Dependency models with bivariate case study

Karolina Wojciechowska

Master's thesis
Delft University of Technology

July 26, 2007

The work was kindly supervised by:
prof.dr.ir. Jan M. van Noortwijk
Delft University of Technology and "HKV Consultants"
and
dr. Chris P.M Geerse
"HKV Consultants"

Acknowledgments

I would like to express my gratitude to my Supervisors for help and devoted time. Also, thanks go to dr. Robbert Fokkink as a member of the thesis committee. I would like to thank to "HKV Consultants" for the opportunity to do this project.
I would also like to thank to my parents and sister for love and support. Special thanks go to my friends Joanna and Wiktoria - thanks to whom time in Delft was a pleasure. I would like to express my appreciation to all my colleagues, especially Iwona and Maria, for always good words.

Contents

Chapter 1

Introduction

This work is motivated and inspired by the report: "Correlaties en meerdimensionale statistiek" (see [7]), written by Chris Geerse, in which several models for bivariate distributions are investigated. These bivariate distributions are needed to determine, e.g. failure of a dike or water barrier on the basis of a combination of random loads such as water level and wind speed. For the probabilistic modeling of random loads, we need to determine the marginal distributions as well as the associated dependence structure.

All considered models are based on the transformation of two random variables, with unknown joint distribution, into the so-called model space. The aim of this transformation is to create a bivariate distribution function with the desired marginal distributions and dependence structure.

The author considers bivariate copulas, which constitute distribution functions on the unit square with uniform margins. The copula is a function, which gathers margin-free dependency of random variables. The notion of copulas became known with the introduction of Sklar's theorem (1959). In [7] the following families of copulas are considered: the Clayton, Gumbel and Frank family. They belong to the Archimedean class, which has useful properties and realizes a variety of dependency forms. Therefore, this class is often used in applications.

In [7] also bivariate conditional models for distribution functions are considered. These models are derived as a product of some conditional distribution and the distribution of the conditioning variable. The constructions of these models are usually straightforward and can be explained by using pictures. Moreover, these models allow to derive the joint upper tail behavior of random variables.

This thesis basically follows the course of report [7]. We will consider some approaches in modeling bivariate distribution functions. We will investigate the Archimedean class of copulas, the mentioned conditional models and newly proposed Rotation Model. In addition, it is our intention to verify whether we can easily judge the fit in the extreme region of the model space in the application study.

The second chapter describes the copula functions as possible models in the bivariate case. We will mainly focus on the description of the Archimedean class of copulas. We will provide an extensive statistical inference, which will be applied in a case study on water levels and wind speeds data.

The third chapter is devoted to the consideration of the three bivariate conditional models, namely the Constant Spread Model, the Variable Spread Model and the Constant Symmetric Spread Model. These models are very interesting and more details about them can be found in [7]. We will prove some tail properties of these models by using the notion of the tail dependence coefficients, and we will verify if the copulas that are related to these models belong to Archimedean class. In order to illustrate the behavior of these models, we will perform a case study on the same dataset as in Chapter 2.
The last chapter is based on a proposal of Chris Geerse. We will construct a new model in the bivariate and 3-variate case. This model originates in the rotation of the coordinate system. Therefore, this model is called Rotation Model. We will prove some useful properties of this model including the exponentiality of the margins and the tail dependence coefficients. At the end of this chapter, a case study will be conducted. It will reveal a serious drawback of this model.
The conclusions from each chapter will be summarized in the final chapter.

Chapter 2

Measures of dependence and copula functions

This chapter is devoted to the introduction of the copula functions. These functions are used to create joint distributions, based on Sklar's theorem.
Copulas enable us to extract the dependence structure from the joint distribution function, and to separate the dependence structure from the marginal distribution functions. This is very helpful, because it gives us a natural way of allowing for dependency that is free of the margins influence.
We will start with the description of the bivariate measures of dependence, which are somewhat related to the bivariate copula notion, since a copula is a distribution used to describe the margin-free dependency. Then, we will discuss the copula functions. We will mainly focus on the Archimedean class of copulas and present three bivariate one-parametric families that belong to this class: the Clayton, Gumbel and Frank family. All theoretical results are based on the books [13] and [11]. At the end of this chapter, we will also provide methods for statistical inference, which will allow us to conduct a case study.

2.1 Measures of dependence

In this section, we introduce the definitions of the most common bivariate measures of dependence. These measures are used in the following theory of copulas. We will start with the definition of independent random variables. Then, we will define the product moment correlation, the rank correlation and Kendall's tau.

Definition 1 *Random variables X and Y are independent if for any Borel sets A and B holds:*

$$P\{X \in A, Y \in B\} = P\{X \in A\}P\{Y \in B\}$$

However, if the random variables are not independent, then we are interested how dependent they are. Hence, we introduce the following measures of dependence between two random variables.

Definition 2 *The product moment correlation of random variables X and Y with finite expectations $E(X), E(Y)$ and finite variances σ_X^2, σ_Y^2, is:*

$$\rho(X,Y) = \frac{E(XY) - E(X)E(Y)}{\sigma_X \sigma_Y}$$

If we are given n pairs of samples (X_i, Y_i) from the random vector (X, Y), we can calculate the so-called population product moment correlation as follows:

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

where $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ and $\overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$. For details we refer to [11].

Definition 3 *The rank correlation of random variables X and Y with cumulative distribution functions F and G respectively, is given as follows:*

$$\rho_r(X,Y) = \rho(F(X), G(Y))$$

Let us assume that we are given n pairs of samples (X_i, Y_i) from the random vector (X, Y). In order to compute the rank correlation, we replace the value of each $X_i, i = 1, ..., n$ by its rank. We apply the same procedure for each $Y_i, i = 1, ..., n$. Then, the population version of the rank correlation is:

$$\rho_r(X,Y) = \frac{\sum_{i=1}^{n}(R_i - \overline{R})(S_i - \overline{S})}{\sqrt{\sum_{i=1}^{n}(R_i - \overline{R})^2}\sqrt{\sum_{i=1}^{n}(S_i - \overline{S})^2}}$$

where R_i is the rank of $X_i, i = 1, ..., n$, S_i is the rank of $Y_i, i = 1, ..., n$, $\overline{R} = \frac{1}{n}\sum_{i=1}^{n} R_i$ and $\overline{S} = \frac{1}{n}\sum_{i=1}^{n} S_i$. For details, we refer to [11].

Definition 4 *Let (X_1, Y_1) and (X_2, Y_2) be two independent pairs of random variables each with joint distribution function F then Kendall's tau is defined as:*

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

For details we refer to [11].

In order to calculate Kendall's tau from the data (X_i, Y_i), for $i = 1, ..., n$, we introduce the notion of *concordance* and *discordance* and follow the reasoning presented in [13]. We say that pairs (X_i, Y_i) and (X_j, Y_j) are concordant if $X_i < X_j$ and $Y_i < Y_j$, or if $X_i > X_j$ and $Y_i > Y_j$. We say that (X_i, Y_i) and (X_j, Y_j) are discordant if $X_i < X_j$ and $Y_i > Y_j$, or if $X_i > X_j$ and $Y_i < Y_j$. Equivalently, we can say that (X_i, X_j) and (Y_i, Y_j) are concordant if $(X_i - X_j)(Y_i - Y_j) > 0$ and discordant if $(X_i - X_j)(Y_i - Y_j) < 0$. Observe that there are ${}_nC_2$ distinct pairs (X_i, Y_i) and (X_j, Y_j) in the sample and each pair is either concordant or discordant, where ${}_nC_2$ is defined as follows:

$${}_nC_2 := \left(\begin{array}{c} n \\ 2 \end{array}\right) = \frac{n!}{2!(n-2)!} = \frac{(n-1)n}{2}$$

Let c stand for the number of concordant pairs and d for the number of discordant pairs. Then, Kendall's tau for the sample is given by:

$$\tau_n = \frac{c-d}{c+d} = \frac{c-d}{{}_nC_2} \tag{2.1}$$

2.2 An introduction to copulas

A d-copula (copula) is a d-variate distribution function defined on the unit product $[0,1]^d$ with uniform marginal distributions. The following theorem is central in the theory of copulas and elucidates the role that copulas play in the relationship between multivariate distribution functions and their univariate margins.

Theorem 1 ***Sklar's theorem*** *Let H be a d-dimensional distribution function with margins $F_1, F_2, ..., F_d$. Then, there exists a d-copula C such that for all $x = (x_1, x_2..., x_d) \in \Re^d$:*

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

If $F_1, F_2, ..., F_d$ are continuous, then C is unique; otherwise, C is uniquely determined on $RanF_1 \times RanF_2... \times RanF_d$. Conversely, if C is a d-copula and $F_1, F_2, ..., F_d$ are distribution functions, then the function H, defined by (2.2), is a joint distribution function with margins $F_1, F_2, ..., F_d$.

The proof can be found in [13].
The name "copula" was chosen to describe the way in which a copula "couples" a multivariate distribution function to its univariate margins. Note that continuous $F_1, ..., F_d$ give an exhaustive description of variables $X_1, ..., X_d$ considered separately, whereas the dependence between these variables is completely characterized by C.
If the d-copula C is absolutely continuous then its density function c is given in the standard form:

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

Moreover, we note that the density function h - if it exists - of the d-dimensional distribution H can be expressed in terms of the density c. Indeed:

$$h(x_1, x_2, ..., x_d) = c(F_1(x_1), F_2(x_2), ..., F_d(x_d)) f_1(x_1) f_2(x_2) ... f_d(x_d)$$

where f_i is the density that corresponds to F_i, for $i = 1, ..., d$.
The notion of copula function entails some useful properties. One of them is the representation of Kendall's tau in terms of a bivariate copula C. This is enclosed in the following proposition:

Proposition 1 *If X and Y are continuous random variables joined by a copula C then Kendall's tau is given by the following formula:*

$$\tau = 4 \int_{[0,1]^2} C(s,t) dC(s,t) - 1$$

For the proof and further details we refer to [13].

2.3 Archimedean copulas

In this section, we introduce the bivariate Archimedean class of copulas. We will begin with the definitions, which give the foundation of this class. Then, we will proceed with the useful properties that arise. The theory below can be found in more details in [13].

Definition 5 *A function $\phi : [0,1] \to [0,\infty]$ is called a generator if it satisfies the following conditions:*

- $\phi(1) = 0$
- *function ϕ is convex*
- *function ϕ is strictly decreasing*

Definition 6 *The pseudo-inverse of the generator ϕ is given by the following formula:*

$$\phi^{[-1]}(t) = \begin{cases} \phi^{-1}(t) & 0 \leq t \leq \phi(0) \\ 0 & \phi(0) \leq t \leq \infty \end{cases}$$

Definition 7 *A bivariate copula C is called Archimedean with generator ϕ if:*

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

Clearly, Archimedean copulas can be constructed from the last definition. We only need to find the functions that can serve as generators. Below, we present three often used one-parametric families of bivariate Archimedean copulas with their generators:

Copula	$C(u,v)$	$\phi(t)$
Clayton	$\max\{(u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}, 0\}$	$\frac{1}{\theta}(t^{-\theta} - 1)$, $\theta \in [-1,\infty)\backslash\{0\}$
Gumbel	$\exp\{-[(-\ln u)^{\theta} + (-\ln v)^{\theta}]^{1/\theta}\}$	$(-\ln t)^{\theta}, \theta \in [1,\infty)$
Frank	$-\frac{1}{\theta}\ln\left(1 + \frac{(e^{-\theta u}-1)(e^{-\theta v}-1)}{e^{-\theta}-1}\right)$	$-\ln\frac{e^{-\theta t}-1}{e^{-\theta}-1}, \theta \in (-\infty,\infty)\backslash\{0\}$

For the bivariate Archimedean class of copulas the following proposition arises:

Proposition 2 *Let us assume that C is a absolutely continuous bivariate Archimedean copula, with generator ϕ, such that $\phi''(t)$ exists and $\phi''(t) > 0$ for $t \in (0,1)$. Then the density function c of copula C is given by the following formula:*

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

Proof The definition of the derivative of the inverse function, the definition of the Archimedean copula and the absolute continuity of a copula C yield:

$$\begin{aligned}
c(u,v) &= \frac{\partial}{\partial v}\frac{\partial}{\partial u}C(u,v) = \frac{\partial}{\partial v}\left(\frac{\partial}{\partial u}\phi^{[-1]}(\phi(u)+\phi(v))\right) \\
&= \frac{\partial}{\partial v}\left(\frac{\phi'(u)}{\phi'(\phi^{[-1]}(\phi(u)+\phi(v)))}\right) = \frac{\partial}{\partial v}\left(\frac{\phi'(u)}{\phi'(C(u,v))}\right) \\
&= \phi'(u)\frac{\partial}{\partial v}\left(\frac{1}{\phi'(C(u,v))}\right) = \phi'(u)\left(-\frac{\phi''(C(u,v))}{(\phi'(C(u,v)))^2}\frac{\partial}{\partial v}C(u,v)\right) \\
&= \phi'(u)\left(-\frac{\phi''(C(u,v))}{(\phi'(C(u,v)))^2}\frac{\partial}{\partial v}(\phi^{[-1]}(\phi(u)+\phi(v)))\right) \\
&= \phi'(u)\left(-\frac{\phi''(C(u,v))}{(\phi'(C(u,v)))^2}\frac{\phi'(v)}{\phi'(\phi^{[-1]}(\phi(u)+\phi(v)))}\right) \\
&= -\frac{\phi''(C(u,v))\phi'(u)\phi'(v)}{(\phi'(C(u,v)))^3}
\end{aligned}$$

Hence, the proof is accomplished.

Under the assumptions of Sklar's theorem let us define function $K(t)$:

$K(t) := P\{H(X_1, X_2, ..., X_d) \leq t\} = P\{C(F_1(X_1), F_2(X_2)..., F_n(X_d)) \leq t\}$ for $t \in [0,1]$

Function K states for the distribution function of the random variable $C(F_1(X_1), ..., F_d(X_d))$. Moreover, we observe that $F_i(X_i), i = 1, ..., d$, are standard uniformly distributed. Indeed:

$$P\{F_i(X_i) \leq x_i\} = P\{X_i \leq F_i^{-1}(x_i)\} = F_i(F_i^{-1}(x_i)) = x_i \quad \text{for } i = 1, ..., n$$

A useful property arises for the bivariate Archimedean class of copulas:

Proposition 3 *Let U and V be standard uniform random variables whose joint distribution function is the Archimedean copula C generated by ϕ. Then, the distribution function of the random variable $C(U,V)$ is given by:*

$$K(t) = P\{C(U,V) \leq t\} = t - \frac{\phi(t)}{\phi'(t^+)} \quad \textit{for } t \in [0,1] \tag{2.3}$$

For the proof, we refer to [13]. $\phi'(t^+)$ denotes the right-side derivative of ϕ at t. If $\phi'(t)$ exists then $\phi'(t^+) = \phi'(t)$.
Applying the above proposition to the bivariate Clayton, Gumbel and Frank copula results in the following table:

Copula	$K(t)$
Clayton	$\frac{t(\theta+1-t^\theta)}{\theta}$
Gumbel	$\frac{t(\theta-\ln t)}{\theta}$
Frank	$\frac{\theta t+(1-e^{\theta t})\ln\left\{\frac{e^{-\theta t}-1}{e^{-\theta}-1}\right\}}{\theta}$

One of the useful properties of the bivariate Archimedean class of copulas is that the corresponding Kendall's tau can be expressed in terms of the generator ϕ. The theory below can be found in [11]. Additionally, we sustain the assumptions of the Proposition 2.

Proposition 4 *If two random variables are joined by the Archimedean copula with generating function ϕ, then Kendall's tau is given by the following formula:*

$$\tau = 1 + 4\int_0^1 \frac{\phi(t)}{\phi'(t)}dt \tag{2.4}$$

Proof Proposition 1 yields:

$$\tau = 4\int_{[0,1]^2} C(s,t)dC(s,t) - 1 = 4\int_{[0,1]^2} C(s,t)c(s,t)dsdt - 1$$

We note that $C(u,v) = 0$ for all (u,v) such that $\phi(u)+\phi(v) = \phi(0)$. This fact and Proposition 2 entail:

$$\int_{[0,1]^2} C(s,t)c(s,t)dsdt = -\int_{\phi(u)+\phi(v)<\phi(0)} C(u,v)\frac{\phi''(C)\phi'(u)\phi'(v)}{[\phi'(C)]^3}dudv$$

We impose the following transformation:

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

with Jacobian $\frac{\partial(s,t)}{\partial(u,v)} = \frac{\phi'(u)}{\phi'(C)}$ and $s \le t$, we also recall that $\phi(1) = 0$, then we obtain:

$$\begin{aligned} -\int_0^1\int_s^1 s\frac{\phi''(s)\phi'(u)\phi'(t)\phi'(s)}{[\phi'(s)]^3}dtds &= -\int_0^1\int_s^1 s\frac{\phi''(s)\phi'(t)}{[\phi'(s)]^2}dtds \\ &= -\int_0^1 s\frac{\phi''(s)\phi'(t)}{[\phi'(s)]^2}\int_s^1 \phi'(t)dtds = \int_0^1 s\frac{\phi''(s)\phi(s)}{[\phi'(s)]^2}ds \end{aligned}$$

Partial integration gives:

$$\left| \begin{array}{cc} w = s\phi(s) & z' = \frac{\phi''(s)}{[\phi'(s)]^2} \\ w' = \phi(s) + s\phi'(s) & z = -\frac{1}{\phi'(s)} \end{array} \right|$$

Hence, we obtain:

$$\begin{aligned} -\int_0^1\int_s^1 s\frac{\phi''(s)\phi'(u)\phi'(t)\phi'(s)}{[\phi'(s)]^3\phi'(u)}dtds &= \left(-s\frac{\phi(s)}{\phi'(s)}\right)_0^1 + \int_0^1\left(\frac{\phi(s)}{\phi'(s)} + s\right)ds \\ &= \int_0^1 \frac{\phi(s)}{\phi'(s)}ds + \frac{1}{2} \end{aligned}$$

Thus, we conclude that $\tau = 1 + 4\int_0^1 \frac{\phi(t)}{\phi'(t)}dt$ and the proof is complete.

The more general case of the above proposition is treated in [13].
The following table presents the values of Kendall's tau for three bivariate Archimedean copulas:

Copula	Kendall's tau
Clayton	$\frac{\theta}{\theta+2}$
Gumbel	$\frac{\theta-1}{\theta}$
Frank	$1+4\int_0^1 \frac{e^{\theta t}-1}{\theta}\ln\left(\frac{e^{-\theta t}-1}{e^{-\theta}-1}\right)dt$

2.4 Estimation

Suppose that some parametric family of copulas is being considered as a model for the dependence structure between some random variables. Given a random sample, we will tackle the problem of estimation of the unknown parameters that arise from the parametric family of copulas but often also from the univariate margins.

2.4.1 Maximum Likelihood Inference

Consider random variables $X_1, X_2, ..., X_d$ joined by the d-dimensional distribution H, with marginal distributions $F_1(x_1;\alpha_1), F_2(x_2;\alpha_2), ..., F_d(x_d;\alpha_d)$. Then, according to Sklar's theorem, the function H can be expressed in terms of the d-copula C:

$$H(x_1, x_2, ..., x_d; \alpha_1, \alpha_2, ..., \alpha_d, \theta) = C(F_1(x_1;\alpha_1), F_2(x_2;\alpha_2), ..., F_d(x_d;\alpha_d); \theta)$$

where α_i, for $i = 1, ..., d$ and θ constitute the vectors of the marginal and copula parameters, respectively. For simplicity, we will call them "parameters".
The d-dimensional density function h of the d-dimensional distribution H is given as follows:

$$h(x_1, x_2, ..., x_d; \alpha_1, \alpha_2, ..., \alpha_d, \theta) = c(F_1(x_1;\alpha_1), F_2(x_2;\alpha_2), ..., F_d(x_d;\alpha_d); \theta) \prod_{i=1}^{d} f_i(x_i;\alpha_i)$$

where $f_i(x_i;\alpha_i)$ is the corresponding density function of distribution $F_i(x_i;\alpha_i)$, for $i = 1, ..., d$ and c is the density function of the d-copula C. We assume that the above theory and notation is valid in the following paragraphs. Moreover, we note that this theory is mathematically less rigorous.

Maximum Likelihood Method

Let us assume that a sample of n d-variate observations $\{(X_1^i, X_2^i, ..., X_d^i)\}_{i=1}^{i=n}$ is at our disposal. Then, the estimates $\hat{\alpha_i}$,$\hat{\theta}$ of the parameters α_i, θ, for $i = 1, ..., d$ maximize the following likelihood function L:

$$L(\alpha_1, \alpha_2, ..., \alpha_d, \theta) = \prod_{i=1}^{n} \left(c(F_1(X_1^i;\alpha_1), F_2(X_2^i;\alpha_2), ..., F_d(X_d^i;\alpha_d); \theta) \prod_{j=1}^{d} f_j(X_j^i;\alpha_j) \right)$$

Since the logarithm is a continuous and strictly increasing function over the range of the likelihood, the values which maximize the likelihood will also maximize its logarithm. Since maximizing the logarithm usually requires simpler calculations, we define the following log-likelihood function:

$$l(\alpha_1,\alpha_2,...,\alpha_d,\theta)=\sum_{i=1}^{n}\ln c(F_1(X_1^i;\alpha_1),F_2(X_2^i;\alpha_2),...,F_d(X_d^i;\alpha_d);\theta)+\sum_{i=1}^{n}\ln\prod_{j=1}^{d}f_j(X_j^i;\alpha_j)$$

Then, the estimates $\hat{\alpha}_i, \hat{\theta}$ of the parameters α_i, θ, for $i = 1, ..., d$ maximize the function l.

