, \phi)$, then*

$$X^{(1)} \sim E_m(\mu^{(1)}, \Sigma_{11}, \phi) \tag{6.4}$$

and

$$X^{(2)} \sim E_{n-m}(\mu^{(2)}, \Sigma_{22}, \phi) \tag{6.5}$$

Theorem 21. *Let* $X =_d \mu + rA^T U^{(n)} \sim E_n(\mu, \Sigma, \phi)$*, where* $\Sigma > 0$*. Then*

$$(X^{(1)}|X^{(2)} = x_0^{(2)}) =_d \mu_{1.2} + r_{q(x_0^{(2)})} A_{11.2}^T U^{(n)} \sim E_m(\mu_{1.2}, \Sigma_{11.2}, \phi_{q(x_0^{(2)})}) \quad (6.6)$$

where

$$\mu_{1.2} = \mu^{(1)} + \Sigma_{12}\Sigma_{22}^{-1}(x_0^{(2)} - \mu^{(2)}),$$
$$\Sigma_{11.2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21},$$
$$q(x_0^{(2)}) = (x_0^{(2)} - \mu^{(2)})^T \Sigma_{22}^{-1}(x_0^{(2)} - \mu^{(2)}).$$

Also the distribution of $r_{q(x_0^{(2)})}$ *is*

$$r_{q(x_0^{(2)})} =_d ((r^2 - q(x_0^{(2)}))^{\frac{1}{2}}|X^{(2)} = x_0^{(2)}).$$

Linear correlation is a natural measure of dependence when working with elliptical distributions. Remember that, when $0 < Var(X_i), Var(X_j) < \infty$, we have the linear correlation $\rho_{ij} = \frac{Cov(X_i,X_j)}{\sqrt{Var(X_i)Var(X_j)}}$. We can extend this to a definition also usable when the variances does not exists, that is $\rho_{ij} = \frac{\Sigma_{ij}}{\sqrt{\Sigma_{ii}\Sigma_{jj}}}$. There is a simple relationship between the linear correlation coefficient and Kendalls tau for elliptical distributed random vectors with continuous marginals. We state it in a theorem, see Hult and Lindskog (2002) for details.

Theorem 22. *Let* $X \sim E_n(\mu, \Sigma, \phi)$*, where* X_i *and* X_j *are continuous for* $i, j \in \{1, \dots, n\}$*. Then*

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

In the world of finance and risk management elliptical distributions is useful because minimizing VaR are equivalent with the portfolio optimizing theory introduced by Markowitz, where the variance is used as a risk measure. VaR is also a coherent risk measure when using elliptical distributions. (Embrechts, McNeil and Strauman (1999)) One should be aware off that the equivalence between uncorrelated random variables and independence found in the multivariate normal distribution does not generalize to other elliptical distributions.

We will now only focus on the bivariate elliptical distribution. An important property with elliptical distributions is that they are radially symmetric, so we would expect the LGC plots to also be radial symmetric. If the principal axis is along $x_1 = x_2$ the distribution will be exchangeable. Let us now assume that $\mu_1 = \mu_2 = 0$ and that $\sigma_1 = \sigma_2$. We will have $[x_1, x_2]\Sigma^{-1}[x_1, x_2]^T = [x_2, x_1]\Sigma^{-1}[x_2, x_1]^T$, which means that the distribution will be exchangeable. This implies a reflection symmetry along the lines $x_1 = x_2$ and $x_1 = -x_2$ for $f(x)$ and $\rho(x)$.

One problem with using this distribution class for multivariate modelling is that the margins are of the same type. For example do the margins for the bivariate student t-distribution have the same degree of freedom. A more flexible way is to use suitable marginal distributions together with an elliptical copula.

6.2 Elliptical copulas

Elliptical copulas are created by using Sklars theorem on elliptical distributions. In this way we can capture the dependence structure from the multivariate elliptical distributions, but choose marginal distributions freely. Elliptical copulas together with chosen margins is also called meta-elliptical distributions. See Fang, Fang and Kotz (2002) and Abdous, Genest and Rèmillard (2005) for properties and applications. If we have an 2 dimensional elliptical distribution with cdf Q and margins Q_1 and Q_2, and choose two univariate cdfs F_1 and F_2, we have a meta-elliptical distribution with cdf $H(x,y) = Q(Q_1^{-1}(F_1(x)), Q_2^{-1}(F_2(y)))$. We will in the following sections only consider continuous elliptical distributions with an existing density function. It is also common practice to standardize the distributions slightly. So without loss of generality we will set $\mu = 0$ and $\Sigma = R = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$. From (6.4) and (6.5) we know that this means that the margins of the elliptical distribution will be the same, and we denote it by Q_g. We can now write the elliptical copula as

$$C(u,v) = \frac{1}{\sqrt{1-\rho^2}} \int_{-\infty}^{Q_g^{-1}(u)} \int_{-\infty}^{Q_g^{-1}(v)} g(\frac{x^2 + y^2 - 2\rho xy}{1-\rho^2}) dxdy$$

6.2.1 Theoretical LGC

We remember that in order to calculate the theoretical LGC we need to find $C_1(u,v) = P(V \leq v | U = u)$ and $C_{11}(u,v)$. If we have a copula function C based on an elliptical distribution $E_2(0, R, g)$ with margins Q_g we know that $C(Q_g(x), Q_g(y))$ will have distribution $E_2(0, R, g)$. This means that $C_1(u,v) = P(V \leq v|U = u) = P(Q_g^{-1}(V) \leq Q_g^{-1}(v)|Q_g^{-1}(U) = Q_g^{-1}(u)) = P(Y \leq y|X = x)$, where $(X,Y) \sim E_2(0, R, g)$, are given by (6.6). The conditional distribution is a one dimensional elliptical distribution, but it is not always obvious which kind of elliptical distribution. In some cases some general rules applies, which makes the calculation of the theoretical LGC a pretty straight forward task. We are going to have a look at some of these examples.

Normal distribution

It is a well known result that given a two dimensional standard normal distributed vector (X,Y) with correlation ρ, we have that $Y|X = x_0 \sim N(\rho x_0, 1 - \rho^2)$. This is a result of (6.6). We note that in the case of the normal distribution, the characteristic generator of the conditional distribution will be $\phi_{q(x_0)} = exp(-\frac{u}{2})$, that is independent of x_0. This is actually a quality that characterizes the normal distribution, see Fang, Kotz and Ng (1987).

Symmetric Pearson type V11 distributions

We say that a $n \times 1$ random vector X is symmetric multivariate Pearson type V11 distributed if it has the following density generator g

$$g(t) = (\pi m)^{-\frac{n}{2}} \frac{\Gamma(N)}{\Gamma(N - \frac{n}{2})} (1 + \frac{t}{m})^{-N}$$

where $N > \frac{n}{2}$ and $m > 0$. In this case $\frac{r^2}{m}$ will be beta type 11 distributed with parameters $\frac{n}{2}$ and $N - \frac{n}{2}$. Here r is the generating variate in the stochastic representation of X, see equation 6.1. A stochastic variable B that is beta type 11 distributed with parameters α and β will be denoted $B \sim Be11(\alpha, \beta)$, and its density will be

$$\frac{1}{B(\alpha, \beta)} b^{\alpha-1}(1 + b)^{-(\alpha+\beta)},$$

where $b > 0$ and $B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$ is the beta function. In the bivariate simplified version we are looking at we get $\frac{\Gamma(N)}{\Gamma(N-\frac{1}{2})} = \frac{\Gamma(N-1)(N-1)}{\Gamma(N-1)} = N - 1$. Using $\mu = 0$ and $\Sigma = R$ we get the density

$$f(x_1, x_2) = \frac{N-1}{\pi m\sqrt{1-\rho^2}}(1 + \frac{x_1^2 + x_2^2 - 2\rho x_1 x_2}{m(1-\rho^2)})^{-N}.$$

We now use (6.6) and find that

$$(X_1|X_2 = x_0) =_d \mu_{1.2} + r\Sigma_{11.2}^{\frac{1}{2}}u, \tag{6.7}$$

where $\mu_{1.2} = \rho x_0$ and $\Sigma_{11.2} = 1-\rho^2$. We also have that $\frac{r^2}{m+x_0^2} \sim Be11(\frac{1}{2}, N-\frac{1}{2})$ (see Fang, Kotz and Ng (1987)). This means that $X_1|X_2 = x_0$ also will be symmetric Pearson type V11 distributed with new parameters $\mu^* = \rho x_0$, $m^* = m + x_0^2$ and $N^* = N$. A special case of the symmetric Pearson type V11 distribution is the Student t distribution, which we get by setting $N = \frac{1}{2}(n+m)$. Here n is the dimension, and m is what we usually call the degree of freedom. Let us now assume that (X_1, X_2) is t distributed (in the standard way with $\mu = 0$) with m degrees of freedom and correlation ρ, and we want to find the distribution of $X_1|X_2 = x_0$. We have the stochastic representation in equation 6.7, but let us now define another random variable $r_* := \sqrt{\frac{m+1}{m+x_0^2}}r$. This gives us

$$\frac{r_*^2}{m+1} = \frac{1}{m+1}\frac{m+1}{m+x_0^2}r^2 = \frac{r^2}{m+x_0^2} \sim Be11(\frac{1}{2}, N - \frac{1}{2}).$$

Thus we can write

$$(X_1|X_2 = x_0) =_d \rho x_0 + r_*\sqrt{\frac{m+x_0^2}{m+1}(1-\rho^2)}u$$

and by standardizing we can write

$$P(X_1 \leq x_1|X_2 = x_0) = t_{m+1}(\frac{x_1 - \rho x_0}{\sqrt{\frac{m+x_0^2}{m+1}(1-\rho^2)}})$$

where t_{m+1} is the cdf of a one dimensional standard t distribution with m+1 degrees of freedom. So we see that the conditional distribution of a two dimensional t distribution can be written as a one dimensional t distribution. We are going to use this result later to develop theoretical LGC expressions.

Let us now take a closer look at the two most used elliptical copulas, namely the Gaussian copula and the t copula.

6.2.2 The Gaussian copula

The bivariate Gaussian copula is defined by

$$\begin{aligned}
C(u,v) &= \Phi_{2,\rho}(\Phi^{-1}(u),\Phi^{-1}(v)) \\
&= \int_{-\infty}^{\Phi^{-1}(u)}\int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi\sqrt{1-\rho}}\exp\{-\frac{x^2-2\rho xy+y^2}{2(1-\rho^2)}\}dxdy
\end{aligned}$$

Here $\Phi_{2,\rho}$ is the the bivariate standard normal distribution function with correlation ρ, and Φ^{-1} is the inverse of the univariate standard normal distribution function. As we can see the only parameter is the linear correlation ρ.

Tail dependence

We remember from (3.5) that we can write the coefficient for the lower tail dependence as $\lambda_l = \lim_{t\to 0^+}\frac{C(t,t)}{t}$. We can use L'Hopitals rule on this expression, and then we use that $\frac{\partial C(u,v)}{\partial u} = P(V \le v|U = u)$. This give us (Embrechts, Lindskog and McNeil (2001)

$$\begin{aligned}
\lambda_l &= \lim_{t\to 0^+}\frac{C(t,t)}{t} \\
&= \lim_{t\to 0^+}\frac{dC(t,t)}{dt} \\
&= \lim_{t\to 0^+}\frac{\partial C(u,v)}{\partial u}|_{u=t,v=t} + \frac{\partial C(u,v)}{\partial v}|_{u=t,v=t} \\
&= \lim_{t\to 0^+} P(V<t|U=t) + P(U<t|V=t) \\
&= 2\lim_{t\to 0^+} P(V<t|U=t),
\end{aligned}$$

where the last equality comes from the fact that the Gaussian copula is exchangeable. If (X,Y) is bivariate standard normal we have that $(X,Y) \sim C(\Phi(x),\Phi(y))$ where C is the Gaussian copula with correlation ρ. From earlier we know that $X|Y = y \sim N(\rho y, 1-\rho^2)$, which gives us

$$\begin{aligned}
\lambda_l &= 2\lim_{x\to-\infty} P(\Phi^{-1}(V)<x|\Phi^{-1}(U)=x) \\
&= 2\lim_{x\to-\infty} P(X<x|Y=x) \\
&= 2\lim_{x\to-\infty} \Phi(\frac{x-\rho x}{\sqrt{1-\rho^2}}) \\
&= 2\lim_{x\to-\infty} \Phi(x\sqrt{\frac{1-\rho}{1+\rho}} = 0
\end{aligned}$$

for $\rho \in (-1,1)$. Because the Gaussian copula is radial symmetric we also have $\lambda_u = 0$. We can say that the Gaussian copula is asymptotic independent, which means that extreme events will happen independently. But it does not mean that it is independent. Coles, Heffernan and Tawn (1999) points out that the convergence $\lambda \to 0$ is very slow for $\rho > 0$. They show that the alternative coefficient $\overline{\lambda}$ (see (3.8)), which is an increasing dependence measure in the case of asymptotic independence, for the Gaussian copula is $\overline{\lambda} = \rho$.

The LGC

Since the Gaussian copula is constructed by the multivariate normal distribution, we know that the correlation ρ will completely determine the dependence structure. From earlier we know that the local Gaussian correlation will be constant the same as ρ in the case of multivariate normal distribution. We will now see that this is a property of the Gaussian copula.

From the calculations above, we know that

$$\begin{aligned} C_1(u,v) &= P(V \leq v | U = u) = P(\Phi^{-1}(V) \leq \Phi^{-1}(v) | \Phi^{-1}(U) = \Phi^{-1}(u)) \\ &= \Phi(\frac{\Phi^{-1}(v) - \rho\Phi^{-1}(u)}{\sqrt{1-\rho^2}}). \end{aligned}$$

Let us set $R(u,v) := \frac{\Phi^{-1}(v)-\rho\Phi^{-1}(u)}{\sqrt{1-\rho^2}}$ to ease notation. If we differentiate one more time we get

$$\begin{aligned} C_{11}(u,v) &= \frac{\partial}{\partial u}\Phi(\frac{\Phi^{-1}(v) - \rho\Phi^{-1}(u)}{\sqrt{1-\rho^2}}) \\ &= \frac{-\frac{\partial}{\partial u}\Phi^{-1}(u)\rho}{\sqrt{1-\rho^2}}\phi(R) \\ &= \frac{-\rho}{\sqrt{1-\rho^2}\phi(\Phi^{-1}(u))}\phi(R). \end{aligned}$$

Let us now choose some arbitrary margins F_X and F_Y, and try to calculate the LGC. We now have $C_1(F_X(x), F_Y(y)) = \Phi(R)$ which gives $\phi(\Phi^{-1}(C_1(F_X(x), F_Y(y))) = \phi(R)$. For simplification let us also denote $\phi_x := \phi(\Phi^{-1}(F_X(x)))$, which gives us $C_{11}(F_X(x), F_Y(y)) = -\frac{\rho}{\sqrt{1-\rho^2}}\frac{\phi(R)}{\phi_x}$. Putting all this into (5.5) gives us

$$\begin{aligned} \rho(x,y) &= \frac{\frac{\rho}{\sqrt{1-\rho^2}}\frac{\phi(R)\phi_x}{\phi_x}}{\sqrt{\phi(R)^2 + \frac{\rho^2}{1-\rho^2}\frac{\phi(R)^2\phi_x^2}{\phi_x^2}}} \\ &= \frac{\frac{\rho}{\sqrt{1-\rho^2}}}{\sqrt{1+\frac{\rho^2}{1-\rho^2}}} = \frac{\rho}{\sqrt{1-\rho^2+\rho^2}} = \rho \end{aligned}$$

We see that the LGC for the Gaussian copula is constant and the same as the linear correlation, and is not depending on the choice of margins. This is the only copula we have encountered with this property. We get the same result for the R2 transformation, which means that this result is valid for all points in $\mathbb{R}^2$.

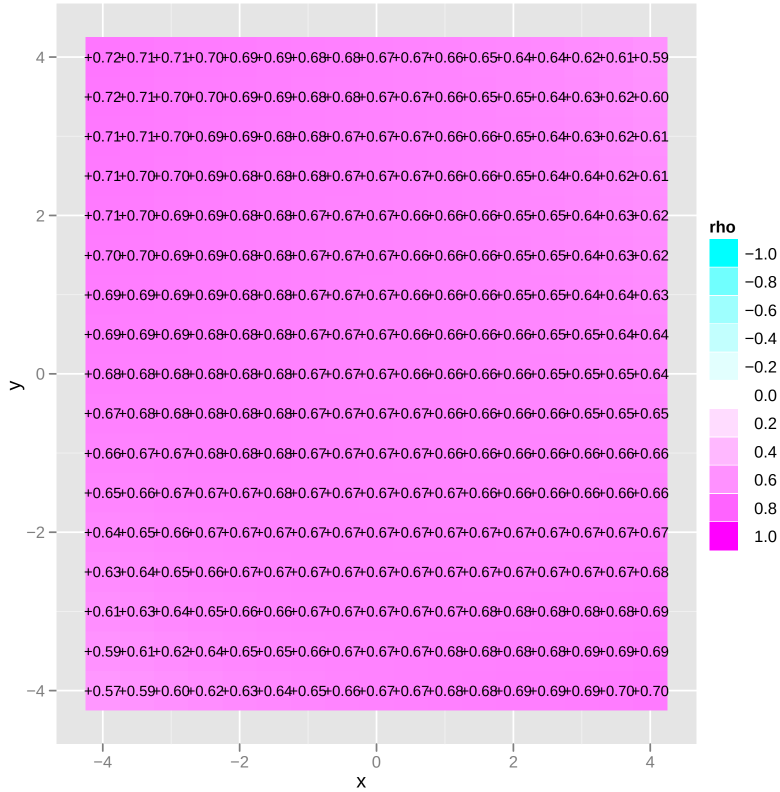

Figure 6.1: Estimated LGC plot of a Gaussian copula with $\rho = 0.7$ and t distributed margins (4 degrees of freedom). Based on 5000 observations and with bandwidth b=2.

6.2.3 t copula

The bivariate t-copula is defined by

$$\begin{aligned} C^t_{\nu,\rho}(u,v) &= t_{\nu,\rho}(t_\nu^{-1}(u), t_\nu^{-1}(v)) \\ &= \int_{-\infty}^{t_\nu^{-1}(u)} \int_{-\infty}^{t_\nu^{-1}(v)} \frac{1}{2\pi(1-\rho^2)^{\frac{1}{2}}} \{1 + \frac{s^2 - 2\rho st + t^2}{\nu(1-\rho^2)}\}^{-\frac{(\nu+2)}{2}} ds dt. \end{aligned}$$

If $\nu > 2$ ρ is just the usual linear correlation coefficient of the corresponding bivariate t-distribution.

Tail dependence

In contrast to the Gaussian copula the t-copula is not asymptotic independent. We can find an expression for the tail dependence by using the fact that if $(X,Y) \sim t_{\nu,\rho}$ we have $P(X_1 \leq x_1 | X_2 = x_0) = t_{m+1}(\frac{x_1 - \rho x_0}{\sqrt{\frac{m+x_0^2}{m+1}(1-\rho^2)}})$. If we now use this in the calculations done for the lower tail dependence in the Gaussian case we get

$$
\begin{aligned}
\lambda_l &= 2 \lim_{x \to -\infty} P(X < x | Y = x) \\
&= 2 \lim_{x \to -\infty} t_{\nu+1}\left(\frac{x - \rho x}{\sqrt{\frac{\nu + x^2}{\nu+1}(1 - \rho^2)}}\right) \\
&= 2 \lim_{x \to -\infty} t_{\nu+1}\left(-\sqrt{\frac{1-\rho}{1+\rho}}\sqrt{\frac{\nu+1}{1+\nu/x^2}}\right) \\
&= 2 t_{\nu+1}\left(-\sqrt{\frac{1-\rho}{1+\rho}}\sqrt{\nu+1}\right)
\end{aligned}
$$

Since the t-copula is exchangeable we know that $\lambda_l = \lambda_u$. We see from the expression above that tail dependence increases with increasing correlation ρ, and decreases with increasing degree of freedom ν. This is illustrated in table(6.1), where the tail dependence coefficient is calculated for some values of ρ and ν. What can be noted is that even when ρ is zero or negative, we still have tail dependence. We also see that as the degrees of freedom tends to infinity, the t-copula tends towards the Gaussian structure.

The LGC

From the calculations above we know that

$$
C_1(u,v) = t_{\nu+1}\left(\frac{t_\nu^{-1}(v) - \rho t_\nu^{-1}(u)}{\sqrt{\frac{(\nu + t_\nu^{-1}(u)^2)(1-\rho^2)}{\nu+1}}}\right).
$$

Set $a = t_\nu^{-1}(v) - \rho t_\nu^{-1}(u)$, $b = \sqrt{\frac{(\nu + t_\nu^{-1}(u)^2)(1-\rho^2)}{\nu+1}}$ and $R = \frac{a}{b}$. Then we get

$$
C_{11} = \frac{\partial R}{\partial u} f_{t_{\nu+1}}(R),
$$

where $f_{t_{\nu+1}}$ is the density of the univariate t-distribution with $\nu + 1$ degrees of freedom.

$$
\frac{\partial a}{\partial u} = \frac{-\rho}{f_{t_\nu}(t_\nu^{-1}(u))}.
$$

$$
\frac{\partial b}{\partial u} = \frac{t_v^{-1}(u)(1-\rho^2}{b(\nu+1) f_{t_\nu}(t_\nu^{-1}(u))}.
$$

$$
\begin{aligned}
\frac{\partial R}{\partial u} &= \frac{a' b - b' a}{b^2} \\
&= -\frac{\frac{\rho b}{f_{t_\nu}(t_\nu^{-1}(u))} - \frac{a t_\nu^{-1}(u)(1-\rho^2)}{b(\nu+1) f_{t_\nu}(t_\nu^{-1}(u))}}{b^2} \\
&= \frac{\frac{-\rho b}{f_{t_\nu}(t_\nu^{-1}(u))} - \frac{R t_\nu^{-1}(u)(1-\rho^2)}{(\nu+1) f_{t_\nu}(t_\nu^{-1}(u))}}{b^2} \\
&= \frac{-\rho b - \frac{R t_\nu^{-1}(u)(1-\rho^2)}{\nu+1}}{b^2 f_{t_\nu}(t_\nu^{-1}(u))} \\
&= \frac{-1}{f_{t_\nu}(t_\nu^{-1}(u)) b^2}\left(\rho b + \frac{1-\rho^2}{\nu+1} t_\nu^{-1}(u) R\right).
\end{aligned}
$$

These expressions can now be put into the LGC expression in (5.5). The final theoretical LGC for the t-copula is not easy to interpret analytically, but we note that the elements described above is only consisting of the known parameters and the density, distribution and quantile functions of the regular one dimensional t distribution. As the t distribution has its own packages in R, it is pretty easy to make R routines which calculates the theoretical LGC and then plot it. For margin choices we have implemented the normal distribution, t distribution and the skewed t distribution.

$\nu\backslash\rho$	-0.9	-0.5	0	0.5	0.9	1
2	0.00	0.06	0.18	0.39	0.72	1.00
4	0.00	0.01	0.08	0.25	0.63	1.00
10	0.00	0.00	0.01	0.08	0.46	1.00
100	0.00	0.00	0.00	0.00	0.02	1.00

Table 6.1: Table of tail dependence for the bivariate t-copula.

Plots

The next couple of pages contains different LGC plots based on the t copula (figures 6.2 to 6.9), and some additional plots concerning the t copula can be found in the appendix (figures A.1 to A.5). Figure 6.2 shows the theoretical LGC on the diagonal for the t copula with normal margins and 4 degrees of freedom, and with different correlations ρ. As expected the LGC is symmetric around zero. The minimum value occurs at the origin, and the LGC at this point is very close to the correlation value. We see that the LGC increases with ρ. With negative correlations an area around zero has negative LGC, but still we see that the tails has positive dependence as expected. Lower ρ only means that the LGC in the tails approaches 1 slower. The estimated LGC plots in figures 6.3 to 6.5 show the same patterns, except some small deviations caused by the estimation. We see there is area around the origin where the LGC is more or less the same as ρ, and that the LGC is increasing when you move towards quadrant one and three. In addition we see that the other two quadrants has negative LGC. The symmetry in the dependence structure is obvious also on these plots. We can see that the LGC is symmetric around $x = y$ and $x = -y$. Figure 6.6 shows theoretical LGC plots for t copula with $\rho = 0.5$ and different values of the degree of freedom ν. Increasing values of ν gives a LGC where the value around the origin gets closer to the correlation value, and in the tails the LGC approaches one with a slower rate when ν gets larger. We know that the t copula converges to the gaussian when $\nu \to \infty$, so it is no surprise that the LGC approaches the constant LGC value of 0.5 when the degree of freedom gets large on this plot. The estimated plots in figures 6.7 and 6.8 shows more or less the same. There is some noise in those plots, especially in 6.8, but still we see that the area with positive dependence close to 0.5 increases as we increase ν. Figure 6.9 shows the theoretical LGC calculated with a t copula with $\rho = 0$ and $\nu = 4$, but different kinds of margins. It is of great interest to see how different choices of margins affects the LGC and the dependence in general, and if it is possible to create an asymmetric dependence structure with the help of the symmetric t copula together with some choice of margins. We have looked at t distributed margins with 4 degrees of freedom, which only gives us the structure of the bivariate t distribution. We have also used the skewed t distribution from the R package sn. This is a generalisation of the regular t distribution, but with an extra shape parameter which controls the skewness. If we set the shape parameter equals to zero we get the t distribution. For details see Azzalini, A. and Capitanio, A. (2003), and the section on skewed elliptical distributions in this thesis. By using the t distributed margins we see that the LGC has similar shape as when we use normal margins, the difference is that with t margins the LGC increases slower in the tails. But with the use of the skewed t margins the LGC gets an asymmetric shape. We see that the point of the minimum is shifted to the right. There is clearly stronger dependence in the left tail than in the right tail. It interesting to note that with increasing shape parameter the left tail gets higher LGC value, while the right tail looks the same. The figures A.1 to A.3 shows the estimated LGC with the different margins choices. We see the same pattern here. With the t margins we have a symmetric pattern, but with a little less dependence in the tails than with normal margins. It might look like there is a little less dependence in quadrant one compared to quadrant three in A.1, but this is probably only estimation errors. We clearly see the asymmetric shape when we use the skewed t margins, with stronger dependence in quadrant three. With larger shape parameter the area with positive dependence is also larger in the third quadrant. Figure A.4 and A.5 shows the estimated LGC for t copula with uniform margins, that is the LGC estimated directly on the copula model without choosing margins. We can perhaps say that these plots shows the dependence structure of the copula model itself. As mentioned in an earlier section we can not calculate a theoretical LGC value when we use uniform margins, at least not with the method we have used on the other models. Still we can estimate the LGC, though we have to be careful cause it can be difficult to get good estimates when we are using the uniform distribution on $[0, 1]$. The plots looks reasonable, and by comparing it with figures 6.3 and 6.4 we see that it is pretty much the same pattern.

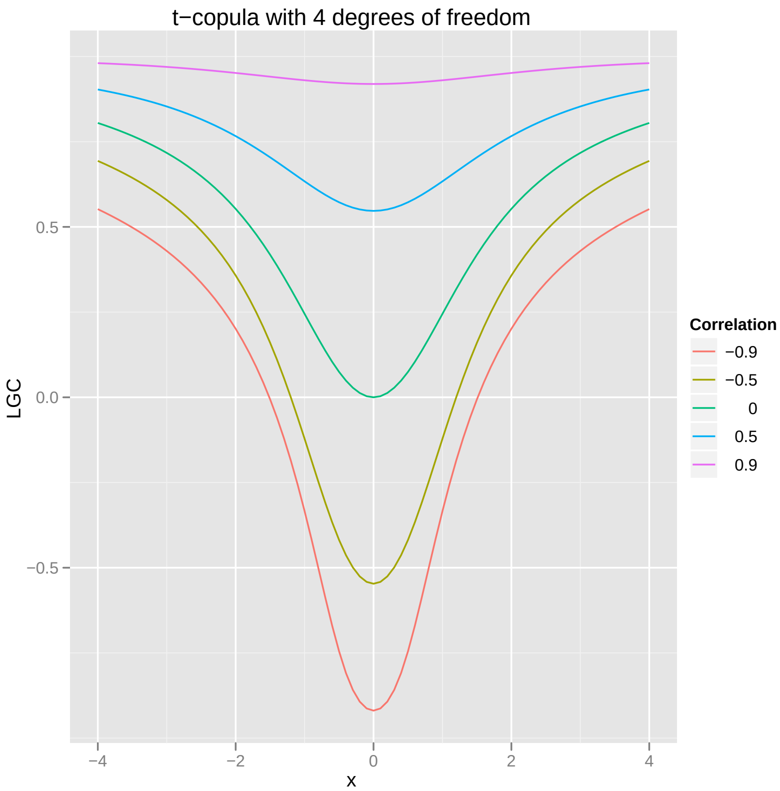

Figure 6.2: Theoretical LGC plot of t-copula with $\nu = 4$, $\rho = -0.9, -0.5, 0, 0.5, 0.9$ and normal margins.