The above method yields consistent estimators, for details we refer to [16].

Remark 1 *Observe that the Maximum Likelihood Method does not separate the marginal distributions from the dependency structure (copula). Hence, it does not incorporate the main feature of the copula function which is the separation of the marginal distributions. The next methods however use this property.*

Inference functions for margins

Let us assume that a sample of n d-variate observations $\{(X_1^i, X_2^i, ..., X_d^i)\}_{i=1}^{i=n}$ is at our disposal. The so-called "Inference functions for margins" is a method of estimation that consists of two steps. The first step requires the estimation of the parameters of the marginal distributions:

$$\hat{\alpha}_i = \arg\max_{\alpha_i}\sum_{j=1}^{n}\ln f_i(X_i^j;\alpha_i) \quad \text{for } i=1,...,d$$

Having estimators $\hat{\alpha}_i$, for $i = 1, ..., d$, we can proceed with the second step, namely the estimation of the copula parameter θ that is given as follows:

$$\hat{\theta}=\arg\max_{\theta}\sum_{j=1}^{n}\ln c(F_1(X_1^j;\hat{\alpha}_1),F_2(X_2^j;\hat{\alpha}_2),...,F_d(X_d^j;\hat{\alpha}_d);\theta)$$

More information about this method can be found in [10].

Remark 2 *The joint estimation of the unknown marginal and copula parameters - as is done by the Maximum Likelihood Method - can be very time consuming in applications, especially when we have to deal with the complicated form of the resulting log-likelihood function. The "Inference functions for margins" method is less time-consuming, because a single numerical optimization with many parameters requires more time compared with several numerical optimizations, each with fewer parameters.*
Observe that in case of the Maximum Likelihood Method the estimation of the dependence parameter θ is affected by the choice of the marginal distributions as well as the marginal parameters are influenced by the choice of the dependence structure (copula). In case of the "Inference functions for margins" method, the copula parameter will be margin dependent. Moreover, in both methods the marginal distribution models can entail the risk of being improperly chosen. In the next paragraph the Canonical Maximum Likelihood Method is proposed, where we estimate the copula parameter without specifying models for the marginal distributions.

Canonical Maximum Likelihood Method

Let us assume that a sample of n d-variate observations $\{(X_1^i, X_2^i, ..., X_d^i)\}_{i=1}^{i=n}$ is at our disposal. This method estimates the copula parameter without specifying precisely the marginal distributions. The estimator $\hat{\theta}$ of the copula parameter θ maximizes the following pseudo log-likelihood function:

$$l(\theta) = \sum_{i=1}^{n} \ln c(\hat{F}_{1,d}(X_1^i), \hat{F}_{2,d}(X_2^i), ..., \hat{F}_{d,d}(X_d^i); \theta)$$

where $\hat{F}_{i,d}$, for $i = 1, ..., d$ are the empirical marginal distributions and are given as follows:

$$\hat{F}_{j,d}(x) = \frac{1}{n} \sum_{i=1}^{n} 1(X_j^i \leq x) \quad \text{for } j = 1, ..., d$$

where $1(A)$ is the indicator of event A.
The resulting estimator of the copula parameter is consistent, for further details we refer to [8] and [9].

Remark 3 *We note that many authors use modified empirical distributions, given as follows:*

$$\hat{F}_{j,d}(x) = \frac{1}{n+1} \sum_{i=1}^{n} 1(X_j^i \leq x) \quad \textit{for } j = 1, ..., d$$

instead of the empirical marginal distributions in the Canonical Maximum Likelihood Method. This is due to the fact that using empirical margins can entail numerical problems in the application.
It is important to remark that the Canonical Maximum Likelihood Method indirectly depends on the empirical distributions. However, it does not mean that we have to assume that our marginal distributions are empirical. In fact, we can specify the marginal models and still perform this method. This is done in [8]. Generally, this method entails an estimation of the copula parameter that is margin-models-free.

2.4.2 Genest-Rivest approach

In this subsection, we consider a bivariate one-parametric Archimedean class of copulas depending on the real parameter θ. We introduce the Genest-Rivest method of estimation. This method does not depend on the choice of the marginal distributions. Moreover, it is especially appropriate for the bivariate Archimedean class of copulas, due to the fact that Kendall's tau for this class is straightforwardly determined by the generator ϕ. Moreover, the generator depends on our parameter of interest θ. Hence, if the estimation of Kendall's tau is known, we can obtain the value of the parameter θ by using the relation (2.4). Thus, the following estimation procedure is proposed:

1. Having the sample of n bivariate observations $(X_1, Y_1), \cdots, (X_n, Y_n)$, compute the estimations of Kendall's tau as follows:

$$\tau_n = \frac{2}{(n-1)n} \sum_{i=1}^{n} \sum_{j=1}^{i-1} Sign[(X_i - X_j)(Y_i - Y_j)]$$

where the function $Sign$ is given by:

$$Sign(x) = \begin{cases} 1 & x > 0 \\ 0 & x = 0 \\ -1 & x < 0 \end{cases}$$

Observe that the above estimator is equivalent to (2.1).

2. Having the estimation of Kendall's tau, compute the estimation of the parameter θ using the relation (2.4).

The above method yields the consistent estimator of the copula parameter, for further details we refer to [8] and [9].

Remark 4 *We note that the above procedure can be generalized for other classes of one-parametric bivariate copulas provided the real parameter θ can be represented in terms of $g(\tau)$, where τ denotes Kendall's tau and g is a smooth function.*

Remark 5 *It is important to remark that an adequate procedure can also be applied for the multivariate one-parametric Archimedean class of copulas. However, instead of Kendall's tau the so-called multivariate version of Kendall's measure of association has to be computed. This; however, will not be discussed. For further details, we refer to [9].*

2.5 Determination of the best copula

In a typical modeling task, the analyst usually has a choice between several different dependence structures for the data at hand. Suppose that several copula models were fitted by some arbitrary estimation method. It is then natural to ask which of the models provides the best fit to the observed dataset. Without loss of generality we can assume that the considered copulas depend on one parameter. In this section the methods of evaluation of goodness-of-fit for copulas will be discussed. Some of the methods are only appropriate for the bivariate case, whereas others can be extended to higher dimensions. To keep things simple, we will consider the bivariate case and we will indicate whether the considered method can be applied for the multivariate case.

2.5.1 Basic approach - Method 1

This approach is introduced in [8].
When we consider the bivariate data $(X_i, Y_i), i = 1, ..., n$, possibly the most natural way of checking the adequacy of a copula model would be to superimpose a scatter plot of the pairs $(F(X_i), G(Y_i)), i = 1, ..., n$ (where F and G denote the marginal distributions) on the sampled dataset of the same size from the copula $C(u, v; \hat{\theta})$ (where $\hat{\theta}$ is the estimation of the copula parameter θ on the basis of the dataset $(X_i, Y_i), i = 1, ..., n$, θ can also be a vector). Since the bivariate Archimedean copulas will be considered in the following case study, we present the sampling algorithm for this class:

1. Generate two independent uniform $(0, 1)$ variates u and t;
2. Set $w = \phi'^{(-1)}(\phi'(u)/t)$;
3. Set $v = \phi^{[-1]}(\phi(w) - \phi(u))$;
4. The desired pair is (u, v).

where ϕ is the generator and $\phi^{[-1]}$ denotes the pseudo-inverse of ϕ. For other algorithms and details,we refer to [13].
It is clear that this simple method is not applicable for dimensions higher than three. Even the 3-dimensional case would result in a cloud of points that would hamper the assessment of the fit.

2.5.2 Pseudo graphical diagnostic based on the distribution function of $C(V|U)$ - Method 2

This pseudo-graphical method is presented in [2] and [12].
In order to avoid ambiguity, we start with the mathematical results that lay the foundation to this method.

Proposition 5 *Let $(X_1, X_2, ..., X_d)$ be a d-variate vector with copula C and marginal distributions $F_1, ..., F_d$. Let $C_k(u_1, ..., u_k) = C(u_1, ..., u_k, 1, ..., 1), k = 1, ..., d$ denote the k-dimensional margin of C, with $C_1(u_1) = u_1$ and $C_d(u_1, ..., u_d) = C(u_1, ..., u_d)$. Then, the following relation holds:*

$$\begin{aligned} P\{X_k \leq x_k | X_1 = x_1, ..., X_{k-1} = x_{k-1}\} &= C_k(u_k|u_1, ..., u_{k-1}) \\ &= \frac{\partial^{k-1} C_k(u_1, ..., u_k)}{\partial u_1 ... \partial u_{k-1}} \bigg/ \frac{\partial^{k-1} C_{k-1}(u_1, ..., u_{k-1})}{\partial u_1 ... \partial u_{k-1}} \end{aligned}$$

where $u_i = F_i(x_i)$, for $i = 1, ..., k$.

Proof This proposition will be proved for the bivariate case ($d = 2$). Hence, assuming $U_1 = F_1(X_1)$, $U_2 = F_2(X_2)$, $u_1 = F_1(x_1)$ and $u_2 = F_2(x_2)$, we obtain the following calculations:

$$\begin{aligned} \frac{\partial}{\partial u_1} C_2(u_1, u_2) &= \lim_{\Delta u_1 \to 0} \frac{C(u_1 + \Delta u_1, u_2) - C(u_1, u_2)}{\Delta u_1} \\ &= \lim_{\Delta u_1 \to 0} \frac{P\{U_1 \leq u_1 + \Delta u_1, U_2 \leq u_2\} - P\{U_1 \leq u_1, U_2 \leq u_2\}}{P\{u_1 \leq U_1 \leq u_1 + \Delta u_1\}} \\ &= \lim_{\Delta u_1 \to 0} P\{U_2 \leq u_2 | u_1 \leq U_1 \leq u_1 + \Delta u_1\} \\ &= P\{U_2 \leq u_2 | U_1 = u_1\} \\ &= P\{F_2(X_2) \leq u_2 | F_1(X_1) = u_1\} \\ &= P\{X_2 \leq F_2^{-1}(u_2) | X_1 = F_1^{-1}(u_1)\} \\ &= P\{X_2 \leq x_2 | X_1 = x_1\} \end{aligned}$$

Since $\frac{\partial}{\partial u_1}C_1(u_1) = 1$, we obtain:

$$\frac{\frac{\partial}{\partial u_1}C_2(u_1,u_2)}{\frac{\partial}{\partial u_1}C_1(u_1)} = P\{X_2 \leq x_2|X_1 = x_1\}$$

Thus, the proof is complete.

The following table gives $\frac{\partial}{\partial u}C(u,v)$ for the bivariate Clayton, Gumbel and Frank copula:

Copula	$\frac{\partial}{\partial u}C(u,v)$
Clayton	$u^{-1-\theta}(u^{-\theta}+v^{-\theta}-1)^{-1-1/\theta}, u^{-\theta}+v^{-\theta}-1>0$
Gumbel	$((-\ln u)^{\theta}+(-\ln v)^{\theta})^{-1+1/\theta}(-\ln u)^{\theta-1}\exp\{-((-\ln u)^{\theta}+(-\ln v)^{\theta})^{1/\theta}\}/u$
Frank	$-(e^{-\theta v}-1)e^{-\theta u}/(-e^{-\theta}-e^{-\theta(u+v)}+e^{-\theta u}+e^{-\theta v})$

The following theory is crucial in the considered method of the evaluation of the fit:

Definition 8 ***Probability Integral Transform*** *Let $X = (X_1, ..., X_k)$ be a random vector with distribution function $F(x_1, ..., x_k)$. Let $z = (z_1, ..., z_k) = Tx = T(x_1, ..., x_k)$, where T is the transformation considered. Then, T is given by:*

$$\begin{aligned}
z_1 &= P\{X_1 \leq x_1\} = F_1(x_1) \\
z_2 &= P\{X_2 \leq x_2|X_1 = x_1\} = F_2(x_2|x_1) \\
&\vdots \\
z_k &= P\{X_k \leq x_k|X_{k-1} = x_{k-1}, ..., X_1 = x_1\} = F_k(x_k|x_{k-1}, ..., x_1)
\end{aligned}$$

Proposition 6 *Under the assumptions of the above definition, the random vector $Z = TX$ is uniformly distributed on the k-dimensional hypercube.*

Proof This proposition will be proved for the bivariate case. The generalization to higher dimensions is straightforward. Thus:

$$\begin{aligned}
P\{Z_1 \leq z_1, Z_2 \leq z_2\} &= \int\int_{\{Z_1\leq z_1, Z_2 \leq z_2\}} dF_2(x_2|x_1)dF_1(x_1) \\
&= \int_0^{z_1}\int_0^{z_2} dz_1 dz_2 = z_1 z_2
\end{aligned}$$

when $0 \leq z_i \leq 1$, for $i = 1, 2$. Hence, Z_1 and Z_2 are uniformly and independently distributed on $[0, 1]$.

Having this theoretical background, let us now proceed with the description of the method in the bivariate case. On the basis of the above proposition, we conclude that $Z_1 = F(X)$ and $Z_2 = C_2(G(Y)|F(X))$ are uniformly and independently distributed on $[0, 1]$. The method uses the fact that $C_2(G(Y)|F(X))$ is standard uniformly distributed random variable. The procedure is given as follows:

1. Compute the estimation $\hat{\theta}$ of the copula parameter θ (it can be a vector) on the basis of the sample of n bivariate observations $(X_i, Y_i), i = 1, ..., n$.

2. Having estimation $\hat{\theta}$, compute $C_2(G(Y_i)|F(X_i); \hat{\theta}) = \frac{\partial}{\partial u} C(F(X_i), G(Y_i); \hat{\theta})$ for all $i = 1, ..., n$.

3. Plot $\left\{\frac{i}{n+1}, C_2(G(Y_i)|F(X_i); \hat{\theta})^*\right\}$, where $i = 1, ..., n$ and $*$ indicates ordering. This is so-called Quantile-Quantile plot (QQ-plot).

The above procedure is repeated for several choices of the copula models. We choose a copula for which the plot most closely resembles a straight line with a slope 1 (an approximately straight line with a slope 1 would indicate that $C_2(G(Y)|F(X))$ is a standard-uniformly distributed).
It is important to notice that this method can in principle be used in the multivariate case as well, due to the definition of the Probability Integral Transform and the resulting properties. However, the application in the 3-dimensional space already becomes less clear. Consider, the 3-variate case with the copula $C(F_1(x_1), F_2(x_2), F_3(x_3))$. It is not stated clearly whether we should examine $C_2(F_2(X_2)|F_1(X_1))$ or $C_3(F_3(X_3)|F_1(X_1), F_2(X_2))$, or maybe both? In this case, it would be better to perform a test that incorporates the information from both quantities. This will be investigated further.

2.5.3 Kolmogorov-Smirnov goodness of fit test

The graphical method introduced in the last subsection gives insight into the behavior of the dataset. However, judging the fit only on the basis of graphical representation can be confusing. Hence, a "measure" that will determine the goodness-of-fit of this method is required. We consider the Kolmogorov-Smirnov test. This test deals with two situations. The first situation arises when we want to determine if the underlying distribution F_1 differs from the hypothesized continuous distribution function F_2. This is the so-called one-sample case. The second situation arises when we want to judge if two underlying distributions F_1 and F_2 are identical. This refers to the so-called two-sample case. In both situations, the null hypothesis assumes that the distributions F_1 and F_2 are the same, whereas the alternative hypothesis states that these two distributions are different (two-tailed test). The test statistic is given as follows:

$$T = \max_x |F_1(x) - F_2(x)|$$

where F_1 and F_2 are appropriately chosen with respect to the considered situation.
We assume that the null hypothesis is rejected if the test is significant at the 5% level. In order to take a correct decision, we compute the p-value, which measures how much evidence we have against the null hypothesis. If $p \leq 5\%$ we reject the null hypothesis, whereas $p > 5\%$ entails no evidence against the null hypothesis. We can also compare the resulting observed test statistic with the critical value. If the observed test statistic is less than the critical value (which depends on the significance level α), then we cannot reject the null hypothesis. Otherwise, we reject it.

We use this test for the bivariate case of the method discussed in the previous subsection. Thus, we test if the distribution function of variable $C_2(G(Y)|F(X))$ is standard uniform. For instance, applying the Kolmogorov-Smirnov test for several copulas would result in several p-values; if all of them are greater than the significance level 5%, we have no evidence to reject any of the considered copulas. Then, in the pragmatic way, we would conclude that the better fit is provided by the copula, for which the corresponding p-value is the highest. This is also suggested in [2]. The rationale for this is that a high p-value corresponds to a low value of the observed test statistic, which in turn corresponds to a good fit. However, we decide to be careful with judging the best fit on the basis of high p-values and we use them more as a support in the graphical method.
It is important to remark that the Kolmogorov-Smirnov test is more sensitive at the points near the median of the distribution than at its tails. More details about this test and hypothesis testing can be found in [17].

2.5.4 Pseudo graphical method based on the distribution of variable $C(U, V)$ - Method 3

This method is described in [8] and it is based on the comparison of the non-parametric and parametric estimations of the distribution function of random variable $C(U, V)$ (in the bivariate case). We will present this method for the bivariate Archimedean copulas. The procedure below is of special interest to the Archimedean class of copulas, because in this case the parametric estimation of the distribution of $C(U, V)$ can be derived by incorporating the generator. Below we describe algorithms that provide non-parametric and parametric estimations.
The non-parametric estimation of a distribution function of the variable $C(U, V)$ can be derived as follows (this estimation procedure holds also for non-Archimedean classes of copulas):

1. Consider the random sample of n bivariate observations $(X_i, Y_i), i = 1, ..., n$.

2. Compute the pseudo-observations for $i = 1, ..., n$ as follows:

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

3. Calculate the non-parametric estimation K_n of the distribution of variable $C(U, V)$:

$$K_n(t) = \frac{\sum_{i=1}^{n} 1(W_i \leq t)}{n} \quad \text{for } t \in [0, 1]$$

The parametric estimation of a distribution function of variable $C(U, V)$ is given by:

1. Compute the estimator $\hat{\theta}$ of the unknown copula parameter θ on the basis of the sample of n bivariate observations $(X_i, Y_i), i = 1, ..., n$.

2. Having estimator $\hat{\theta}$, compute the generator ϕ (let us recall that the generator of the Archimedean copula depends on the parameter θ).

3. Compute the distribution function of the copula according to the following formula:

$$K(t;\hat{\theta}) = t - \frac{\phi(t,\hat{\theta})}{\phi'(t^+,\hat{\theta})} \quad \text{for } t \in [0,1]$$

We perform the above method of estimation for several choices of the bivariate Archimedean copulas. The best choice of the copula function corresponds to the parametric estimate that most closely resembles the non-parametric estimate. One possibility is to plot $K_n(t)$ and $K(t;\hat{\theta})$ on the same graph to see how well they agree. If the empirical plot approximately follows the parametric estimation we can conclude that the considered copula is a good model.

Remark 6 *Let us consider the following quantity:*

$$E = \int_0^1 |K_n(t) - K(t;\hat{\theta})|^2 dt$$

Clearly, E measures the distance between the parametric and non-parametric estimations. We choose a copula function for which this quantity is minimized. The distance E should be calculated analytically. However for simplicity, in the following case study the solution will be calculated numerically.

A formal methodology for testing the goodness-of-fit of a copula on the basis of the empirical and parametric estimations of the distribution function of the copula, is introduced in [8] and [9]. The authors propose the following test statistic:

$$S_n = n \int_0^1 |K_n(t) - K(t;\hat{\theta})|^2 k(t;\hat{\theta}) dt$$

where $k(t;\hat{\theta}) = \frac{d}{dt}K(t;\hat{\theta})$. The above expression looks a bit complicated. However, it can be simplified as follows:

$$\begin{aligned} S_n &= \frac{n}{3} + n \sum_{j=1}^{n-1} K_n^2\left(\frac{j}{n}\right) \left\{ K\left(\frac{j+1}{n};\hat{\theta}\right) - K\left(\frac{j}{n};\hat{\theta}\right) \right\} \\ &- n \sum_{j=1}^{n-1} K_n\left(\frac{j}{n}\right) \left\{ K^2\left(\frac{j+1}{n};\hat{\theta}\right) - K^2\left(\frac{j}{n};\hat{\theta}\right) \right\} \end{aligned}$$

For the proof, we refer to Appendix A.

Remark 7 *Note that the proposed statistic is of the Cramer-von Mises type, which is used for judging the goodness-of-fit of a probability distribution F^* compared to a given distribution F, with the test statistic defined as:*

$$W^2 = \int_{-\infty}^{\infty} (F(x) - F^*(x))^2 dF(x)$$

In applications F is the theoretical distribution and F^* is the empirical distribution function.

The testing procedures based on this statistic would consist of rejecting the composite null hypothesis (H_0: the data follows some copula $C \in \mathcal{C}_\theta$) if the value of the observed test statistic is greater than the $100(1-\alpha)$th percentile of its distribution under the null hypothesis, where α is the significance level. However, the authors in [8] and [9] observe that this distribution depends on the unknown copula parameter θ even in the limit (so when the sample size increase). In order to circumvent this, the Bootstrap method is proposed. Hence, the critical value and the associated p-value are computed as follows:

1. Calculate estimator $\hat{\theta}$ of the unknown copula parameter θ from the original observations $(X_i, Y_i), i = 1, ..., n$ and compute the value of the test statistic S_n. The estimator must be consistent.

2. Generate B random samples of size n from the copula of interest with the estimator $\hat{\theta}$.

3. Estimate parameter θ (it can be a vector of parameters) for each sample (by the same method as before) and determine the value of the test statistics $S_i^*, i = 1, ..., B$.

4. If $S_{1:B}^* \leq ... \leq S_{B:B}^*$ denote the ordered values of the test statistics, then the critical value of the sampling distribution arises by computing the $(1-\alpha)$-quantile of the values $S_{1:B}^*, ..., S_{B:B}^*$ (then, the significance level is α) whereas the associated p-value is computed from:
$$\frac{1}{B}\#\{i : S_i^* \geq S_n\}$$

We compare the resulting observed test statistic with the critical value. If the observed test statistic is less than the critical value, then we cannot reject the null hypothesis. Otherwise, we reject it. Alternatively, we can compare the resulting p-value with the significance level α. If the p-value is less than α, we reject the null hypothesis. If the p-value is greater than α, then we have no evidence against the null hypothesis. It can be observed empirically that the model for which the test statistic is the smallest generally has the highest p-value (see [9]). As we will discover, the highest p-value will usually correspond to a good model fit. However, we have to be careful with the selection of the best model only on the basis of high p-values. They can be misleading when the sample size is small or large. Therefore, it is good to support our decision, regarding the best model, by using a graphical representation - i.e. the comparison of $K(t; \hat{\theta})$ and $K_n(t)$.

Remark 8 *Let us observe that the above method does not depend on the choice of the marginal distributions. However, the estimator $\hat{\theta}$ can be affected by the choice of the margins.*

Remark 9 *The above method can be extended to the multivariate case. Moreover, it can also be applied for non-Archimedean classes of copulas. This is discussed in [9]. However, the theoretical formula for $K(t)$ should be available. Otherwise, a laborious numerical approximation would have to be implemented.*

2.5.5 A Goodness-of-fit Test for Copulas based on Rosenblatt's Transformation - Method 4

This method is described in [3] and it uses the result of [15].
In this subsection we introduce the goodness-of-fit test based on the Rosenblatt's Transformation. We start with a discussion of the bivariate case. Hence, let X and Y denote two random variables with joint distribution function $H(x,y) = P\{X \leq x, Y \leq y\}$ for $(x,y) \in \Re^2$, and marginal distribution functions $F(x) = P\{X \leq x\}$ and $G(y) = P\{Y \leq y\}$ for $x, y \in \Re$. So we have from Sklar's theorem:

$$H(x,y) = C(F(x), G(y))$$

We note that according to Proposition 6, the random variables:

$$Z_1 = U = F(X) \quad \text{and} \quad Z_2 = C(V|U) = C(G(Y)|F(X))$$

are independent and standard uniformly distributed. The above transformation of variables X and Y into variables Z_1 and Z_2 is called Rosenblatt's Transformation. Then, the following proposition arises:

Proposition 7 *If variable U is uniformly distributed on the interval $[0,1]$ and Φ is the standard normal distribution, then the variable $\Phi^{-1}(U)$ is standard normally distributed.*

Proof Using the definition of a distribution function, we obtain the following calculations:

$$P\{\Phi^{-1}(U) \leq x\} = P\{U \leq \Phi(x)\} = \Phi(x)$$

Thus, $\Phi^{-1}(U)$ is a standard normally distributed random variable and the proof is accomplished.

Definition 9 *If $X_1, X_2 \cdots, X_k$ are independent random variables, such that:*

$$X_i \sim N(0,1), \qquad i = 1, \cdots, k$$

then, the random variable $Z = \sum_{i=1}^{k} X_i^2$ is chi-square distributed with k degrees of freedom. This distribution is denoted by χ_k^2.

Since $\Phi^{-1}(Z_1) = \Phi^{-1}(F(X))$ and $\Phi^{-1}(Z_2) = \Phi^{-1}(C(G(Y))|F(X))$ are independent and standard normally distributed, we conclude that the random variable:

$$S(X,Y) = [\Phi^{-1}(F(X))]^2 + [\Phi^{-1}(C(G(Y)|F(X)))]^2$$

has a χ_2^2-distribution.
The test uses the fact that $S(X,Y)$ has a χ_2^2 distribution. Hence, if $(X_1, Y_1), ..., (X_n, Y_n)$ is a random sample from (X,Y), then $S(X_1,Y_1), ..., S(X_n,Y_n)$ is a random sample from a χ_2^2-distributed random variable. The null hypothesis of interest is given as follows:

$$H_0 : (X,Y) \sim C(F(x), G(y))$$

when the marginal distribution functions are known. This hypothesis can be equivalently expressed in terms of the variable S:

$$H_0 : S(X, Y) \sim \chi^2_2$$

We test the null hypothesis using the Anderson-Darling test. This is suggested in [3]. However, the authors point out that the Kolmogorov-Smirnov test can also be applied. The test statistic of the Anderson-Darling test is:

$$AD = -n - \sum_{i=1}^{n} \frac{2i-1}{n} [\ln(F_0(S_{(j)})) + \ln(1 - F_0(S_{(n-j+1)}))]$$

where $S_i = S(X_i, Y_i)$, $i = 1, ..., n$ and $S_{(1)} \leq ... \leq S_{(n)}$. The function F_0 is the distribution function of χ^2_2-distributed random variable.

Remark 10 *The Anderson-Darling test is applied to test if a sample comes from a population with a specific distribution. It is a modification of the Kolmogorov-Smirnov test and assigns more weight to the tails than the Kolmogorov-Smirnov test does. The Anderson-Darling test makes use of the specific distribution in calculating critical values. This has the advantage of imposing a more sensitive test and the disadvantage that the critical values have to be calculated for each distribution.*

Extensive research devoted to the goodness-of-fit test for copulas based on the Rosenblatt's Transformation is performed in [3]. The authors show that this test works well if the marginal distributions are known and are used in the test statistic, whereas if the marginal distributions are unknown and are replaced by their empirical counterparts the test's properties change dramatically. They consider two cases. The first case corresponds to the theoretical situation when the marginal distributions and the parameters of the considered copula are known. The second case is more relevant for applications, namely when the marginals are empirical and are used to estimate the copula parameters. The authors compare these two cases by checking the performance of the test for each of them. The results indicate that the power of the test in the second situation is significantly lower than in the first case. However, as the authors state, this problem can be fixed by introducing the Bootstrap method (given below). It is important to mention that if the marginal distributions are known and the copula parameter has to be estimated, then the performance of the test is adequate to the first case. The authors claim that if the "Inference for margins" method is implemented, then the pre-testing (i.e. checking the goodness-of-fit for margins) has a negative impact on the properties of the goodness-of-fit test for copulas; the authors do not know how to modify the test in order to solve this problem. They suspect that the Bootstrap method will not solve this problem.
The Bootstrap method:

1. Calculate estimator $\hat{\theta}$ of the unknown copula parameter θ (it can be a vector of parameters) from the original observations $(X_i, Y_i), i = 1, ..., n$ and compute the value of the test statistic AD.

2. Generate B random samples of size n from the copula of interest with the estimator $\hat{\theta}$.

3. Estimate parameter θ for each sample (by the same method as before) and determine the value of the test statistics $AD_i^*, i = 1, ..., B$.

4. If $AD_{1:B}^* \leq ... \leq AD_{B:B}^*$ denote the ordered values of the test statistics, then the critical value of the sampling distribution arises by computing the $(1-\alpha)$-quantile of the values $AD_{1:B}^*, ..., AD_{B:B}^*$ (the significance level is α), whereas the associated p-value is computed from:

$$\frac{1}{B}\#\{i : AD_i^* \geq AD\}$$

We compare the resulting observed test statistic with the critical value. If the observed test statistic is less than the critical value, then we cannot reject the null hypothesis. Otherwise, we reject it. Alternatively, if the resulting p-value is less than the significance level α, the null hypothesis is rejected. Otherwise, we have no evidence against the null hypothesis.

Remark 11 *The extension of this test to the multivariate case is possible. It is due to the fact that the definition of the Probability Integral Transform, Proposition 6 and Definition 9 are stated in the multivariate context. Thus, in the d-dimensional case S becomes:*

$$\begin{aligned} S(X_1, X_2, ..., X_d) &= [\Phi^{-1}(F_1(X_1))]^2 + [\Phi^{-1}(C_2(F_2(X_2)|F_1(X_1)))]^2 \\ +...+ & [\Phi^{-1}(C_d(F_d(X_d)|F_1(X_1), ..., F_{d-1}(X_{d-1})))]^2 \end{aligned}$$

where $(X_1, ..., X_d)$ is the d-dimensional random vector with continuous marginal distributions $F_1, ..., F_d$ and d-copula C. Consequently, we test if S is χ_d^2-distributed. Again, the Bootstrap method is required if the marginal distributions are empirical.

2.5.6 Percentile lines - Method 5

In this subsection we introduce the concept of the so-called percentile lines, which are proposed as one of the methods of the graphical evaluation of the fit. Let us consider the bivariate case. Assume that X and Y are two random variables with joint distribution function $H(x, y) = P\{X \leq x, Y \leq x\}$ for $(x, y) \in \Re^2$ and the marginal distribution functions $F = P\{X \leq x\}$ and $G(y) = P\{Y \leq y\}$ for $x, y \in \Re$. On the basis of Sklar's theorem we have:

$$H(x, y) = C(F(x), G(y))$$

We call the space $(F(X), G(Y))$ a copula or transformed space. Whereas the space (X, Y) is called the original or physical space.

Definition 10 *The p-percentile line in the copula space is defined by the following function:*

$$v = f(u; p), \quad u \in [0, 1]$$

where p is a percentage and $f(u; p)$ is obtained by solving equation $C(v|u) = p$ with respect to v.

Remark 12 *We are usually interested in the percentile lines for $p = 10\%, 50\%, 90\%$ or $p = 5\%, 50\%, 95\%$. These are the most interesting.*

We are also interested in the percentile lines in the physical space. This is obtained by the following transformation:

$$\{(u, f(u; p)) : u \in [0, 1]\} \quad \rightarrow \quad \{(F^{-1}(u), G^{-1}(f(u; p))) : u \in [0, 1]\}$$

The percentile lines can be used as a tool for determining the best copula fit to the dataset. If the copula model fits well then the percentile lines should reflect the trend of the dataset. Moreover, if the considered copula fits well, then approximately $r\%$ of the dataset should fall below the percentile lines corresponding to $p = r\%$ (both in the copula and in the original space). Moreover, this should also be satisfied in the subregions of the physical and copula space.
This method is not applicable in the multivariate case. First of all, the definition of the percentile lines in the multivariate case could be a problem, and moreover the visual assessment would be difficult.

Remark 13 *We note that the percentile lines are the most interesting in the physical space, since they visualize the model in comparison with the original dataset.*
Assume that we do not specify the marginal models F and G - thus, they are empirical. Under special conditions it is possible to invert them. However, the resulting percentile lines in the physical space would not be interesting. Since the empirical margins do not provide a nice upper tail picture, the upper corner in the physical space will be cut. Hence, it is better when the marginal distributions have some tail prescribed.

2.6 Case study

This section presents the application of the above methods. Thus, given a dataset, we will try to investigate which copula function can serve as the best model.

2.6.1 Description of the dataset

In the case study we use a dataset provided by "HKV Consultants". This dataset consists of $n = 89$ bivariate observations $(V_i, W_i), i =, ..., n$ of the water levels and wind speeds. To get a feeling of the dependence between V and W, it is traditional to look at the scatter plot of pairs $(V_i, W_i), i = 1, ..., n$, which is given in Figure 2.1. Note that the scatter plot of the original dataset does not only incorporate information about the dependence between V and W, but also about their marginal behavior. We can obtain an informal, margin-free picture of dependence by transforming the data to have uniform $(0, 1)$ marginal distributions using the modified empirical marginal distribution functions $\hat{F}_n$ and $\hat{G}_n$. If we consider $(V_i, W_i), i = 1, ..., n$, as the realizations of the independent random variables with joint distribution function H, then $(\hat{F}_n(V_i), \hat{G}_n(W_i)), i = 1, ..., n$ can be interpreted as realizations from the copula C corresponding to H. The points $(\hat{F}_n(V_i), \hat{G}_n(W_i)), i = 1, ..., n$, are presented in Figure 2.2.

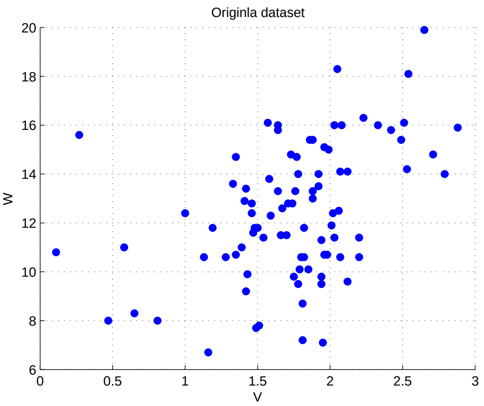


Figure 2.1: Original dataset

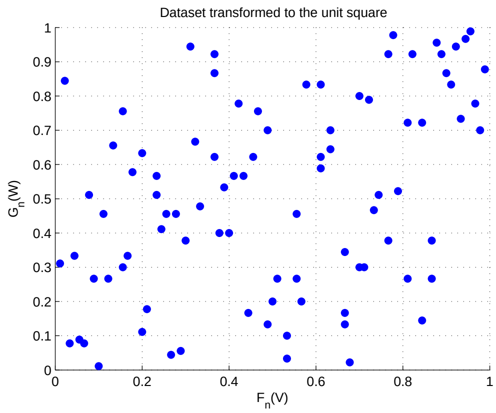


Figure 2.2: Dataset transformed to the unit square

The covariance matrix Σ of the original dataset is:

$$\Sigma = \begin{bmatrix} 0.26 & 0.61 \\ 0.61 & 7.50 \end{bmatrix}$$

Since the off-diagonal elements in the covariance matrix are positive, we should expect that both variables V and W increase jointly - as indeed confirmed by Figure 2.1. Moreover, we use the population versions of product moment correlation and rank correlation in order to check the dependency in our dataset. The numerical computations yield:

ρ	ρ_r
0.43815	0.39838

Both quantities are positive, showing indeed a positive dependency.

2.6.2 Modified empirical marginal distributions

In this subsection, we consider the modified empirical distributions of variables V and W - they will be used in the following case study. As it was already mentioned, using the exact empirical distributions entails some numerical problems in the application. Therefore, we decide to use the modified empirical distributions instead, let us recall their definitions:

$$\hat{F}_V(v) = \frac{1}{n+1}\sum_{i=1}^{n} 1(V_i \leq v) \quad \text{and} \quad \hat{F}_W(w) = \frac{1}{n+1}\sum_{i=1}^{n} 1(W_i \leq w)$$

Figure 2.3 illustrates the comparison between the empirical and modified empirical margins. As we can see the difference is very slight (the largest in the upper corners). Thus, we expect that using the modified empirical distributions will not change the final results much.

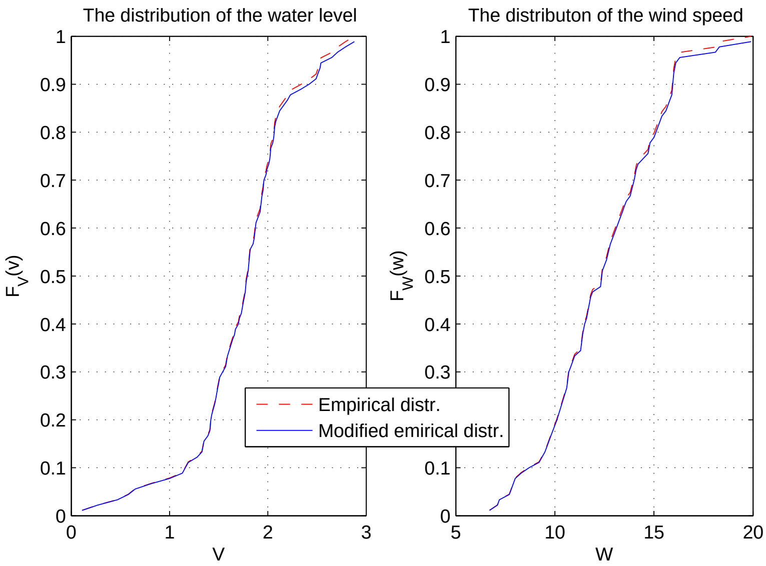


Figure 2.3: The empirical and modified empirical marginal distributions.

2.6.3 Prescribed marginal distributions

In this subsection, we construct the marginal distribution functions $F_V(v)$ and $F_W(w)$. This construction is presented in [7]. We assume that the distribution $F_V(v)$ is given as follows:

$$F_V(v) = \begin{cases} 0 & \text{for } v < a_1 \\ 1 - \exp\left\{\frac{b_{i+1}-b_i}{a_{i+1}-a_i}v + \left(b_i - \frac{b_{i+1}-b_i}{a_{i+1}-a_i}a_i\right)\right\} & \text{for } a_i \leq v < a_{i+1} \text{ and } i = 1, ..., 6 \\ 1 - \frac{\mu_1\Delta}{N}\exp\left\{-\left(\frac{v}{\sigma_1}\right)^{\alpha_1} + \left(\frac{\omega_1}{\sigma_1}\right)^{\alpha_1}\right\} & \text{for } v \geq a_7 \end{cases}$$

whereas the distribution $F_W(w)$ takes the subsequent form:

$$F_W(w) = \begin{cases} 0 & \text{for } w < c_1 \\ 1 - \exp\left\{\frac{d_{i+1}-d_i}{c_{i+1}-c_i}w + \left(d_i - \frac{d_{i+1}-d_i}{c_{i+1}-c_i}c_i\right)\right\} & \text{for } c_i \leq w < c_{i+1} \text{ and } i = 1, ..., 5 \\ 1 - \frac{\mu_2\Delta}{N}\exp\left\{-\left(\frac{w}{\sigma_2}\right)^{\alpha_2} + \left(\frac{\omega_2}{\sigma_2}\right)^{\alpha_2}\right\} & \text{for } w \geq c_6 \end{cases}$$

where the parameters are given in the following tables:

i	a_i	b_i	c_i	d_i
1	0.1	$\ln(1.000e+00)$	6	$\ln(1.000e+00)$
2	1.2	$\ln(9.100e-01)$	8	$\ln(9.600e-01)$
3	1.4	$\ln(8.400e-01)$	10	$\ln(8.400e-01)$
4	1.6	$\ln(6.800e-01)$	12	$\ln(5.600e-01)$
5	1.7	$\ln(6.000e-01)$	14	$\ln(3.200e-01)$
6	1.8	$\ln(5.000e-01)$	16	$\ln(1.133e-01)$
7	1.97	$\ln(3.500e-01)$	-	-

i	α_i	μ_i	σ_i	ω_i	Δ	N
1	0.74	1.298	0.0913	1.97	24	89
2	2.05	0.42	8.66	16	24	89

The Figure 2.4 presents the functions $1 - F_V(v)$ and $1 - F_W(w)$ on a logarithmic scale.

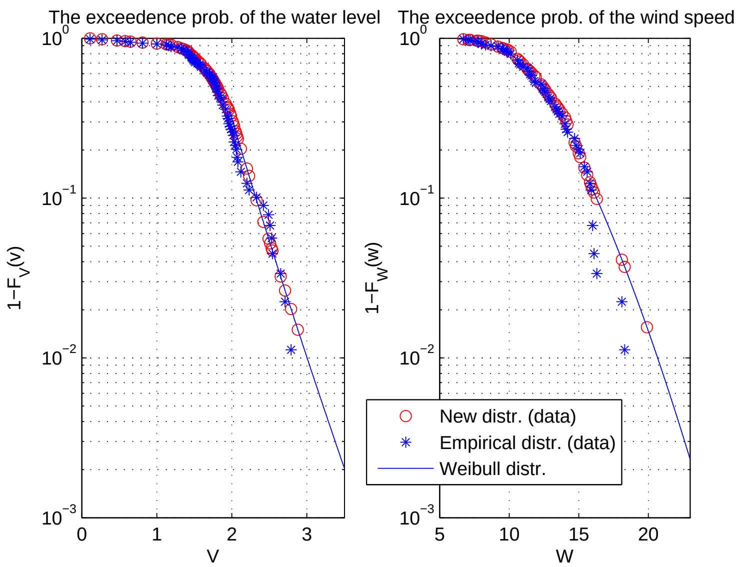


Figure 2.4: The functions $1 - F_V(v)$ and $1 - F_W(w)$ on a logarithmic scale.

Remark 14 *We observe that the survival function:*

$$\frac{\mu\Delta}{N} \exp\left\{-\left(\frac{x}{\sigma}\right)^{\alpha} + \left(\frac{\omega}{\sigma}\right)^{\alpha}\right\} \quad \textit{for } x \geq \omega$$

is the modified Weibull. The term $\mu\Delta/N$ is a frequency that $X \geq \omega$. Indeed, the Weibull cumulative distribution function is given as follows:

$$P(X \leq x) = 1 - \exp\left\{-\left(\frac{x}{\lambda}\right)^{k}\right\} \quad \textit{for } x \geq 0$$

where $\lambda > 0$ is a scale parameter and $k > 0$ is a shape parameter. Then, it is straightforward that for $x > y$ the conditional distribution $P\{X \leq x|X > y\}$ is:

$$
\begin{aligned}
P\{X \leq x|X > y\} &= 1 - P\{X > x|X > y\} = 1 - \frac{\exp\left\{-\left(\frac{x}{\lambda}\right)^k\right\}}{\exp\left\{-\left(\frac{y}{\lambda}\right)^k\right\}} \\
&= 1 - \exp\left\{-\left(\frac{x}{\lambda}\right)^k + \left(\frac{y}{\lambda}\right)^k\right\}
\end{aligned}
$$

We observe that the distributions $F_V(v)$ and $F_W(w)$ consists of two parts. The first part consists of the interpolated exponential "pieces", whereas the second part is the modified conditional Weibull distribution. This construction of margins allows for a description of the tail behavior of the considered variables.
We remark that in this thesis $F_V(v)$ and $F_W(w)$, constructed above, are considered as "known", although in [7] they are actually fitted to the data.