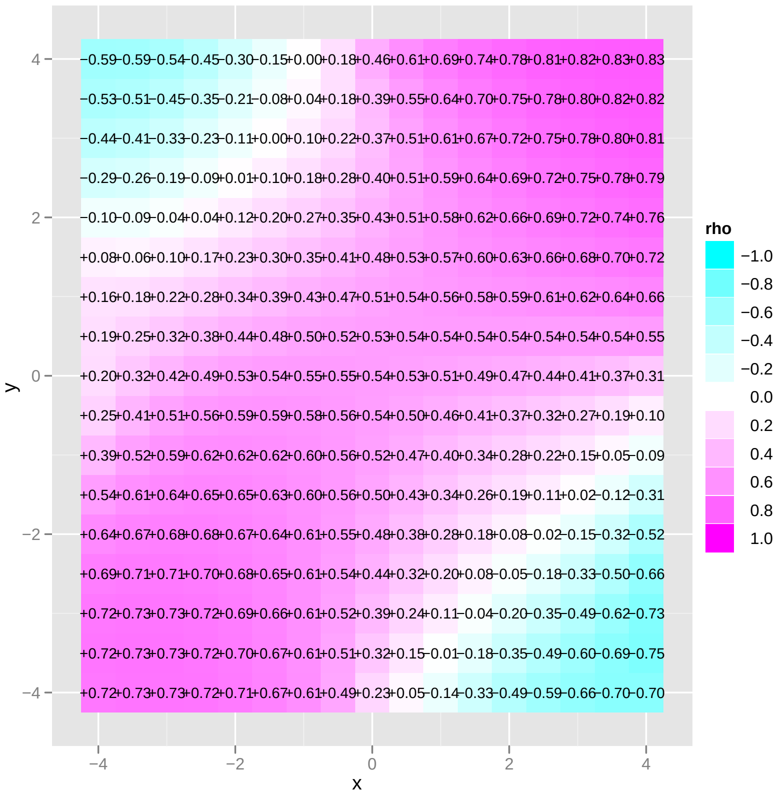

Figure 6.3: Estimated LGC plot of t-copula with $\rho = 0.5$, $\nu = 4$ and normal margins. Based on 5000 generated observations and bandwidth b=1.

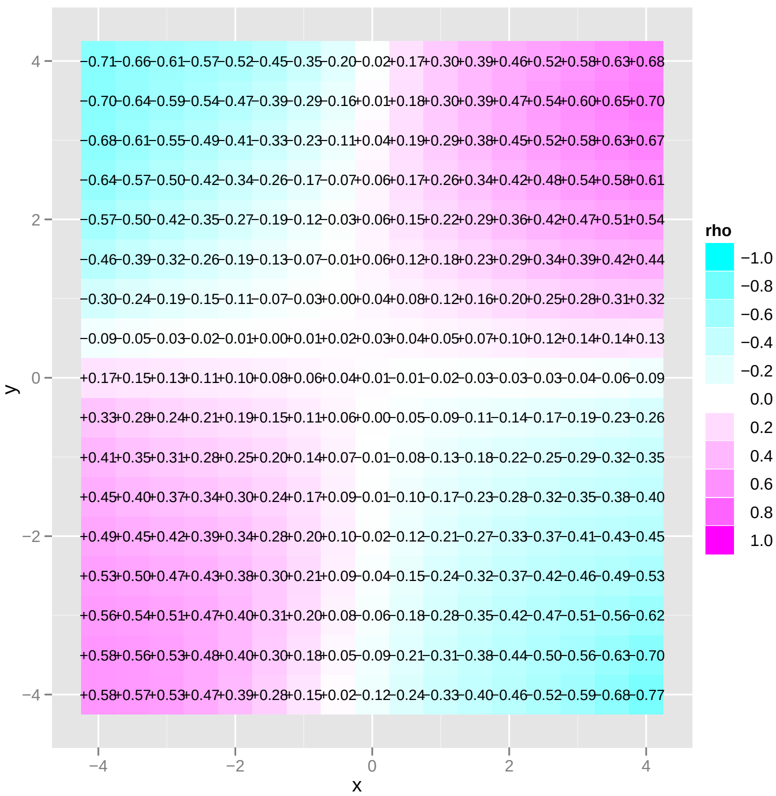

Figure 6.4: Estimated LGC plot of t-copula with $\rho = 0$, $\nu = 4$ and normal margins. Based on 5000 generated observations and bandwidth b=1.

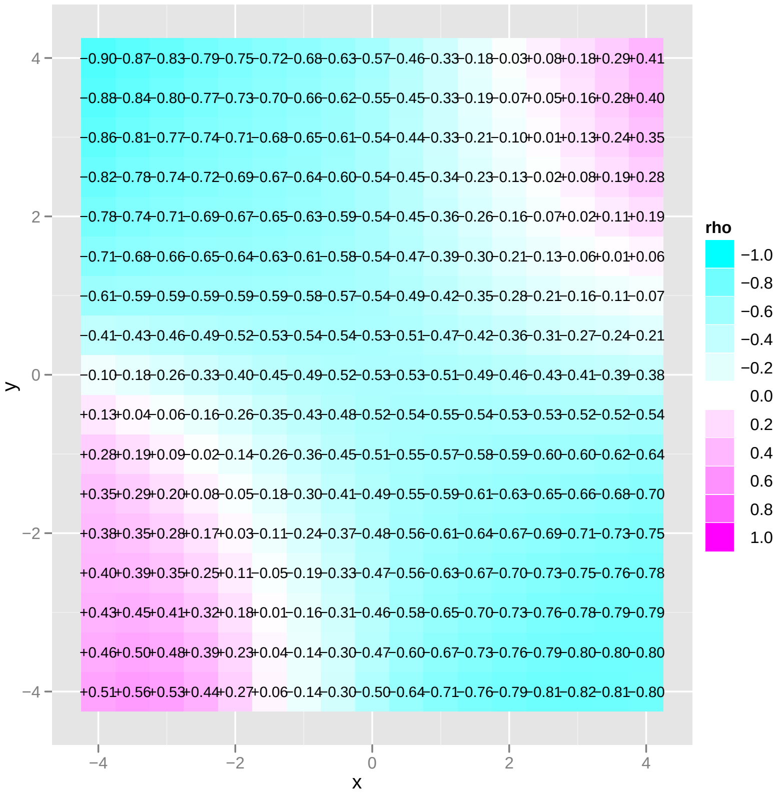

Figure 6.5: Estimated LGC plot of t-copula with $\rho = -0.5$, $\nu = 4$ and normal margins. Based on 5000 generated observations and bandwidth b=1.

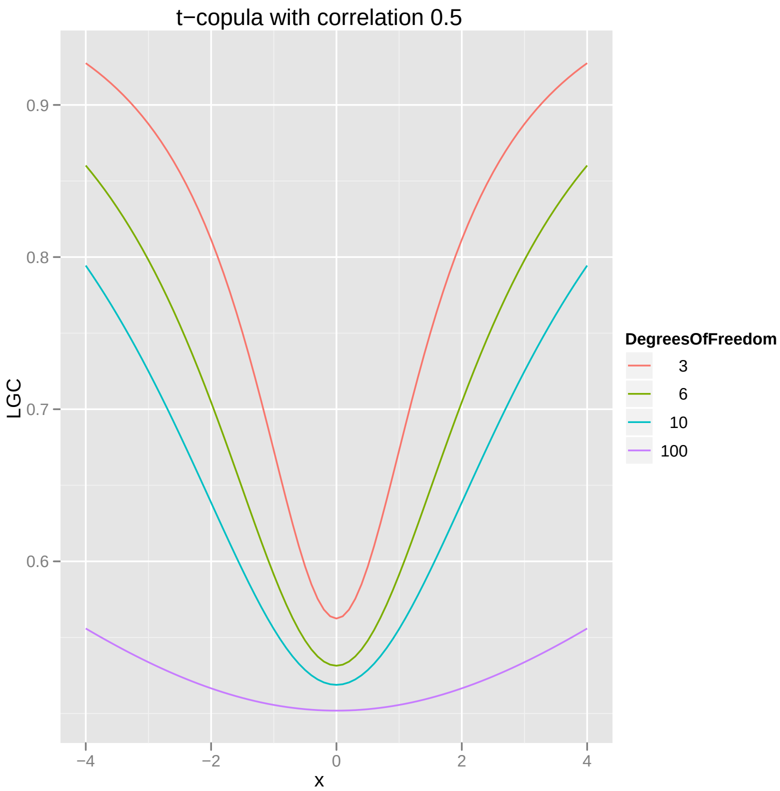

Figure 6.6: Theoretical LGC plot of t-copula with $\rho = 0.5$, $\nu = 3, 6, 10, 100$, and normal margins.

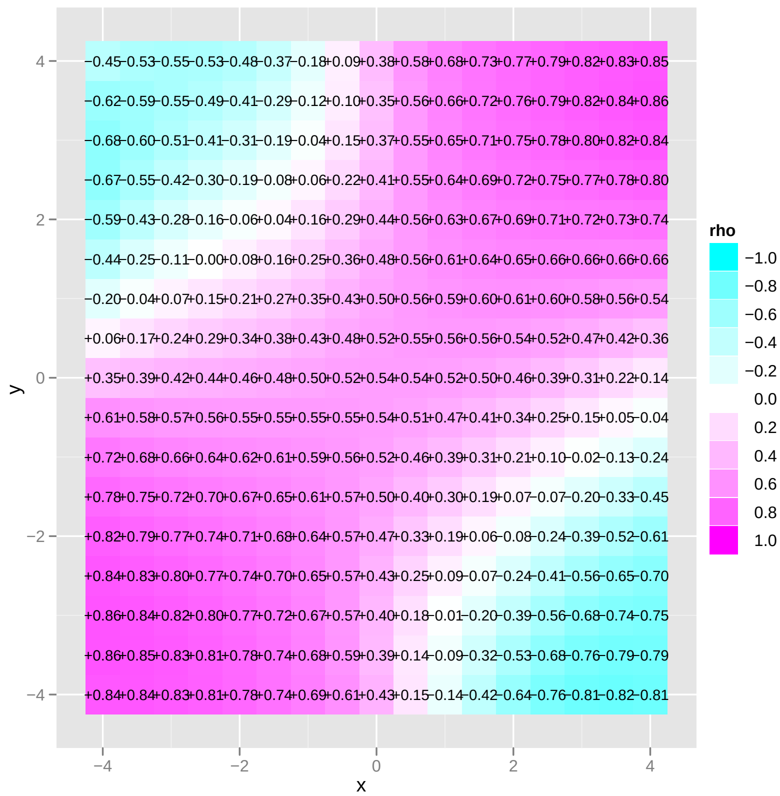

Figure 6.7: Estimated LGC plot of t-copula with $\rho = 0.5$, $\nu = 3$ and normal margins. Based on 5000 generated observations and bandwidth b=1.

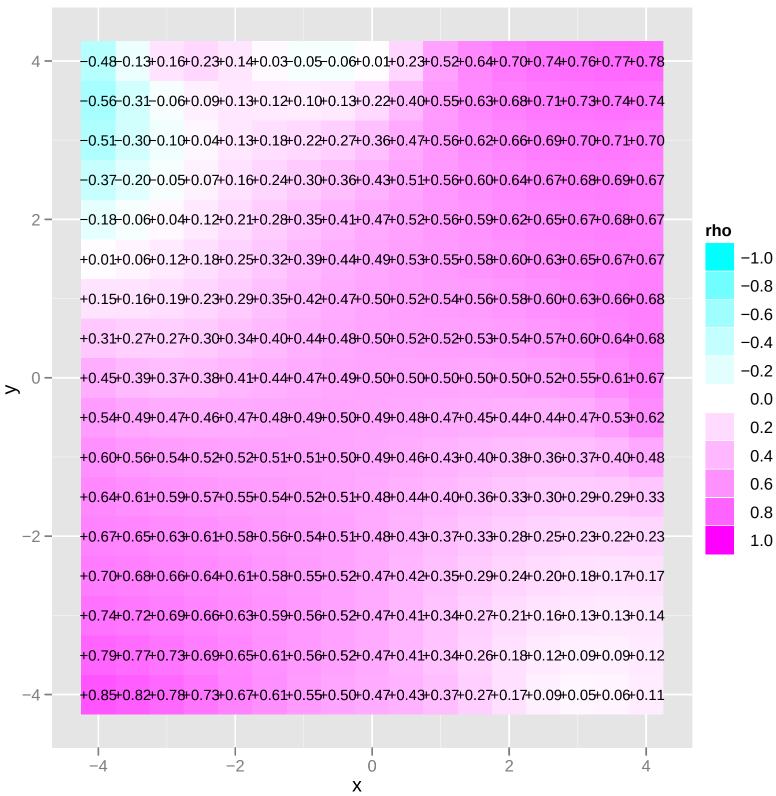

Figure 6.8: Estimated LGC plot of t-copula with $\rho = 0.5$, $\nu = 6$ and normal margins. Based on 5000 generated observations and bandwidth b=1.

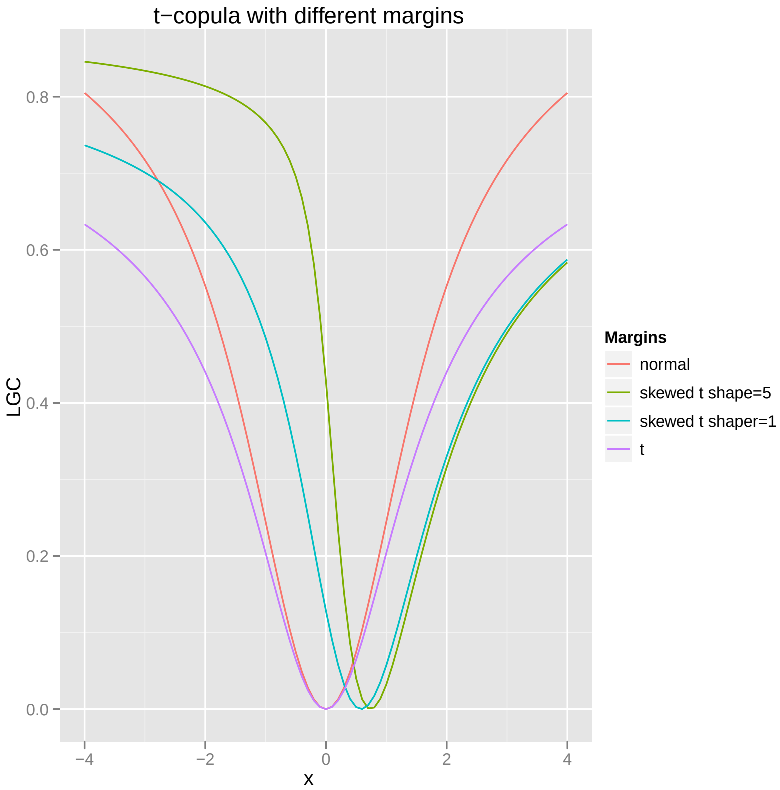

Figure 6.9: Theoretical LGC plot of t-copula with four different choices for margins: Normal margins, t distributed margins with 4 degrees of freedom, skewed t distributed margins with 4 degrees of freedom and shape parameter 1, and skewed t distributed margins with 4 degrees of freedom and shape parameter 5.

Chapter 7

Skew elliptical distributions and copulas

As we have seen we can get skewed LGC plots from putting skewed marginals into the symmetric t copula. Now we are going to look at copulas created from distributions with asymmetric dependence structures. The last couple of decades a lot of effort has been put into the work of finding skewed alternatives to the symmetric distributions. A lot of different alternatives has been proposed. Let us now present one approach which seems like a reasonable way to make skew versions of elliptical distributions. To start with we state a theorem from Azzalini (2005).

Theorem 23. *If f_0 is a d-dimensional pdf such that $f_0(x) = f_0(-x)$ for $x \in \mathbb{R}^d$, G is a one-dimensional differentiable function such that $G^{'}$ is a density symmetric about 0, and w is a real-valued function such that $w(-x) = -w(x)$ for all $x \in \mathbb{R}^d$, then*

$$f(z) = 2f_0(z)G\{w(z)\},\ z \in \mathbb{R}^d$$

is a density function on $\mathbb{R}^d$.

By choosing elliptical distributions for f_0 and G we get what we will call skew elliptical distributions. An interesting example is the skewed t distribution which we get by inserting a t density for f_0 and a t distribution function for G. This is useful because it gives us the possibility to regulate both the skewness and tail thickness. It is the one dimensional version of this distribution we have used as margins in some of the LGC plots.

7.1 Skewed normal distribution

Let us now choose $f_0(x) = \phi_d(x; \Omega)$, that is the density of the d-dimensional normal distribution $N_d(0, \Omega)$. Also choose $G(x) = \Phi$, and let w be a linear function. This gives us the skewed normal distribution, also denoted $SN(\xi, \Omega, \alpha)$, with density

$$f(x) = 2\phi_d(x - \xi; \Omega)\Phi(\alpha^T w^{-1}(x - \xi)) \tag{7.1}$$

where $x \in \mathbb{R}^d$. Here ξ is a location parameter vector, w is a diagonal matrix formed by the standard deviations of Ω, α is a shape parameter vector, and $\overline{\Omega} = w^{-1}\Omega w^{-1}$ is a correlation matrix associated to Ω. Immediately we note that the skewed normal distribution is reduced to the regular normal distribution when $\alpha = 0$. The SN distribution have many nice properties, see for example Azzalini (1985), Azzalini and Valle (1996), Azzalini and Capitanio (1999) and Azzalini (2005) for detailed descriptions. We are only going to consider a couple of properties which are going to be useful when we want to create a SN copula and calculate the LGC. Firstly we note that the skewed normal distribution are closed under marginalization. Now suppose that $X \sim SN_d(\xi, \Omega, \alpha)$, and let us partition it into $X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$, where X_1 has dimension h. Let also $\Omega = \begin{bmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{bmatrix}$ and $\alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}$. Then (Azzalini (2005))

$$X_1 \sim SN_h(\xi, \Omega_{11}, \overline{\alpha_1})$$

where $\overline{\Omega}_{22} - \overline{\Omega}_{21}\overline{\Omega}_{11}^{-1}\overline{\Omega}_{12}$ and $\overline{\alpha}_1 = \frac{\alpha_1 + \overline{\Omega}_{11}^{-1}\overline{\Omega}_{12}\alpha_2}{\sqrt{1+\alpha_2^T\overline{\Omega}_{22.1}\alpha_2}}$. Let us now look at the conditional distribution $X_2|X_1 = x_1$, which density can be written like (Azzalini and Capitanio (1999))

$$\frac{\phi_{d-h}(x_2 - \xi_2^c; \Omega_{22.1})\Phi(\alpha_2^T w_2^{-1}(x_2 - \xi_2^c) + x_0^{'})}{\Phi(x_0)}. \tag{7.2}$$

Here $\xi_2^c = \xi_2 + \Omega_{21}\Omega_{11}^{-1}(x_1 - \xi_1)$, $x_0 = \overline{\alpha}_1^T w_1^{-1}(x_1 - \xi_1)$ and $x_0^{'} = x_0\sqrt{1 + \alpha_2^T\overline{\Omega}_{22.1}\alpha_2}$. We see that it is not on the form (7.1), which means that the SN family is not closed under conditioning like the regular normal distribution. But it is a proper density function, and in Azzalini and Capitanio (1999) it is pointed out that (7.2) in most cases resembles the density of the skew normal distribution. They proposes to approximate the conditional distribution with a SN distribution which matches the cumulants up to the third order. It is also worth mentioning that an extended SN distribution has been proposed, which has density on the form (Azzalini (2005))

$$f(x) = \frac{\phi_d(x - \xi; \Omega)\Phi(\alpha_0 + \alpha^T w{-}1(x - \xi)}{\Phi(\psi)}.$$

Here $\alpha_0 = \psi\sqrt{1 + \alpha^T\overline{\Omega}\alpha}$ and ψ is a new parameter. We see that when $\psi = 0$ the density is reduced to (7.1). Clearly (7.2) is of this form, so the extended SN distribution has the advantage of being closed under conditioning. But as far as we know no R routine is yet implemented for the extended skew normal, only for the regular SN distribution (and the skewed t distribution).

7.2 SN copula

We have not seen copula functions constructed by the skewed normal distribution many places, and there is to our knowledge no standard way of doing so. Like we did with the elliptical distribution copulas we are going to standardize the distribution in a natural way and use Sklars theorem to construct the copula. So we choose $\xi = 0$ and $\Omega = R = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$, which gives us the density

$$q(x_1, x_2) = 2\phi_2(x_1, x_2; R)\Phi(\alpha_1 x_1 + \alpha_2 x_2). \tag{7.3}$$

We will denote the distribution function as Q, with margins $Q_{\overline{\alpha}_1}$ and $Q_{\overline{\alpha}_2}$. This give us the copula function

$$C(u, v) = Q(Q_{\overline{\alpha}_1}^{-1}(u), Q_{\overline{\alpha}_2}^{-1}(v)).$$

We would like the copula to be symmetric when we calculate the theoretical LGC. From the section on elliptical distributions we know that the density $\phi_2(x_1, x_2; R)$ is exchangeable, and looking at (7.3) we see that $q(x_1, x_2) = q(x_2, x_1)$ if $\alpha_1 x_1 + \alpha_2 x_2 = \alpha_2 x_1 + \alpha_1 x_2$. This is obviously true if $\alpha_1 = \alpha_2$. In this case we have for the marginal parameter $\overline{\alpha}_1 = \frac{\alpha(1+\rho)}{\sqrt{1+\alpha_2^2(1-\rho^2}} = \overline{\alpha}_2$, which implies that the marginal distribution in this case has the same distribution. That is $Q_{\overline{\alpha}_1} = Q_{\overline{\alpha}_2} = Q_{\overline{\alpha}}$. Now we have a continuous and symmetric copula model with two parameters α and ρ

$$C(u, v) = Q(Q_{\overline{\alpha}}^{-1}(u), Q_{\overline{\alpha}}^{-1}(v))$$

7.2.1 Theoretical LGC

Like before we are going to to develop the theoretical LGC by conditioning. From (7.2) and by using our simplified parameters we get that the density function of the distribution $X_2|X_1 = x_1$ is

$$\frac{\frac{1}{\sqrt{1-\rho^2}}\phi(\frac{x_2-\rho x_1}{\sqrt{1-\rho^2}})\Phi(\alpha(x_1+x_2))}{\Phi(\overline{\alpha}x_1)}.$$

To ease the calculations let us define

$$K(x_1, x_2) = \frac{1}{\sqrt{1-\rho^2}}\phi(\frac{x_2-\rho x_1}{\sqrt{1-\rho^2}})\Phi(\alpha(x_1+x_2))$$

such that the cdf of the conditional density can be written

$$F_K(x_1, x_2) = \frac{1}{\Phi(\overline{\alpha}x_1)}\int_{-\infty}^{x_2} K(x_1, t)dt.$$

Since no R routines to our knowledge exists for this density, we have chosen to integrate numerically the necessary integrals. Now we have

$$C_1(u, v) = F_K(Q_{\overline{\alpha}}^{-1}(u), Q_{\overline{\alpha}}^{-1}(v)). \tag{7.4}$$

Let us now differentiate $F_K(x_1, x_2)$ with respect to x_1, and denote it f_K, that is

$$f_K(x_1, x_2) = \frac{\partial}{\partial x_1}F_K(x_1, x_2) = \frac{\Phi(\overline{\alpha}x_1)\frac{\partial}{\partial x_1}\int_{-\infty}^{x_2} K(x_1, t)dt - \overline{\alpha}\phi(\overline{\alpha}x_1)\int_{-\infty}^{x_2} K(x_1, t)dt}{\Phi(\overline{\alpha}x_1)^2}.$$

Move the derivation sign inside the integral, and note that

$$\frac{\partial}{\partial x_1}K(x_1, x_2) = \frac{\rho(x_2-\rho x_1)}{1-\rho^2}K(x_1, x_2) + \frac{\alpha}{\sqrt{1-\rho^2}}\phi(\frac{x_2-\rho x_1}{\sqrt{1-\rho^2}})\phi(\alpha(x_1+x_2)).$$

We now have

$$C_{11}(u,v) = \frac{1}{q_{\overline{\alpha}}(Q_{\overline{\alpha}}^{-1}(u))} f_K(Q_{\overline{\alpha}}^{-1}(u), Q_{\overline{\alpha}}^{-1}(v)). \tag{7.5}$$

Now put (7.4) and (7.5) into the theoretical LGC expression and choose margins. For the SN pdfs and cdfs we have used the R package SN.

7.2.2 Plots

The next couple of pages contains different LGC plots based on the skewed normal copula (figures 7.1 to 7.5), with an additional scatterplot in the appendix (figure A.6). Figure 7.1 to 7.3 shows theoretical LGC plots on the diagonal. In some of the plots there have been some problems in the tails for very high and very low α values, probably because of problems with the numerical integration in those cases. We have decided to cut off the problematic parts since they do not give any information regarding the LGC. As we would expect the LGC is constant ρ when $\alpha = 0$ since the SN distribution in that case is reduced to the regular normal distribution. It seems like ρ is an upper limit for the LGC. For positive parameter values we see that the LGC is approximately ρ for positive x values, and with sinking values in the left tail. The difference between the constant LGC of the normal distribution and the LGC for the SN copula gets larger for increasing values of the shape parameter, as we would expect. For negative shape parameters the situation is opposite, with almost constant LGC for negative x values and low LGC in the right tail. So the dependence structure for the SN copula is highly asymmetric, but in practice we would probably be more interested in a dependence structure with stronger dependence in the tail compared to the normal distribution, not weaker dependence as given by the SN copula. Comparing 7.1 and 7.4, and 7.2 and 7.5 we see that the estimated LGC plots resembles the theoretical LGC pretty good.

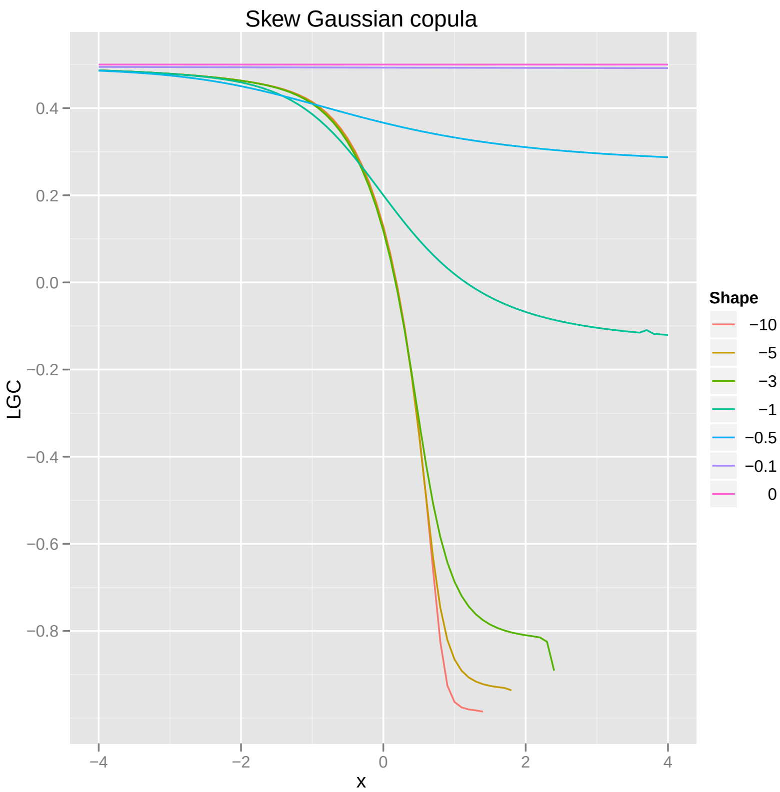

Figure 7.1: Theoretical LGC for SN copula with $\rho = 0.5$ and normal margins.

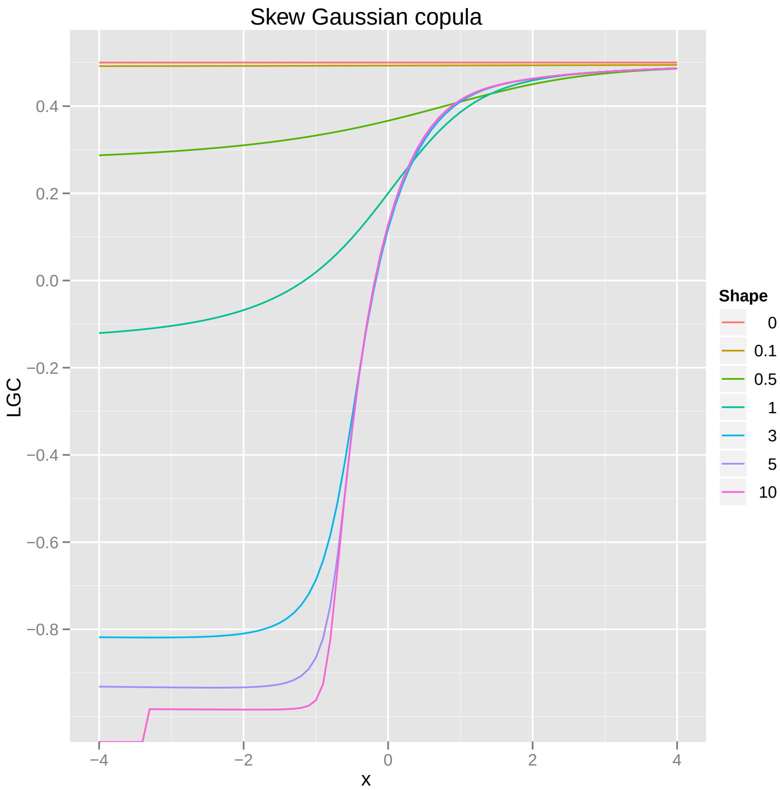

Figure 7.2: Theoretical LGC for SN copula with $\rho = 0.5$ and normal margins.

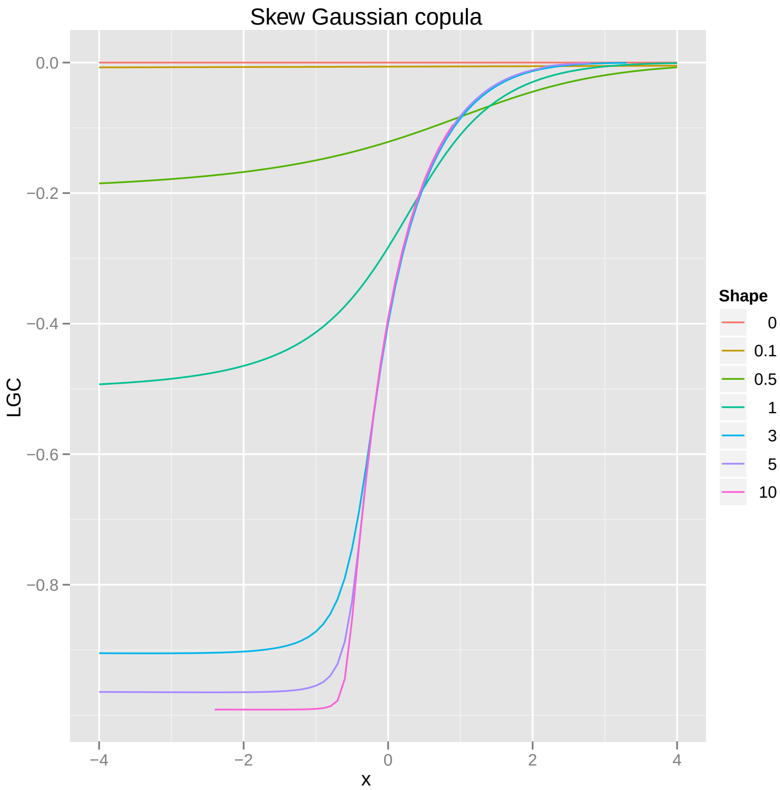

Figure 7.3: Theoretical LGC for SN copula with $\rho = 0$ and normal margins.

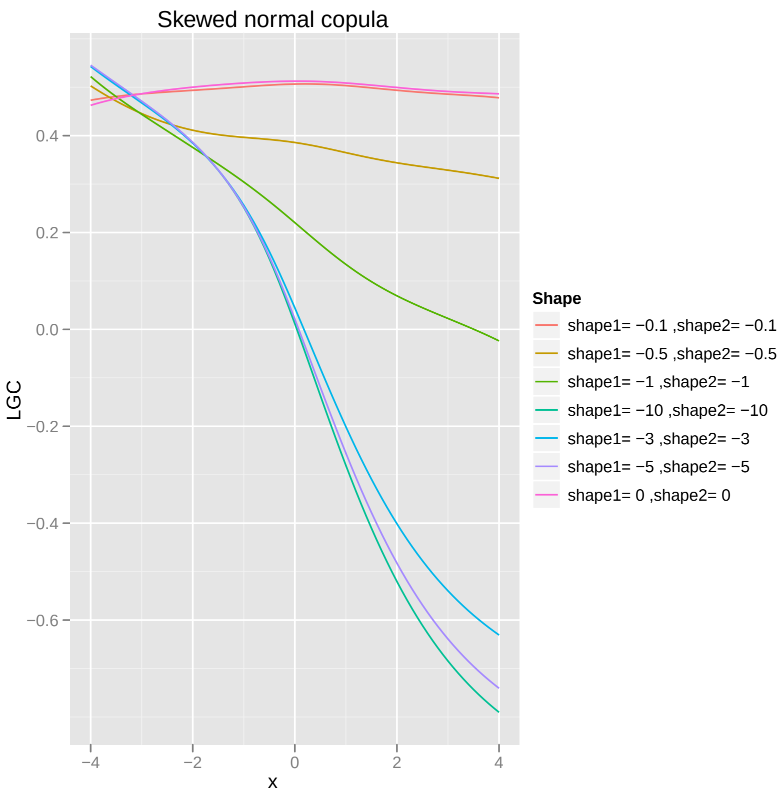

Figure 7.4: Estimated LGC for SN copula with $\rho = 0.5$ and normal margins. Based on 5000 generated observations and bandwidth b=1.5. We could have estimated the LGC for a SN copula with different shape parameters, but for comparison with the theoretical LGC we have chosen shape1=shape2.

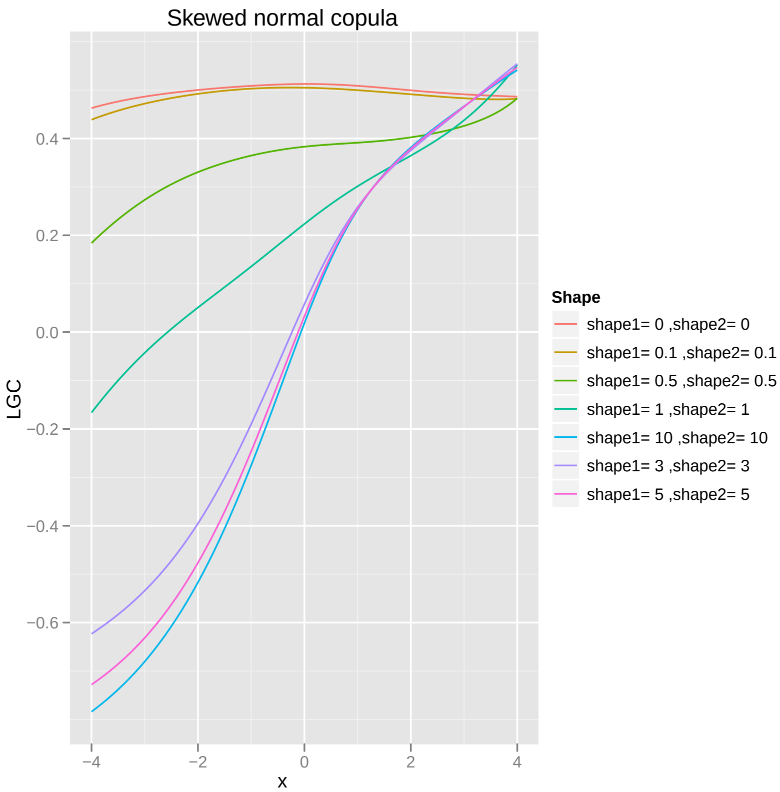

Figure 7.5: Estimated LGC for SN copula with $\rho = 0.5$ and normal margins. Based on 5000 generated observations and bandwidth b=1.5. We could have estimated the LGC for a SN copula with different shape parameters, but for comparison with the theoretical LGC we have chosen shape1=shape2.

Chapter 8

Archimedean copulas

We have seen that copulas constructed from skewed versions of the elliptical distributions offers an opportunity to model asymmetric dependence structures. Now we are going to look at a family offering many different dependence structures with the help of, in many cases, only one parameter. We call them Archimedean Copulas and they are not constructed from well known distributions, but from a special type of functions. These functions need to be imposed some conditions to make sure that we get proper distributions functions, and this will be addressed shortly. Most of these copulas can be written in closed form expression. A lot have been written about Archimedean copulas over the years (see for example Nelsen (2006) for a thorough introduction). We hope that we with the help of LGC plots will be able to shred some new light over the dependence structure of some of the most useful copulas from this family. Because of the easy expression of the Archimedean copulas we can in many cases easily calculate the theoretical value of the locale Gaussian correlation when marginal distribution has been chosen. This will also give us a chance to explore more of the properties of the LGC.

8.1 Definitions and properties

8.1.1 Definition

We are going to follow the same procedure as in Embrechts, Lindskog and McNeil(2001) and Nelsen(2006). Let ϑ be a continuous, strictly decreasing function from $[0,1]$ to $[0,\infty]$ such that $\vartheta(1)=0$. We define the pseudo-inverse of ϑ, $\vartheta^{[-1]}$, to be the function from $[0,\infty]$ to $[0,1]$ given by

$$\vartheta^{[-1]}(t) = \begin{cases} t, & 0 \le t \le \vartheta(0) \\ 0, & \vartheta(0) \le t \le \infty. \end{cases} \quad (8.1)$$

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

$$\vartheta(\vartheta^{[-1]}(t)) = \begin{cases} t, & 0 \le t \le \vartheta(0) \\ \vartheta(0), & \vartheta(0) \le t \le \infty. \end{cases} \quad (8.2)$$

We now state a theorem which will show how we can use ϑ to create copula functions, and we call these copulas Archimedean.

Theorem 24. *Let ϑ be a continuous, strictly decreasing function from $[0,1]$ to $[0,\infty]$ such that $\vartheta(1) = 0$, and let $\vartheta^{[-1]}$ be the pseudo-inverse of ϑ. Then the function C from $[0,1]^2$ to $[0,1]$ given by*

$$C(u,v) = \vartheta^{[-1]}(\vartheta(u) + \vartheta(v)) \tag{8.3}$$

is a copula if and only of ϑ is convex.

Proof can be found in Nelsen(2006).
The function ϑ is called the generator of the copula. We note that in case of $\vartheta(0) = \infty$, $\vartheta^{[-1]}$ will just be the regular inverse function, and in this case we say that $\vartheta(t)$ is a strict generator and the copula is called a strict Archimedean copula.

8.1.2 Properties

Theorem 25. *Let C be an Archimedean copula with generator ϑ. Then*

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

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

Proof. Property 1 follows directly from the definition. The other is reached by straight forward calculations. See Embrechts, Lindskog and McNeil (2001).

That C is symmetric will, as we have discussed earlier, affect the LGC plots. We will expect the plots to show rotation symmetry around u=v. This property will also able us to calculate the theoretical LGC.

8.1.3 Kendall's tau

From the section on dependence measures we remember that Kendall's tau could be expressed with the help from the expected value of the random variable C(U,V). This can be used to derive an easy relationship between the Archimedean copula and Kendall's tau. We state it in a theorem.

Theorem 26. *Let X and Y be random variables with Archimedean copula C and generator ϑ. Kendall's tau is then given by*

$$\tau = 1 + 4\int_0^1 \frac{\vartheta(t)}{\vartheta'(t)}dt \tag{8.4}$$

For proof see Embrechts, Lindskog and McNeil (2001).

8.1.4 Tail dependence

Also the coefficients of tail dependence can be written with the help of the generator function. This theorem is from Nelsen (2006)

Theorem 27. *Let C be an Archimedean copula with generator ϑ, then the upper and lower tail dependence coefficients can be written*

$$\lambda_u = 2 - \lim_{t \to 1^-} \frac{1 - \vartheta^{[-1]}(2\vartheta(t))}{1 - t} = 2 - \lim_{x \to 0^+} \frac{1 - \vartheta^{[-1]}(2x)}{1 - \vartheta^{[-1]}(x)} \tag{8.5}$$

$$\lambda_l = \lim_{t \to 0^+} \frac{\vartheta^{[-1]}(2\vartheta(t))}{t} = \lim_{x \to \infty} \frac{\vartheta^{[-1]}(2x)}{\vartheta^{[-1]}(x)} \tag{8.6}$$

We could also have used the L'Hopital and expressed the tail dependence by the derivative of the generator functions. Actually this theorem can be stated in a slightly more useful way if we restrict ourself to "nice" generator functions. This theorem is from Embrechts, Lindskog and McNeil (2001) and the proof can be found there.

Theorem 28. *Let ϑ be a strict generator such that ϑ^{-1} belongs to the class of Laplace transforms of strictly positive random variables. Let C be copula generated by ϑ. If $\vartheta^{-1\prime}(0)$ is finite then C does not have upper tail dependence. If C has upper tail dependence, then $\vartheta^{-1\prime}(0) = -\infty$ and*

$$\lambda_u = 2 - 2 \lim_{t \to 0^+} \frac{\vartheta^{-1\prime}(2t)}{\vartheta^{-1\prime}(t)}. \tag{8.7}$$

We also have the following expression for the lower tail dependence.

$$\lambda_l = 2 \lim_{t \to \infty} \frac{\vartheta^{-1\prime}(2t)}{\vartheta^{-1\prime}(t)}. \tag{8.8}$$

8.1.5 Theoretical LGC

As we have seen Kendall's tau and the tail dependence coefficient can be stated pretty nicely by using the generator function, so let us now try to state the expression for the LGC with the help off the generator function and its derivatives. We remember we can write the LGC as

$$-\frac{C_{11}(F_X(x), F_Y(y))\phi(\Phi^{-1}(F_X(x)))}{\sqrt{(C_1(F_X(x), F_Y(y)))^2 + (C_{11}(F_X(x), F_Y(y)))^2(\phi(\Phi^{-1}(F_X(x))))^2}}, \tag{8.9}$$

which means that we need to find $C_{11}(u, v)$ and $C_1(u, v)$, where C_1 is also required to be invertible. To avoid getting in trouble we will now look at strict absolute continuous Archimedean copulas, where $\vartheta''(t) > 0$. First we differentiate C given by equation (8.3) one time with respect to u, which gives us

$$C_1(u, v) = \frac{\vartheta'(u)}{\vartheta'(C(u, v))}. \tag{8.10}$$

If we differentiate again, now with respect to v, we get the density

$$c(u, v) = C_{12}(u, v) = -\frac{\vartheta''(C(u, v))\vartheta'(u)\vartheta'(v)}{[\vartheta'(C(u, v))]^3}.$$

We note that $C_1(u, v)$ is invertible with respect to v if $c(u, v) > 0$, which is guaranteed by the condition that $\vartheta''(t) > 0$ ($\vartheta'(t) < 0$ is true because $\vartheta(t)$ is strictly decreasing). Now we can differentiate equation (8.10) with respect to u, and we get

$$C_{11} = \frac{\vartheta''(u)\vartheta'(C(u,v))^2 - \vartheta'(u)^2\vartheta''(C(u,v))}{\vartheta'(C(u,v))^3}. \tag{8.11}$$

Now we can put (8.10) and (8.11) into (8.9) together with margins of our choice. This gives us an expression for the LGC of an Archimedean copula, but it is not a particular easy expression. The good thing is that it is easily implemented on the computer. In R there is available several packages where the most famous Archimedean copulas can be chosen, and where it possible to extract the generator function together with its first and second derivative, and also its inverse. So in this case we can calculate C_1 and C_{11} with the help from such a package, and choose marginal distribution functions from one of the many packages available for that. Too illustrate this method I have made a little function in R working together with the package "fCopulae", where one can choose between the 22 copulas presented in Nelsen (2006), though it will not be possible to calculate the theoretical LGC for all of them. We have picked out four different Archimedean copulas to analyse a bit more. This is Clayton (number 1), Gumbel (number 4), Frank (number 5) and CG (number 12). The first three are the most common Archimedean copulas presented in literature, both because they are easy to analyse and represents three different kinds of dependence structures. Hopefully we will be able to investigate these models a bit more with the help from the LGC plots. The CG copula is also presented in Nelsen (2006) but we have not seen it discussed other places.

8.2 Clayton

The Clayton family are the copulas given by $C(u,v) = (u^{-\theta} + v^{-\theta} - 1)^{-\frac{1}{\theta}}$ for $\theta > 0$. It has generator $\vartheta(t) = \frac{t^{-\theta}-1}{\theta}$, and since $\lim_{t\to 0}\vartheta(t) = \infty$ it is a strict generator with inverse $\vartheta^{-1}(s) = (1+\theta s)^{\frac{-1}{\theta}}$. The derivative of the generator is $\vartheta'(t) = -t^{-(\theta+1)}$, and the second derivative is

$$\vartheta''(t) = (\theta+1)t^{-(\theta+2)} > 0$$

for $t > 0$. If we use the expression for Kendall's tau given by the generator function we get

$$\tau = 1 + 4\int_0^1 \frac{\vartheta(t)}{\vartheta'(t)}dt = 1 + 4\int_0^1 \frac{t^{\theta+1}-t}{\theta} = 1 + \frac{4}{\theta}(\frac{1}{\theta+2} - \frac{1}{2}) = \frac{\theta}{\theta+2}. \tag{8.12}$$

We can see that the coefficient goes towards 0 when θ approaches 0, which would make us think that C approaches the independence copula when θ gets small. On the other end we can see that $\lim_{\theta\to\infty}\tau = 1$, that is perfect dependence, which means that $C \to M$ when θ gets big. Since $\theta > 0$ we see that the Clayton copula only offers positive dependence. For the lower tail dependence we get

$$\lambda_l = \lim_{t\to\infty}\frac{\vartheta^{-1}(2t)}{\vartheta^{-1}(t)} = \lim_{t\to\infty}(\frac{1+2\theta t}{1+\theta t})^{\frac{-1}{\theta}} = \lim_{t\to\infty}(\frac{\frac{1}{t}+2\theta}{\frac{1}{t}+\theta})^{\frac{-1}{\theta}} = 2^{\frac{-1}{\theta}}. \tag{8.13}$$

Figure (8.1) shows a plot of the lower tail dependence for different parameter values, and it shows that the lower tail dependence approaches zero as θ approaches zero, while it goes to one as θ gets larger. We have that $\vartheta^{-1\prime}(0) = -1$, and since it is finite this implies that the Clayton copula is asymptotic independent in the upper tail. The alternative tail dependence coefficient in (3.8) can now be analysed for potentially more information. This turned out to be difficult analytically, so instead we can write

$$\overline{\lambda}(u) = \frac{2\log(1-\vartheta^{-1}(u))}{\log(1-2\vartheta^{-1}(u)+\vartheta^{-1}(2u))}, \tag{8.14}$$

and make a plot for decreasing values of u ($u \to 0$) to get an approximate value. After done this for several different θ values it seems that the alternative upper tail coefficient also is 0, which indicates independence in the upper tail for the Clayton copula. Figure (8.2) show a plot for $\theta = 1$.

Some places the Clayton family is also defined for parameter values in the interval $[-1, 0)$. Then we need to write the copula as $C(u,v) = [max(u^{-\theta} + v^{-1\theta} - 1, 0)]^{\frac{-1}{\theta}}$. For $\theta < 0$ this copula is no longer strict, and for $\theta = -1$ it becomes W (Nelsen 2006). Allowing for negative parameters will open for modelling negative dependence with the Clayton family.

8.2.1 Plots

The next couple of pages contains different LGC plots based on the Clayton copula (figures 8.1 to 8.7), and some additional plots concerning the Clayton copula can be found in the appendix (figures A.7 to A.13). Figure 8.3 show plots of the theoretical LGC for different parameter values. As expected we see strong lower tail dependence, and independence in the upper tail. With increasing values of Kendall's tau we see that the LGC value also increases. Figures 8.4 to 8.6 shows estimated LGC plots for Clayton copulas with normal margins for three different Kendall's tau values, while figures A.7 to A.9 shows scatterplots for the same observations. The scatter plots clearly gets significantly narrower in the left tail which indicates strong dependence, while in the right tail the points look more independent. We also see that the scatter plots in general gets more narrow with increasing values of τ, but in the far right tail it could look like there is independence for every value of τ. This is as expected from the theoretical LGC, and we also see that the estimated LGC plots shows the same patterns. Figure 8.5 and 8.6 indicates some dependence in the first quadrant, which does not show on the theoretical LGC plot. It is not surprising that the LGC is a bit overestimated in the first quadrant since in the estimation process we uses points from closer to the origin where the dependence is higher. With a smaller bandwidth the plot would probably have shown less dependence in the first quadrant, though it might be difficult to distinguish this effect from the noise we get from a smaller bandwidth. As we see from the scatter plots there is not that many points at the boundaries, which clearly makes the estimates here more uncertain. The estimated LGC plots also shows us that the symmetric property of the Archimedean copulas gives us plots which is rotational symmetric around the diagonal. Figure 8.7 shows theoretical LGC for Clayton copula with different margins. As we can see the choice of margins affects the dependence structure very little. The skewed t margins gives a little more dependence, while the t margins has a little less dependence in the lower tail and a little more in the upper tail than the normal margins. This we can also see from the estimated

LGC plots in figure A.10 to A.12. The plots with the skewed t margins gives a little more dependence, but in general it looks very similar. They are also still symmetric around the diagonal. One curious thing is that all these three estimated plots gives negative LGC values at the far upper tail. This is not in accordance with the theoretical plot, which shows slightly more dependence in the right tail for the alternative margins. This is probably just a boundary effect caused by very few observations in this area. Figure A.13 shows an estimated LGC plot of the Clayton copula with uniform margins. From experience with the t copula we would expect this plot to look like figure 8.4, and it looks similar except from the first quadrant. So it seems like the estimation procedure has problems capturing the independence in the right tail. There might be to few observations in this area, or perhaps it is possible to find a more suitable bandwidth to get it more right.

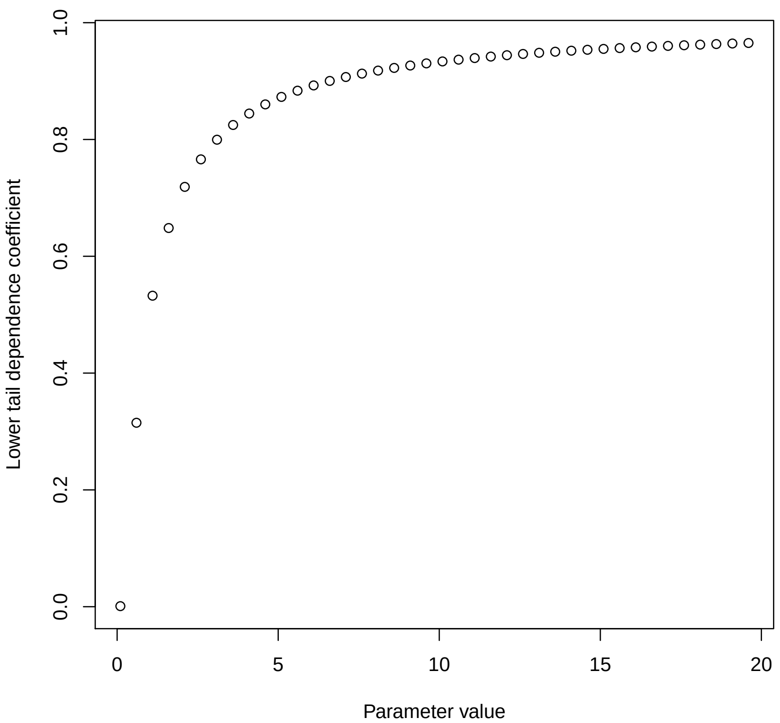

Figure 8.1: The lower tail dependence coefficient for different parameter values for Clayton copula.

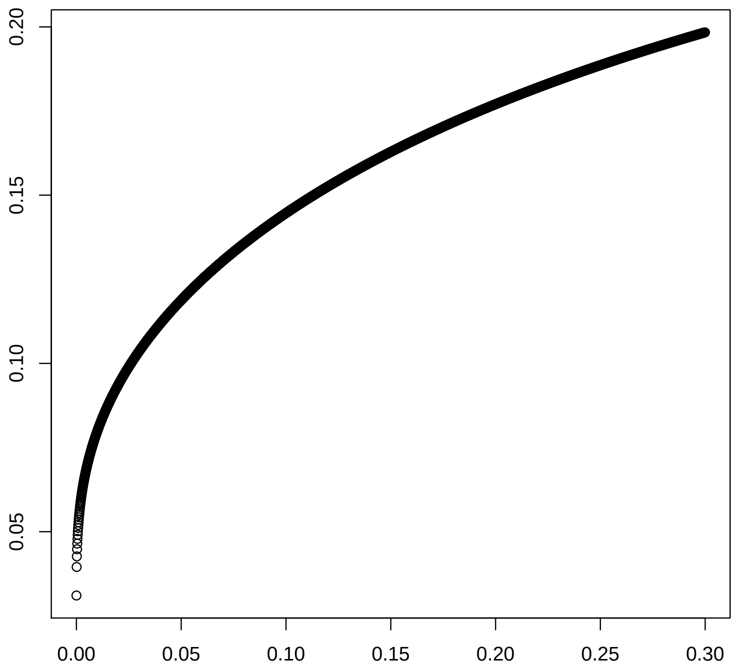

Figure 8.2: Plot of equation (8.14) for decreasing u values.

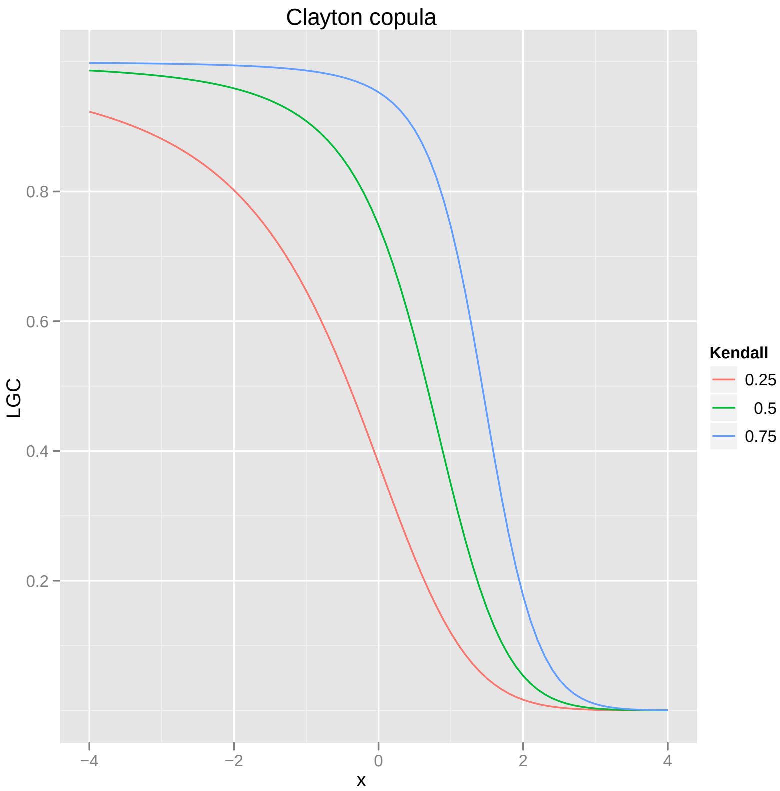

Figure 8.3: Theoretical LGC of Clayton copula with $\tau = \frac{1}{4}, \frac{1}{2}, \frac{3}{4}$, and normal margins.

Figure 8.4: Estimated LGC plot of a Clayton copula with $\tau = \frac{1}{4}$ and normal margins. Based on 5000 generated observations and bandwidth b=1.2.

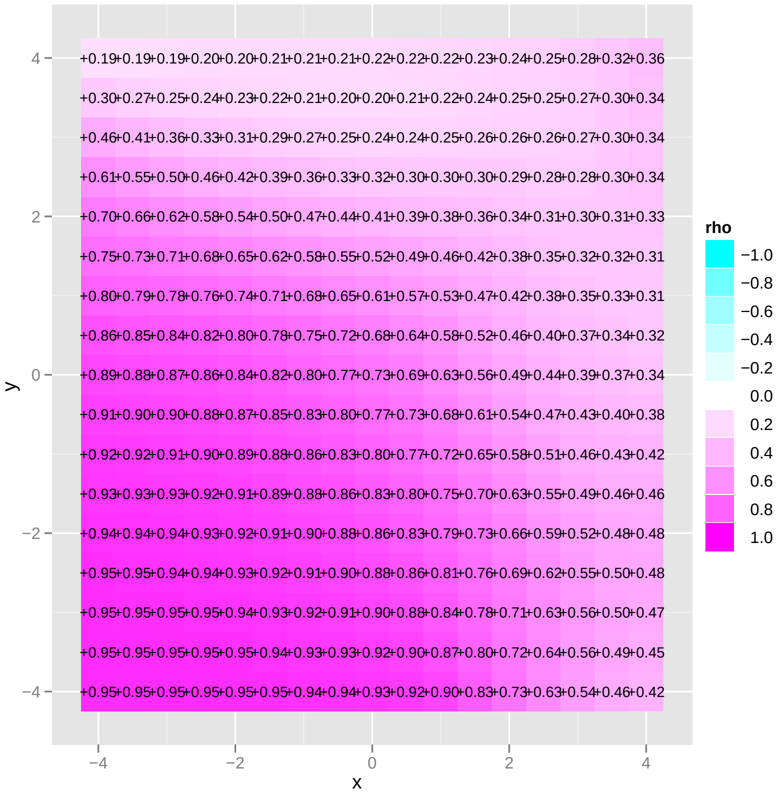

Figure 8.5: Estimated LGC plot of a Clayton copula with $\tau = \frac{1}{2}$ and normal margins. Based on 5000 generated observations and bandwidth b=1.