2.6.4 Determination of the best copula - the modified empirical marginal distributions

Our aim is to fit a copula function that best reflects the dependency in our dataset. We consider three Archimedean families of copulas: Clayton, Gumbel and Frank. Moreover, in this subsection we do not specify the marginal distribution models. We decide to approximate them by modified empirical distributions.

Estimation of the unknown parameters

Each of the considered copula functions depends on some parameter θ that needs to be estimated from the dataset. The modified empirical margins entails the use of the Canonical Maximum Likelihood Method (CML). Moreover, we also perform the estimation according to the Genest-Rivest approach (GR). Both methods are described in section 2.4. The resulting estimators of the unknown parameters and associated approximate 95% confidence intervals are given in the following table:

Copula	$\hat{\sigma}$ CML	Confidence interval for CML	$\hat{\sigma}$ GR	Confidence interval for GR
Clayton	0.49	$[0.1439, 0.8361]$	0.75968	$[0.238, 1.2813]$
Gumbel	1.41	$[1.1852, 1.6348]$	1.3798	$[1.119, 1.64067]$
Frank	2.67	$[1.3494, 3.9906]$	2.6	$[1.3492, 3.9908]$

The construction of the approximate confidence intervals is described in [8].

Method 1

Having the estimated parameters for each considered copula model, we can proceed with the evaluation of the fit. In this paragraph we consider the method introduced in subsection 2.5.1. This approach requires drawing the samples of size n from the

Clayton, Gumbel and Frank copula. Recall, that in our case $n = 89$. Since the sample from the copula is random, the "small" size $n = 89$ can lead to a quite arbitrary picture of the copula. Therefore, we decide to draw samples of size 2000. The resulting plots are presented in Figure 2.5. We consider only the estimators derived according to the CML Method, because the plots that correspond to the estimations derived under the GR approach are very similar. We conclude that the plots do not

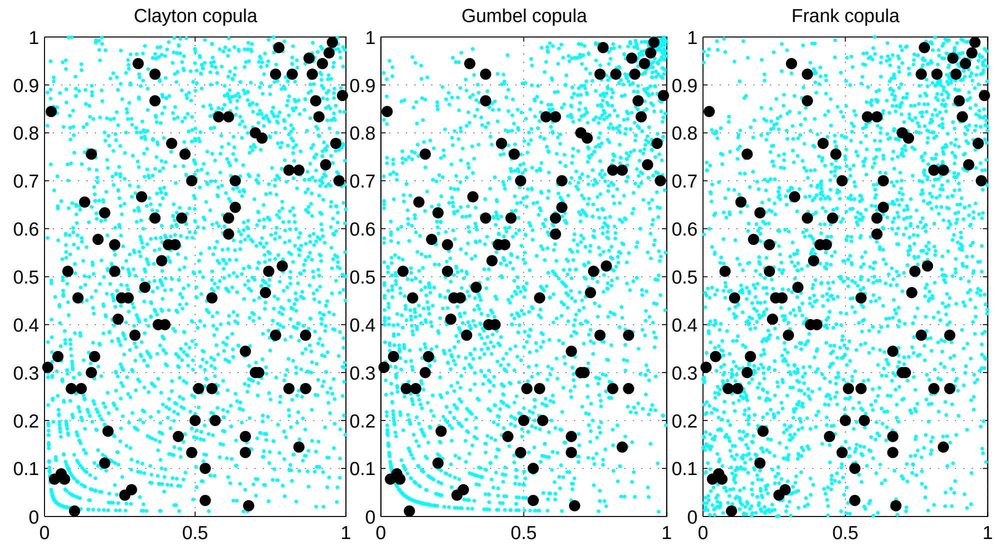


Figure 2.5: Scatter plots of Clayton, Gumbel and Frank copula with superimposed transformed dataset (CML). The copula parameters are estimated by the CML method. The black dots correspond to the transformed dataset, whereas the cyan dots represent the copula models.

give a straightforward answer. The pictures show how difficult the problem of the selection of the best copula is.

Method 2

The aim of this paragraph is to determine the best copula model by using the approach introduced in subsection 2.5.2. Let us first consider this method applied to the case when parameters are estimated according to the CML Method. Then, the QQ-plots in Figure 2.6 emerge. According to the previous section, we should choose the copula for which the plot most closely resembles a straight line with slope 1. Since for the Frank copula this is not the case, we may suspect that this copula does not provide a good fit. However, for the Clayton and Gumbel copula we cannot unambiguously say which one is better. Additionally, we apply the Kolmogorov-Smirnov test discussed in subsection 2.5.3. The resulting p-values are 0.6584, 0.6525 and 0.38 (the order of the p-values corresponds to the order Clayton, Gumbel and Frank copula). All p-values are greater than the significance level 5%. Hence, we cannot reject any of the considered copulas. We observe that the magnitude of p-values confirm our previous observations. The p-value is the smallest for Frank copula, whereas the Clayton and Gumbel copula seem both reasonable.

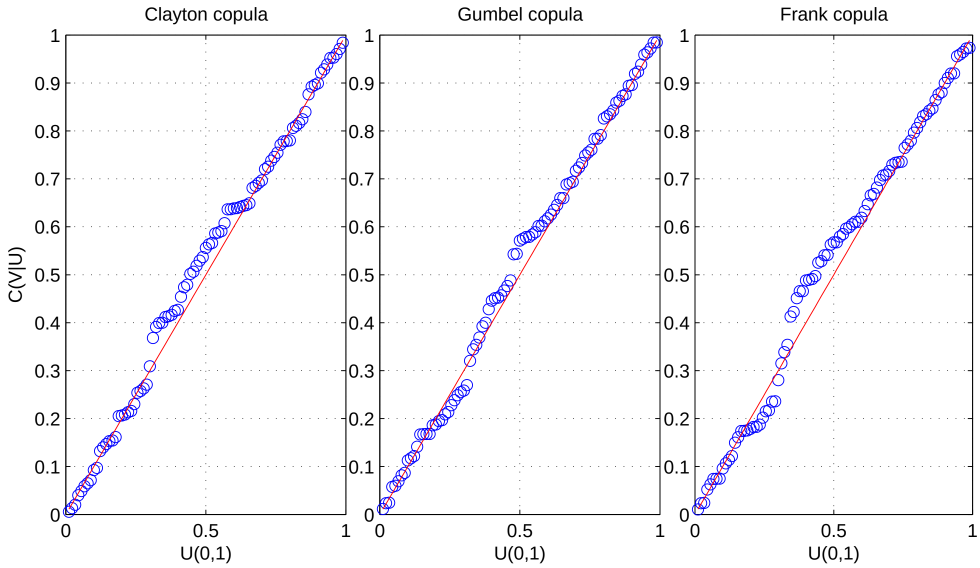


Figure 2.6: QQ-plots for three copulas fitted with Canonical Maximum Likelihood Method (CML)

Now, let us repeat the analysis for the GR approach. Then, the QQ-plots in Figure 2.7 emerge. We observe that the plot for the Gumbel copula most closely resembles

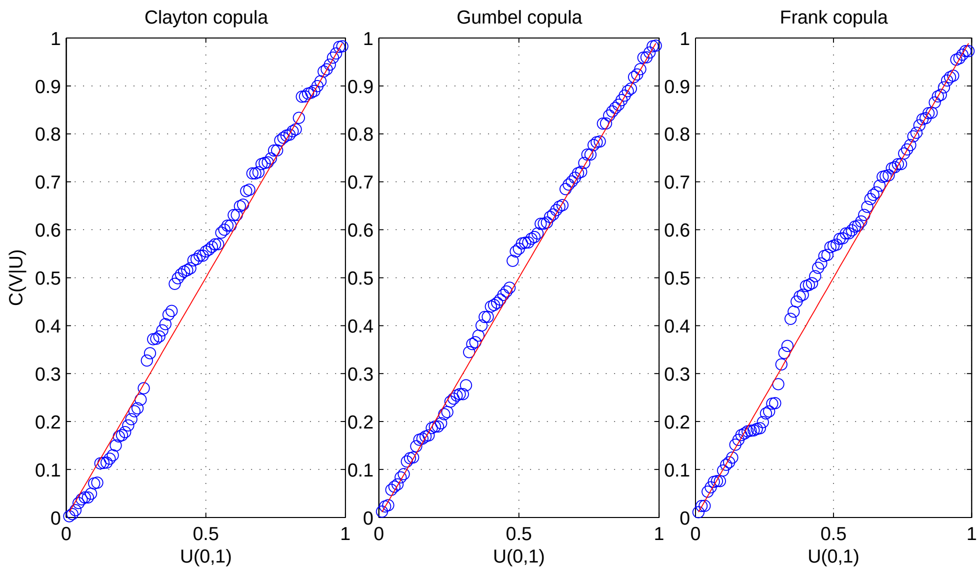


Figure 2.7: QQ-plots for three copulas fitted with the Genest-Rivest method (GR)

a straight line with slope 1. Hence, we may conclude that here this copula fits best. Additionally, we perform the Kolmogorov-Smirnov test. The resulting p-values are 0.2604, 0.7399 and 0.4409. All p-values are greater than the significance level 5%. Hence, we cannot reject any considered copula function. The highest p-value is obtained for the Gumbel copula. For convenience, we present all p-values in the following table:

Copula	p-value	
	CML	GR
Clayton	**0.6584**	0.2604
Gumbel	**0.6525**	**0.7399**
Frank	0.38	0.4409

Method 3

In this paragraph, we apply the method described in subsection 2.5.4. First, we consider the estimators of the unknown parameters derived according to the CML Method. The panels of Figure 2.8 illustrate the empirical estimation of the distribution function $K_n(t)$ with the comparison of the parametric estimations $K(t;\hat{\theta})$. Figure 2.9 presents the plots with respect to the estimations derived according to the GR approach. We observe that in both cases the highest agreement between $K_n(t)$

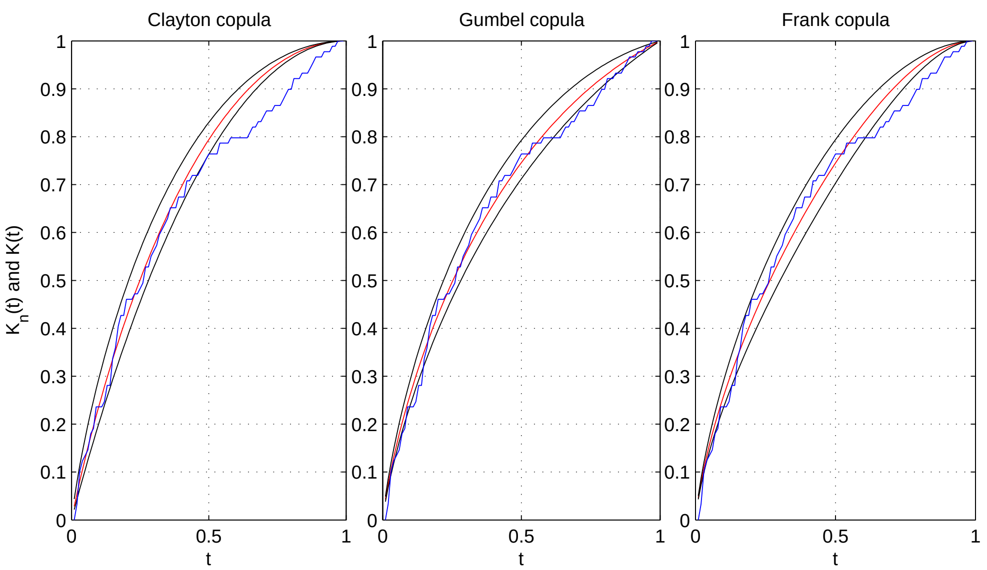


Figure 2.8: $K_n(t)$ and $K(t;\hat{\theta})$ (CML): The red line corresponds to $K(t;\hat{\theta})$, the black lines constitute the 95% confidence interval and the blue line the empirical estimation $K_n(t)$. Observe that the highest agreement is obtained for the Gumbel copula.

and $K(t;\hat{\theta})$ is obtained for the Gumbel copula. Moreover, invoking Remark 6 results in the following table:

Copula	$E = \int_0^1 \lvert K_n(t) - K(t;\hat{\theta})\rvert^2 dt$ CML	$E = \int_0^1 \lvert K_n(t) - K(t;\hat{\theta})\rvert^2 dt$ GR
Clayton	0.0019864	0.0027444
Gumbel	**0.00080452**	**0.00069542**
Frank	0.0015108	0.00144558

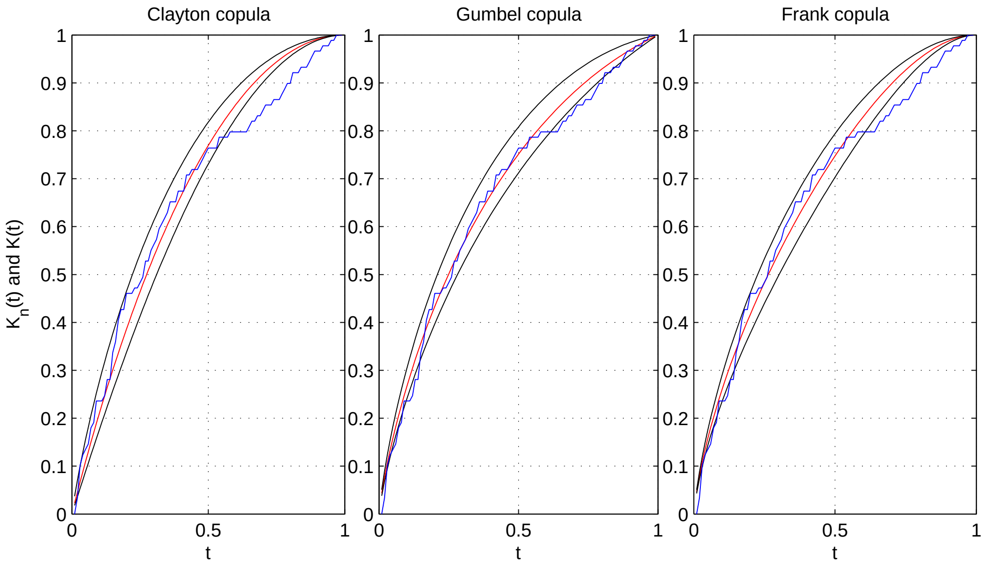


Figure 2.9: $K_n(t)$ and $K(t;\hat{\theta})$ (GR): The red line corresponds to $K(t;\hat{\theta})$, the black lines constitute the 95% confidence interval and the blue line the empirical estimation $K_n(t)$. Observe that the highest agreement between the parametric and non-parametric estimation is obtained for the Gumbel copula.

From the table, we can see that E is the smallest for the Gumbel copula. This is another rationale that according to this method the Gumbel copula is the best (of course in the considered group of copulas). In addition, applying the Bootstrap procedure with $B = 500$ at the 5%-significance level, results in:

Copula	CML estimation			GR estimation		
	Test stat.	Crit. val.	p-value	Test stat.	Crit. val.	p-value
Clayton	0.11918	0.162	0.138	0.13632	0.165	0.094
Gumbel	0.05706	0.157	**0.634**	0.0629	0.140	**0.490**
Frank	0.08862	0.141	0.248	0.08841	0.153	0.308

Consider the CML estimation. All resulting p-values are above the 5%-significance level. Hence, we cannot reject any model. We observe that the highest p-value is obtained for the Gumbel copula. This is consistent with the previous observations. Consider the GR estimation. All p-values are higher than the significance level 5%. Thus, we cannot reject any considered copula. The highest p-value is obtained for Gumbel copula. Moreover, the magnitudes of the resulting p-values agree with the outcomes of Figure 2.9. Note that the Clayton copula would be rejected, if the significance level was 10%.

Method 4

In this paragraph, we implement the method introduced in subsection 2.5.5. Let us recall that the marginal distributions $F_V(v)$ and $F_W(w)$ are modified empirical. However, these functions are only slightly different from the empirical distributions, as is shown in Figure 2.3. Therefore, we decide to apply the Bootstrap method, with $B = 500$ repetitions. The following table arises:

Copula	CML estimation			GR estimation		
	Test stat.	New crit. val.	New p-val.	Test stat.	New crit. val.	New p-val.
Clayton	0.5926	0.87	0.242	0.6809	0.94	0.196
Gumbel	0.549	0.85	**0.338**	0.5722	0.90	**0.328**
Frank	0.7042	0.90	0.134	0.6994	0.94	0.156

We note that all resulting p-values are grater than the significance level 5%. Clearly, we cannot reject any of the considered copula models.

Final results

For convenience, we summarize the results of the previous paragraphs in the following table:

	CML estimation	GR estimation
Method 1	The results are difficult to interpret.	The results are difficult to interpret.
Method 2	The plots favor the Clayton and Gumbel copula. On the basis of the p-values we cannot reject any of the considered copulas. The highest p-values are obtained for the Clayton and Gumbel copula.	The plots favor the Gumbel copula. On the basis of the p-values we cannot reject any of the considered copulas. The highest p-value is obtained for Gumbel copula.
Method 3	The plots and error E favor the Gumbel copula. All resulting p-values are greater than the significance level 5%. Thus, we cannot reject any of the considered copulas. The highest p-value is obtained for the Gumbel copula.	The plots and error E favor the Gumbel copula. All resulting p-values are greater than the significance level 5%.Thus, we cannot reject any of the considered copulas. The highest p-value is obtained for the Gumbel copula. Note that the Clayton copula would be rejected, if the significance level was 10%.
Method 4	Resulting p-values are greater than the significance level 5%. Thus, non of the considered copulas can be rejected.	Resulting p-values are greater than the significance level 5%. Thus, non of the considered copulas can be rejected.

On the basis of the above table, we can conclude that "on the average" the Gumbel copula gives a good fit in each method of estimation. It was indicated by the graphical methods and not rejected by the statistical tests. However, it is important to notice that the statistical tests did not lead to the rejection of any considered copula.

2.6.5 Determination of the best copula - prescribed marginal distributions

Let us again consider the three Archimedean families Clayton, Gumbel and Frank. We assume that the marginal distribution functions $F_V(v)$ and $F_W(w)$ are given (the construction of the margins is presented in the previous section). We repeat the estimation of the unknown parameters and the evaluation of the fit.

Estimation of the unknown parameters

In this paragraph, we use the modification of CML Method (we call it MCML), in order to estimate the copula parameter. To be more precise, we use the CML Method, where the empirical distributions are replaced by the known margins, which applied on the dataset are "almost" empirical. This kind of estimation can also be viewed as the Maximum Likelihood Method. We obtain the following table of estimations of parameter θ and the associated approximate 95% confidence intervals:

Copula	$\hat{\theta}$ MCML	Confidence interval for MCML
Clayton	0.42	$[0.1009, 0.7391]$
Gumbel	1.44	$[1.2094, 1.6706]$
Frank	2.82	$[1.4986, 4.1414]$

Method 1

In this paragraph, we consider the method discussed in subsection 2.5.1 with prescribed marginal distributions and the estimations derived according to the MCML Method. Again, we sample from each considered copula a set of bivariate points of size 2000. Thus, Figure 2.10 emerges. The results are not significantly different

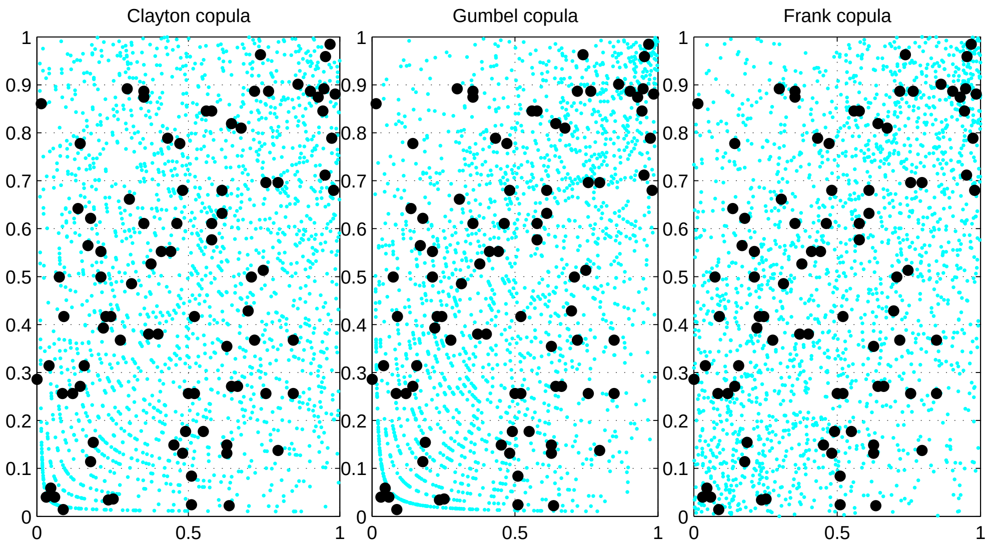


Figure 2.10: The Clayton, Gumbel and Frank copula with the dataset transformed to the unit square. The copula parameters are estimated by the MCML method. The black dots correspond to the transformed dataset, whereas the cyan dots represent the copula functions.

than those for the modified empirical marginal distributions. This is not surprising, due to the fact that the prescribed distributions, $F_V(v)$ and $F_W(w)$, applied on the considered dataset are very close to their empirical counterparts. Hence, again the resulting plots do not give an unambiguous answer.

Method 2

In this paragraph, we consider the method discussed in subsection 2.5.2 with prescribed marginal distributions and the estimations derived according to the MCML Method. The corresponding plots are presented in Figure 2.11. We observe that

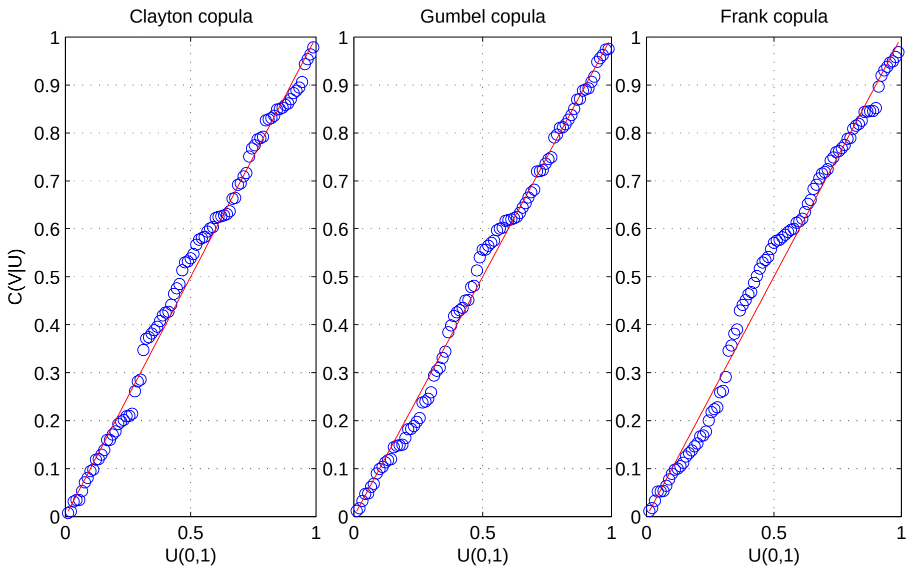


Figure 2.11: QQ-plot (MCML)

the highest agreement is obtained for the Clayton and Gumbel copula. The p-values derived from the Kolmogorov-Smirnov test are given:

Copula	p-value MCML
Clayton	**0.9160**
Gumbel	**0.8702**
Frank	0.5982

All p-values are above the 5%-significance level. Thus, none of the considered copulas can be rejected. The highest p-values are obtained for the Clayton and Gumbel copula.

Method 3

In this paragraph, we consider the method discussed in subsection 2.5.4 with the prescribed marginal distributions and the estimations derived according to the MCML Method. Hence, we obtain Figure 2.12 where the empirical estimation $K_n(t)$ together with the parametric estimation $K(t;\hat{\theta})$ are presented as a functions of t. Let us note that the highest agreement is obtained for the Gumbel copula. Moreover, Remark 6 results in the following table:

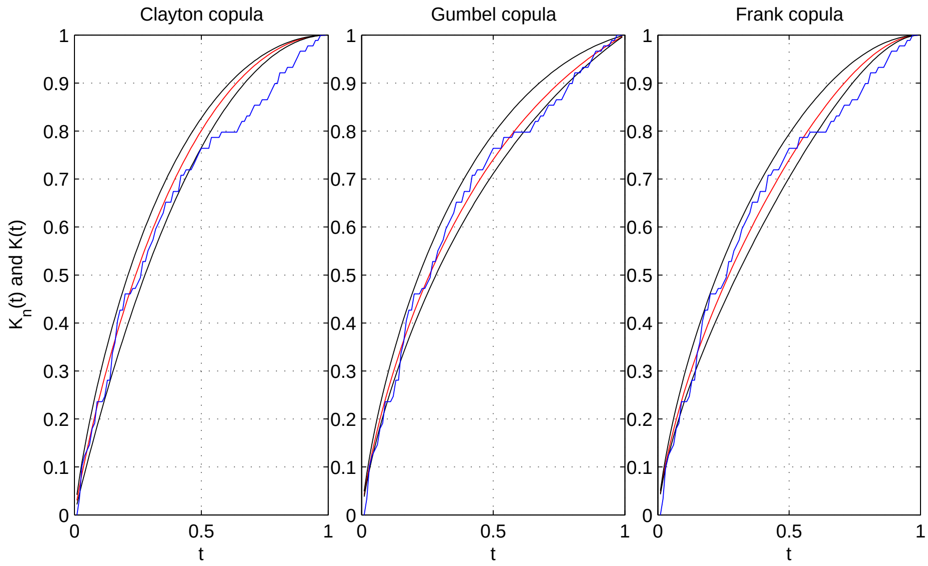


Figure 2.12: $K_n(t)$ and $K(t;\hat{\theta})$ (MCML) The red line corresponds to $K(t;\hat{\theta})$, the black lines constitute the 95% confidence interval and the blue line the empirical estimation $K_n(t)$. Observe that the highest agreement between the empirical and parametric estimations is obtained for the Gumbel copula.

Copula	$E = \int_0^1 \lvert K_n(t) - K(t;\hat{\theta})\rvert^2 dt$ MCML
Clayton	0.0019568
Gumbel	**0.0009416**
Frank	0.001669

From the above table, we can see that E is the smallest for the Gumbel copula. This is another reason which confirms that the Gumbel copula constitutes a good fit according to this method. In addition, applying the Bootstrap procedure, with $B = 500$ at the 5%-significance level, results in:

Copula	MCML estimation		
	Test stat.	Crit. val.	p-value
Clayton	0.13838	0.159	0.106
Gumbel	0.054426	0.158	**0.650**
Frank	0.091038	0.136	0.214

All p-values are greater than the significance level 5%. Hence, we cannot reject any model. The highest p-value is assigned to the Gumbel copula. Moreover, the magnitude of the p-values agrees with the previous results, i.e. with outcomes of plots and error E.

Method 4

In this paragraph, we consider the Rosenblatt Transformation Method, to evaluate the fit introduced in subsection 2.5.5. This method is developed for known or empirical marginal distributions (in case of the empirical margins, the Bootstrap simulation is required). In our case, the marginal distributions are somewhat "between" known and empirical. We consider them as known; however, applied to the considered dataset they lead to "almost" empirical values. In order to avoid bias in our derivations, we decide to not implement this test. We believe that other methods of evaluation of the fit are enough to make conclusions about the fit.

Method 5

In this paragraph, we consider the method of percentile lines. The theory is introduced in subsection 2.5.6. We derive the percentile lines in the physical and transformed spaces for MCML estimation for each considered copula model. Recall that the assumption of the prescribed margins $F_V(v)$ and $F_W(w)$ have tails with infinite support (contrary to the empirical margins), which makes it possible to judge the model fit in the extreme region of the original space. We obtain Figures 2.13, 2.14 and 2.15.

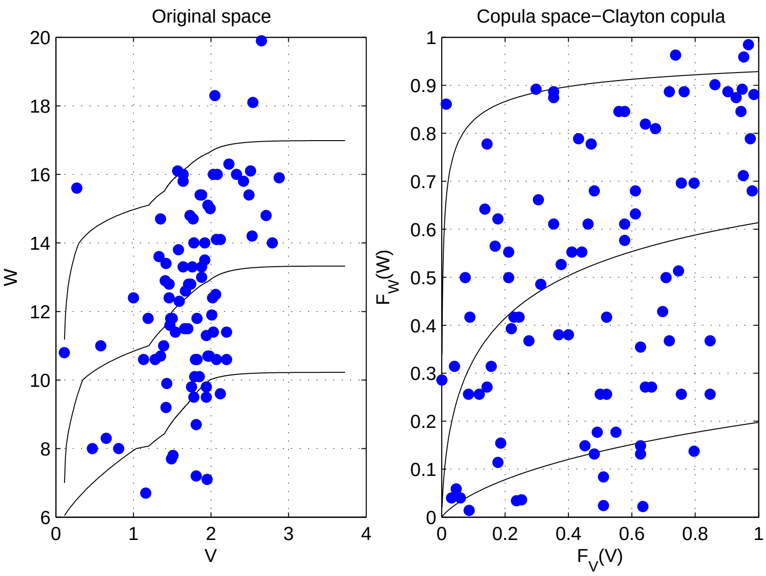


Figure 2.13: Percentile lines $(10\%, 50\%, 90\%)$ in original and transformed spaces for Clayton copula with the MCML estimator $\hat{\theta} = 0.42$: We observe that approximately 5 data points fall above the 90%-percentile line and approximately 10 data points fall below the 10%-percentile line when we consider the whole dataset. Recall that it should be approximately 9 data points (10% of the dataset) when we consider the whole dataset.

From these graphs (physical and copula space), we conclude that the Clayton copula does not reflect the increasing tendency of the dataset. We observe that the increasing behavior of the data is best described by the Gumbel copula; also the number of points that fall between the percentile lines agrees with the theory in the physical and copula space.

We observe that it is a bit easier to judge the fit in the extreme region in the original space than in the copula space. This is due to the fact that the copula function is focused on the whole transformed dataset. In the original space, we can "evaluate" the fit in the extreme area - this is because of the tails of functions F_V and F_W.

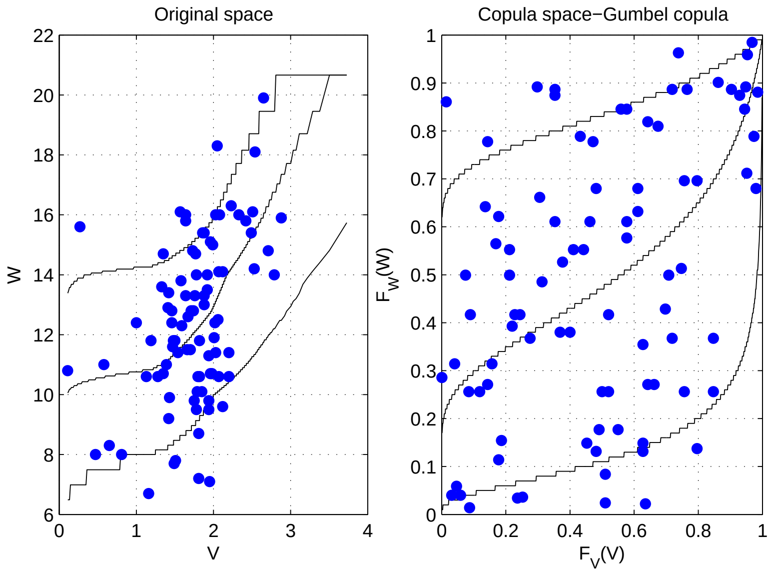


Figure 2.14: Percentile lines $(10\%, 50\%, 90\%)$ in original and transformed spaces for Gumbel copula with the MCML estimation $\hat{\theta} = 1.44$: We observe that approximately 7 data points fall above the 90%-percentile line and approximately 8 data points fall below the 10%-percentile line when we consider the whole dataset. Moreover, the percentile lines follow increasing the trend of the data points.

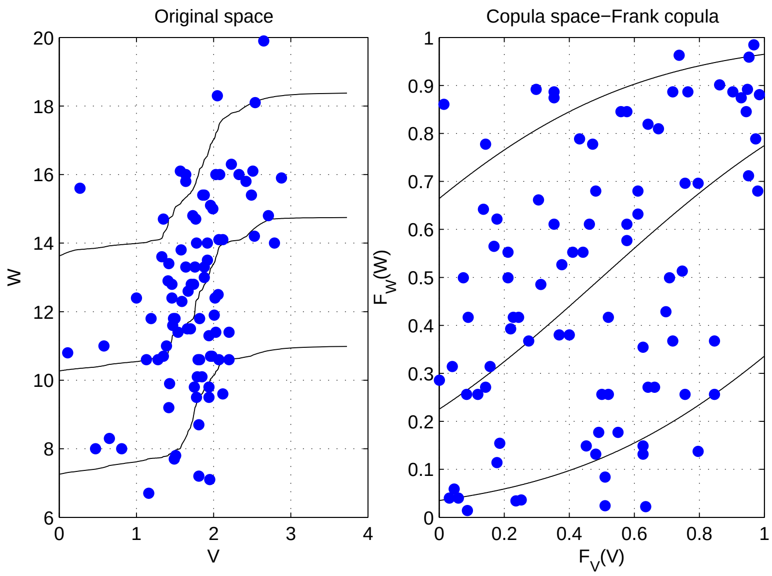


Figure 2.15: Percentile lines $(10\%, 50\%, 90\%)$ in original and transformed spaces for Frank copula with the MCML estimation $\hat{\theta} = 2.82$: We observe that approximately 8 data points fall above the 90%-percentile line and below the 10%-percentile line when we consider the whole dataset.

Final results

For convenience, we summarize the results of the previous paragraphs in the following table:

	MCML estimation
Method 1	The results are difficult to interpret.
Method 2	The plots favor the Clayton and Gumbel copula. On the basis of the p-values we cannot reject any of the considered copulas. All p-values are very high.
Method 3	The plots and error E favor the Gumbel copula. All resulting p-values are greater than the significance level 5%. Thus, we cannot reject any of the considered copulas.
Method 5	The increasing tendency of the dataset is best described by the Gumbel copula.

Generally, the best model selection is very difficult. However, we observe that "on the average" the Gumbel copula can be considered as a good model. This copula was indicated by the graphical methods and was not rejected by the statistical tests. It is important to notice that the resulting p-values did not lead to the rejection of any considered copula.

2.7 Conclusions

A bivariate copula is a cumulative distribution function defined on the unit square. The main property of the copula is that it describes the margin-free dependency between two random variables. There are many classes of copulas. In our considerations, we focused on the Archimedean class, which is often applied in hydrology and finance. Moreover, this class is "friendly" in studies, because the corresponding density functions and the derivatives of the copulas are usually easily obtained.

Copulas can be relatively easily extended to the multivariate case. This is due to Sklar's theorem, which is formulated in the multivariate context. This is very useful in applications if one wants to know the relation between more than two random variables.

A case study is always very helpful in understanding the underlying methodology. In case of the copula function the statistical inference is large; however, it is usually poorly described in the literature. Most authors focus on the final result rather than on the mathematical explanation. There are many goodness-of-fit tests proposed, which often require the Bootstrap method. The main drawback of this kind of simulation is that it is time consuming especially when we deal with complicated algorithms for copula sampling and Maximum Likelihood Method of estimation. Of course the higher the number of repetitions (samples) the better the final result. Thus, there is a trade-off between the computation time and the quality of our outcomes.

Looking at the percentile lines, it is a bit difficult to judge the fit in the extreme region of the copula space. The copula space focuses on the whole transformed dataset. Therefore, if we are interested in the extremes, it is recommended to evaluate the fit in the physical space, where the picture of the fit in the extreme region becomes more clear - provided we have information about the tails of distributions F_V and F_W.

In our case study we considered three bivariate Archimedean copulas: the Clayton, Gumbel and Frank. We concluded that the Gumbel copula provides the best fit to the dataset of bivariate observations of water levels and wind speeds. The results are affected by the choice of the estimation method and marginal distributions.

In conclusion, we have to be careful with the application of goodness-of-fit tests for copulas and experience is required.

Chapter 3

Bivariate dependency models

In this chapter, a number of bivariate dependency models will be introduced. A detailed description of these models can be found in [5] and [7]. The models are created using a transformation of random variables into the so-called model space. The transformation, here described for an arbitrary number of variables is as follows: Let us assume that $V_1, V_2, ..., V_n$ are random variables with known distribution functions $F_{V_1}(v_1), F_{V_2}(v_2), ..., F_{V_n}(v_n)$, respectively. The transformation of $V_1, V_2, ..., V_n$ into the variables $X_1, X_2, ..., X_n$ is defined as follows:

$$\begin{aligned} F_{V_i}(v_i) &= F_{X_i}(x_i) \quad \text{for } i = 1, ..., n \\ &\Downarrow \\ x_i &= F_{X_i}^{-1}(F_{V_i}(v_i)) \quad \text{for } i = 1, ..., n \end{aligned}$$

where $F_{X_i}(x_i)$ is the distribution of variable X_i, for $i = 1, ..., n$. Moreover, the density function of vector $X = (X_1, X_2, ..., X_n)$ is known and is denoted by $f_X(x_1, x_2, ..., x_n)$. We observe that variables $X_1, ..., X_n$ are functions of $V_1, ..., V_n$, respectively. Hence, the above transformation and assumptions entail the formula for the density function $f_V(v_1, v_2, ..., v_n)$ of vector $V = (V_1, ..., V_n)$:

$$f_V(v_1, ..., v_n) = \frac{f_X(F_{X_1}^{-1}(F_{V_1}(v_1)), ..., F_{X_n}^{-1}(F_{V_n}(v_n))) \prod_{i=1}^{n} f_{V_i}(v_i)}{\prod_{i=1}^{n} f_{X_i}(F_{X_i}^{-1}(F_{V_i}(v_i)))} \tag{3.1}$$

where $f_{X_i}(x_i)$ is the density function of X_i and $f_{V_i}(v_i)$ is the density function of V_i, for $i = 1, ..., n$.

The above transformation is used to construct the joint density function $f_V(v_1, ..., v_n)$ - if it is unknown. The aim is to derive f_V such that dependency in the dataset is sufficiently described. Moreover, the margins of the density $f_V(v_1, ..., v_n)$ are $f_{V_i}(v_i)$, for $i = 1, ..., n$ - this is inherent from the used transformation. The variables $X_1, ..., X_n$, defined in terms of the joint density function $f_X(x_1, ..., x_n)$, constitute the model. If the model density f_X is appropriately chosen, then density f_V describes well the dependency in the data.

In the following subsection we will consider three bivariate models, namely the Constant Spread Model (Model CS), Variable Spread Model (Model VS) and the Constant

Symmetric Spread Model (Model CSS). The main purpose of these models is the creation of a bivariate distribution function, which gives the information of the upper tail behavior. This will be obtained by a somewhat modified regression analysis. Models will be supported by pictures, which help to understand the underlying methodology. Moreover, we will introduce the notion of the tail dependence coefficients and verify the relation of these models with the Archimedean class of copulas.

Remark 15 *Consider the special case $F_{X_i}(x_i) = x_i$, for $i = 1, ..., n$. Hence, the marginal distributions are uniform. From transformation (3.1) follows:*

$$F_V(v_1, ..., v_n) = F_X(F_{V_1}(v_1), ..., F_{V_n}(v_n))$$

so, in this case F_X corresponds to the copula of F_V.

3.1 Constant Spread Model

In this section, definitions and necessary details concerning the Constant Spread Model will be provided. We start with the description of the variable X. We assume that X is a standard exponentially distributed random variable. Hence, its cumulative distribution and corresponding density function are given as follows:

$$F_X(x) = \begin{cases} 1 - e^{-x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

$$f_X(x) = \begin{cases} e^{-x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

For further purpose, we define a cumulative distribution function Λ_σ and the corresponding density function λ_σ. We assume that if random variable T satisfies:

$$P\{T \leq t\} = \Lambda_\sigma(t)$$

then $\mu := E(T) = 0$ and $\sigma := \sqrt{E(T^2)} > 0$, where E denotes the expected value. Let us define the "positive" domain D of the function λ_σ, on which it takes positive values:

$$D = \{t | \lambda_\sigma(t) > 0\}$$

We assume that only functions λ_σ with open and "unbroken" domain D will be considered. This domain will be denoted by the open interval (y_b, y_e), where indexes b and e indicate the beginning and the ending of the set D, respectively. In addition, we assume that the density function λ_σ satisfies the following conditions:

1. λ_σ is bounded and continuous except maybe for finitely many points of discontinuity.