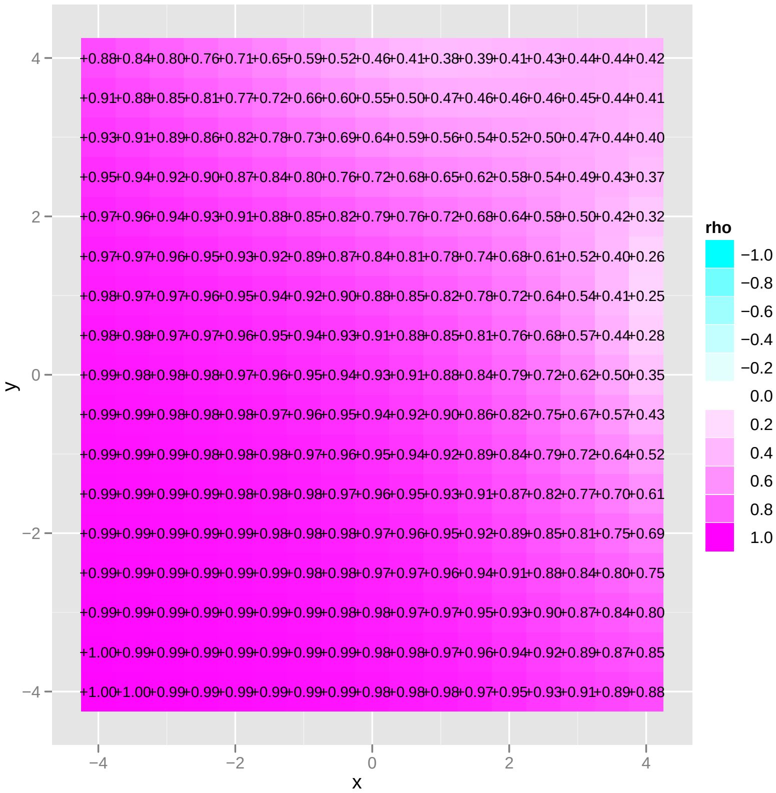

Figure 8.6: Estimated LGC plot of a Clayton copula with $\tau = \frac{3}{4}$ and normal margins. Based on 5000 generated observations and bandwidth b=1.

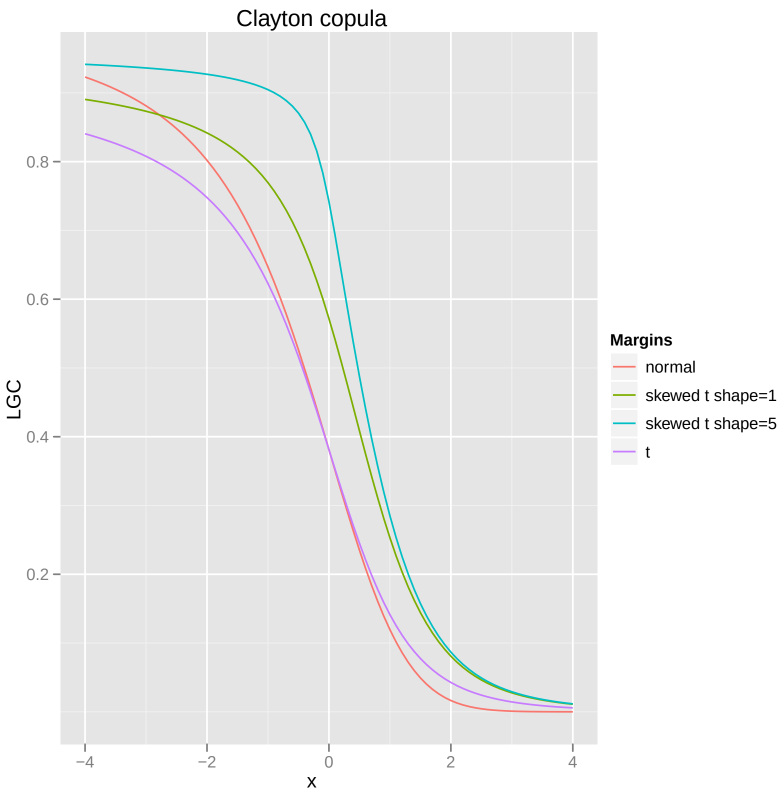

Figure 8.7: Theoretical LGC of Clayton copula with $\tau = \frac{1}{4}$ and with four different choices for margins: Normal margins, t distributed margins with 4 degrees of freedom, skewed t distributed margins with 4 degrees of freedom and shape parameter 1, and skewed t distributed margins with 4 degrees of freedom and shape parameter 5.

8.3 Gumbel copula

The Gumbel copula is given by

$$C(u,v) = \exp(-((-\log u)^\theta + (-\log v)^\theta)^{\frac{1}{\theta}}), \tag{8.15}$$

for $\theta \in [1,\infty)$. The generator function is $\vartheta(t) = (-\log t)^\theta$, with inverse $\vartheta^{-1}(t) = \exp(-t^{\frac{1}{\theta}})$. $\lim_{t\to 0}\vartheta(t) = \infty$ which shows that ϑ is strict. We have $\vartheta'(t) = -\frac{\theta}{t}(-\log t)^{\theta-1}$ and $\vartheta''(t) = \frac{\theta}{t^2}(-\log t)^{\theta-1} + \frac{\theta}{t^2}(\theta-1)(-\log t)^{\theta-2}$. Since $\log t < 0$ for $t \in (0,1)$, we have that $\vartheta''(t) > 0$ for $t \in (0,1)$. If we put $\theta = 1$ we get $C(u,v) = \exp(-((-\log u) + (-\log v))) = \exp(\log u)\exp(\log v) = uv$, which is the independence copula. Using (8.4) we can calculate Kendall's tau, that is

$$\begin{aligned}
\tau &= 1 + 4\int_0^1 \frac{\vartheta(t)}{\vartheta'(t)}dt = 1 - 4\int_0^1 \frac{t(-\log t)^\theta}{(-\log t)^{\theta-1}}dt \\
&= 1 + 4\int_0^1 \frac{t\log t}{\theta}dt = 1 + \frac{4}{\theta}[\frac{1}{2}t^2\log t|_0^1 - \frac{1}{2}\int_0^1 tdt] \\
&= 1 + \frac{4}{\theta}(-\frac{1}{4}) = 1 - \frac{1}{\theta},
\end{aligned} \tag{8.16}$$

where we used partial integration. We notice that $\lim_{\theta\to\infty}\tau = 1 - 0 = 1$, which implies that $C \to M$ when θ gets big. In addition we can see that $\tau \geq 0$ for all parameter values. If we use (8.7) we can find the upper tail dependence coefficient.

$$\begin{aligned}
\lambda_u &= 2 - 2\lim_{t\to 0}\frac{\vartheta^{-1\prime}(2t)}{\vartheta^{-1\prime}(t)} = 2 - 2\lim_{t\to 0} 2^{\frac{1}{\theta}-1}\exp(-(2t)^{\frac{1}{\theta}} + t^{\frac{1}{\theta}}) \\
&= 2 - 2^{\frac{1}{\theta}}\lim_{t\to 0}\exp(t^{\frac{1}{\theta}}(1 - 2^{\frac{1}{\theta}}) = 2 - 2^{\frac{1}{\theta}}.
\end{aligned} \tag{8.17}$$

This shows us that the Gumbel copula has positive upper tail dependence, which increases with increasing θ values. By using (8.8) we get that

$$\lambda_l = 2\lim_{t\to\infty} 2^{\frac{1}{\theta}-1}\exp(t^{\frac{1}{\theta}}(1 - 2^{\frac{1}{\theta}}) = 0, \tag{8.18}$$

which tells us that the Gumbel copula is asymptotic independent in the lower tail. So the dependence structure is pretty opposite to the Clayton copula. As usual we would like to make a more thorough investigation of the tail, to see if there might be some dependence despite the fact that $\lambda_l = 0$. We use that the coefficient of lower tail dependence for C is the coefficient of upper tail dependence for $\widehat{C}$, which enable us to use the alternative tail coefficient in (3.8). This gives us

$$\begin{aligned}
\overline{\lambda} &= \lim_{t\to 1}\frac{2\log(1-t)}{\log\overline{\widehat{C}}(t,t)} - 1 = \lim_{t\to 1}\frac{2\log(1-t)}{\log C(1-t,1-t)} - 1 \\
&= 2\lim_{u\to 0}\frac{\log u}{\log C(u,u)} - 1 = -2\lim_{u\to 0}\frac{\log u}{((-\log u)^\theta + (-\log u)^\theta)^{\frac{1}{\theta}}} - 1 \\
&= -2\lim_{u\to 0}\frac{\log u}{(2(-\log u)^\theta)^{\frac{1}{\theta}}} - 1 = 2^{1-\frac{1}{\theta}}\lim_{u\to 0}\frac{\log u}{\log u} - 1 \\
&= 2^{1-\frac{1}{\theta}} - 1.
\end{aligned} \tag{8.19}$$

When $\theta = 1$ we get $\overline{\lambda} = 0$ which make sense since this is the independence copula, but as θ increases so does the dependence in the lower tail.

8.3.1 Plots

The next couple of pages contains different LGC plots based on the Gumbel copula (figures 8.8 to 8.12), and some additional plots concerning the Gumbel copula can be found in the appendix (figures A.14 to A.20). We have used the same type of plots and with the same Kendall's tau value as for the Clayton copula to better compare them. As expected 8.8 shows that the Gumbel copula has strong right tail dependence, but also some dependence in the left tail. The dependence in both tails are, as shown above, increasing with increasing parameter values. For $\tau = \frac{3}{4}$ there is not that much difference left between the tails. If we compare with figure 8.3 of the Clayton copula, we see that they are pretty much opposite, though the dependence structure for the Gumbel copula is less asymmetric between the tails. For the estimated plots in figures 8.9, 8.10 and 8.11 we see the same patterns, that is increasing from left to right, and increasing with τ. We see that most of these plots are clearly overestimated in the left tail and underestimated in the right tail compared to the theoretical LGC. This is, as explained before, an anticipated effect of the estimation since we use points closer to the middle of the plot when estimating the tails. In this case, when the LGC value is increasing from left to right we get to high levels in the left tail and to small in the right tail. It also important to notice from the scatterplots how few observations there are at the boundaries, which makes the information from these areas on the LGC plot not very exact. Figures 8.12 and A.17 to A.19 show us that changing the margins does not change the dependence structure that much. It seems like the other choices of margins gives a slightly smaller LGC value, especially in the right tail. The estimated plots have a bit of noise at the boundaries, but we can still see that there is increasing dependence from left to right, and symmetry around the diagonal. Figure A.20 shows the estimated LGC for the copula model with uniform margins, and by comparison with 8.9 we see that they are really similar. Only in the third quadrant is the LGC estimate a bit higher with the uniform margins.

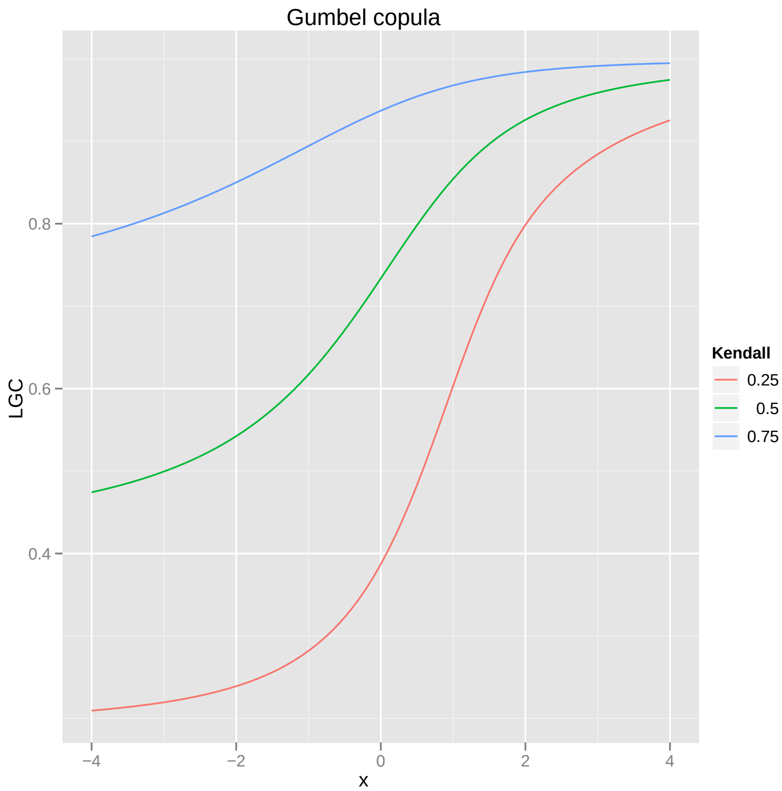

Figure 8.8: Theoretical LGC of Gumbel copula with $\tau = \frac{1}{4}, \frac{1}{2}, \frac{3}{4}$, and normal margins.

Figure 8.9: Estimated LGC plot of a Gumbel copula with $\tau = \frac{1}{4}$ and normal margins. Based on 5000 generated observations and bandwidth b=1.2.

Figure 8.10: Estimated LGC plot of a Gumbel copula with $\tau = \frac{1}{2}$ and normal margins. Based on 5000 generated observations and bandwidth b=1.

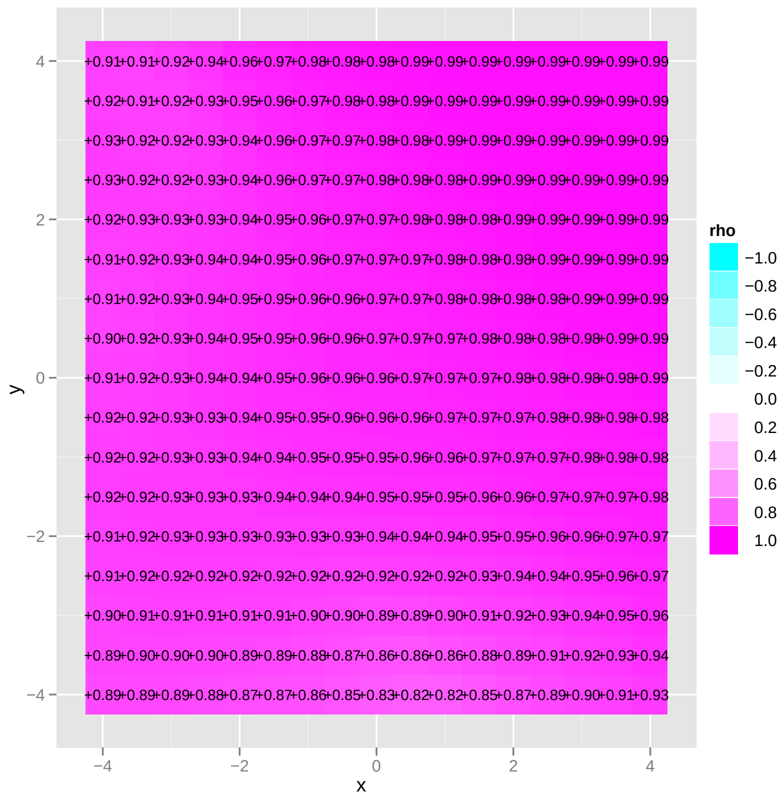

Figure 8.11: Estimated LGC plot of a Gumbel copula with $\tau = \frac{3}{4}$ and normal margins. Based on 5000 generated observations and bandwidth b=0.8.

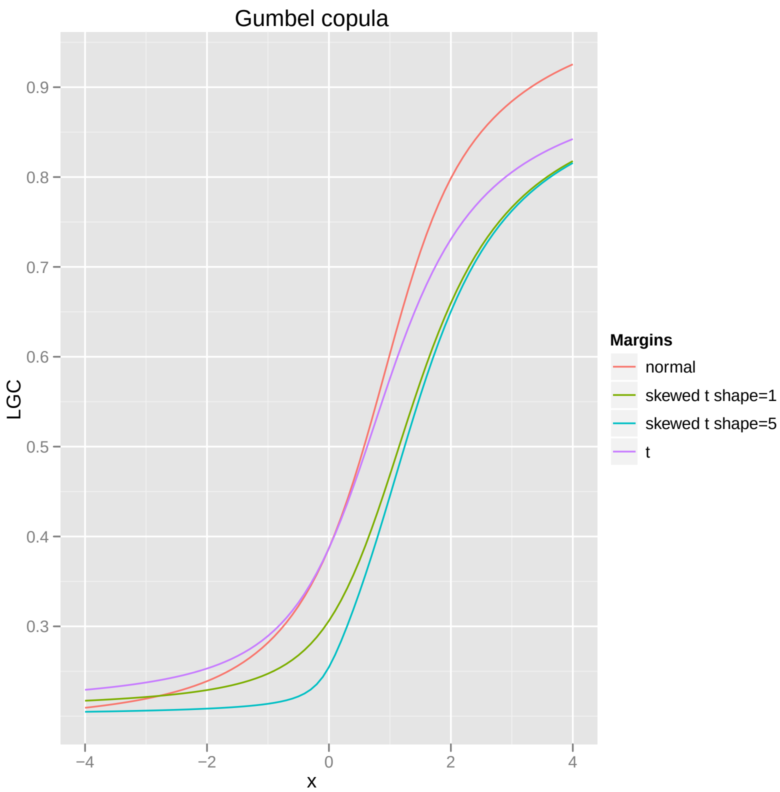

Figure 8.12: Theoretical LGC of Gumbel copula with $\tau = \frac{1}{4}$ and with four different choices for margins: Normal margins, t distributed margins with 4 degrees of freedom, skewed t distributed margins with 4 degrees of freedom and shape parameter 1, and skewed t distributed margins with 4 degrees of freedom and shape parameter 5.

8.4 Frank copula

The Frank copula is given by

$$C(u,v) = -\frac{1}{\theta}\log(1+\frac{(e^{-\theta u}-1)(e^{-\theta v}-1)}{e^{-\theta}-1}), \tag{8.20}$$

where the generator function is $\vartheta(t) = -\log(\frac{e^{-\theta t}-1}{e^{-\theta}-1})$ and its inverse is $\vartheta^{-1}(t) = -\frac{1}{\theta}\log(e^{-t}(e^{-\theta}-1)+1)$. The parameter θ must be in the set $\mathbf{R}\setminus\{0\}$. It is easily seen that $\lim_{t\to 0}\vartheta(t) = \infty$, which shows that it is a strict copula. The derivatives of the generator function are $\vartheta'(t) = \frac{\theta}{1-e^{\theta t}}$ and $\vartheta''(t) = \frac{\theta^2 e^{\theta t}}{(1-e^{\theta t})^2} > 0$. It can be shown that the Frank copula is the only Archimedean copula with radial symmetry (Frank 1979). In some ways the Frank copula is more difficult to analyse analytical than the other two families we have described. Its relationship with Kendall's tau can be expressed with a integral, which is as follows (Genest 1987)

$$\tau = 1 - \frac{4}{\theta}(1 - \frac{1}{\theta}\int_0^\theta \frac{t}{e^t-1}dt). \tag{8.21}$$

When $\theta \to 0$ C approaches the independence copula, while it approaches W when $\theta \to -\infty$ and M when $\theta \to \infty$. So the Frank copula gives us possibility of modelling both negative and positive dependence. If we differentiate the inverse generator function we get $\vartheta^{-1\prime}(t) = -\frac{e^{-t}(1-e^{-\theta})}{\theta(1-e^{-t}(1-e^{-\theta}))}$. This gives us $\vartheta^{-1\prime}(0) = -\frac{e^{-\theta}-1}{\theta}$, which is finite and implies that the Frank copula is asymptotic independent in the upper tail, and by radial symmetry also in the lower tail.

8.4.1 Plots

The next couple of pages contains different LGC plots based on the Frank copula (figures 8.13 to 8.17), and some additional plots concerning the Frank copula can be found in the appendix (figures A.21 to A.27). We continue with the same plots, but since the Frank copula also allow negative dependence we include some example of that also. On figure 8.13 we clearly see that the radial symmetry of the Frank copula is inherited by the LGC. As expected there is no dependence in the tail for any parameter value. We can notice that for positive τ values the LGC graphs is more round. The scatter plots gives us an indication of the general dependence, if its positive or negative, and also how strong the dependence is. A.22 clearly shows more positive dependence than A.21. Maybe we also can sense some stronger dependence in the middle of the plots than in the tails, which coincides with the LGC plots. Clearly there is few points in the tails, and we see from all the estimated LGC plots for the Frank copula that a lot of funny things happens. These estimates are only reliable closer to the middle of the plot. Here we see that they coincide with the theoretical LGC. We have increasing LGC towards the middle, and symmetry around the origin. As anticipated the estimates around the origin is a little lower than in the theoretical plot, this is because the LGC reaches its maximum around the origin. For the plots with negative τ values it is opposite. These plots also shows that the LGC estimation has special problems dealing with independence in low density areas. Figure 8.17 shows the Frank copula with different margins. We see that the use of t margins gives some more dependence in the tails, though the symmetric property around zero is still there. As for the t copula, the use of skewed t margins also creates asymmetries in the LGC plot for the Frank copula. The maximum point is shifted to the right. Also there is a larger interval with independence in the left tail, and more dependence in the right tail. As we might expect the LGC plot is more skewed when we use the most skewed margins (that is shape=5). It is hard to see the effect of the t margins on the estimated LGC plot in figure A.24 when you compare with figure 8.14, because the theoretical difference is pretty small and also in the area where there are few points and therefore a lot of noise in the plots. In figure A.25 you can perhaps see a shifting of the dependence to the right. While figure A.26 is quite opposite of what we expected, and does not really make any sense at all.

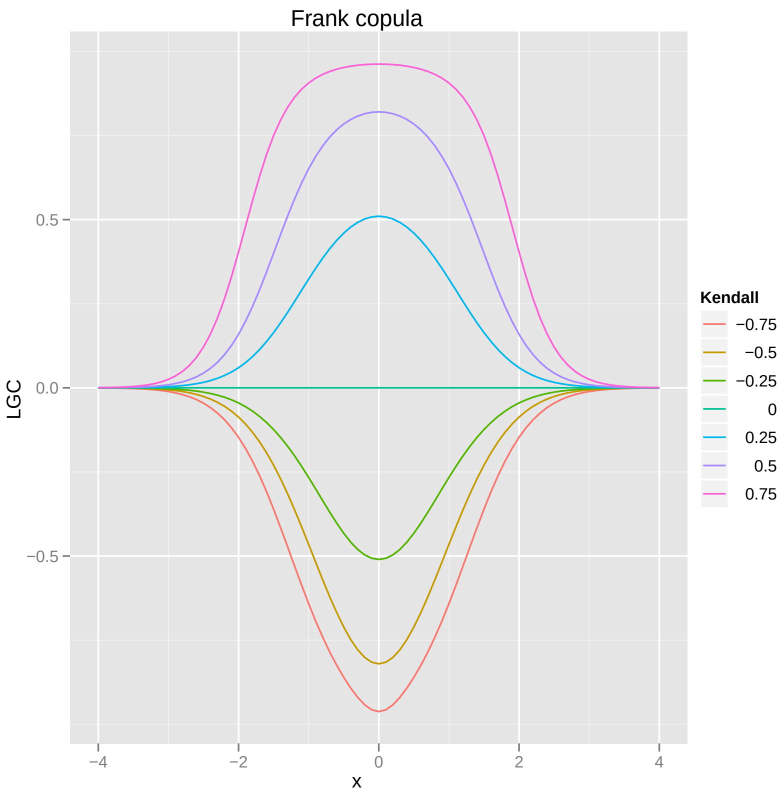

Figure 8.13: Theoretical LGC of Frank copula with $\tau = \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 0, -\frac{1}{4}, -\frac{1}{2}, -\frac{3}{4}$, and normal margins.

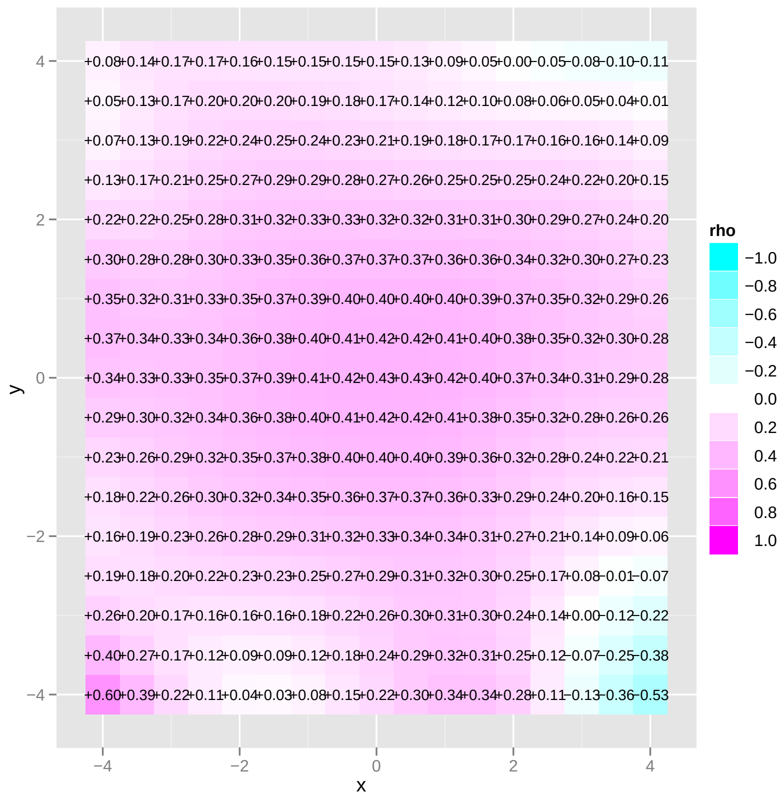

Figure 8.14: Estimated LGC plot of a Frank copula with $\tau = \frac{1}{4}$ and normal margins. Based on 5000 generated observations and bandwidth b=1.1.

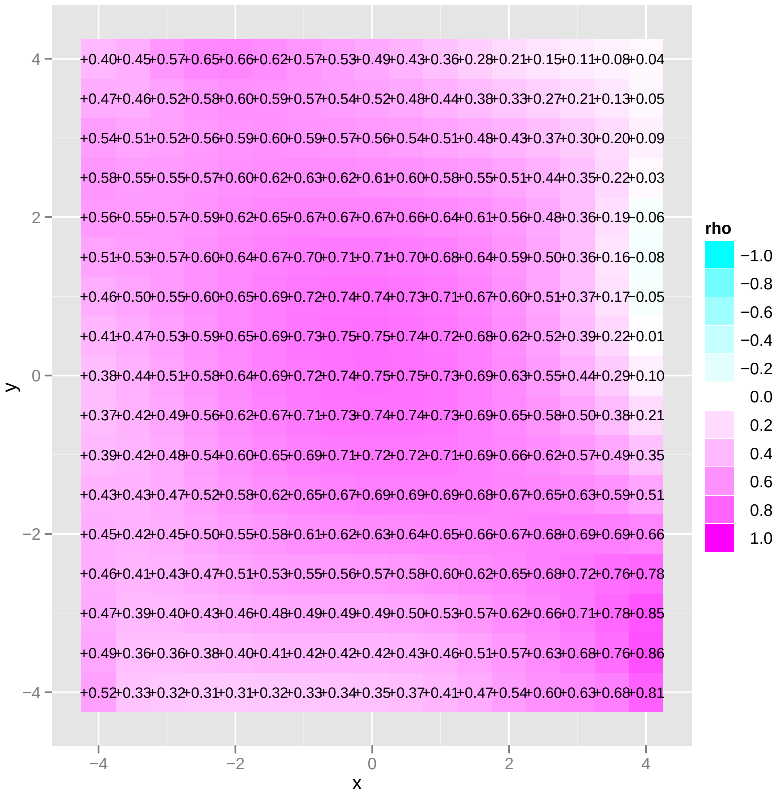

Figure 8.15: Estimated LGC plot of a Frank copula with $\tau = \frac{1}{2}$ and normal margins. Based on 5000 generated observations and bandwidth b=1.