2. If $P\{T \leq t\} = \Lambda_\sigma(t)$ then:

$$E(e^T) = \int_{-\infty}^{\infty} e^t \lambda_\sigma(t) dt < \infty \tag{3.2}$$

Remark 16 *In [5] another condition is mentioned, which however is not necessary for proofs in this report.*

We note that if Λ_σ is a normal distribution function with corresponding zero mean and standard deviation $\sigma > 0$, and $P\{T \le t\} = \Lambda_\sigma(t)$, then the following relations hold:

$$\begin{aligned} E(e^T) &= \int_{-\infty}^{\infty} e^t \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{t^2}{2\sigma^2}\right) dt \\ &= \exp\left(\frac{\sigma^2}{2}\right) \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left(-\frac{(t-\sigma^2)^2}{2\sigma^2}\right) dt \\ &= \exp\left(\frac{\sigma^2}{2}\right) \end{aligned} \tag{3.3}$$

Thus, condition 2 given by (3.2) is satisfied for every value of the parameter σ. This is not the case if we consider Λ_σ as an exponential distribution function with corresponding mean zero and standard deviation $\sigma > 0$:

$$\lambda_\sigma(t) = \frac{1}{\sigma} e^{-\frac{t+\sigma}{\sigma}} \quad \text{for } t > -\sigma$$

Then, the expectation $E(e^T)$ is infinite for $\sigma \ge 1$. However, for $0 < \sigma < 1$ the expectation $E(e^T)$ is equal to $e^{-\sigma}/(1-\sigma)$ and thus, condition (3.2) is fulfilled for every $0 < \sigma < 1$.
Let us define the conditional density of variable $Y|X = x$ as follows:

$$f_{Y|X=x}(y|x) = \lambda_\sigma(y - x - \delta) \tag{3.4}$$

where δ is a real number.
Relation (3.4) provides a formula for the corresponding density function $f_{X,Y}(x,y)$ of vector (X,Y) in terms of the function λ_σ:

$$f_{X,Y}(x,y) = f_X(x) f_{Y|X=x}(y|x) = e^{-x}\lambda_\sigma(y - x - \delta) \quad \text{for} \quad x > 0 \tag{3.5}$$

The resulting model (X,Y) is presented in Figure 3.1. The picture helps to visualize the main idea and origin of this model. The construction of this model entails that we can easily observe what happens in the upper tail of the distributions. This is mainly determined by the function λ_σ and the corresponding standard deviation σ.
Formula (3.4) entails that the conditional density $f_{Y|X=x}(y|x)$ is equal to the function λ_σ shifted over $x + \delta$. Since the random variable with density λ_σ has mean 0, we expect that $Y|X = x$ has mean $x + \delta$. Indeed, this statement is enclosed in the following proposition:

Proposition 8 *If the density of variable $Y|X = x$ is given by:*

$$f_{Y|X=x}(y|x) = \lambda_\sigma(y - x - \delta) \quad \textit{for } x > 0$$

then:

$$E(Y|X = x) = x + \delta \quad \textit{for } x > 0$$

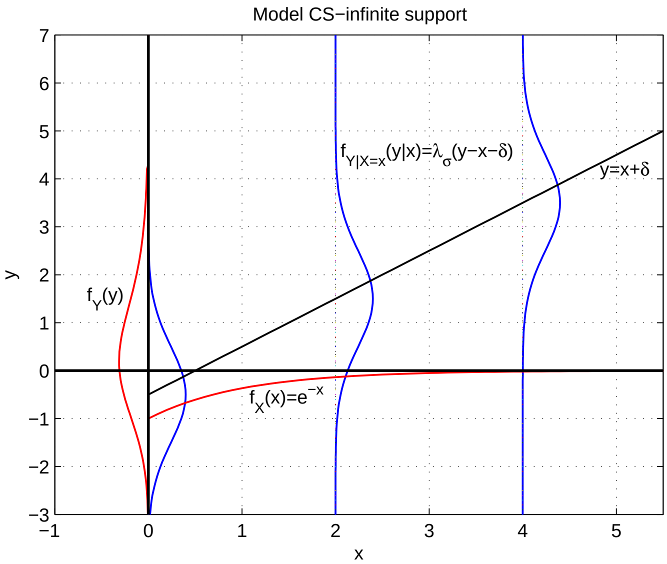


Figure 3.1: The Constant Spread Model.

Proof Using the "zero mean-property" of density function λ_σ we obtain the following calculations:

$$
\begin{aligned}
E(Y|X=x) &= \int_{-\infty}^{\infty} y f_{Y|X=x}(y|x)dy = \int_{-\infty}^{\infty} y\lambda_\sigma(y-x-\delta)dy \\
&= \int_{-\infty}^{\infty} (z+x+\delta)\lambda_\sigma(z)dz = \int_{-\infty}^{\infty} z\lambda_\sigma(z)dz + \int_{-\infty}^{\infty} (x+\delta)\lambda_\sigma(z)dz \\
&= 0 + (x+\delta) = x+\delta
\end{aligned}
$$

Thus, the proof is accomplished.

The last proposition entails that the successive means of variables $Y|X=x$, for $x>0$ follow a linear function $x+\delta$. It is worth noticing that $f_X(x)$ and $f_{Y|X=x}(y|x)$ describe completely a dependency model. In fact, knowing these densities is equivalent to knowing the bivariate density function $f_{X,Y}(x,y)$ (see equation (3.5)).

3.1.1 The distribution function of variable Y

In this section the distribution function F_Y of variable Y will be discussed. This distribution can be derived from (3.5) as follows:

$$
\begin{aligned}
F_Y(y) &= \int_{-\infty}^{y} \left\{ \int_0^{\infty} f_{X,Y}(x,t)dx \right\} dt = \int_{-\infty}^{y} \left\{ \int_0^{\infty} e^{-x}\lambda_\sigma(t-x-\delta)dx \right\} dt \\
&= \int_0^{\infty} \left\{ \int_{-\infty}^{y} e^{-x}\lambda_\sigma(t-x-\delta)dt \right\} dx \\
&= \int_0^{\infty} e^{-x}\Lambda_\sigma(y-x-\delta)dx \qquad (3.6)
\end{aligned}
$$

Defining $\bar{F}_Y = 1 - F_Y$, $\bar{\Lambda}_\sigma = 1 - \Lambda_\sigma$ and substituting $t = y - x - \delta$ yields:

$$\begin{aligned}
\bar{F}_Y(y) &= 1 - F_Y(y) = 1 - \int_0^\infty e^{-x}\Lambda_\sigma(y-x-\delta)dx = \int_0^\infty e^{-x}dx - \int_0^\infty e^{-x}\Lambda_\sigma(y-x-\delta)dx \\
&= \int_0^\infty e^{-x}(1-\Lambda_\sigma(y-x-\delta))dx = \int_0^\infty e^{-x}\bar{\Lambda}_\sigma(y-x-\delta)dx \\
&= e^{\delta-y}\int_{-\infty}^{y-\delta} e^t\bar{\Lambda}_\sigma(t)dt \qquad (3.7)
\end{aligned}$$

The relation (3.7) entails that the density f_Y of variable Y can be expressed as follows:

$$\begin{aligned}
f_Y(y) &= \frac{d}{dy}F_Y(y) = -\frac{d}{dy}\bar{F}_Y(y) = -\frac{d}{dy}\left(e^{\delta-y}\int_{-\infty}^{y-\delta} e^t\bar{\Lambda}_\sigma(t)dt\right) \\
&= -\left(-e^{\delta-y}\int_{-\infty}^{y-\delta} e^t\bar{\Lambda}_\sigma(t)dt + e^{\delta-y}e^{y-\delta}\bar{\Lambda}_\sigma(y-\delta)\right) \\
&= e^{\delta-y}\int_{-\infty}^{y-\delta} e^t\bar{\Lambda}_\sigma(t)dt - \bar{\Lambda}_\sigma(y-\delta) \qquad (3.8)
\end{aligned}$$

Using relation (3.7) and the fact that $\bar{\Lambda}_\sigma(y) = 1$ for $y \le y_b$, we can conclude that for $y \le y_b + \delta$, $\bar{F}_Y(y) = 1$ and hence $f_Y(y) = 0$. Before we proceed with the next proposition, we will prove the following lemma:

Lemma 1 *If Λ_σ is a distribution function with corresponding density λ_σ and T is a random variable such that $P\{T \le t\} = \Lambda_\sigma(t)$ then:*

$$E(e^T) = \int_{-\infty}^\infty e^t\lambda_\sigma(t)dt = \int_{-\infty}^\infty e^t\bar{\Lambda}_\sigma(t)dt$$

where $\bar{\Lambda}_\sigma = 1 - \Lambda_\sigma$.

Proof Here we only treat the case in which λ_σ has no discontinuity points. The more general case is not really more complicated and is treated in [5]. Partial integration yields:

$$\int_{-\infty}^z e^t\bar{\Lambda}_\sigma(t)dt = e^t\bar{\Lambda}_\sigma(t)|_{-\infty}^z + \int_{-\infty}^z e^t\lambda_\sigma(t)dt$$

To finish the proof, it is enough to show that $\lim_{z\to\infty} e^z\bar{\Lambda}_\sigma(z) = 0$. Indeed, using de l'Hopital's rule we obtain:

$$\lim_{z\to\infty} e^z\bar{\Lambda}_\sigma(z) = \lim_{z\to\infty}\frac{\bar{\Lambda}_\sigma(z)}{e^{-z}} = \lim_{t\to\infty}\frac{\lambda_\sigma(t)}{e^{-z}} = \lim_{t\to\infty}\lambda_\sigma(z)e^z = 0$$

where the last step follows from condition (3.2). Hence, the proof is accomplished.

Proposition 9 *The condition $E(e^T) = \int_{-\infty}^\infty e^t\lambda_\sigma(t)dt < \infty$ is equivalent to:*

$$\int_{-\infty}^\infty e^t\bar{\Lambda}_\sigma(t)dt < \infty$$

Proof From Lemma 1 it follows that:

$$E(e^T) = \int_{-\infty}^{\infty} e^t \lambda_\sigma(t)dt = \int_{-\infty}^{\infty} e^t \bar{\Lambda}_\sigma(t)dt$$

The assumption $E(e^T) < \infty$ yields that $\int_{-\infty}^{\infty} e^t \bar{\Lambda}_\sigma(t)dt < \infty$. Thus, the proof is accomplished.

Below, we introduce the definition of a strictly convex function and discuss a modification of "Jensen's inequality". Compared with the ordinary inequality, the modification deals with the strictness of the convexity.

Definition 11 *Function $f : \Re \to \Re$ is called strictly convex, if:*

$$\forall_{x,y\in\Re}\forall_{\alpha\in(0,1)} f((1-\alpha)x + \alpha y) < (1-\alpha)f(x) + \alpha f(y)$$

Proposition 10 ***Modification of Jensen's inequality*** *If f is a strictly convex function, then:*

$$E(f(T)) > f(E(T))$$

provided the expectations exist.

Proof Let us assume that T is a random variable with density function λ_σ. We note that if function f is strictly convex, then for all $x_0 \in \Re$ there exists a linear function $k(x) = ax + b$ for $a, b \in \Re$ such that:

$$f(x_0) = k(x_0) \quad \text{and} \quad k(x) = ax + b < f(x) \qquad \text{for} \quad x \neq x_0 \tag{3.9}$$

Suppose $x_0 = E(T)$. Then, observation (3.9) entails that:

$$\int_{-\infty}^{\infty} (f(x) - ax - b)\lambda_\sigma(x)dx > 0$$

which leads to:

$$E(f(T)) > aE(T) + b = f(E(T))$$

Thus, the proof is accomplished.

Let us assume that T is a random variable with density function λ_σ. We observe that function $f(t) = e^t, t \in \Re$ is strictly convex. Hence, applying Proposition 10 results in:

$$E(e^T) > e^{E(T)} = 1 \tag{3.10}$$

since T is a random variable with zero mean. For further use, we introduce the following quantity:

$$\delta_0 = -\ln(E(e^T)) \tag{3.11}$$

We observe that thanks to inequality (3.10) quantity δ_0 is strictly less than 0. Moreover, if we consider λ_σ as a normal density function with zero mean and standard deviation σ, by relation (3.3) we obtain:

$$\delta_0 = -\frac{\sigma^2}{2}$$

3.1.2 Situation $y_e < \infty$

In this section we will show that in case of a density λ_σ with finite y_e, the distribution of Y becomes a exactly shifted exponential for sufficiently large values of y. Consider $y_e < \infty$ and $y \geq y_e + \delta$. Because $\bar{\Lambda}_\sigma(t) = 0$ for $t > y_e$, we observe that the function $\bar{F}_Y$ given by relation (3.7) simplifies to:

$$\bar{F}_Y(y) = e^{\delta - y} \int_{-\infty}^{y_e} e^t \bar{\Lambda}_\sigma(t) dt$$

Then by Lemma 1 for $y_e < \infty$ and $y \geq y_e + \delta$, we obtain:

$$\bar{F}_Y(y) = e^{\delta - y} \int_{-\infty}^{y_e} e^t \bar{\Lambda}_\sigma(t) dt = e^{\delta - y} E(e^T) = e^{-y+\delta-\delta_0}$$

This implies that for large y the random variable Y becomes exponential with scale parameter 1 and location parameter $\delta - \delta_0$. If we substitute $\delta = \delta_0$ then, the location parameter becomes 0 and Y becomes a standard exponential random variable. Note that the value y_e implicitly depends on the standard deviation σ of variable Y. Larger values of σ lead to larger values of y_e.

3.1.3 Situation $y_e = \infty$

In this subsection the situation $y_e = \infty$ will be discussed. By Lemma 1 the following relation holds:

$$\lim_{y \to \infty} \int_{-\infty}^{y-\delta} e^t \bar{\Lambda}_\sigma(t) dt = E(e^T) \tag{3.12}$$

Then, for sufficiently large y, we can write:

$$\int_{-\infty}^{y-\delta} e^t \bar{\Lambda}_\sigma(t) dt \approx E(e^T)$$

Combining this result with the exact form of $\bar{F}_Y$, we obtain for large y that:

$$\bar{F}_Y(y) \approx e^{\delta - y} E(e^T) = e^{-y+\delta-\delta_0} \tag{3.13}$$

The accuracy of approximation (3.13) is determined by the rate at which the function on the left hand side of (3.12) approaches $E(e^T)$. This happens faster if the tail of density λ_σ decreases more rapidly. If it decreases slowly the usefulness of this approximation might be questionable. Therefore, the quality of this approximation depends on the type of the considered density function. It is worth noticing that large values of the standard deviation σ provide a slower decay of the tail than smaller values of σ.

Clearly, differentiation of relation (3.13) yields an approximation of the density function f_Y of variable Y:

$$f_Y(y) \approx e^{-y+\delta-\delta_0}, \quad \text{for large} \quad y \tag{3.14}$$

Example 1 ***λ_σ-Exponential random variable with mean zero and standard deviation*** $\sigma \geq 1$ *Let us recall that the exponential random variable with zero mean and standard deviation $\sigma \geq 1$ does not satisfy condition (3.2). If density function λ_σ is given by:*

$$\lambda_\sigma(t) = e^{-(t+1)}, \quad for \ \ t > -1$$

then, as can be shown by explicit calculations, the density f_Y of variable Y is given as follows:

$$f_Y(y) = (y - \delta + 1)e^{-y+\delta-1}, \quad for \ \ y > \delta - 1$$

Definition 12 *The gamma density h with shape parameter α and scale parameter β is given as follows:*

$$h(x; \alpha, \beta) = \frac{x^{\alpha-1}\beta^{\alpha}e^{-\beta x}}{\Gamma(\alpha)}, \quad for \ \ x > 0$$

where function Γ is:

$$\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}dt$$

It is easy to check that, if z is a positive integer, then $\Gamma(z) = (z-1)!$.
We observe that:

$$f_Y(y) = h(y - \delta + 1; 2, 1)$$

3.1.4 The conditional variable X given $Y = y$

In the previous section the conditional variable Y given $X = x$ has been discussed, whereas now, we would like to focus on the opposite case. In order to treat $f_{X|Y=y}(x|y)$, it is useful to introduce the characteristic function χ of the interval $(0, \infty)$:

$$\chi(x) = \begin{cases} 1 & x > 0 \\ 0 & x \leq 0 \end{cases}$$

Since the joint density function $f_{X,Y}(x, y)$ is given by relation (3.5), the conditional density $f_{X|Y=y}(x|y)$ is given by:

$$f_{X|Y=y}(x|y) = \frac{f_{X,Y}(x, y)}{f_Y(y)} = \frac{e^{-x}\chi(x)\lambda_\sigma(y - x - \delta)}{f_Y(y)}, \quad \text{for} \ \ y > y_b + \delta$$

We introduce an additional condition for function λ_σ:

$$E(Te^T) = \int_{-\infty}^{\infty} te^t\lambda_\sigma(t) < \infty \tag{3.15}$$

The above restriction is satisfied by a normal distribution with arbitrary standard deviation σ, as well as by the exponential distribution with standard deviation $\sigma < 1$. In practice the requirements (3.2) and (3.15) mean approximately the same.
We introduce a function γ as follows:

$$\gamma(t) = e^{\delta_0 - t}\lambda_\sigma(-t)$$

Since the functions $e^{\delta_0 - t}$, $\lambda_\sigma(-t)$ are nonnegative, using $\delta_0 = -\ln E(e^T)$, it is straightforward to check that γ integrates to 1 and hence constitutes a probability density function. The corresponding average value μ_γ is then given by:

$$\mu_\gamma = \int_{-\infty}^{\infty} t\gamma(t)dt = -e^{\delta_0}\int_{-\infty}^{\infty} te^t\lambda_\sigma(t)dt$$

The above result makes clear why condition (3.15) is required.
Moreover, it can be shown that:

$$\mu_\gamma < 0$$

For the proof we refer to [5].
We observe that the joint density $f_{X,Y}(x,y) = e^{-x}\chi(x)\lambda_\sigma(y-x-\delta)$ can be equivalently expressed in terms of the function γ, indeed:

$$\begin{aligned} f_{X,Y}(x,y) &= e^{-x}\chi(x)\lambda_\sigma(y-x-\delta) \\ &= e^{-x}\chi(x)e^{x+\delta-y-\delta_0}\gamma(x-y+\delta) \\ &= e^{-y+\delta-\delta_0}\chi(x)\gamma(x-y+\delta) \end{aligned} \tag{3.16}$$

From relation (3.16) we obtain a formula for the conditional density function $f_{X|Y=y}(x|y)$ in terms of function γ:

$$f_{X|Y=y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{e^{-y+\delta-\delta_0}}{f_Y(y)}\chi(x)\gamma(x-y+\delta), \quad \text{for} \quad y > y_b + \delta \tag{3.17}$$

3.1.5 Situation $y_e < \infty$

In this subsection we consider the situation when $y_e < \infty$ and $y \geq y_e + \delta$. Since $\bar{F}_Y(y) = e^{-y+\delta-\delta_0}$ for $y_e < \infty$ and $y \geq y_e + \delta$ we have:

$$f_Y(y) = -\frac{d}{dy}\bar{F}_Y(y) = e^{-y+\delta-\delta_0}$$

Substituting f_Y into equation (3.17) results in:

$$f_{X|Y=y}(x|y) = \gamma(x-y+\delta)\chi(x), \quad \text{for} \quad y \geq y_e + \delta, y_e < \infty$$

The assumption $x < 0$ leads to $x - y + \delta \leq x - y_e < -y_e$. Then, the definition of function γ entails that $\gamma(x-y+\delta) = 0$. Thus, the presence of the characteristic function χ is redundant and we can write:

$$f_{X|Y=y}(x|y) = \gamma(x-y+\delta), \quad \text{for} \quad y \geq y_e + \delta, y_e < \infty$$

This implies that for $y \geq y_e + \delta$ the average values that correspond to density $f_{X|Y=y}(x|y)$ lie on a straight line. Indeed, for $y \geq y_e + \delta$ we obtain the following calculations:

$$\begin{aligned} E(X|Y=y) &= \int_0^\infty x f_{X|Y=y}(x|y)dx \\ &= \int_0^\infty x\gamma(x-y+\delta)dx = \int_{-y-\delta}^{\infty} t\gamma(t)dt + y - \delta \\ &= \mu_\gamma + y - \delta \end{aligned}$$

For $y < y_e + \delta$ the above relation is not satisfied.
Let us recall that $E(Y|X = x) = x + \delta$. Then, we observe that $E(Y|X = x)$ and $E(X|Y = y)$ are parallel and the distance between them is $|\mu_\gamma|$, for $y \geq y_e + \delta$.

3.1.6 Situation $y_e = \infty$

In this subsection we assume that $y_e = \infty$. Let us recall that in this case the density function f_Y of variable Y can be approximated by the right-hand side of relation (3.14). Using this fact and the observation (3.17) results in the approximation of the conditional density function $f_{X|Y=y}(x|y)$:

$$f_{X|Y=y}(x|y) \approx \gamma(x - y + \delta)\chi(x), \quad \text{for large} \quad y \tag{3.18}$$

We note that for $x < 0$ and for large values of y the function $\gamma(x - y + \delta)$ becomes close to zero. Therefore, the term χ can be safely omitted. Hence, we can write:

$$f_{X|Y=y}(x|y) \approx \gamma(x - y + \delta), \quad \text{for large} \quad y \tag{3.19}$$

In fact, the relation (3.19) is more practical than result (3.18), because the right-hand side of relation (3.18) is not a density function. The integration over all x values does not appear to be 1. In order to show this we will first prove that $f_Y(y) < e^{-y+\delta-\delta_0}$. Indeed, using relations (3.8) and (3.11) we obtain:

$$\begin{aligned}
f_Y(y) &= e^{\delta-y}\int_{-\infty}^{y-\delta} e^t\bar{\Lambda}_\sigma(t)dt - \bar{\Lambda}_\sigma(y-\delta) \\
&\leq e^{\delta-y}\int_{-\infty}^{\infty} e^t\bar{\Lambda}_\sigma(t)dt - \bar{\Lambda}_\sigma(y-\delta) \\
&= e^{\delta-y}E(e^T) - \bar{\Lambda}_\sigma(y-\delta) = e^{\delta-y-\delta_0} - \bar{\Lambda}_\sigma(y-\delta) \\
&< e^{\delta-y-\delta_0}
\end{aligned}$$

The above result and relation (3.16) yield:

$$\begin{aligned}
\int_0^\infty \chi(x)\gamma(x-y+\delta)dx &= \int_0^\infty \frac{f_Y(y)f_{X|Y=y}(x|y)}{e^{-y+\delta-\delta_0}}dx = \frac{f_Y(y)}{e^{-y+\delta-\delta_0}}\int_0^\infty f_{X|Y=y}(x|y)dx \\
&= \frac{f_Y(y)}{e^{-y+\delta-\delta_0}} < 1
\end{aligned}$$

Due to approximation (3.19), we obtain the approximation of $E(X|Y = y)$, for large values of y:

$$E(X|Y = y) \approx \mu_\gamma + y - \delta$$

Hence, for large y the average values tend to lay on a straight line, what is consistent with the situation $y_e < \infty$.