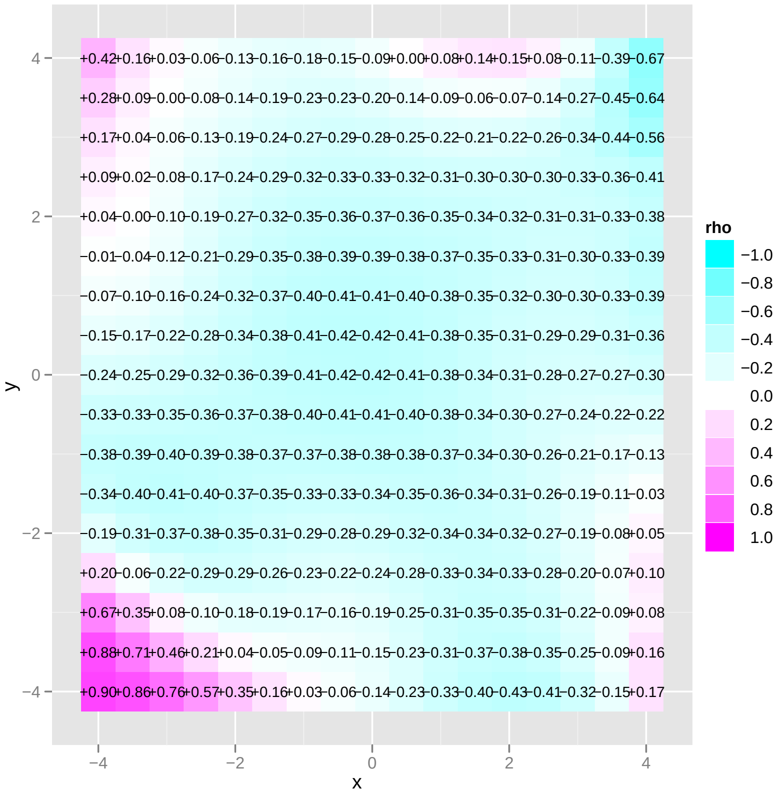

Figure 8.16: Estimated LGC plot of a Frank copula with $\tau = -\frac{1}{4}$ and normal margins. Based on 5000 generated observations and bandwidth b=1.

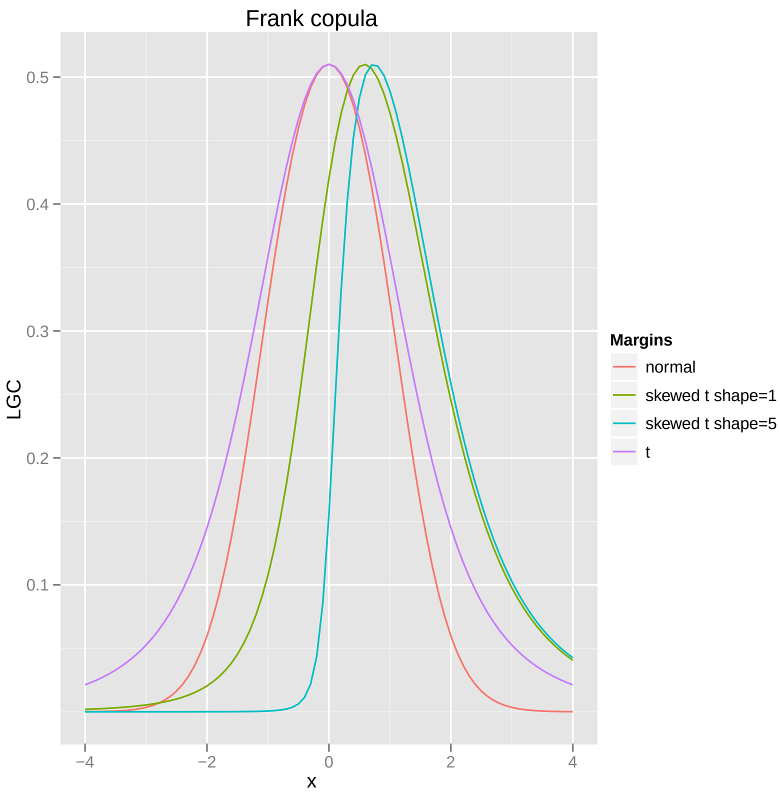

Figure 8.17: Theoretical LGC of Frank copula with $\tau = \frac{1}{4}$ and with four different choices for margins: Normal margins, t distributed margins with 4 degrees of freedom, skewed t distributed margins with 4 degrees of freedom and shape parameter 1, and skewed t distributed margins with 4 degrees of freedom and shape parameter 5.

8.5 CG-copula

We denote the CG-copula as

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

for $\theta \geq 1$. The generator function is $\varphi(t) = (\frac{1}{t} - 1)^{\theta}$. This copula does not have any particular name in Nelsen (2006), but we will denote it as the CG copula because it has the same lower tail dependence coefficient as the Clayton copula and the upper tail dependence coefficient as the Gumbel copula. That is $\lambda_l = 2^{-\frac{1}{\theta}}$ and $\lambda_u = 2 - 2^{\frac{1}{\theta}}$. Kendalls tau for the CG copula is given by $\tau = 1 - \frac{2}{3\theta}$. We see that for the limiting parameter value $\theta = 1$ we have $\tau = \frac{1}{3}$, and when $\theta \to \infty$ we get $\tau \to 1$. This means that the CG copula only can model dependence for Kendall's tau values in the interval $[\frac{1}{3}, 1)$. For $\theta = 1$, the CG copula looks like $C(u,v) = \frac{1}{\frac{1}{u} + \frac{1}{v} - 1}$, which is the same as the Clayton copula for $\theta = 1$. This also fits together with the fact that the upper tail dependence becomes zero for this parameter value.

8.5.1 Plots

The next couple of pages contains different LGC plots based on the CG copula (figures 8.18 to 8.22), and some additional plots concerning the CG copula can be found in the appendix (figures A.28 to A.34). We see from figure 8.18 that the left tail dependence is strong and approaches one quickly for every parameter value, while the right tail has a slower convergence towards one. The tail dependence decreases with decreasing parameter values, but it is in the right tail the differences are biggest. For the limiting case $\tau = \frac{1}{3}$ we get as expected that the right tail dependence is zero The minimum LGC value also decreases and is shifted to the right for lower values of τ. It is also possible to see these patterns on the scatterplots. Clearly there is strong dependence in the left tail for all parameter values, while in the right tail the dependence increases significantly with increasing values of τ. One can also sense that there is a bit more spreading in the points around the middle of the plot, which coincides with the LGC. The estimated LGC plots also shows the same patterns. Namely a general increase in LGC values for increasing τ, where the left tail has a small increase and the right tail a significantly increase. The estimates in figure 8.20 and 8.21 also shows that the dependence decreases in the middle of the plot (or a little bit to the right to be more precise). Surprisingly this effect is really small (almost not visible) in figure 8.19. It looks like the right tail is really underestimated here. The use of the other margins shifts the minimum point a bit to the right, and gives lower values in the right tail. Especially with the skewed t margins with shape parameter 5 we get a bigger difference between left and right tail. The estimated LGC plot with uniform margins in figure A.34 seems to capture the dependence structure nicely, though the right tail is a bit lower than the theoretical LGC with normal margins.

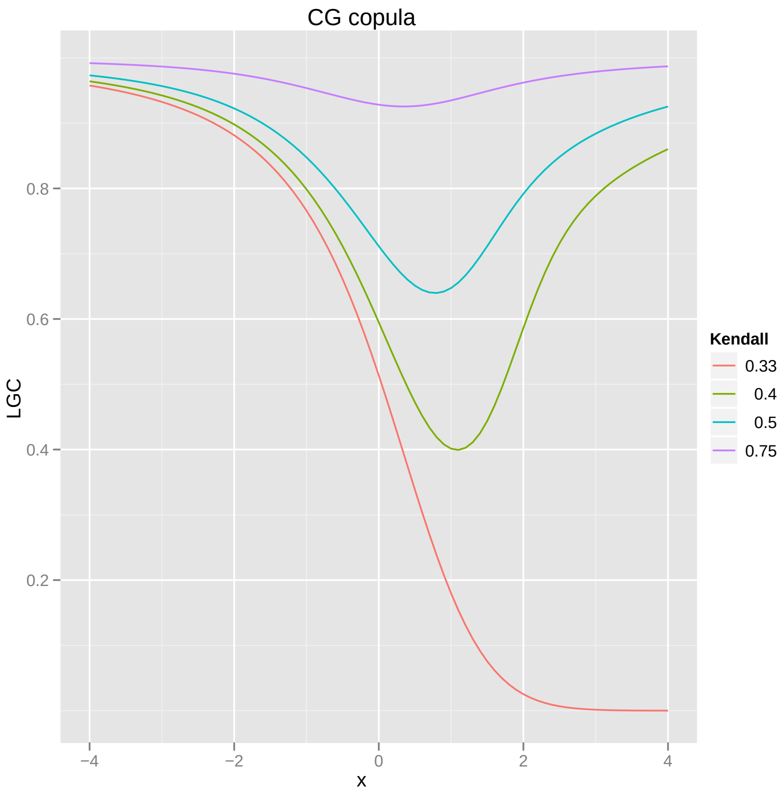

Figure 8.18: Theoretical LGC of CG copula with $\tau = \frac{1}{3}, \frac{2}{5}\frac{1}{2}, \frac{3}{4}$, and normal margins.

Figure 8.19: Estimated LGC plot of a CG copula with $\tau = \frac{2}{5}$ and normal margins. Based on 5000 generated observations and bandwidth b=1.

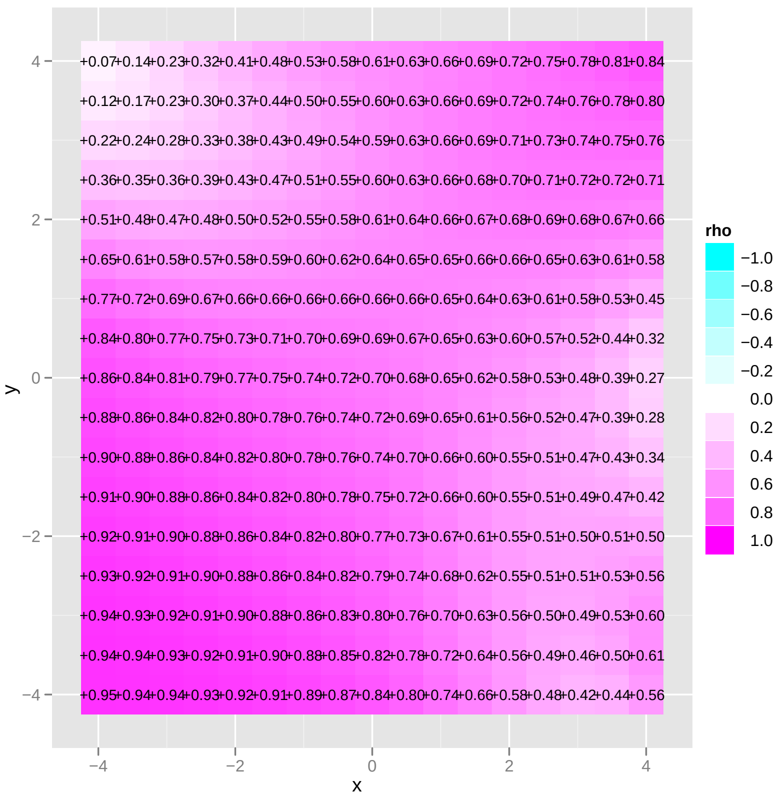

Figure 8.20: Estimated LGC plot of a CG copula with $\tau = \frac{1}{2}$ and normal margins. Based on 5000 generated observations and bandwidth b=1.

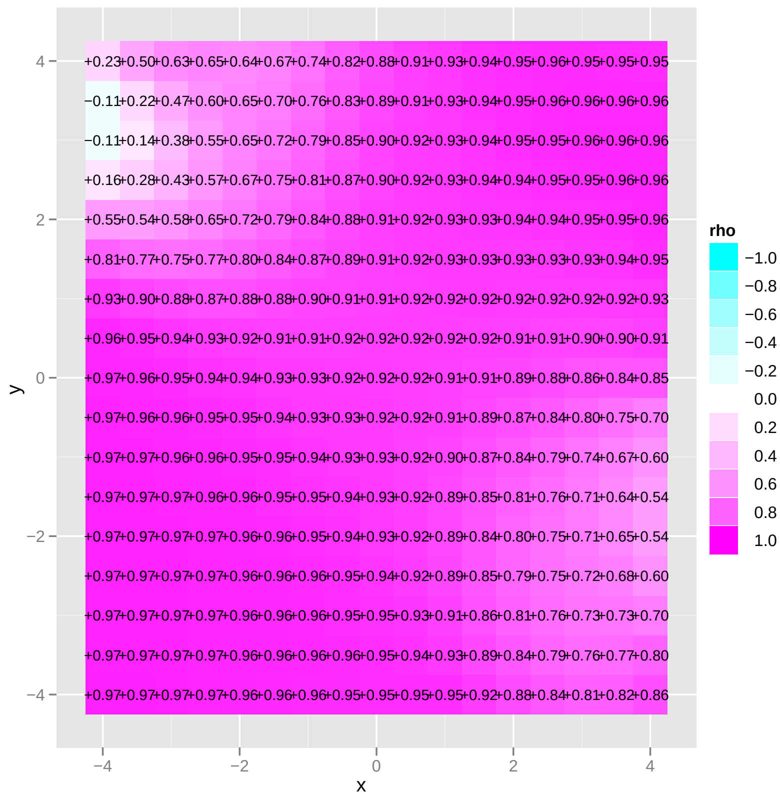

Figure 8.21: Estimated LGC plot of a CG copula with $\tau = \frac{3}{4}$ and normal margins. Based on 5000 generated observations and bandwidth b=0.8.

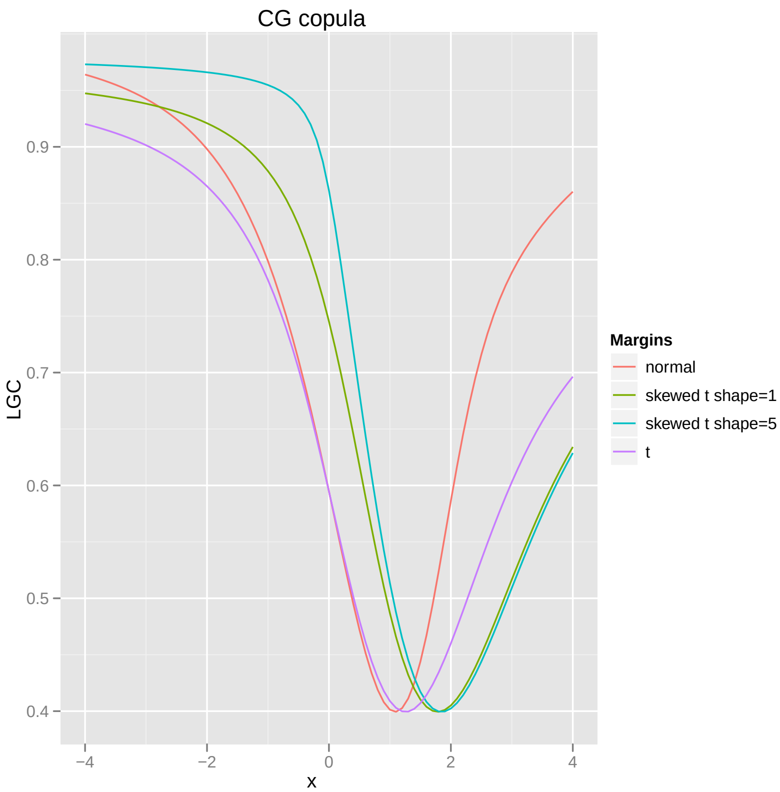

Figure 8.22: Theoretical LGC of CG copula with $\tau = \frac{2}{5}$ and with four different choices for margins: Normal margins, t distributed margins with 4 degrees of freedom, skewed t distributed margins with 4 degrees of freedom and shape parameter 1, and skewed t distributed margins with 4 degrees of freedom and shape parameter 5.

Chapter 9

Estimation, selection and goodness of fit

Modelling dependence by using the copula framework can be fruitful, and there is an abundant amount of different copula functions to choose from. The big problem is often to find the best suitable copula function for a given data set. No real procedure has been established, but several different approaches has been presented over the years. We will now first have a look at some ways of estimating copula functions, and then present a couple of model selection tools and Goodness-of-Fit test. This will all be done in a general setting. Some of the copula families have special properties that can simplify the estimation. We are mainly interested in the dependence structure, not the marginal distributions. This means that the estimation and validation of the margins will not be presented in any detail here. One of the good qualities about the copula approach is that the margins and the copula itself can be dealt with separately.

9.1 Estimation

9.1.1 Parametric estimation

Let $x_t = (x_{1,t}, \dots, x_{n,t})$ for $t = 1, 2, \dots, k$ be a random sample of iid n-dimensional vectors. Consider a n-dimensional distribution with distribution function F and density f, and univariate margins $F_1, \dots, F_n$ with densities respectively $f_1, \dots, f_n$. Let $C(u_1, \dots, u_n)$ be its copula with density $c(u_1, \dots, u_n)$. Now we remember that we can write the joint density as

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

Let θ be the parameter vector for the copula, and α_i be the parameter vector for marginal distribution i for $i = 1, 2, \dots, n$. Now we can write the log-likelihood as

$$
\begin{aligned}
L(\alpha_1, \ldots, \alpha_n, \theta) &= \sum_{j=1}^{k} \log f(x_{1,j}, \ldots, x_{n,j}) \\
&= \sum_{j=1}^{k} \log c(F_1(x_{1,j}; \alpha_1), \ldots, F_n(x_{n,j}; \alpha_n); \theta) + \sum_{j=1}^{k} \sum_{i=1}^{n} \log f_i(x_i, j; \alpha_i).
\end{aligned} \tag{9.1}
$$

We can notice that the log-likelihood is split into two parts, one representing the dependence through the copula and one part consisting of the marginal densities. The MLE (maximum likelihood estimate) for the paramters $(\alpha_1, \ldots, \alpha_n, \theta)$ is found by maximising the log likelihood function L for all parameters simultaneously. That is we can write

$$
(\widehat{\alpha_1}^{MLE}, \ldots, \widehat{\alpha_n}^{MLE}, \widehat{\theta}^{MLE}) = argmax_{\alpha_1, \ldots, \alpha_n, \theta} L(\alpha_1, \ldots, \alpha_n, \theta).
$$

This method can often be very computationally intensive, especially in higher dimensions. So instead we can use the noted fact about the copula representation of the log-likelihood (9.1), that is we can split the expression into different parts where we can estimate the different parameters alone. We call the following method IFM (inference for the margins). Firstly we estimate every marginal parameter α_i by maximizing $\sum_{j=1}^{k} \log f_i(x_{i,j}; \alpha_i)$, that is

$$
\widehat{\alpha_i}^{IFM} = argmax_{\alpha_i} \sum_{j=1}^{k} \log f_i(x_{i,j}; \alpha_i).
$$

Then we use the estimated marginal parameters when we maximize

$$
\sum_{j=1}^{k} \log c(F_1(x_{1,j}), \ldots, F_n(x_{n,j}); \theta)
$$

to find the estimate for the copula parameter. That is

$$
\widehat{\theta}^{IFM} = argmax_{\theta} \sum_{j=1}^{k} \log c(F_1(x_{1,j}; \widehat{\alpha_1}^{IFM}), \ldots, F_n(x_{n,j}; \widehat{\alpha_n}^{IFM}); \theta).
$$

This method is computationally more simple. Both the MLE and the IFM estimators are consistent and asymptotic normal for both the multivariate model and for the margins under regularity conditions (see for example Triverdi and Zimmer 2005).

9.1.2 Semi-parametric estimation

The IFM estimator tends to work well in many cases, but because it involves estimates of the marginal distribution we always have the risk that the model selected for the margins will be bad and then effect the estimation of the dependence parameter. This is among the reasons why in practise probably the most used technique is a semi-parametric method some places called PML/CML. The idea is to put the empirical distribution functions, $\widehat{F}_i$, in for the margins in the copula log-likelihood. We remember that the empirical distribution functions for a sample $X_1, \dots X_d$ is given by

$$\widehat{F} = \frac{1}{d} \sum_{i=1}^{d} \mathbf{1}(X_i \leq x). \tag{9.2}$$

If we have a sample $X_t = (X_{1,t}, \dots, X_{n,t})$, we say that

$$U_t = (U_{1,t}, \dots, U_{n,t}) = (\widehat{F}_1(X_{1,t}), \dots, \widehat{F}_n(X_{n,t})) \tag{9.3}$$

is pseudo observations. The CML method gives us the following estimate for the copula parameter

$$\widehat{\theta}^{CML} = argmax_\theta \sum_{j=1}^{k} \log c(\widehat{F}_1(x_{1,j}), \dots, \widehat{F}_n(x_{n,j}); \theta). \tag{9.4}$$

Under suitable regularity conditions $\widehat{\theta}^{CML}$ is consistent and asymptotically normal (see Genest, Ghoudi and Rivest (1995)).

9.1.3 Empirical copula

From any copula of an empirical distribution function we could create an empirical copula, but it would not be unique. We will use the definition from Deheuvels (1981), that is

Definition 29. *The empirical copula $\widehat{C}$ is any copula defined on the lattice*

$$\{(\frac{t_1}{T}, \dots, \frac{t_n}{T}) : 1 \leq i \leq n, t_i = 0, 1, \dots, T\}$$

by

$$\widehat{C}(\frac{t_1}{T}, \dots, \frac{t_n}{T}) = \frac{1}{T} \sum_{t=1}^{T} \prod_{i=1}^{n} \mathbf{1}(r_i^t \leq t_i).$$

Here $\{x_1^{(t)}, \dots, x_n^{(t)}\}$ is the order statistic and $\{r_1^t, \dots, r_n^t\}$ is the rank statistic, where their relationship is given by $x_n^{r_n^t} = x_n, t$ for $t = 1, \dots, T$. The Radon-Nikodym density for the empirical copula $\widehat{C}$ is given by

$$\widehat{c}(\frac{t_1}{T}, \dots, \frac{t_n}{T}) = \sum_{j_1=1}^{2} \cdots \sum_{j_n=1}^{2} (-1)^{j_1+\cdots+j_n} \widehat{C}(\frac{t_1 - j_1 + 1}{T}, \dots, \frac{t_n - j_n + 1}{T}).$$

So we can write

$$\widehat{C}(\frac{t_1}{T}, \dots, \frac{t_n}{T}) = \sum_{j_1=1}^{t_1} \cdots \sum_{j_n=1}^{t_n} \widehat{c}(\frac{j_1}{T}, \dots, \frac{j_n}{T}).$$

We say that $\widehat{c}$ is the empirical copula frequency. Deheuvels (1978,1981) proves that $\widehat{C}$ converges uniformly to the underlying copula C. The empirical copula gives us sample versions of some of the dependence measures described earlier.

For example can we define the sample version of the coefficient for upper tail dependence as (Durrleman, Nikeghbali and Roncalli (2000))

$$\widehat{\lambda}_u = \lim_{t \to T} 2 - \frac{1 - \widehat{C}(\frac{t}{T}, \frac{t}{T})}{1 - \frac{t}{T}}.$$

The sample versions of Spearmans rho $\widehat{\rho}_s$ and Kendalls tau $\widehat{\tau}$ will be (Nelsen (2006))

$$\widehat{\rho}_s = \frac{12}{T^2 - 1} \sum_{t_1=1}^{T} \sum_{t_2=1}^{T} (\widehat{C}(\frac{t_1}{T}, \frac{t_2}{T}) - \frac{t_1 t_2}{T^2})$$

and

$$\widehat{\tau} = \frac{2T}{T-1} \sum_{t_1=2}^{T} \sum_{t_2=2}^{T} \sum_{i_1=1}^{t_1-1} \sum_{i_2=1}^{t_2-1} (\widehat{c}(\frac{t_1}{T}, \frac{t_2}{T}) \widehat{c}(\frac{i_1}{T}, \frac{i_2}{T}) - \widehat{c}(\frac{t_1}{T}, \frac{i_2}{T}) \widehat{c}(\frac{i_1}{T}, \frac{t_2}{T})).$$

For many parametric copula models we can find closed expressions linking the copula parameter with Kendalls tau, which means that we can use the sample version of the dependence measure to estimate the copula parameter. This is the case for many Archimedean copula families.

9.1.4 Kernel methods

Because off the differentiable property of the copula, using the smooth kernel functions in estimation can be a way to go. In this case we can estimate the dependence structure without having to make initially guess on certain parametric copula families. Fermanian and Scaillet (2002) presented in their paper a kernel approach in the context of multivariate stationary processes satisfying strong mixing conditions. We will denote the observation as before, that is $(X_{1,t}, \dots, X_{n,t})$ for $t = 1, \dots, n$, where F is the joint distribution function and F_j the margins. We remember the formula

$$C(u_1, \dots, u_n) = F(F_1^{-1}(u_1), \dots, F_n^{-1}(u_n)), \tag{9.5}$$

which comes from Sklars theorem. We assume that F_j is such that we get a unique solution ξ_j from the equation $F_j(x) = u_j$. The right kernel function has to be chosen, and then we use the regular kernel methods to get estimates for f_j, which again leads to estimates for f, F_j and F. Let now $\widehat{\xi} = (\widehat{\xi}_1, \dots, \widehat{\xi}_n)$ where $\widehat{\xi}_j = inf_{x \in \mathbf{R}} \{x : \widehat{F}_j(x) \geq u_j\}$. Then we get the estimate by plugging $\widehat{F}$ and $\widehat{\xi}$ into equation 9.5, that is we get

$$\widehat{C} = \widehat{F}(\widehat{\xi}).$$

We refer to Fermanian and Scaillet (2002) for a more thorough presentation and for discussion about properties and applications.

9.2 Detecting dependence

To use the parametric estimation methods we need to have an initial idea of what the underlying copula is, we need to find out as much as possible about the dependence structure of the data. The first thing we could try is to make a scatter-plot of the observed data, but it is difficult to separate the dependence structure from the effect of the margins on the plot. Then it is better to plot the ranks or the pseudo observations. These plots will be invariant under strictly increasing transformations of the margins, and gives more insight into the dependence between the data. We could compare the plot with simulated scatter-plots from known copula families, and then choose those that look most similar. Still it can be hard to distinguish patterns from just random variation on such plots. So we will present two other graphical procedures which is suitable to detect dependence, and then discuss how we can use the LGC-plots to choose the right copula.

9.2.1 Chi plots

Say we have a bivariate random sample $(X_1, Y_1), \ldots, (X_n, Y_n)$. For a given pair (X_i, Y_i), let

$$H_i = \frac{1}{n-1} \sum_{j=1, j \neq i}^{n} \mathbf{1}(X_j \leq X_i, Y_j \leq Y_i),$$

$$F_i = \frac{1}{n-1} \sum_{j=1, j \neq i}^{n} \mathbf{1}(X_j \leq X_i)$$

$$G_i = \frac{1}{n-1} \sum_{j=1, j \neq i}^{n} \mathbf{1}(Y_j \leq Y_i).$$

We note the resemblance with the empirical distribution function, and that H_i, F_i and G_i depend on the ranks of the observations. Further let

$$\chi_i = \frac{H_i - F_i G_i}{\sqrt{F_i(1 - F_i) G_i(1 - G_i)}}$$

and

$$\lambda_i = 4 sign(\tilde{F}_i \tilde{G}_i) max(\tilde{F}_i^2, \tilde{G}_i^2),$$

where $\tilde{F}_i = F_i - \frac{1}{2}$ and $\tilde{G}_i = G_i - \frac{1}{2}$. We get the Chi plot by plotting the pairs (λ_i, χ_i). This was proposed by Fisher and Switzer (1985,2001), and they recommended to only plot pairs for which $|\lambda_i| \leq 4(\frac{1}{n-1} - \frac{1}{2})^2$ to avoid outliers. We have that λ_i and χ_i will be in the interval $[-1, 1]$. λ_i will be a measure of the distance between the pair (X_i, Y_i) and the center of the scatter plot. Under independence we would expect $H_i \approx F_i G_i$ which means that values of χ_i far from zero indicates non-independence. It is useful to plot horizontal guidelines, $\chi = \pm \frac{c_p}{\sqrt{n}}$, where c_p is chosen such that $p \times 100\%$ of the pairs (λ_i, χ_i) lays between the lines. Monte Carlo simulations can be used to determine the correct c_p. We can compare Chi plot of our data with Chi plot of different copula functions, and then choose the copula function with the most similar plot. This is not always a practical approach though because of the amount of different copula functions and the fact that there is no simple connection between the concept of copulas and the Chi plots.