Example 2 *Because of the importance of the normal density in applications, we consider the situation in which λ_σ is a normal density function with zero mean and* *standard deviation $\sigma > 0$. In this case δ_0 takes the value $-\sigma^2/2$ and the function γ, as can be verified, is given by:*

$$\gamma(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(t+\sigma^2)^2}{2\sigma^2}\right)$$

We note that γ is a normal density with mean $\mu_\gamma = -\sigma^2$ and standard deviation σ. Therefore, γ and λ_σ belong to the same class of density functions. In general, this will not be the case. The above result for function γ and approximation (3.14) leads to:

$$f(x|y) \approx \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{-[x-(y-\delta-\sigma^2)]^2}{2\sigma^2}\right), \quad \textit{for large} \quad y$$

Hence, for large values of y the density $f(x|y)$ tends to be a normal density with mean $(y-\delta-\sigma^2)$ and standard deviation σ.
In applications, δ equal to δ_0 is usually chosen. In this case the function g becomes asymptotically standard exponential and $E(X|Y=y) \approx y-\delta-\sigma^2 = y-\sigma^2/2$. Thus, the line $E(X|Y=y)$ tends to lie above the line $y=x$ for large values of y, whereas, the function $E(Y|X=x) = x-\sigma^2/2$ lies below the line $y=x$.

3.1.7 The role of parameter δ

In this subsection, we investigate the role of parameter δ. Consider two Constant Spread Models:

1. $f_{X,Y}(x,y;\sigma,\delta_1)$ with transformation: $\underbrace{F_V(v) = F_X(x), F_W(w) = F_Y(y;\sigma,\delta_1)}_{\text{System 1}}$
2. $f_{X,Y}(x,y;\sigma,\delta_2)$ with transformation: $\underbrace{F_V(v) = F_X(x), F_W(w) = F_Y(y;\sigma,\delta_2)}_{\text{System 2}}$

Solving the Systems 1 and 2 with respect to x and y we obtain:

1. For System 1: $x(v) = F_X^{-1}(F_V(v))$ and $y_1(w) = F_Y^{-1}(F_W(w);\delta_1)$
2. For System 2: $x(v) = F_X^{-1}(F_V(v))$ and $y_2(w) = F_Y^{-1}(F_W(w);\delta_2)$

Since $y_1(w)$ and $y_2(w)$ must satisfy the second equations of Systems 1 and 2, we obtain:

$$\begin{aligned}
F_W(w) &= F_Y(y_1(w);\delta_1) = F_Y(y_2(w);\delta_2) \\
F_W(w) &= \int_0^\infty e^{-t}\lambda_\sigma(y_1(w)-t-\delta_1)dt = \int_0^\infty e^{-t}\lambda_\sigma(y_2(w)-t-\delta_2)dt
\end{aligned}$$

Then $y_1(w)-\delta_1 = y_2(w)-\delta_2$. Using this result and definitions of $f_{X,Y}(x,y)$, $f_X(x)$ and $f_Y(y)$ we obtain:

$$\begin{aligned}
f_{V,W}(v,w;\sigma,\delta_1) &= \frac{e^{-x(v)}\lambda_\sigma(y_1(w)-x(v)-\delta_1)f_V(v)f_W(w)}{e^{-x(v)}\int_0^\infty e^{-t}\lambda_\sigma(y_1(w)-t-\delta_1)dt} \\
&= \frac{e^{-x(v)}\lambda_\sigma(y_2(w)-x(v)-\delta_2)f_V(v)f_W(w)}{e^{-x(v)}\int_0^\infty e^{-t}\lambda_\sigma(y_2(w)-t-\delta_2)dt} \\
&= f_{V,W}(v,w;\sigma,\delta_2)
\end{aligned}$$

Hence, the Constant Spread Model in the original space does not depend on the value of parameter δ - it is invariant under the choice of δ. Therefore, δ can be safely omitted, for simplicity we propose to assume $\delta = 0$.

3.2 Variable Spread Model

In the previous section we have discussed the Constant Spread Model, where the spread of the variable $Y|X = x$ stays constant for all $x > 0$. In this section, a more general model will be considered. This model allows for changes of the spread of variable $Y|X = x$ for all values of $x > 0$. It is called the Variable Spread Model. This model was first considered in [6].
We assume that X and Y are two random variables with cumulative distribution functions $F_X(x)$ and $F_Y(y)$, respectively. The random variable X is standard exponentially distributed, that is:

$$F_X(x) = 1 - e^{-x} \quad \text{for } x \geq 0$$

Hence, the corresponding density function is given by:

$$f_X(x) = e^{-x} \quad \text{for } x \geq 0$$

It is assumed that the conditional density function $f_{Y|X=x}(y|x)$ of variable $Y|X = x$ takes the following form:

$$f_{Y|X=x}(y|x) = \lambda_{\sigma(x)}(y - x - \delta)$$

where $\sigma(x)$ is a strictly positive function defined on the interval $[0, \infty)$, δ is some chosen constant, and $\lambda_\sigma(t)$ is a density function with mean zero and standard deviation $\sigma > 0$.
We note that the above assumptions characterize the whole model, since knowing them, the joint density of variables X and Y can be derived:

$$f_{X,Y}(x, y) = f_X(x) f_{Y|X=x}(y|x) = e^{-x} \lambda_{\sigma(x)}(y - x - \delta)$$

The construction of this model is illustrated in Figure 3.2. The picture helps us to understand the idea behind this model. Moreover, due to the construction we can easily observe what occurs in the upper tail of the distribution. This is determined by the form of the "spread function" $\sigma(x)$ and density $\lambda_{\sigma(x)}$, for $x \in [0, \infty)$.
The distribution function $F_Y(y)$ of variable Y can be calculated as follows:

$$F_Y(y) = \int_{-\infty}^{y} f_{X,Y}(x, t) dt = \int_0^\infty e^{-x} \Lambda_{\sigma(x)}(y - x - \delta) dx$$

where $\Lambda_\sigma(t)$ is the cumulative distribution function that corresponds to λ_σ.
Moreover, as in case of the Constant Spread Model the parameter δ is redundant, and in the applications in this report we put $\delta = 0$. Moreover, in the case study we will consider a normal distribution function for $\lambda_\sigma(t)$.

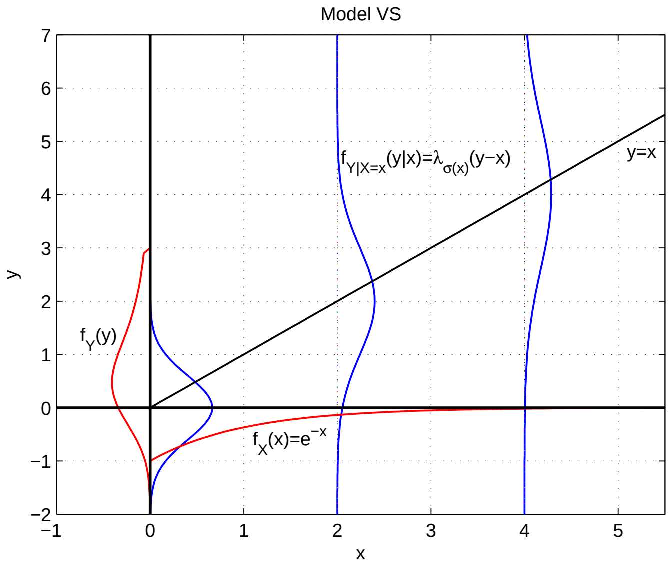


Figure 3.2: The Variable Spread Model.

3.3 Constant Symmetric Spread Model

In this section, the Constant Symmetric Spread Model will be introduced. It was developed by Duits and van Noortwijk (see [4]). Moreover, this model was also discussed in [5].

Assume that X and Y are two random variables with distribution functions $F_X(x)$ and $F_Y(y)$, respectively. Suppose that $\Lambda_\sigma(t)$ is a cumulative distribution function of a normally distributed random variable with mean zero and standard deviation $\sigma > 0$. Let us define the quantity s_0 as follows:

$$s_0 = \ln\{2(1 - \Lambda_{2/\sigma}(1))\} \leq 0$$

and for $x \geq 0$:

$$\mu(x) = x - \frac{\sigma^2}{2}$$

$$A(x) = 1 - \Lambda_\sigma(s_0 - \mu(x))$$

The model is obtained by defining $f_X(x)$ and $f_{Y|X=x}(y|x)$. The density function $f_X(x)$ of variable X is given by:

$$f_X(x) = A(x)e^{-x}, x \geq s_0$$

We note that the choice of s_0 indeed implies that $f_X(x)$ integrates to 1, for $x \geq s_0$ (this can be shown using partial integration).

The conditional density $f_{Y|X=x}(y|x)$ of variable $Y|X = x$ is:

$$f_{Y|X=x}(y|x) = \frac{1}{A(x)} \frac{1}{\sigma\sqrt{2\pi}} \exp\{-(y - \mu(x))^2/2\sigma^2\}, x \geq s_0, y \geq s_0$$

This information is enough to build the complete model, since knowing $f_X(x)$ and $f_{Y|X=x}(y|x)$, is equivalent with having $f_{X,Y}(x,y)$. Figure 3.3 illustrates this model. Observe that we get an impression of the tail behavior of the vector (X,Y), due to

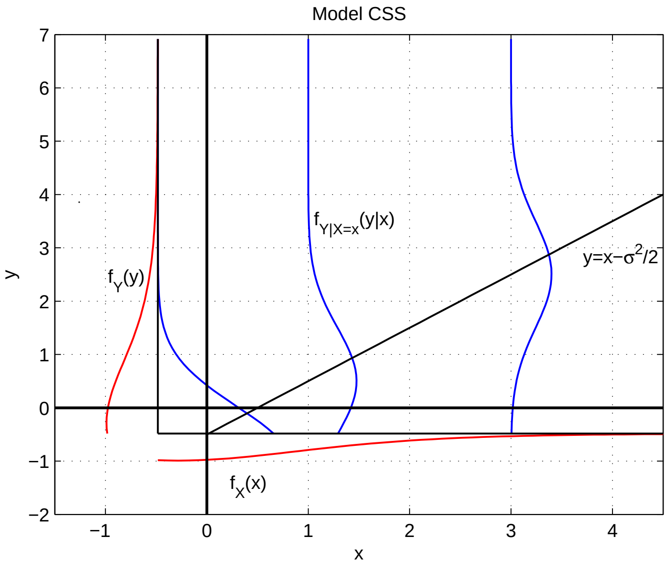


Figure 3.3: The Constant Symmetric Spread Model.

the definition of the conditional density function $f_{Y|X=x}(y|x)$.

Remark 17 *Note that function $A(x) \to 1$ as $x \to \infty$. Hence, the random variable X is asymptotically standard exponential. Moreover, the formula for $f_{Y|X=x}(y|x)$ yields that $Y|X = x$ can be considered as a left-truncated modified normal variable.*

Combining expressions for $f_{Y|X=x}(y|x)$ and $f_X(x)$ leads to the density function $f_Y(y)$ of variable Y. As can be shown in a straightforward way:

$$f_Y(y) = \int_{s_0}^{\infty} f_X(x) f_{Y|X=x}(y|x)dx = A(y)e^{-y}, y \geq s_0$$

We note that density $f_Y(y)$ has the same functional form as $f_X(x)$. Thus, the random variable Y is asymptotically standard exponentially distributed as well. Moreover, it can be shown that $f_{X|Y=y}(x|y)$ is the same as $f_{Y|X=x}(y|x)$ (in case x and y are exchanged). It confirms that the model is symmetric with respect to X and Y.

Remark 18 *The symmetry entails that $F_X = F_Y$ as well as $F_X^{-1} = F_Y^{-1}$.*

3.4 Tail dependence

In this section the definitions and propositions concerning the so-called tail dependence coefficients will be introduced. The tail dependence coefficients provide some kind of measure of the dependence in the upper and lower tails of the bivariate distribution. The tail dependence coefficients are defined in the bivariate case. More

information about this theory can be found in [14].
The tail dependence coefficients will be computed for the Constant Spread Model, the Variable Spread Model and the Constant Symmetric Spread Model.

Definition 13 *Let (X, Y) be a bivariate vector of continuous random variables with marginal distribution functions F_X and F_Y. The lower tail dependence coefficient is defined as follows:*

$$\lambda^L = \lim_{u \downarrow 0} P\{Y \leq F_Y^{-1}(u) | X \leq F_X^{-1}(u)\} = \lim_{u \downarrow 0} P\{X \leq F_X^{-1}(u) | Y \leq F_Y^{-1}(u)\}$$

provided that this limit exists.

We say that X and Y are lower tail dependent if $\lambda^L \in (0, 1]$ and lower tail independent if $\lambda^L = 0$. Lower tail dependence arises when there is a positive probability of small outliers occurring jointly.

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

$$\lambda^U = \lim_{u \uparrow 1} P\{Y > F_Y^{-1}(u) | X > F_X^{-1}(u)\} = \lim_{u \uparrow 1} P\{X > F_X^{-1}(u) | Y > F_Y^{-1}(u)\}$$

provided that this limit exists.

We say that X and Y are upper tail dependent if $\lambda^L \in (0, 1]$ and upper tail independent if $\lambda^U = 0$. Upper tail dependence arises when there is a positive probability of large outliers occurring jointly.

Proposition 11 *Let (X, Y) be a bivariate vector of continuous random variables with marginal distribution functions F_X and F_Y and the associated copula $C(u, v)$, such that it is differentiable with respect to u and v. Then the lower tail dependence coefficient is given as follows:*

$$\lambda^L = \lim_{u \downarrow 0}(P\{Y \leq F_Y^{-1}(u) | X = F_X^{-1}(u)\} + P\{X \leq F_X^{-1}(u) | Y = F_Y^{-1}(u)\}) \quad (3.20)$$

whereas the upper tail dependence coefficient is:

$$\lambda^U = \lim_{u \uparrow 1}(P\{Y > F_Y^{-1}(u) | X = F_X^{-1}(u)\} + P\{X > F_X^{-1}(u) | Y = F_Y^{-1}(u)\}) \quad (3.21)$$

Provided that the above limits exist.

Proof Let C denote the copula corresponding to (X, Y). Applying the definition of copula function and de l'Hopital's rule we obtain:

$$\begin{aligned}
\lambda^L &= \lim_{u \downarrow 0} P\{Y \leq F_Y^{-1}(u) | X \leq F_X^{-1}(u)\} \\
&= \lim_{u \downarrow 0} \frac{P\{X \leq F_X^{-1}(u), Y \leq F_Y^{-1}(u)\}}{P\{X \leq F_X^{-1}(u)\}} \\
&= \lim_{u \downarrow 0} \frac{C(u, u)}{u} \\
&= \lim_{u \downarrow 0} \left(\frac{\partial}{\partial s} C(s, t)|_{s=t=u} + \frac{\partial}{\partial t} C(s, t)|_{s=t=u} \right)
\end{aligned}$$

$$
\begin{aligned}
\lambda^U &= \lim_{u\uparrow 1} P\{Y > F_Y^{-1}(u) | X > F_X^{-1}(u)\} \\
&= \lim_{u\uparrow 1} \frac{P\{X > F_X^{-1}(u), Y > F_Y^{-1}(u)\}}{P\{X > F_X^{-1}(u)\}} \\
&= \lim_{u\uparrow 1} \frac{1 - P\{X \le F_X^{-1}(u)\} - P\{Y \le F_Y^{-1}(u)\} + P\{X \le F_X^{-1}(u), Y \le F_Y^{-1}(u)\}}{1 - P\{X \le F_X^{-1}(u)\}} \\
&= \lim_{u\uparrow 1} \frac{1 - 2u + C(u,u)}{1-u} \\
&= \lim_{u\uparrow 1} \left(2 - \frac{\partial}{\partial s} C(s,t)|_{s=t=u} - \frac{\partial}{\partial t} C(s,t)|_{s=t=u}\right)
\end{aligned}
$$

We note that:

$$
\begin{aligned}
&P\{Y \le F_Y^{-1}(u) | X = F_X^{-1}(u)\} = P\{F_Y(Y) \le u | F_X(X) = u\} \\
&= \frac{P\{F_X(X) = u, F_Y(Y) \le u\}}{P\{F(X) = u\}} \\
&= \lim_{\Delta u \downarrow 0} \frac{P\{u \le F_X(X) < u + \Delta u, F_Y(Y) \le u\}}{P\{u \le F_X(X) < u + \Delta u\}} \\
&= \lim_{\Delta u \downarrow 0} \frac{P\{F_X(X) < u + \Delta u, F_Y(Y) \le u\} - P\{F_X(X) \le u, F_Y(Y) \le u\}}{P\{F_X(X) < u + \Delta u\} - P\{F_X(X) \le u\}} \\
&= \lim_{\Delta u \downarrow 0} \frac{C(u + \Delta u, u) - C(u,u)}{\Delta u} \\
&= \frac{\partial}{\partial s} C(s,t)|_{(s,t)=(u,u)}
\end{aligned}
$$

analogously:

$$
P\{X \le F_X^{-1}(u) | Y = F_Y^{-1}(u)\} = \frac{\partial}{\partial t} C(s,t)|_{(s,t)=(u,u)}
$$

Rearranging the expressions above, we obtain:

$$
\lambda^L = \lim_{u\downarrow 0} (P\{Y \le F_Y^{-1}(u) | X = F_X^{-1}(u)\} + P\{X \le F_X^{-1}(u) | Y = F_Y^{-1}(u)\})
$$

$$
\lambda^U = \lim_{u\uparrow 1} (P\{Y > F_Y^{-1}(u) | X = F_X^{-1}(u)\} + P\{X > F_X^{-1}(u) | Y = F_Y^{-1}(u)\})
$$

Hence, the proof is accomplished.

Corollary 1 *If $P\{X \le x, Y \le y\} = P\{X \le y, Y \le x\}$ (symmetric) then Proposition 11 is equivalent with:*

$$
\lambda^L = \lim_{u\downarrow 0} 2P\{Y \le F_X^{-1}(u) | X = F_X^{-1}(u)\}
$$

and

$$
\lambda^U = \lim_{u\uparrow 1} 2P\{Y > F_X^{-1}(u) | X = F_X^{-1}(u)\}
$$

The symmetry implies that $C(s,t) = C(t,s)$ and hence:

$$\frac{\partial}{\partial s}C(s,t)|_{(s,t)=(u,u)} = \frac{\partial}{\partial s}C(t,s)|_{(s,t)=(u,u)} = \frac{\partial}{\partial t}C(s,t)|_{(s,t)=(u,u)}$$

The corollary then easily follows from the proof of Proposition 11.

Remark 19 *It is well known that if the limits $\lim_{x\to c} f(x)$ and $\lim_{x\to c} g(x)$ exist then:*

$$\lim_{x\to c}(f(x)+g(x)) = \lim_{x\to c} f(x) + \lim_{x\to c} g(x)$$

If Proposition 11 is applied for the upper tail dependence coefficient, we will usually start with splitting the limit (3.21). As we will find out in all considered cases the limits:

$$\lim_{u\uparrow 1} P\{Y > F_Y^{-1}(u)|X = F_X^{-1}(u)\}$$
$$\lim_{u\uparrow 1} P\{X > F_X^{-1}(u)|Y = F_Y^{-1}(u)\}$$

exist.