9.2.2 K-plot

Genest and Boies (2003) suggest another plot, which in their opinion is easier to interpret. This is because the curvature these plots displays is related in a definite way to the copula in case of association. They call it Kendall plots or just K-plots, and are similar to the well known QQ-plot. Let us continue to consider the random sample $(X_1, Y_1), \ldots, (X_n, Y_n)$, where H is the joint distribution function, and F and G are the margins. The K-plot will be the plot of the pairs $(W_{i:n}, H_{(i)})$. Here $H_{(1)} \leq H_{(2)} \leq \cdots \leq H_{(n)}$ is the order statistic of the H_is defined in the previous section. Let

$$K(w) = P(W \leq w) = P(H(X,Y) \leq w) \tag{9.6}$$

for $w \in [0,1]$, and let K_0 be the distribution function of H(X,Y) under the null hypothesis of independence between X and Y. Now $W_{i:n}$ represents the expectation of the ith order statistic in a random sample of size n from K_0. From regular techniques regarding order statistics we get that

$$W_{i:n} = n\binom{n-1}{i-1}\int_0^1 w(K_0(w))^{i-1}(1-K_0(w))^{n-i}dK_0(w).$$

What we need is some sort of expression for K_0 to put into the equation above. This is done by some simple calculations, where we remember that X and Y (and therefore also U and V) is considered to be independent

$$\begin{aligned}K_0(w) &= P(H(X,Y) \leq w) = P(C(U,V) \leq w) = P(UV \leq w)\\ &= \int_0^1 P(U \leq \frac{w}{v})dv = \int_0^1 \frac{w}{v}dv = \int_0^w dv + \int_w^1 \frac{w}{v}dv\\ &= w - w\log(w).\end{aligned}$$

We can interpret K to be a univariate summary of the dependence embodied in C, and by this define a stochastic ordering in the following way. Let us consider two random pairs (X,Y) and (X',Y') with distribution H and H' respectively. We say that (X,Y) is less positive dependent than (X',Y'), and write $(X,Y) \prec_K (X',Y')$, if and only if $K(w) \geq K'(w)$ for all $w \in [0,1]$. K is connected to Kendalls tau in the following way

$$\tau(X,Y) = 4E[C(U,V)] - 1 = 3 - 4\int_0^1 K(w)dw$$

When we have independence the plot will be linear and follow the line x=y. Positive dependence will give plots over this line, while negative dependence will give plots under. The stronger the dependence the greater the distance will be between the line x=y and the plots. For perfect positive dependence the data points will follow the curve $K_0(w)$, while perfect negative dependence gives points along the x-axis. One problem is that different copulas does not necessary imply different K's, which means that it can be difficult to precisely decide the dependence structure from the K-plot. This can be seen by the relation between K and Kendalls tau when we now that Kendalls tau do not give a complete description of the dependence. While the K-plots may be easier to interpret than the Chi-plots, it may give less information about the dependence structure do to its univariate nature. The K-plots are extendable to the multivariate case. For more properties and some examples of K-plots see Genest and Boies (2003).

9.2.3 LGC plot

The LGC plots gives an easy interpretable method for detecting dependence. Even though estimating the LGC can be a bit computer intensive, it is still pretty fast, specially if we restrict our self to one line like the diagonal. We can compare the LGC estimated from the data with the LGC for different copula models. If we can use the theoretical LGC, like the ones derived in this thesis, this can be done quickly. For comparing it is probably easiest to draw all lines in one plot. It is important to chose bandwidth with care, and to be aware that the estimated LGC is not that reliable in low density areas. Otherwise it is easy to interpret estimation noise as changes in the underlying dependence.

9.3 Model diagnostics and Goodness-of-Fit

Let us now assume that we have fitted the copula C_{θ_n} by one of the methods described above. It is crucial to figure out if our estimated copula really do describe our data in a good way. Let us first consider some quick graphical diagnostic methods.

9.3.1 Graphical diagnostic

One easy way to get a first impression of our fit is to simulate random observations from the copula C_{θ_n} and plot them together with the pseudo observations we got from our data. If the pseudo observations differs a lot from the simulated data we know that our model may need improvement. By using the estimated marginal distributions $\widehat{F}$ and $\widehat{G}$ we can also transform our simulated data back into original units. That is we get the scatter plots $(X_i, Y_i) = (\widehat{F}^{-1}(U_i), \widehat{G}^{-1}(V_i))$ which we compare with the observed data. Another way to go is to use the K introduced in the K-plots (see Genest and Boies (2003). We have

$$W_i = \frac{1}{n}\sum_{j=1}^{n} \mathbf{1}(X_j \leq X_i, Y_j \leq Y_i),$$

for $i = 1, \ldots, n$. Let K_n be the empirical distribution function to $W_1, \ldots, W_n$. Now we can compare K_n with the distribution function of $C_{\theta_n}(U, V)$, denoted by K_{θ_n}, where (U,V) is drawn from C_{θ_n}. There is two different approaches for the comparison. The first one is to plot K_n and K_{θ_n} in the same plot. Here K_n will be a step function, but we know that $K_n \to K$ and $C_{\theta_n} \to C_\theta$, such that if we have a large sample the two graphs should be close to each other if the model is correct. Another way is to make a QQ-plot by using the pairs $(W_{i:n}, W_{(i)})$, where in this case $W_{i:n}$ will be the expected value of the ith order statistic from a random sample consisting of n observations from K_{θ_n}. If the copula we have fitted is correct the points should lay approximately along the diagonal straight line.

9.3.2 GoF

Let us consider an iid sample $X_t = (X_{1,t}, \ldots, X_{d,t})$ for $t = 1, \ldots, n$, with distribution function H and marginal distributions $F_1, \ldots, F_d$. We are interested in establishing some formal test procedures in order to get some idea if our estimated copula is any good. That is we are considering problems like

$$\mathbf{H}_0 : C = C_0 \text{ against } \mathbf{H}_a : C \neq C_0 \tag{9.7}$$

$$\mathbf{H}_0 : C = C_\theta \text{ against } \mathbf{H}_a : C \notin C_\theta. \tag{9.8}$$

Here C_0 is a known copula, and in this case we say that $\mathbf{H}_0$ is a simple zero-assumption. C_θ is known parametric copula family, and we say the zero-assumption in this case is composite. It is in general difficult to construct such goodness off fit tests for copulas because the marginal distributions are unknown. Some authors have proposed to test for the fit of the marginal and joint behaviour at the same time, that is construct tests for H itself in the way usually done for multivariate distributions. But this is probably not the most fruitful approach, we would prefer to find ways to test the fit of the dependence structure alone. Usually empirical margins are used. A natural way to start would be to consider the distance between the empirical copula C_n and our chosen copula model C_0. But the limiting distribution of $\sqrt{n}(C_n - C_0)$ is very complex, which leads to extensive use of computer intensive bootstrap methods.

9.3.3 Probability integral transformation

There has been several suggestions of GoF tests based on the probability integral transformation (PIT). This is a way of transforming a set of dependent variables into independent U(0,1) variables, and we can use a test for multivariate independent uniformity and use it on any model. We see that this transformation is the same as the Rosenblatt transformation discussed earlier. Let us look at a random vector $X = (X_1, \ldots, X_d)$ with marginal distributions $F_i(x_i)$. We define $T(X) = (T_1(X_1), \ldots, T_d(X_d))$ as (Berg and Bakken (2005))

$$\begin{aligned}
T_1(X_1) &= P(X_1 \leq x_1) = F_1(x_1) \\
T_2(X_2) &= P(X_2 \leq x_2 | X_1 = x1) = F_{2|1}(x_2|x_1) \\
&\vdots \\
T_d(X_d) &= P(X_d \leq x_d | X_1 = x_1 \ldots X_{d-1} = x_{d-1}) = F_{d|1\ldots d-1}(x_d|x_1, \ldots, x_{d-1}).
\end{aligned} \tag{9.9}$$

Now the variables $T_i(X_i)$ for $i = 1, \ldots, d$ are independent and uniformly distributed on $[0,1]^d$, and we say that $T(X)$ is the PIT of the random variable X. If we now have a parametric copula family as null hypothesis, we can use the PIT transformation on the data set, and test for multivariate independence. There are several different test statistic that can be used.

9.3.4 Dimension reduction approaches

We are first going to look at a couple of different dimension reduction approaches. The common idea is to reduce the data to one dimension, and then use a univariate test statistic. Let $Y = (Y_1, \ldots, Y_d)$ be the iid uniformly distributed variables we get from performing the PIT on a data set $X = (X_1, \ldots, X_n)$. Breymann,Dias and Embrechts (2003) proposes the following dimension reduction approach

$$M_G = \sum_{i=1}^{d} \Phi^{-1}(Y_i)^2.$$

M_G is χ^2_d distributed since $\Phi^{-1}(Y_i)$ is independent standard normal. Now

$$W_G = F_{\chi^2_d}(M_G)$$

is U(0,1) distributed. Let

$$F_G(w) = P(W_G \leq w), w \in [0,1].$$

Under H_0 we will have $F_G(w) = w$ and density function $f_G(w) = 1$. Now we can apply a univariate test statistic to F_G or f_G. This approach is computationally efficient, but not consistent. Berg and Bakken (2005) proposes an extension which is consistent and also with more flexible weighting. Let $\widehat{Y} = (\widehat{Y_1}, \ldots, \widehat{Y_d})$ be the sorted counterpart of Y and r_i is rank variable i. We now define Y^* as

$$Y_i^* = P(r_i \leq \widehat{Y}_i | r_1, \ldots, r_{i-1}) = 1 - (\frac{1 - \widehat{Y}_i}{1 - r_{i-1}})^{d-(i-1)}.$$

Let γ be a weighting function depending on Y_i and weighting parameters α. We now get the dimension reduction

$$M_B = \sum_{i=1}^{d} \gamma(Y_i; \alpha)\Phi^{-1}(Y_i^*)^2.$$

Let

$$W_B = F_{Y_B}(Y_B)$$

where the distribution of Y_B can be found numerically or by simulation. Under H_0 W_B will be uniformly distributed. A different dimension reduction approach is presented in Genest, Quessy and Remillard (2006), and it is based on the K function from (9.6). The expression for the K function have to be found for the specific copula model under H_0, and then we can make a test based on the difference between this expression and the empirical K function. This approach is most suited for Archimedean copulas where there is a nice analytic expression for K. Berg and Bakken (2005) presents an alternative way to this last approach, see their paper for details and also test results for the methods presented.

9.3.5 Other GoF

Fermanian (2005) proposes a distribution free goodness of fit test based on a kernel estimate of the copula which leads to a chi-square type test procedure. We use the L2 norm to measure the proximity between the smoothed copula density and the estimated parametric density. They will be near each other under $\mathbf{H}_0$. Panchenko (2005) also presents a full multivariate approach. Among the most recent papers presenting new copula selection methods are Huard et al. (2006), which proposes an Bayesian method for choosing the copula family most probable in a given set, and Karlis and Nikoloupoulos (2008), which uses the Mahalanobis squared distance between original and simulated data to construct a GoF test.

Several of the GoF tests presented here is evaluated and tested in Berg (2009).

9.3.6 GoF based on the LGC

In Berentsen et al. (2012) a GoF test for copulas based on the LGC is proposed. Let $\rho_\theta(\cdot)$ be the LGC for the distribution $C_\theta(F_1(x_1), F_2(x_2))$. If we have a set of iid observations $X_1, \ldots, X_n$ the copula parameter θ can be estimated by θ_n as described above. We let $\rho_{\theta_n}(\cdot)$ be the estimate of $\rho_\theta(\cdot)$, that is $\rho_\theta(\cdot)$ where we have used θ_n and the empirical distribution functions $\widehat{F}_j$ for $j = 1, 2$. The local likelihood estimate of the LGC from the observations is denoted $\rho_{n,b}$. The proposed GoF test in Berentsen et al. (2012) is now based on the process $P_n(\cdot) = \rho_{n,b}(\cdot) - \rho_{\theta_n}(\cdot)$. We specify a grid $(x_1, \ldots, x_n)$ and calculate

$$T_n = \sum_{i=1}^{p} P_n(x_i)^2.$$

H_0 is now rejected if T_n gets to big. To obtain approximate p-values Berentsen et al. (2012) suggests using a parametric bootstrap method similar to the one used in Genest, Remillard and Beaudoin (2009), see their papers for details. Preferably we want to use an analytic expression for ρ_{θ_n}. But as we remember the theoretical LGC in (5.5) is only consistent with the local likelihood estimate $\rho_{n,b}$ along $F_1(x_1) = F_2(x_2)$. This means that we are restricted to choosing gridpoints where $\widehat{F}_1(x_{i1}) \approx \widehat{F}_2(x_{i2})$. To be able to perform a GoF test also when the theoretical LGC (5.5) is not available, and to get to chose the grid $(x_1, \ldots, x_p)$ freely, Berentsen et al. (2012) also proposes a double parametric bootstrap procedure where ρ_θ is estimated by monte carlo approximation. This is computationally considerably more demanding.

As a graphical diagnostic tool and a way to investigate the departures from the null hypothesis, Berentsen et al. (2012) proposes two different plots. First the curves $\rho_{n,b}(\cdot)$ and $\rho_\theta(\cdot)$ can be drawn in the same plot, together with bootstrap confidence intervals. Also a plot based on local "goodness-of-fit" tests performed over a grid on $\mathbb{R}^2$ is proposed. If we have $\rho_\theta(\cdot)$ under the original H_0, the idea is to test the null hypothesis $\rho(x_j) = \rho_\theta(x_j)$ for every point x_j on the grid $(x_1, \ldots, x_p)$. The test statistic used is $\rho_{n,b}(x_j)$. By sampling from $F^*(x) = C_{\theta_n}(\widehat{F}_1(x_1), \widehat{F}_2(x_2))$ we generate R bootstrap samples of size n. Now the LGC estimate at x_j based on bootstrap sample number p is denoted $\rho^{*,p}_{n,b}(x_j)$. By significance level α we now reject the hypothesis $\rho(x_j) = \rho_\theta(x_j)$ if $\rho_{n,b}(x_j)$ is smaller or bigger than the $(\alpha/2)\%$ or $(1-\alpha/2)\%$ quantile of $\rho^{*,1}_{n,b}(x_j), \ldots, \rho^{*,R}_{n,b}(x_j)$ respectively. This is then presented in plots similar to the $\mathbb{R}^2$ LGC plots shown earlier in this thesis. Here magenta is used where $\rho_{n,b}(x_j)$ is significantly larger and cyan is used when it is significantly smaller. White is used if the null hypothesis is not rejected. Examples of these plots, a simulation study and a real data study can be found in Berentsen et al. (2012).

Chapter 10

Conclusion

We have seen that we can find an explicit expression for the theoretical LGC in some cases, and that the estimated LGC mostly agrees with the theoretical LGC. Of course some predictable (and a few surprising) estimation errors occurs. Especially problems arise at the boundaries and other low density areas. The choice of bandwidth do play a big part when we are trying to get good estimates, and we have seen that bandwidth has to be chosen with care. It shows the importance of creating good algorithms for choosing bandwidth, preferably one that allows local bandwidth. All in all the LGC has shown itself to be a useful tool for analysing copulas. Even though the theoretical LGC usual is not easy interpretable analytical, it is for the most part easily implemented on a computer, and the diagonal plots are fast to produce. In the future we will hopefully find a more general way to define a theoretical LGC, which will open up for plotting larger areas than one line. But it is clear that the diagonal gives us a lot of information, and it is often easier to extract the structure from the diagonal than from a full LGC plot. Plotting the diagonal (or other lines) also gives us the possibility to plot several different models in one plot, for easy comparison and for showing trends when for example we vary a parameter. It has been positive to see how well the LGC plots for the different copula models have coincided with our expectations built on other dependence measures, like the tail dependence. It is always reassuring when many approaches points in the same direction. Using the LGC for selection and goodness-of-fit also seems very promising, and the development of these applications will continue. We should in the future develop theoretical LGC algorithm for more copula models. The skewed t copula would for example be an interesting case to explore.

Chapter 11

Bibliography

Abdous, B., Genest, C., and Rèmillard, B. (2005). *Dependence Properties of Meta-Elliptical Distributions.* Statistical Modeling and Analysis for Complex Data Problems. Springer US. Pages 1-15.

Azzalini, A. (1985). *A Class of Distributions Which Includes the Normal Ones.* Scandinavian Journal of Statistics, 12, 2, pages 171-178.

Azzalini, A. (2005). *The skew-normal distribution and related multivariate families (with discussion).* Scandinavian Journal of Statistics, 32, 2, pages 159-188.

Azzalini, A. (2010). *R package sn: The skew-normal and skew-t distributions (version 0.4-16).* URL http://azzalini.stat.unipd.it/SN

Azzalini, A. and Capitanio, A. (1999). *Statistical applications of the multivariate skew-normal distributions.* J. Roy. Statist. Soc. Ser. B 61, pages 579-602.

Azzalini, A. and Capitanio, A. (2003). *Distributions generated by perturbation of symmetry with emphasis on a multivariate skew-t distribution.* J.Roy. Statist. Soc. B 65, pages 367-389.

Azzalini, A. and Dalla Valle, A. (1996). *The multivariate skew-normal distribution.* Biometrika 83, pages 715-726.

Berg, D. (2009). *Copula goodness-of-fit testing: an overview and power comparison.* The European Journal of Finance, 15, pages 675-701.

Berg, D. and Bakken H. (2005). *A goodness-of-fit test for copulae based on the probability integral transform.* Technical Report SAMBA/41=05, Norsk Regnesentral, Oslo, Norway.

Berentsen, G., Nordbø, T., Støve, B., and Tjøstheim, D. (2012). *The copula and local Gaussian correlation: Vizualization of local dependence structure and a goodness-of-fit test.* Manuscript, Department of Mathematics, University of Bergen.

Breymann, W., Dias, A., and Embrechts, P. (2003). *Dependence structures for multivariate high-frequency data in finance.* Quantitative Finance 1, pages 1-14.

Coles, S., Heffernan, J., and Tawn, J. (1999). *Dependence Measures for Extreme Value Analyses.* Extremes 2:4, pages 339-365.

Deheuvels, P. (1978). *Caractèrisation complète des lois extrêmes multivarièes et de la convergence des types extrêmes.* Publications de l'Institut de Statistique del'Universite de Paris, 23, pages 1-36

Deheuvels, P. (1981). *A non parametric test for independence.* Publications de l'Institut de Statistique del'Universite de Paris, 26, pages 29-50

Durrleman, V., Nikeghbali, A., and Roncalli, T. (2000). *Which copula is the right one?* Groupe de Recherche Operationnelle, Credit Lyonnais, Working Paper.

Embrechts, P., Lindskog, F. og McNeil, A. (2001). *Modelling Dependence with Copulas and Applications to Risk Management.* Published in *Handbook of Heavy Tailed Distributions in Finance*, (2003) ed. S. Rachev, Elsevier, Chapter 8, pages 329-384.

Embrechts, P., McNeil, A. og Straumann, D. (1999). *Correlation and dependence in risk management: properties and pitfalls.* I Risk Management: Value at Risk and Beyond, ed. M.A.H. Dempster, Cambridge University Press, Cambridge, pages 176-223

Fang, K.-T., Kotz, S. and Ng, K.-W. (1987). *Symmetric Multivariate and Related Distributions.* Chapman and Hall, London.

Fang, H.-B., Fang, K.-T., and Kotz, S. (2002). *The meta-elliptical distributions with given marginals.* Journal of Multivariate Analysis, 82, pages 1-16.

Fermanian, J. (2005). *Goodness of fit tests for copulas.* Journal of Multivariate Analysis, 95, 1, pages 119-152.

Fermanian, J-D. and Scaillet, O. (2002). *Nonparametric estimation of copulas for time series.* Journal of Risk 5, 4, pages 25-54.

Fisher, N. I. and Switzer, P. (1985). *Chi-Plots for Assessing Dependence.* Biometrika, 72, pages 253-265.

Fisher, N. I. and Switzer, P. (2001). *Graphical Assessment of Dependence: Is a Picture Worth 100 Tests?.* The American Statistician, 55, pages 233-239.

Frank, M. J. (1979). *On the simultaneous associativity of F(x,y) and x+y-F(x,y).* Aequationes Math. 19, pages 194-226.

Frèchet, M. (1957). *Les tableaux de corrèlation dont les marges et des bornes sont donnèes.* Annales de l'Universitè de Lyon, Sciences Mathêmatiques et Astronomie, 20, pages 13-31.

Genest, C. (1987). *Frank's family of bivariate distributions.* Biometrika, 74, 3, pages 549-555.

Genest, C. and Boies, J-C. (2003). *Detecting dependence with Kendall plots.* The American Statistician, 57, 4, pages 275-284.

Genest, C., Quessy,J.-F., and Rèmillard, B. (2006). *Goodness-of-fit procedures for copula models based on the probability integral transform.* Scandinavian Journal of Statistics 33, pages 337-366.

Genest, C. and Rèmillard, B. (2006). *Discussion of "Copulas: Tales and Facts," by Thomas Mikosch.* Extremes, 9, pages 27-36.

Genest, C., Remillard, B., and Beaudoin, D. (2009). *Goodness-of-fit tests for copulas: A review and a power study.* Insurance Mathematics and Economics, 44, pages 199-213.

Genest, C., Ghoudi K., and Rivest, L. (1995). *A semi-parametric estimation procedure of dependence parameters in multivariate families of distributions.* Biometrika 82, pages 543-552.

Hjort, N. and Jones, M. (1996). *Locally parametric nonparametric density estimation.* Annals of Statistics, 24, pages 1619-1647

Huard, D., Evin, G., and Favre, A.C. (2006). *Bayesian copula selection.* Journal of Computational Statistics and Data Analysis, 51, pages 809-822.

Hult, H. and Lindskog, F. (2002). *Multivariate extremes, aggregation and dependence in ellitpical distributions.* Advances in Applied Probability, 34, pages 587-608.

Karlis, D. and Nikoloupoulos, A.K. (2008). *Copula model evaluation based on parametric boostrap.* Journal of Computational Statistics and Data Analysis, 52, 7, pages 3342-3353.

Malevergne,Y. and D. Sornette. (2006). *Extreme financial risks: From dependence to risk management.* Berlin:Springer.

Mikosch, T. (2005). *Copulas: Tales and facts.* Extremes 9, pages 3-20.

Nelsen, R.B. (2006). *An Introduction to Copulas.* Second Edition. Springer, New York.

Panchenko, V. (2005). *Goodness-of-fit test for copulas.* Physica A 355(1), pages 176-182.

R Development Core Team (2010). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing,Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Rosenblatt, M. (1952). *Remarks on a multivariate transformation.* The Annals of Mathematical Statistics, 23, pages 470-472.

Sklar, A. (1959). *Fonctions de rèpartition à n dimensions et leurs marges.* Publications de l'Institut de Statistique de l'Universitè de Paris, 8, pages 229-231.

Tjøstheim D. and Hufthammer K.O. (2012). *Local Gaussian correlation: A new measure of dependence.* Unpublished manuscript.

Triverdi, P.K. and Zimmer, D.M. (2005). *Copula modeling: An introduction for practitioners.* Foundations and trends in econometrics, 1,1, pages 1-111

Wuertz, D. and many others, see the SOURCE file (2009). *fCopulae: Rmetrics - Dependence Structures with Copulas.* R package version 2110.78. http://CRAN.R-project.org/package=fCopulae

Appendix A

Plots

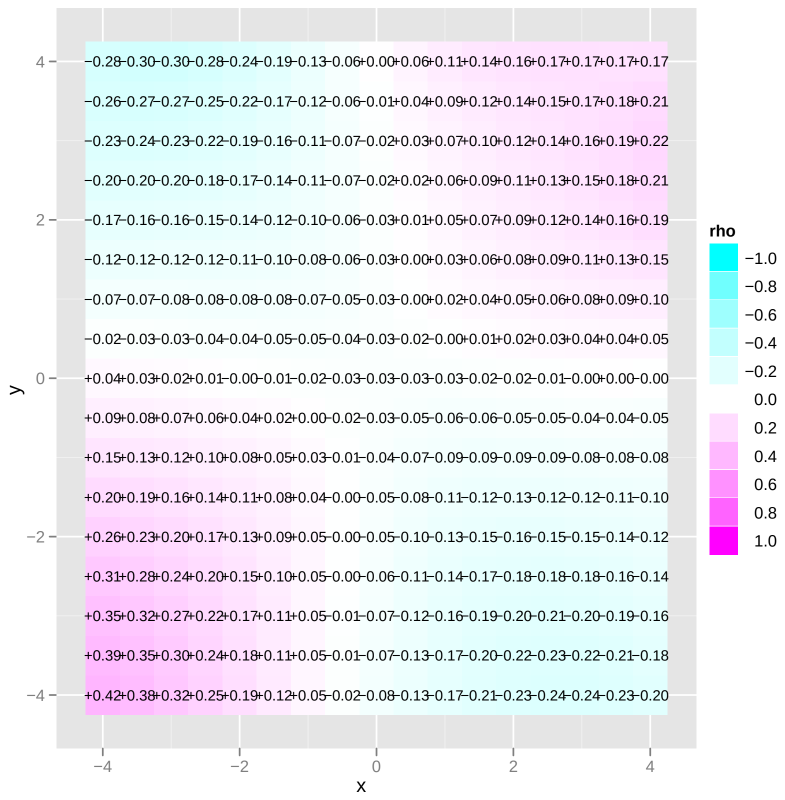

Figure A.1: Estimated LGC plot of t-copula with $\rho = 0$, $\nu = 4$ and t distributed margins with 4 degrees of freedom. Based on 5000 generated observations and bandwidth $b = 1.5$.

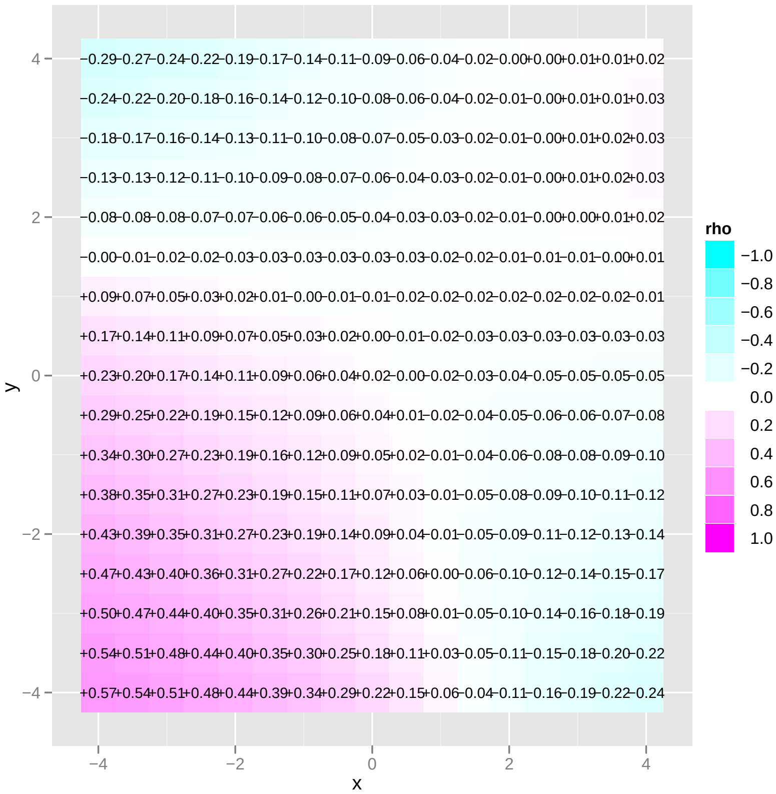

Figure A.2: Estimated LGC plot of t-copula with $\rho = 0$, $\nu = 4$ and skewed t distributed margins with 4 degrees of freedom and shape parameter 1. Based on 5000 generated observations and bandwidth $b = 1.5$.

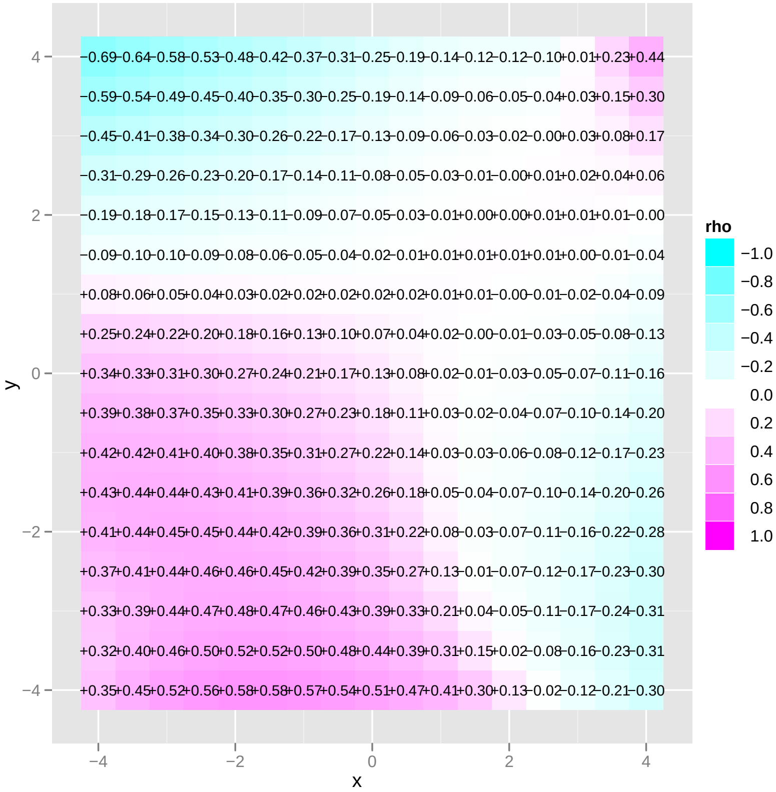

Figure A.3: Estimated LGC plot of t-copula with $\rho = 0.5$, $\nu = 4$ and skewed t distributed margins with 4 degrees of freedom and shape parameter 5. Based on 5000 generated observations and bandwidth $b = 1.5$.

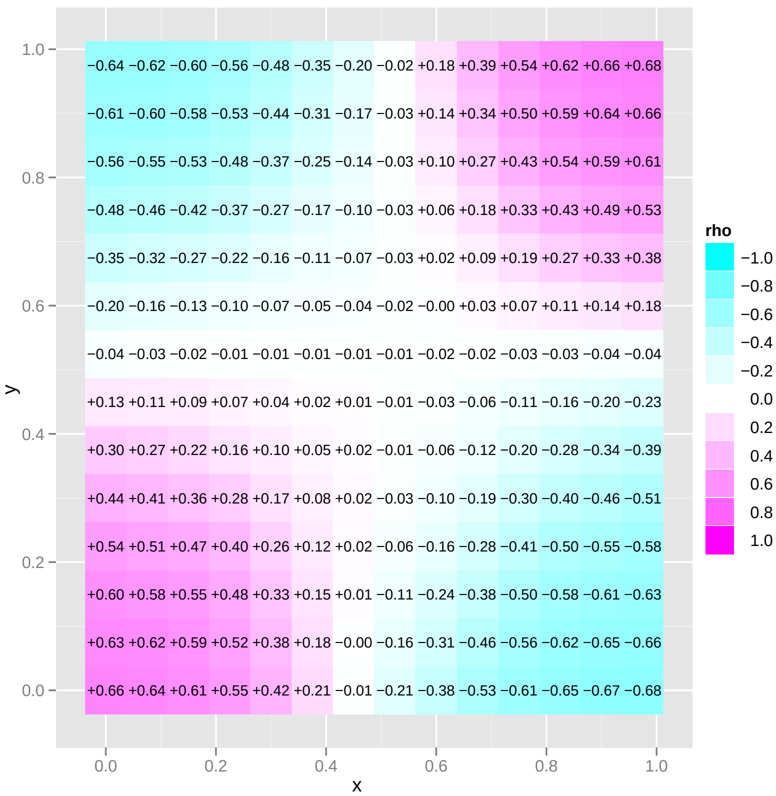

Figure A.4: Estimated LGC plot of t-copula with $\rho = 0$, $\nu = 4$ and uniform margins. Based on 5000 generated observations and bandwidth $b = 0.2$.

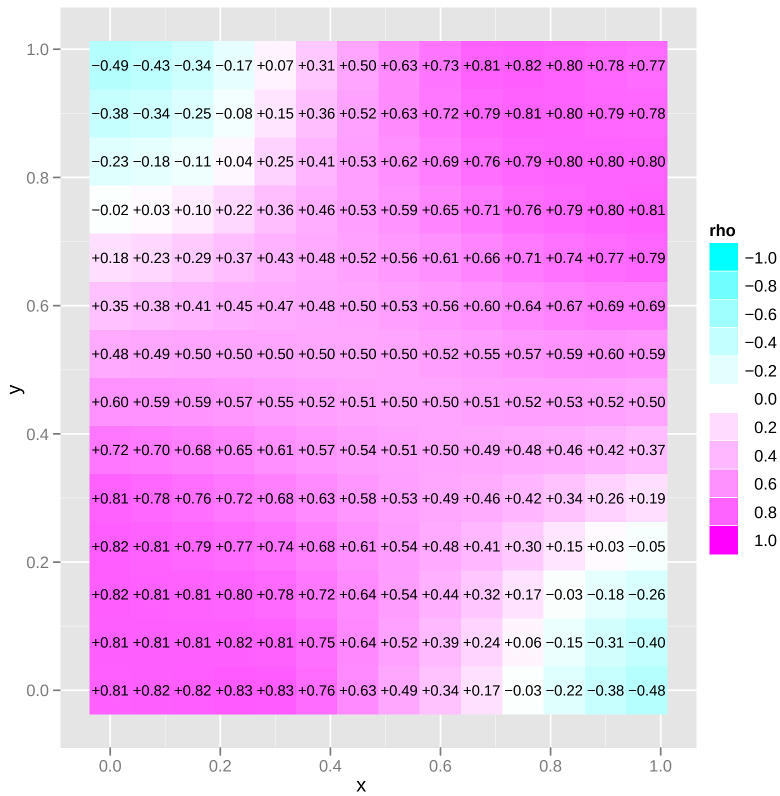

Figure A.5: Estimated LGC plot of t-copula with $\rho = 0.5$, $\nu = 4$ and uniform margins. Based on 5000 generated observations and bandwidth $b = 0.2$.

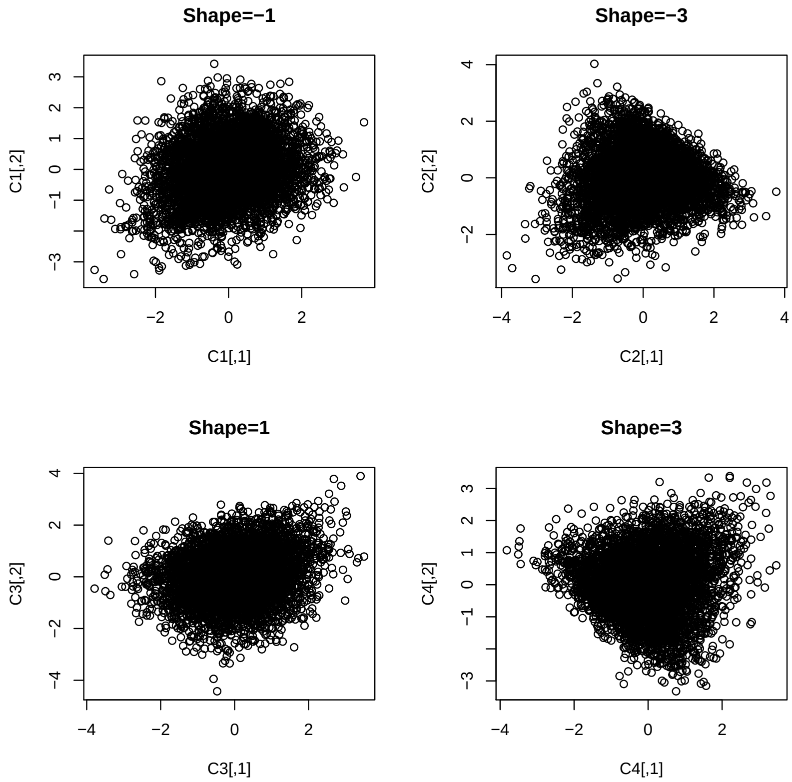

Figure A.6: 5000 generated observations from SN copula with $\rho = 0.5$ and normal margins.

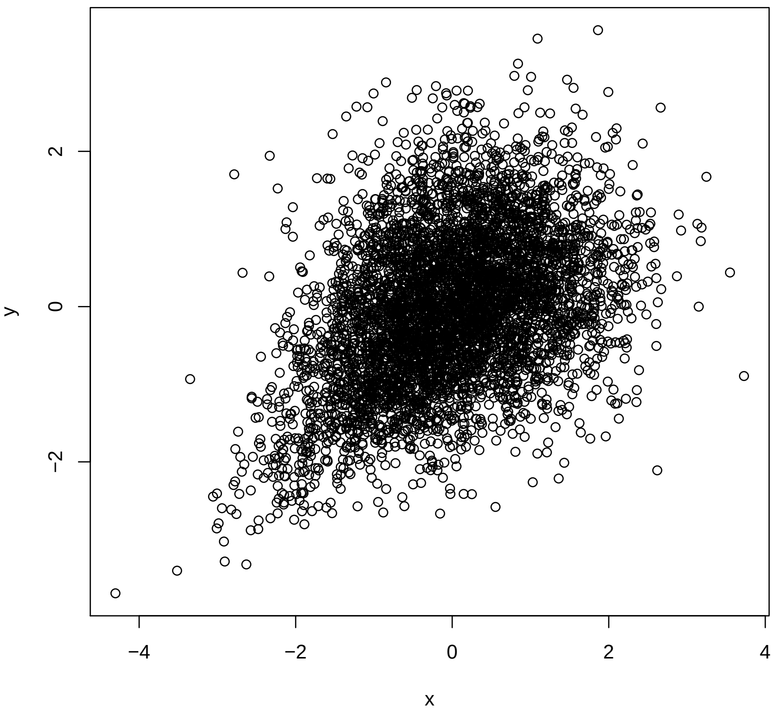

Figure A.7: 5000 generated observation from Clayton copula with $\tau = \frac{1}{4}$ and normal margins.

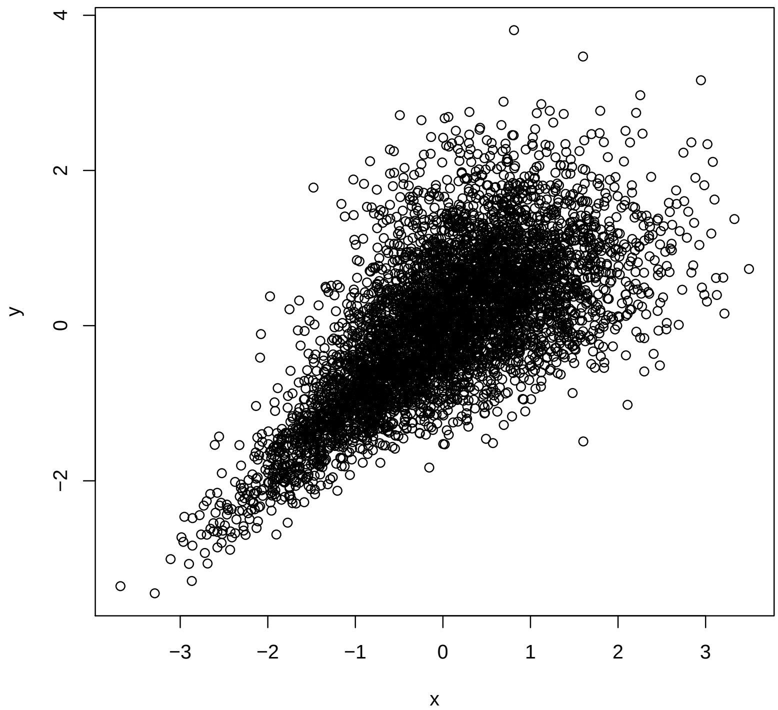

Figure A.8: 5000 generated observation from Clayton copula with $\tau = \frac{1}{2}$ and normal margins.

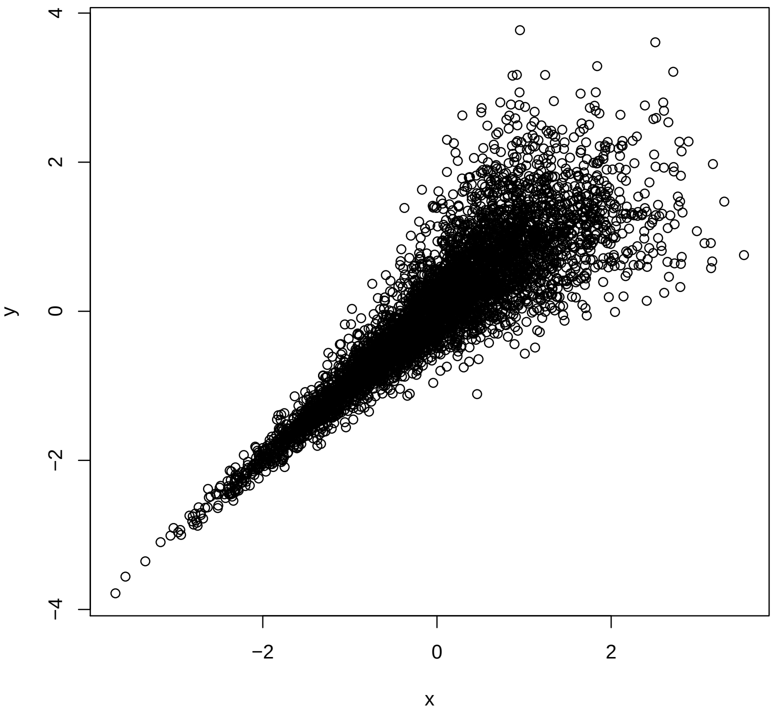

Figure A.9: 5000 generated observation from Clayton copula with $\tau = \frac{3}{4}$ and normal margins.

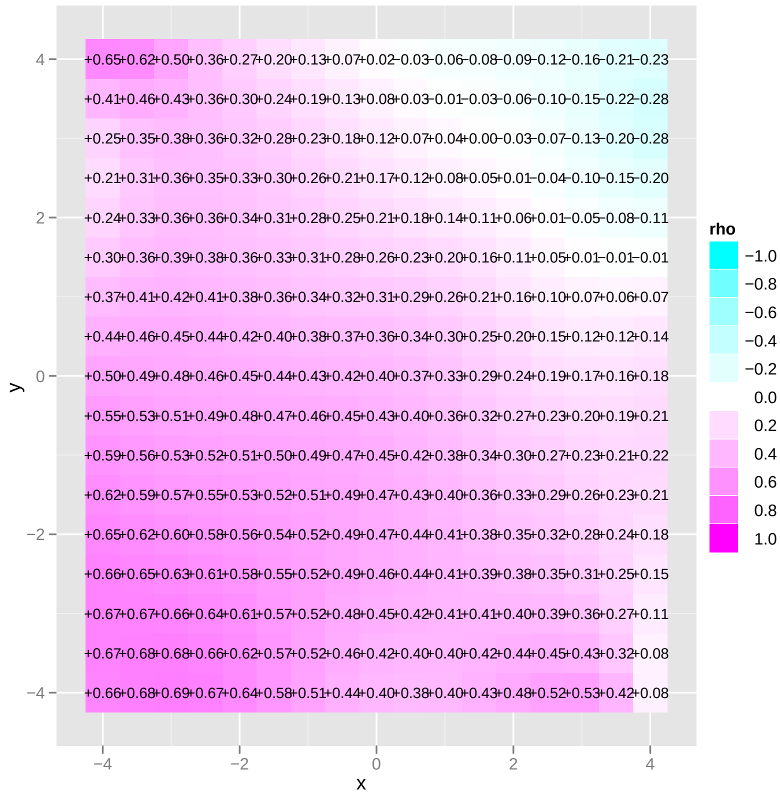

Figure A.10: Estimated LGC plot of a Clayton copula with $\tau = \frac{1}{4}$ and t distributed margins with 4 degrees of freedom. Based on 5000 generated observations and bandwidth b=1.2.

Figure A.11: Estimated LGC plot of a Clayton copula with $\tau = \frac{1}{4}$ and skewed t distributed margins with 4 degrees of freedom and shape parameter 1. Based on 5000 generated observations and bandwidth b=1.2.

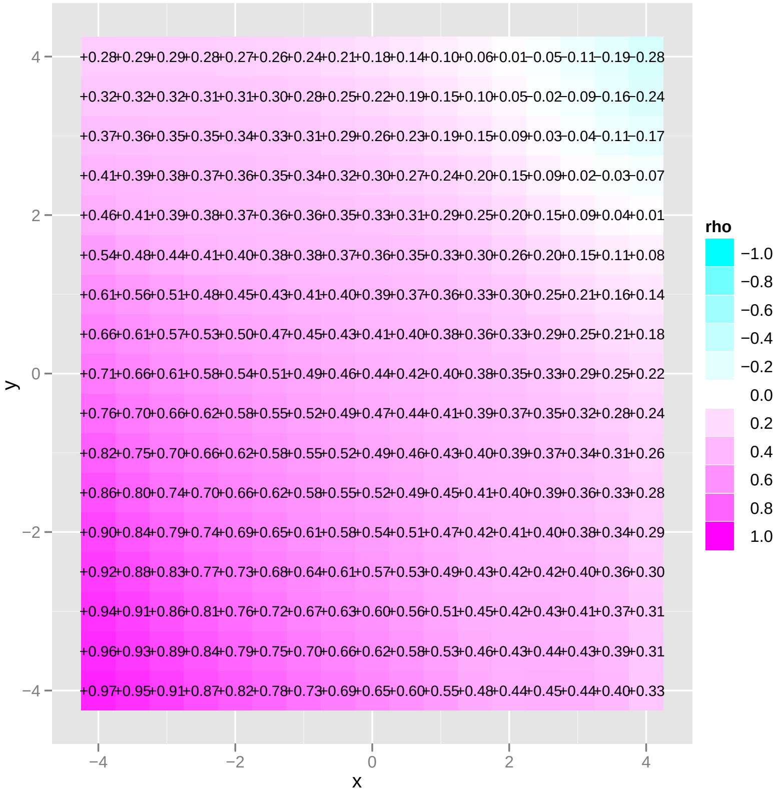

Figure A.12: Estimated LGC plot of a Clayton copula with $\tau = \frac{1}{4}$ and skewed t distributed margins with 4 degrees of freedom and shape parameter 5. Based on 5000 generated observations and bandwidth b=1.2.

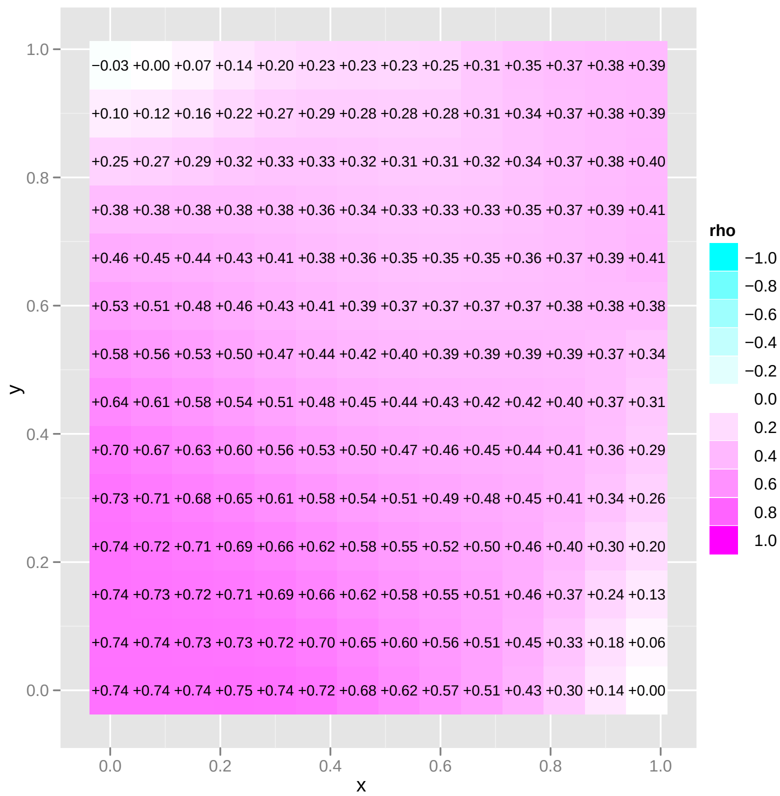

Figure A.13: Estimated LGC plot of a Clayton copula with $\tau = \frac{1}{4}$ and uniform margins. Based on 5000 generated observations and bandwidth b=0.25.

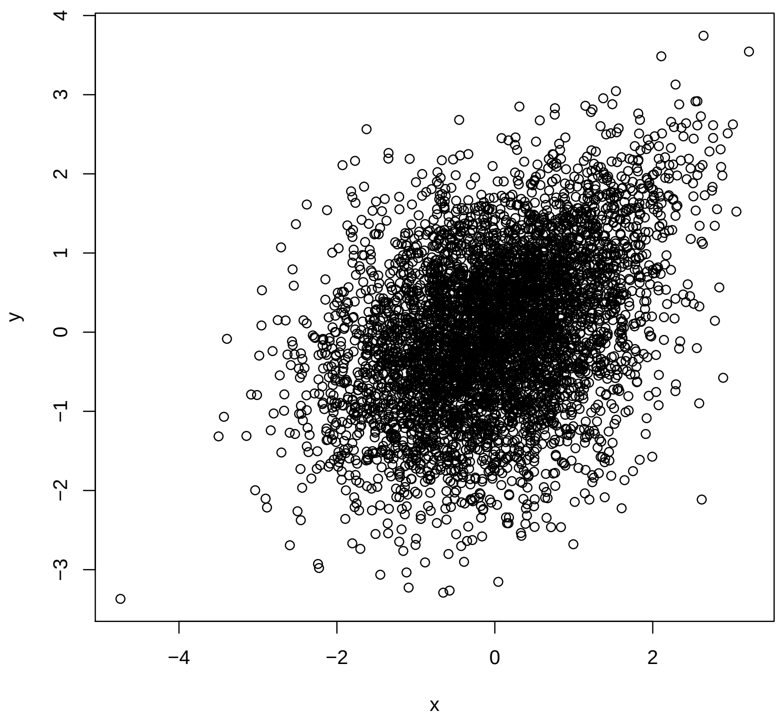

Figure A.14: 5000 generated observation from Gumbel copula with $\tau = \frac{1}{4}$ and normal margins.

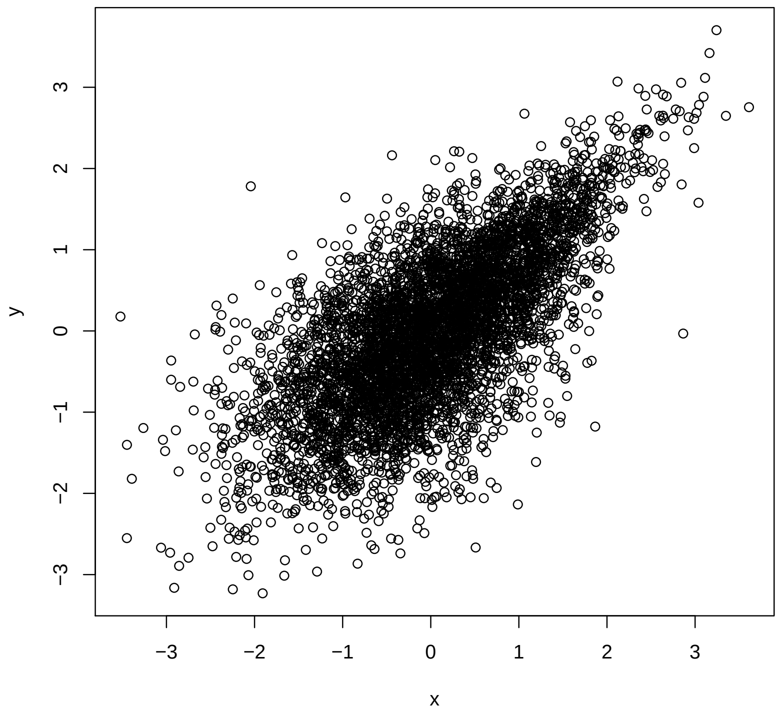

Figure A.15: 5000 generated observation from Gumbel copula with $\tau = \frac{1}{2}$ and normal margins.

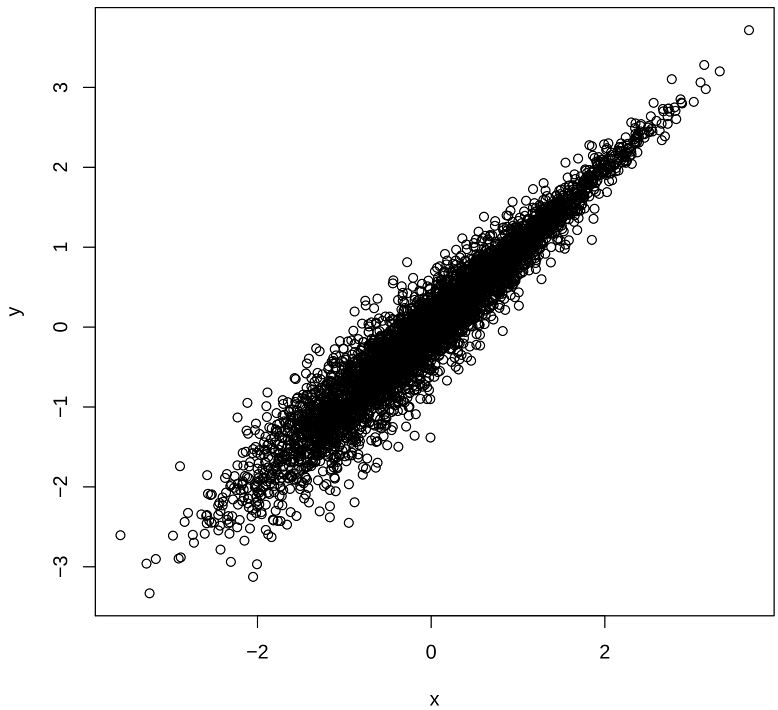

Figure A.16: 5000 generated observation from Gumbel copula with $\tau = \frac{3}{4}$ and normal margins.

Figure A.17: Estimated LGC plot of a Gumbel copula with $\tau = \frac{1}{4}$ and t distributed margins with 4 degrees of freedom. Based on 5000 generated observations and bandwidth b=1.2.

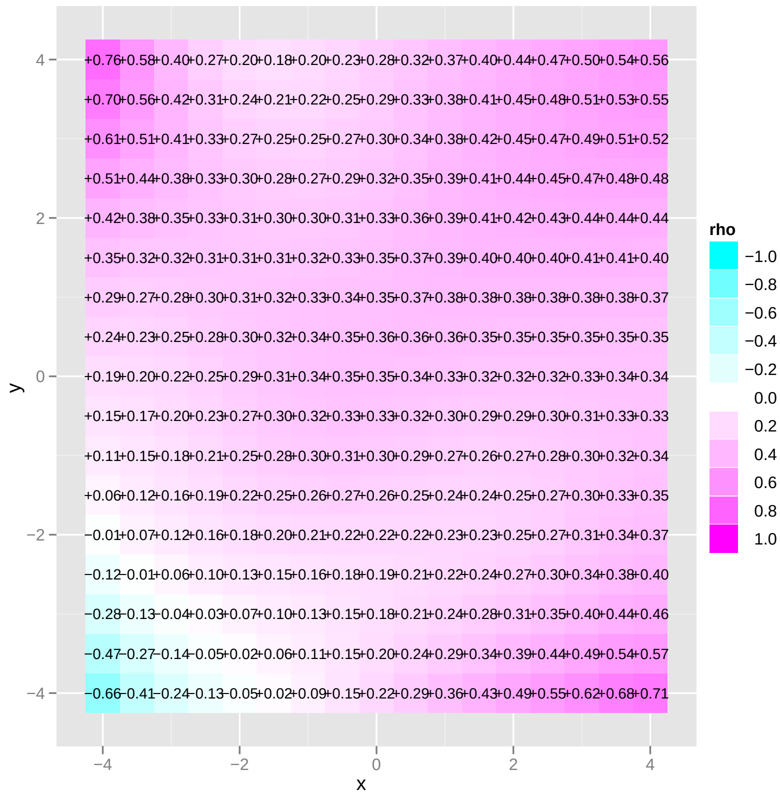

Figure A.18: Estimated LGC plot of a Gumbel copula with $\tau = \frac{1}{4}$ and skewed t distributed margins with 4 degrees of freedom and shape parameter 1. Based on 5000 generated observations and bandwidth b=1.2.

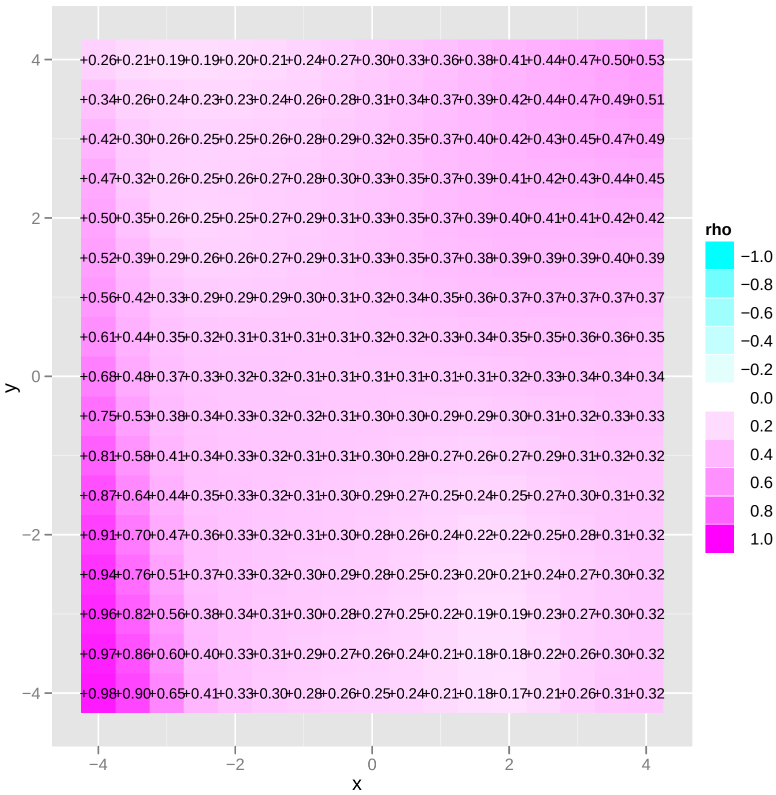

Figure A.19: Estimated LGC plot of a Gumbel copula with $\tau = \frac{1}{4}$ and skewed t distributed margins with 4 degrees of freedom and shape parameter 5. Based on 5000 generated observations and bandwidth b=1.2.

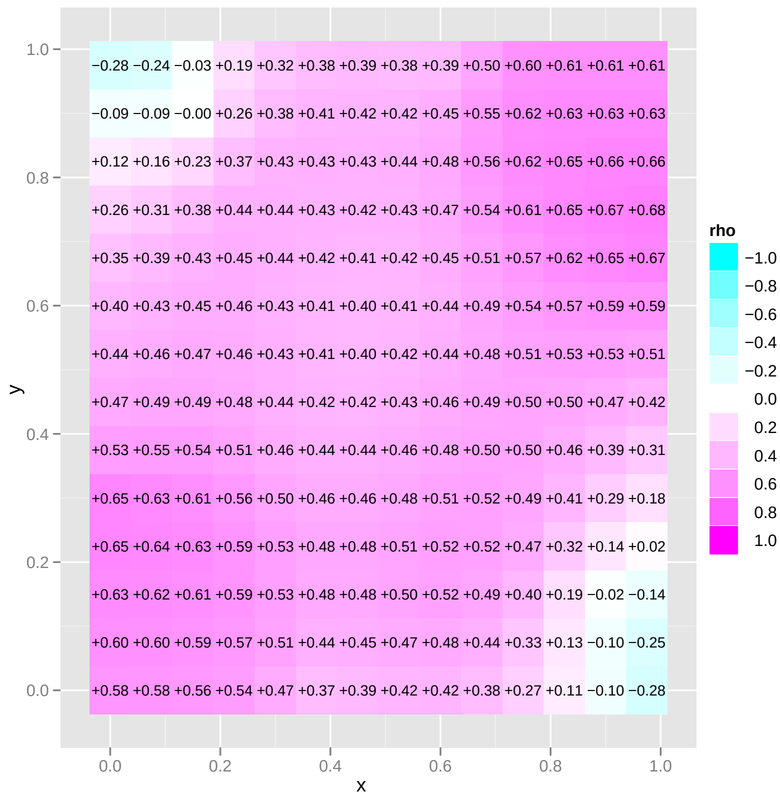

Figure A.20: Estimated LGC plot of a Gumbel copula with $\tau = \frac{1}{4}$ and uniform margins. Based on 5000 generated observations and bandwidth b=0.25.

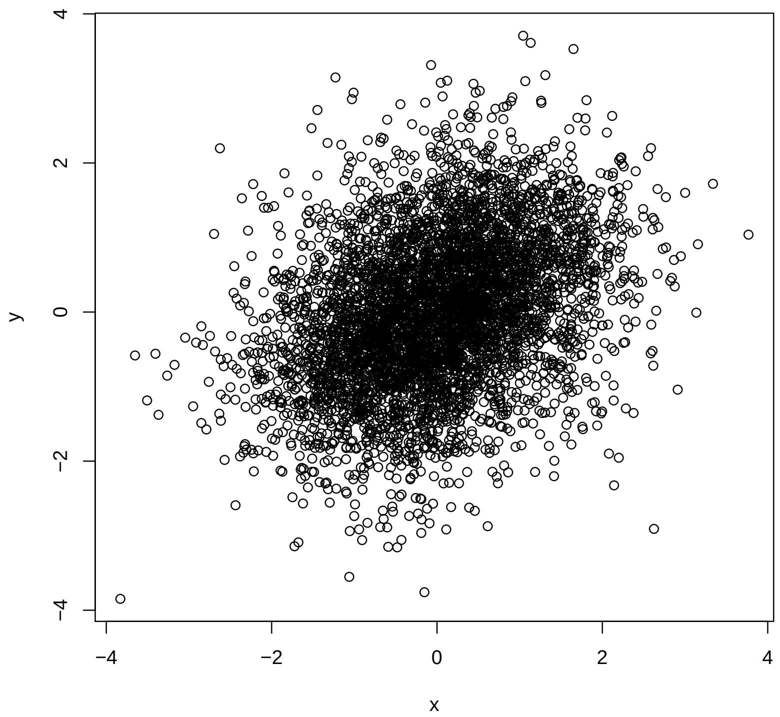

Figure A.21: 5000 generated observation from Frank copula with $\tau = \frac{1}{4}$ and normal margins.

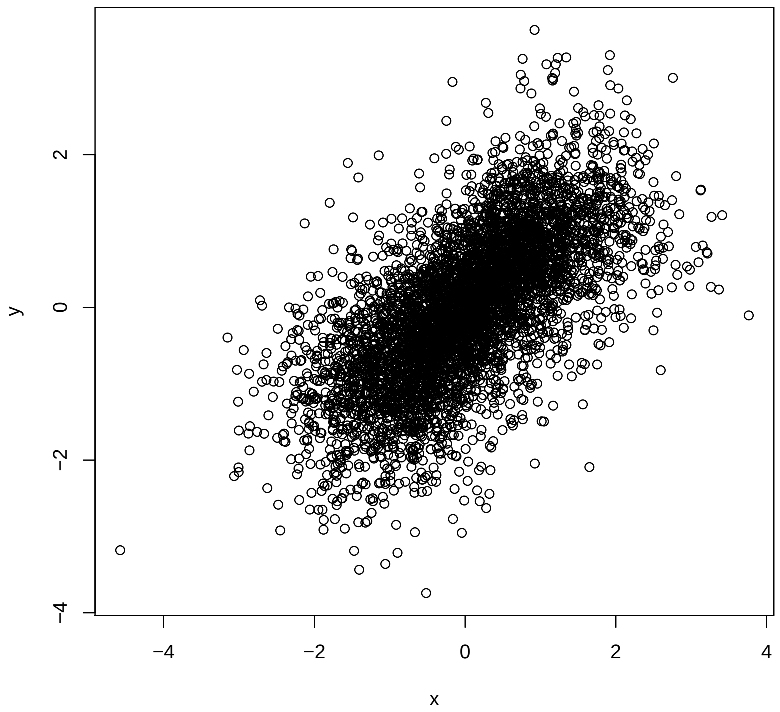

Figure A.22: 5000 generated observation from Frank copula with $\tau = \frac{1}{2}$ and normal margins.

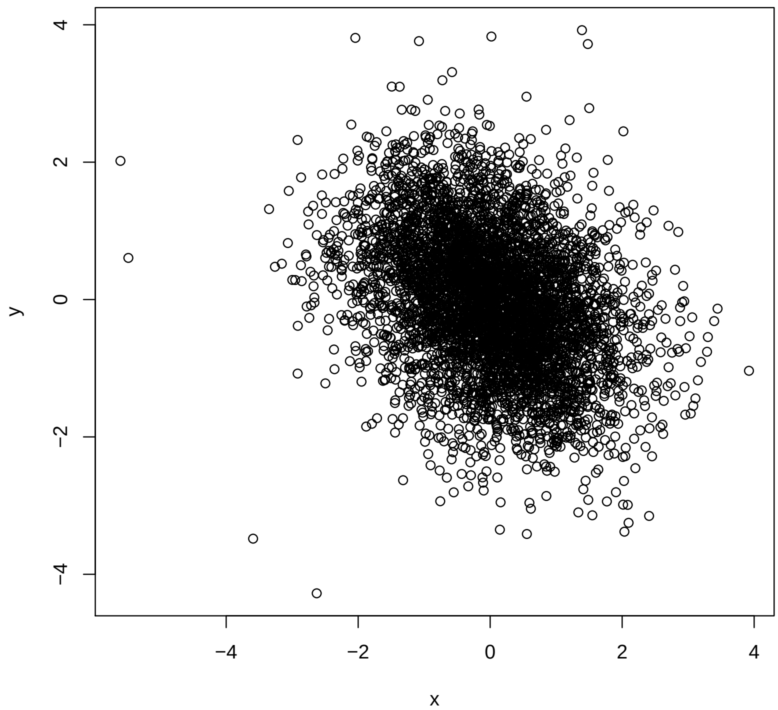

Figure A.23: 5000 generated observation from Frank copula with $\tau = -\frac{1}{4}$ and normal margins.

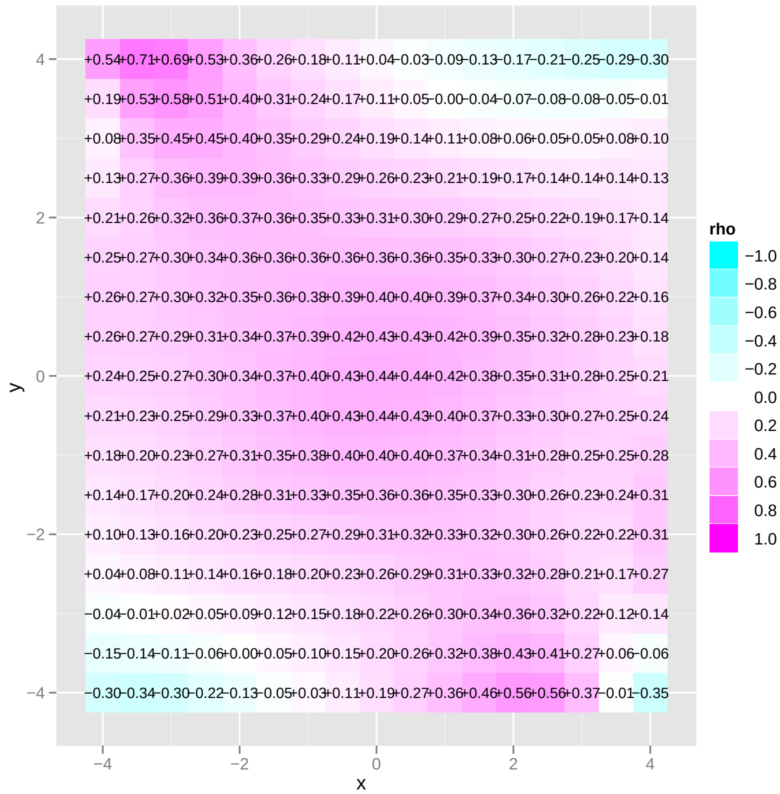

Figure A.24: Estimated LGC plot of a Frank copula with $\tau = \frac{1}{4}$ and t distributed margins with 4 degrees of freedom. Based on 5000 generated observations and bandwidth b=1.1.

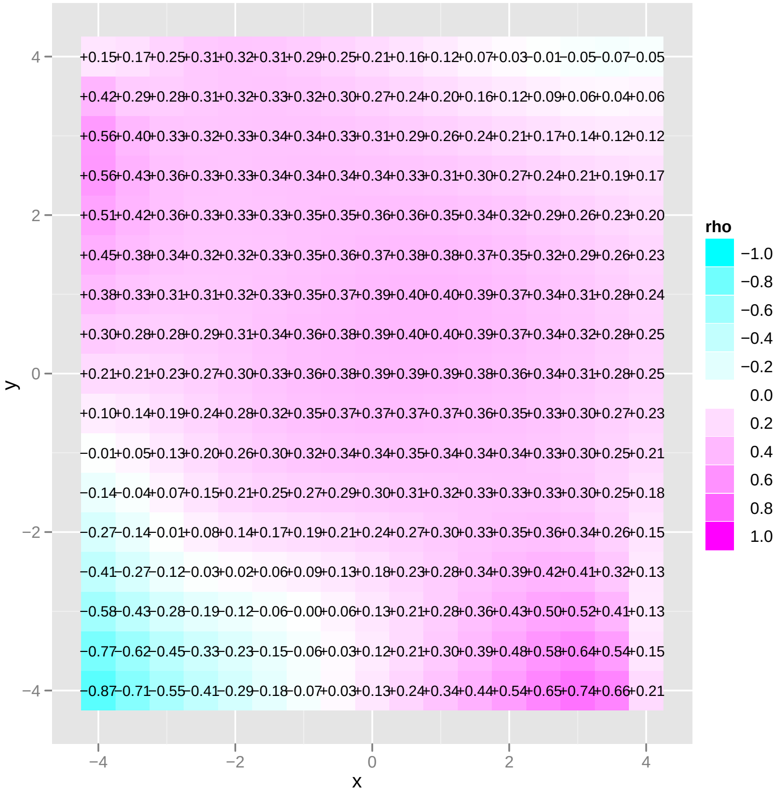

Figure A.25: Estimated LGC plot of a Frank copula with $\tau = \frac{1}{4}$ and skewed t distributed margins with 4 degrees of freedom and shape parameter 1. Based on 5000 generated observations and bandwidth b=1.2.

Figure A.26: Estimated LGC plot of a Frank copula with $\tau = \frac{1}{4}$ and skewed t distributed margins with 4 degrees of freedom and shape parameter 5. Based on 5000 generated observations and bandwidth b=1.5.

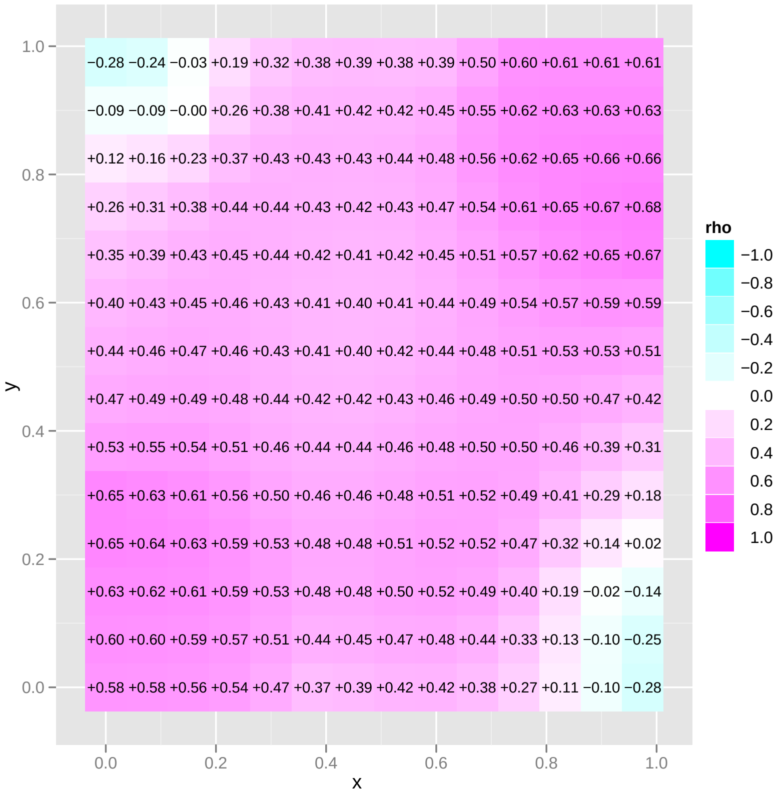

Figure A.27: Estimated LGC plot of a Frank copula with $\tau = \frac{1}{4}$ and uniform margins. Based on 5000 generated observations and bandwidth b=1.2.

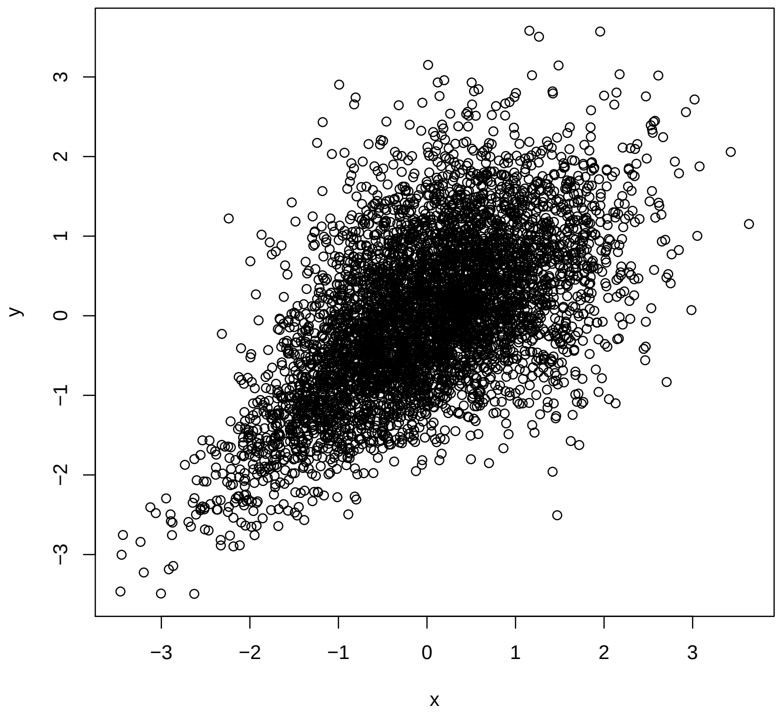

Figure A.28: 5000 generated observation from CG copula with $\tau = \frac{2}{5}$ and normal margins.

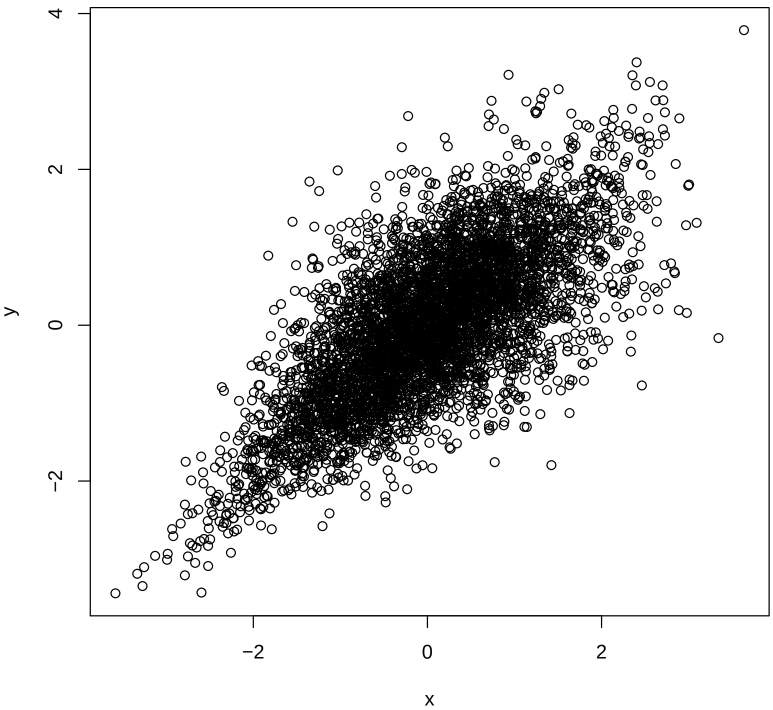

Figure A.29: 5000 generated observation from CG copula with $\tau = \frac{1}{2}$ and normal margins.

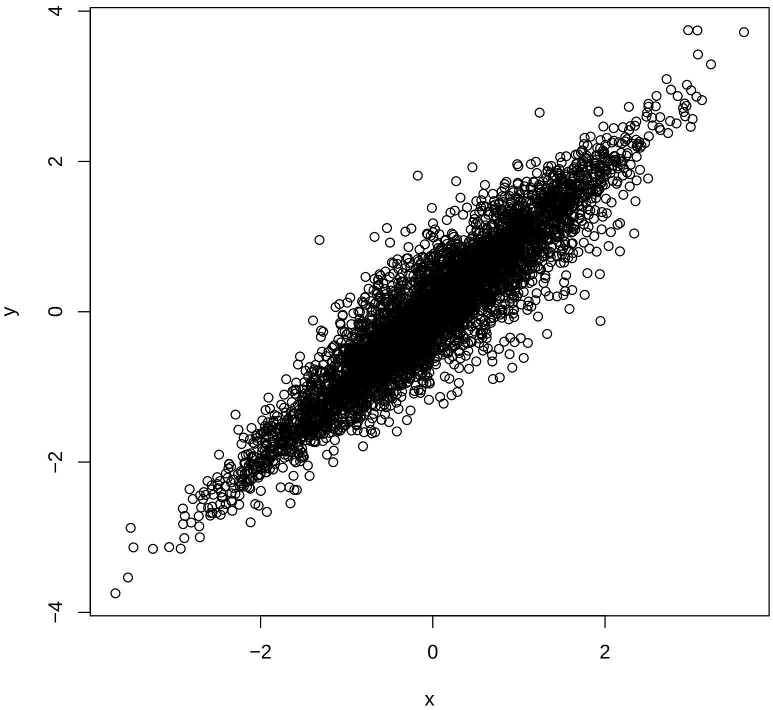

Figure A.30: 5000 generated observation from CG copula with $\tau = \frac{3}{4}$ and normal margins.

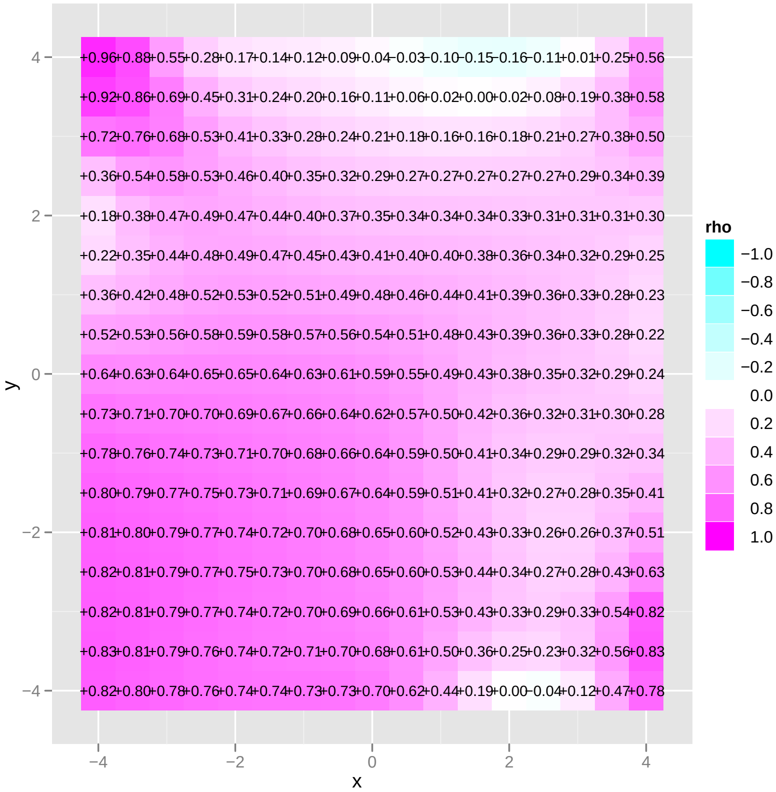

Figure A.31: Estimated LGC plot of a CG copula with $\tau = \frac{2}{5}$ and t distributed margins with 4 degrees of freedom. Based on 5000 generated observations and bandwidth b=1.

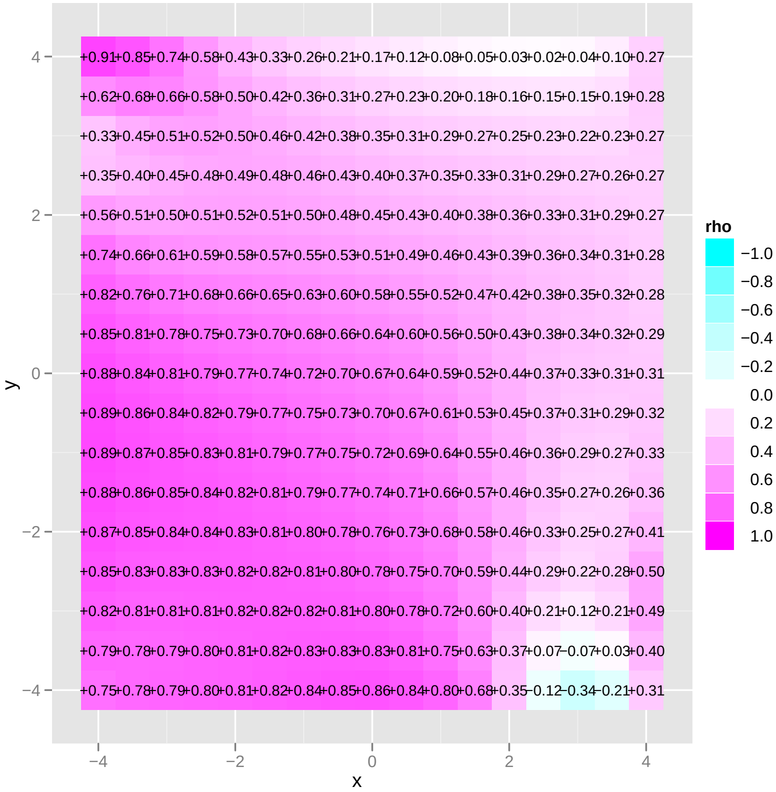

Figure A.32: Estimated LGC plot of a CG copula with $\tau = \frac{2}{5}$ and skewed t distributed margins with 4 degrees of freedom and shape parameter 1. Based on 5000 generated observations and bandwidth b=1.

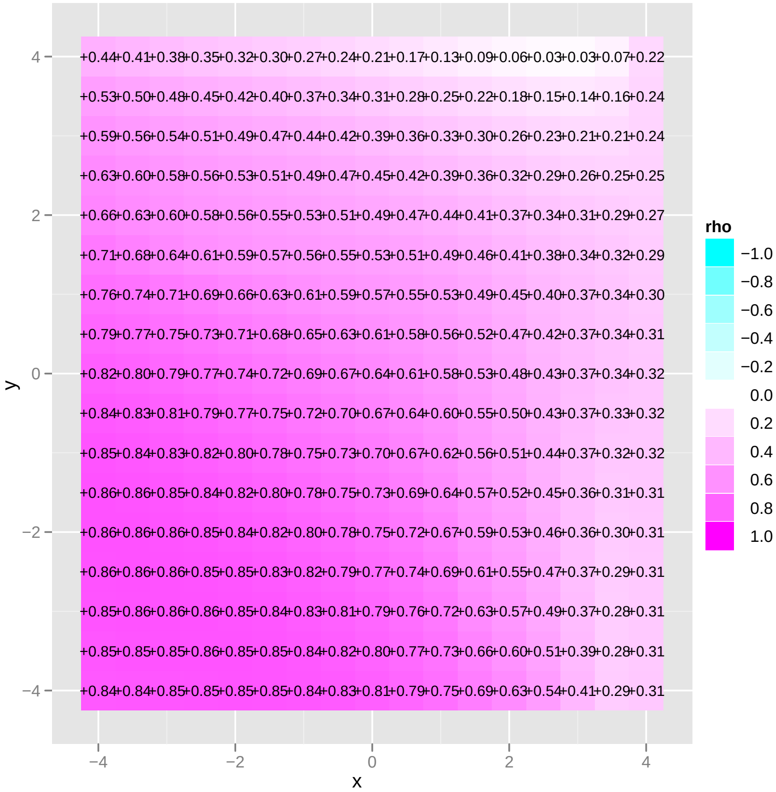

Figure A.33: Estimated LGC plot of a CG copula with $\tau = \frac{2}{5}$ and skewed t distributed margins with 4 degrees of freedom and shape parameter 5. Based on 5000 generated observations and bandwidth b=1.

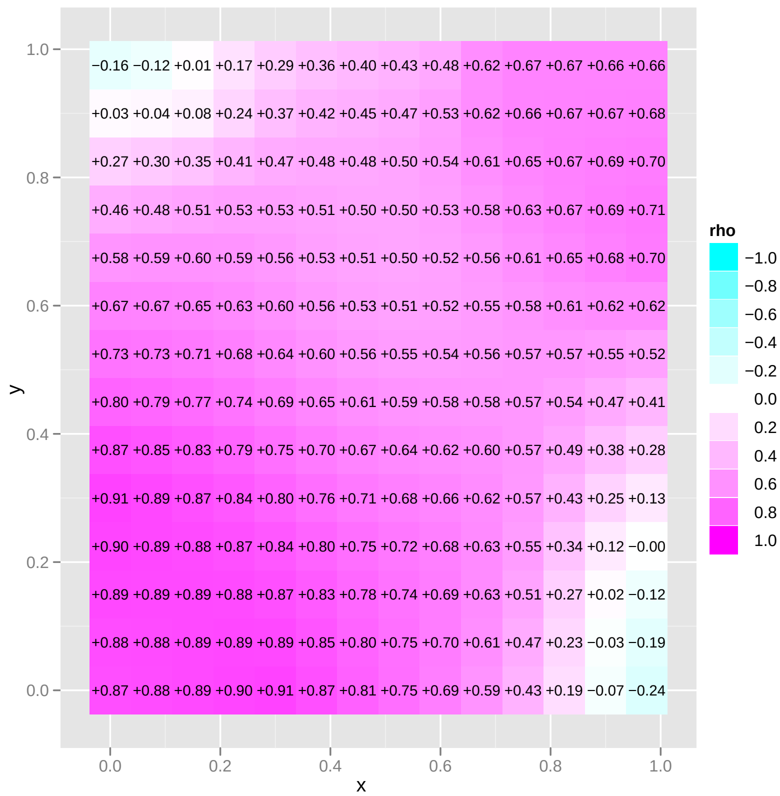

Figure A.34: Estimated LGC plot of a CG copula with $\tau = \frac{2}{5}$ and uniform margins. Based on 5000 generated observations and bandwidth b=0.2.