3.4.1 The tail dependence for Constant Spread Model with finite endpoints

In this subsection, the tail dependence coefficients for the Constant Spread Model with finite endpoints will be derived. Let us denote $x = F_X^{-1}(u)$ and $y = F_Y^{-1}(u)$ for $u \in [0,1]$. We observe that $x \downarrow 0$ and $y \downarrow \delta + y_b$ as $u \downarrow 0$. We start with the calculation of the probability:

$$\begin{aligned}
P\{Y \le y|X \le x\} &= \frac{P\{Y \le y, X \le x\}}{P\{X \le x\}} = \frac{\int_0^x \left\{\int_{w+\delta+y_b}^{y} e^{-w}\lambda_\sigma(v-w-\delta)dv\right\}dw}{1-e^{-x}} \\
&= \frac{\int_0^x \left\{\int_{y_b}^{y-w-\delta} e^{-w}\lambda_\sigma(z)dz\right\}dw}{1-e^{-x}} \le \frac{\int_0^x e^{-w}dw \int_{y_b}^{y-\delta}\lambda_\sigma(z)dz}{1-e^{-x}} \\
&= \frac{(1-e^{-x})\int_{y_b}^{y-\delta}\lambda_\sigma(z)dz}{1-e^{-x}} = \int_{y_b}^{y-\delta}\lambda_\sigma(z)dz
\end{aligned}$$

Taking the limit with respect to $u \downarrow 0$ yields that the lower tail dependence coefficient λ^L is:

$$\lambda^L = \lim_{u\downarrow 0} P\{Y \le y|X \le x\} = 0$$

Thus, we can conclude that lower tail independence occurs.
In order to compute the upper tail dependence coefficient λ^U, we apply Proposition

11. Let us first consider the following limit, where it is understood that x and y depend on u:

$$\begin{aligned}
\lim_{u\uparrow 1} P\{X > x|Y = y\} &= \lim_{u\uparrow 1} \int_{x}^{y-\delta-y_b} f_{X|Y=y}(k|y)dk \\
&= \lim_{u\uparrow 1} \frac{1}{f_Y(y)} \int_{x}^{y-\delta-y_b} e^{-k}\lambda_\sigma(y-k-\delta)dk \\
&= \lim_{u\uparrow 1} \frac{e^{-y+\delta}}{f_Y(y)} \int_{y_b}^{y-x-\delta} e^{z}\lambda_\sigma(z)dz
\end{aligned}$$

Since $x \to \infty$ and $y \to \infty$ as $u \uparrow 1$, we consider this limit for $y \geq y_e + \delta$, where the function $f_Y(y)$ becomes a shifted exponential, thus:

$$\begin{aligned}
\lim_{u\uparrow 1} P\{X > x|Y = y\} &= \lim_{u\uparrow 1} \frac{e^{-y+\delta}}{e^{-y+\delta-\delta_0}} \int_{y_b}^{y-x-\delta} e^{z}\lambda_\sigma(z)dz \\
&= \lim_{u\uparrow 1} \frac{1}{e^{-\delta_0}} \int_{y_b}^{y-x-\delta} e^{z}\lambda_\sigma(z)dz
\end{aligned}$$

To calculate the above limit, first observe that for $y \geq y_e + \delta$:

$$\begin{aligned}
\lim_{u\uparrow 1}(y-x-\delta) &= \lim_{u\uparrow 1}(F_Y^{-1}(u) - F_X^{-1}(u) - \delta) \\
&= \lim_{u\uparrow 1}(\delta - \delta_0 - \ln(1-u) + \ln(1-u) - \delta) = -\delta_0
\end{aligned}$$

Using this result, we obtain that:

$$\lim_{u\uparrow 1} P\{X > x|Y = y\} = e^{\delta_0} \int_{y_b}^{-\delta_0} e^{z}\lambda_\sigma(z)dz \tag{3.22}$$

An analogical procedure yields:

$$\begin{aligned}
\lim_{u\uparrow 1} P\{Y > y|X = x\} &= \lim_{u\uparrow 1} \int_{y}^{x+\delta+y_e} f_{Y|X=x}(k|x)dk \\
&= \lim_{u\uparrow 1} \frac{1}{f_X(x)} \int_{y}^{x+\delta+y_e} e^{-x}\lambda_\sigma(k-x-\delta)dk \\
&= \lim_{u\uparrow 1} \frac{1}{e^{-x}} \int_{y}^{x+\delta+y_e} e^{-x}\lambda_\sigma(k-x-\delta)dk \\
&= \lim_{u\uparrow 1} \int_{y}^{x+\delta+y_e} \lambda_\sigma(k-x-\delta)dk = \lim_{u\uparrow 1} \int_{y-x-\delta}^{y_e} \lambda_\sigma(z)dz \\
&= \int_{\lim_{u\uparrow 1}(y-x-\delta)}^{y_e} \lambda_\sigma(z)dz = \int_{-\delta_0}^{y_e} \lambda_\sigma(z)dz
\end{aligned} \tag{3.23}$$

Adding (3.22) and (3.23) entails that the upper tail dependence coefficient λ^U for the considered model is:

$$\lambda^U = e^{\delta_0} \int_{y_b}^{-\delta_0} e^{z}\lambda_\sigma(z)dz + \int_{-\delta_0}^{y_e} \lambda_\sigma(z)dz$$

The upper tail dependence coefficient λ^U should (from its definition) be smaller or equal to 1. Indeed, recalling that $\delta_0 < 0$, we conclude that $e^{\delta_0+z} < 1$ for $z \in [y_b, -\delta_0)$. Thus, the upper tail dependence coefficient can be bounded by:

$$\lambda^U < \int_{y_b}^{-\delta_0} \lambda_\sigma(z)dz + \int_{-\delta_0}^{y_e} \lambda_\sigma(z)dz = 1$$

where we have used the fact that λ_σ is a density function and takes only positive values on the interval $[y_b, y_e]$.

3.4.2 The tail dependence for Constant Spread Model with infinite endpoints

In this subsection the lower and upper tail dependence coefficients for the Constant Spread Model with the assumption of the infinite endpoints will be derived. For simplicity we denote $x = F_X^{-1}(u)$ and $y = F_Y^{-1}(u)$. We observe that $x \downarrow 0$ and $y \downarrow -\infty$ as $u \downarrow 0$. First, we consider the following probability:

$$\begin{aligned} P\{Y \leq y | X \leq x\} &= \frac{P\{X \leq x, Y \leq y\}}{P\{X \leq x\}} = \frac{\int_0^x \left\{\int_{-\infty}^{y} e^{-u} \lambda_\sigma(v-u-\delta)dv\right\} du}{1-e^{-x}} \\ &= \frac{\int_0^x \left\{\int_{-\infty}^{y-u-\delta} e^{-u} \lambda_\sigma(z)dz\right\} du}{1-e^{-x}} \leq \frac{\int_0^x \left\{\int_{-\infty}^{y-\delta} e^{-u} \lambda_\sigma(z)dz\right\} du}{1-e^{-x}} \\ &= \frac{(1-e^{-x}) \int_{-\infty}^{y-\delta} \lambda_\sigma(z)dz}{1-e^{-x}} = \int_{-\infty}^{y-\delta} \lambda_\sigma(z)dz \end{aligned}$$

Taking the limit with respect to $u \downarrow 0$ results in:

$$\lambda^L = \lim_{u \downarrow 0} P\{Y \leq y | X \leq x\} = 0$$

Hence, we conclude that lower tail independence occurs.
In order to derive the upper tail dependence coefficient λ_U we recall Proposition 11. First, we compute the following probability:

$$P\{Y > y | X = x\} = 1 - \Lambda_\sigma(y - x - \delta) \tag{3.24}$$

For the probability $P\{X > x | Y = y\}$ we obtain:

$$\begin{aligned} P\{X > x | Y = y\} &= \int_x^\infty f_{X|Y=y}(k|y)dk = \frac{\int_x^\infty e^{-k} \lambda_\sigma(y-k-\delta)dk}{\int_0^\infty e^{-k} \lambda_\sigma(y-k-\delta)dk} \\ &= \frac{\int_{-\infty}^{y-x-\delta} e^{z-y+\delta} \lambda_\sigma(z)dz}{\int_{-\infty}^{y-\delta} e^{z-y+\delta} \lambda_\sigma(z)dz} = \frac{\int_{-\infty}^{y-x-\delta} e^{z} \lambda_\sigma(z)dz}{\int_{-\infty}^{y-\delta} e^{z} \lambda_\sigma(z)dz} \end{aligned} \tag{3.25}$$

The calculation of the upper tail dependence coefficient λ^U requires the following proposition:

Proposition 12 *Under the assumptions of the Constant Spread Model the following relation holds:*

$$\lim_{u\uparrow 1}(y-x) = \lim_{u\uparrow 1}\{F_Y^{-1}(u) - F_X^{-1}(u)\} = \ln(e^{\delta}E(e^{Z}))$$

In order to prove Proposition 12, we apply Lemma 2, which is given as follows:

Lemma 2 *Under the assumption of the Constant Spread Model the following relation is satisfied:*

$$\lim_{y\to\infty}\frac{1-F_Y(y)}{e^{-y}} = e^{\delta}E(e^{Z})$$

Proof of Lemma 2 We observe that $\lim_{y\to\infty}(1-F_Y(y)) = 0$ and $\lim_{y\to\infty} e^{-y} = 0$. Hence, using de l'Hospital's rule we obtain:

$$\begin{aligned}
\lim_{y\to\infty}\frac{1-F_Y(y)}{e^{-y}} &= \lim_{y\to\infty}\frac{f(y)}{e^{-y}} \\
&= \lim_{y\to\infty}\frac{\int_0^{\infty} e^{-x}\lambda_{\sigma}(y-x-\delta)dx}{e^{-y}} \\
&= \lim_{y\to\infty}\frac{\int_{-\infty}^{y-\delta} e^{z-y+\delta}\lambda_{\sigma}(z)dz}{e^{-y}} \\
&= e^{\delta}\lim_{y\to\infty}\int_{-\infty}^{y-\delta} e^{z}\lambda_{\sigma}(z)dz = e^{\delta}E(e^{Z})
\end{aligned}$$

Hence, the proof is complete.

Proof of Proposition 12 We express the limit $\lim_{u\uparrow 1}\{F_Y^{-1}(u) - F_X^{-1}(u)\}$ in terms of variable y as follows:

$$F_Y^{-1}(u) = y \Rightarrow u = F_Y(y)$$

then:

$$F_X^{-1}(u) = -\ln(1-u) = -\ln(1-F_Y(y))$$

Hence, we can write:

$$\begin{aligned}
\lim_{u\uparrow 1}\{F_Y^{-1}(u) - F_X^{-1}(u)\} &= \lim_{y\to\infty}\{y + \ln(1-F_Y(y))\} \\
&= \lim_{y\to\infty}\{-\ln(e^{-y}) + \ln(1-F_Y(y))\} \\
&= \lim_{y\to\infty}\ln\left(\frac{1-F_Y(y)}{e^{-y}}\right) \qquad (3.26) \\
&= \ln(e^{\delta}E(e^{Z})) \qquad (3.27)
\end{aligned}$$

where the equality between equations (3.26) and (3.27) follows from Lemma 2. Thus, the proof is finished.

Adding results (3.24) and (3.25), taking the limit with respect to $u \uparrow 1$ and applying Proposition 12 entail that the upper tail dependence coefficient λ^U is:

$$\lambda^U = 1 - \Lambda_\sigma(\ln(E(e^Z))) + \frac{\int_{-\infty}^{\ln(E(e^Z))} e^z \lambda_\sigma(z)dz}{E(e^Z)} \tag{3.28}$$

Example 3 *Consider λ_σ as a normal density function with the corresponding zero mean and standard deviation $\sigma > 0$. In this case:*

$$E(e^Z) = e^{\sigma^2/2}$$

Then, as can be shown by a straightforward calculation, the upper tail dependence coefficient (given by the general formula (3.28)) diminishes to:

$$\lambda^U = 2(1 - \Phi(\sigma/2)) = 2\Phi(-\sigma/2)$$

where $\Phi(t) = \Lambda_\sigma(t/\sigma)$.

We note that the upper tail dependence coefficient given by (3.28) is less than or equal to 1. Observe that λ^U can be written as follows:

$$\lambda^U = \int_{\ln(E(e^Z))}^{\infty} \lambda_\sigma(z)dz + \int_{-\infty}^{\ln(E(e^Z))} e^{-\ln(E(e^Z))+z}\lambda_\sigma(z)dz \tag{3.29}$$

Hence, if $z \in (-\infty, \ln(E(e^Z)))$ then $e^{-\ln(E(e^Z))+z} < 1$. Thus, we conclude that the upper tail dependence coefficient can be bounded by:

$$\lambda^U < \int_{\ln(E(e^Z))}^{\infty} \lambda_\sigma(z)dz + \int_{-\infty}^{\ln(E(e^Z))} \lambda_\sigma(z)dz = 1$$

where we have used the fact that λ_σ is a strictly positive density function on the real line.

3.4.3 Tail dependence for the Constant Symmetric Spread Model

In this subsection the tail dependence coefficients for the Constant Symmetric Spread Model will be derived. The symmetry property of this model immediately entails the application of Corollary 1. Let us denote $x = F_X^{-1}(u)$ for $u \in [0,1]$. We observe that $x \downarrow s_0$ as $u \downarrow 0$. Recall the definitions of $f_{Y|X=x}(y|x)$ and $A(x)$ from section 3.3 and compute the following probability:

$$\begin{aligned}
2P\{Y \leq x | X = x\} &= 2\int_{s_0}^{x} f_{Y|X=x}(k|x)dk \\
&= \frac{2}{A(x)} \int_{s_0}^{x} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\{-(k-\mu(x))^2/2\sigma^2\}dk \\
&= \frac{2}{A(x)} \int_{\frac{s_0-\mu(x)}{\sigma}}^{\frac{x-\mu(x)}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \\
&= 2\frac{\Phi((x-\mu(x))/\sigma) - \Phi((s_0-\mu(x))/\sigma)}{A(x)} \\
&= 2\frac{\Phi((x-\mu(x))/\sigma) - \Phi((s_0-\mu(x))/\sigma)}{1 - \Phi((s_0-\mu(x))/\sigma)}
\end{aligned}$$

Recalling that $\mu(x) = x - \sigma^2/2$ and taking the limit with respect to $u \downarrow 0$ entail that the lower tail dependence coefficient λ^L is:

$$\lambda^L = \lim_{u \downarrow 0} 2P\{Y \leq x | X = x\} = 0$$

Thus, we can conclude that lower tail independence occurs.
To deal with the upper tail dependence coefficient λ^U, note that $x \to \infty$ as $u \uparrow 1$. Hence, using Corollary 1 we obtain:

$$\begin{aligned}
\lambda^U &= \lim_{u \uparrow 1} 2(1 - P\{Y \leq x | X = x\}) \\
&= \lim_{u \uparrow 1} 2\left(1 - \frac{\Phi((x - \mu(x))/\sigma) - \Phi((s_0 - \mu(x))/\sigma)}{1 - \Phi((s_0 - \mu(x))/\sigma)}\right) \\
&= 2(1 - \Phi(\sigma/2)) = 2\Phi(-\sigma/2)
\end{aligned}$$

Remark 20 *Observe that the upper tail dependence coefficient for the Constant Symmetric Spread Model is the same as in case of the Constant Spread Model with the assumption of λ_σ as a normal density function with zero mean and standard deviation $\sigma > 0$.*

3.4.4 Tail dependence for a special case of the Variable Spread Model

In this subsection the tail dependence coefficients for a special case of the Variable Spread Model will be investigated. We assume that λ_σ is the normal density function with mean zero and standard deviation $\sigma > 0$. In addition, we assume that the "spread function" $\sigma(x) : [0, \infty) \to \Re$, which plays a crucial role in this model, satisfies the following conditions:

1. $\sigma(x)$ is an increasing function
2. $\sigma(0) = \sigma_0 > 0$
3. $\lim_{x \to \infty} \sigma(x) = \sigma_1 < \infty$

Based on the previous results we expect that $\lambda^L = 0$ and $\lambda^U = 2\Phi(-\sigma_1/2)$. This is reasonable, since as $\sigma(x) \to \sigma_1$, the Variable Spread Model tends to the Constant Spread Model with the standard deviation σ_1. It is relatively easy to show that $\lambda^L = 0$; however, a mathematically rigorous proof for the statement concerning λ^U is more difficult.

Proposition 13 *Under the above assumptions the lower and upper tail dependence coefficients are:*

$$\lambda^L = 0 \quad \textit{and} \quad \lambda^U = 2\Phi\left(-\frac{\sigma_1}{2}\right)$$

Proof Let us denote $x = F_X^{-1}(u)$ and $y = F_Y^{-1}(u)$ for $u \in [0,1]$. We observe that $y \to -\infty$ as $u \downarrow 0$. Consider the following probability:

$$\begin{aligned} P\{Y \leq y | X \leq x\} &= \frac{\int_0^x \left\{ \int_{-\infty}^y e^{-u} \lambda_{\sigma(u)}(v-u-\delta)dv \right\} du}{1-e^{-x}} = \frac{\int_0^x \left\{ \int_{-\infty}^{\frac{y-u-\delta}{\sigma(u)}} e^{-u}\lambda_1(z)dz \right\} du}{1-e^{-x}} \\ &\leq \frac{(1-e^{-x}) \int_{-\infty}^{\frac{y-\delta}{\sigma_0}} \lambda_1(z)dz}{1-e^{-x}} = \int_{-\infty}^{\frac{y-\delta}{\sigma_0}} \lambda_1(z)dz \end{aligned}$$

where we have used the fact that $\sigma(x)$ is an increasing function with $\sigma(0) = \sigma_0 > 0$. Taking the limit with respect to $u \downarrow 0$ results in:

$$\lambda^L = \lim_{u \downarrow 0} P\{Y \leq y | X \leq x\} = 0 \tag{3.30}$$

Thus, we conclude that lower tail independence occurs.
In order to derive the upper tail dependence coefficient λ^U we apply Proposition 11. Moreover, the following theorem is needed. It is known as Lebesgue's dominated convergence theorem and it is stated in terms of the Lebesgue integrals. We note that implicitly all integrals in this section are in the sense of Lebesgue.

Theorem 2 ***Lebesgue's dominated convergence theorem*** *If $|f_n| \leq g$ almost everywhere, where g is integrable, and if $f_n \to f$ almost everywhere, then f and the f_n are integrable and:*

$$\lim_{n \to \infty} \int f_n(x)dx = \int f(x)dx$$

For a proof, we refer to [1].
Denote $x = F_X^{-1}(u)$ and $y = F_Y^{-1}(u)$, for $u \in [0,1]$. We want to apply Proposition 11. Let us compute the following limit (further it will appear that this limit indeed exists):

$$\begin{aligned} \lim_{u \uparrow 1} P\{Y > y | X = x\} &= \lim_{u \uparrow 1} \int_y^\infty f_{Y|X=x}(k|x)dk \\ &= \lim_{u \uparrow 1} \frac{1}{e^{-x}} \int_y^\infty e^{-x} \lambda_{\sigma(x)}(k-x-\delta)dk \\ &= \lim_{u \uparrow 1} \int_{\frac{y-x-\delta}{\sigma(x)}}^\infty \lambda_1(z)dz \end{aligned} \tag{3.31}$$

This limit exists if the following limit exists:

$$\lim_{u \uparrow 1} \frac{y-x-\delta}{\sigma(x)} \tag{3.32}$$

where $x = -\ln(1-u)$ and $y = F_Y^{-1}(u)$. We can write:

$$u = F_Y(y) \quad \text{and} \quad x = -\ln[1-F_Y(y)]$$

Moreover, we note that $y \to \infty$ as $u \uparrow 1$. Then we obtain that limit (3.32) is equivalent with:

$$\lim_{y\to\infty} \frac{y + \ln[1 - F_Y(y)] - \delta}{\sigma(-\ln[1 - F_Y(y)])} = \lim_{y\to\infty} \frac{\ln\left(\frac{1-F_Y(y)}{e^{-y}}\right) - \delta}{\sigma(-\ln[1 - F_Y(y)])}$$

Consider the numerator of the above expression. Using de l'Hopital's rule, we obtain:

$$\begin{aligned}
\lim_{y\to\infty} \frac{1 - F_Y(y)}{e^{-y}} &= \lim_{y\to\infty} \frac{\int_0^\infty e^{-z}\lambda_{\sigma(z)}(y - z - \delta)dz}{e^{-y}} = e^{\delta} \lim_{y\to\infty} \int_{-\infty}^{y-\delta} e^k \lambda_{\sigma(y-k-\delta)}(k)dk \\
&= e^{\delta} \lim_{y\to\infty} \int_{-\infty}^{\infty} 1_{(-\infty,y-\delta)}(k) e^k \lambda_{\sigma(y-k-\delta)}(k)dk
\end{aligned}$$

The derivation of the above limit requires some careful steps. Let us first consider the following function:

$$g(k) = \frac{e^k}{\sigma_0\sqrt{2\pi}} e^{-\frac{k^2}{2\sigma_1^2}}$$

We note that g is nonnegative and Lebesgue integrable. Now, consider a sequence y_n such that $y_n \to \infty$ as $n \to \infty$, and define:

$$f_n(k) = 1_{(-\infty,y_n-\delta)}(k) e^k \lambda_{\sigma(y_n-k-\delta)}(k)$$

Due to the conditions imposed on the "spread function" $\sigma(x)$, we can derive the following limit:

$$|f_n(k)| \le g(k)$$

Moreover, we note that $\forall k \in \Re$ the following relation holds:

$$\lim_{n\to\infty} f_n(k) = e^k \lambda_{\sigma_1}(k) \tag{3.33}$$

Then from Lebesgue's dominated convergence theorem we obtain:

$$\lim_{n\to\infty} \frac{1 - F_Y(y_n)}{e^{-y_n}} = e^{\delta} \lim_{n\to\infty} \int_{-\infty}^{\infty} f_n(k)dk = e^{\delta} \int_{-\infty}^{\infty} e^k \lambda_{\sigma_1}(k)dk = e^{\delta} E(e^K)$$

Since this limit exists and gives the same answer for every y_n such that $y_n \to \infty$ as $n \to \infty$, we conclude that:

$$\lim_{y\to\infty} \frac{1 - F_Y(y)}{e^{-y}} = e^{\delta} E(e^K)$$

So, the limit (3.32) is:

$$\lim_{y\to\infty} \frac{y + \ln[1 - F_Y(y)] - \delta}{\sigma(-\ln[1 - F_Y(y)])} = \frac{\ln(e^{\delta} E(e^K)) - \delta}{\sigma_1} = \frac{\ln(E(e^K))}{\sigma_1}$$

Thus, the limit (3.31) is equal to:

$$\lim_{u \uparrow 1} P\{Y > y | X = x\} = \int_{\frac{\ln(E(e^K))}{\sigma_1}}^{\infty} \lambda_1(z) dz \tag{3.34}$$

Next, to apply Proposition 11, let us compute the following probability:

$$\begin{aligned} P\{X > x | Y = y\} &= \frac{\int_x^\infty e^{-k} \lambda_{\sigma(k)}(y - k - \delta) dk}{\int_0^\infty e^{-k} \lambda_{\sigma(k)}(y - k - \delta) dk} \\ &= \frac{\int_{-\infty}^{y-x-\delta} e^{z-y+\delta} \lambda_{\sigma(y-z-\delta)}(z) dz}{\int_{-\infty}^{y-\delta} e^{z-y+\delta} \lambda_{\sigma(y-z-\delta)}(z) dz} \\ &= \frac{\int_{-\infty}^{\infty} 1_{(-\infty, y-x-\delta)} e^z \lambda_{\sigma(y-z-\delta)}(z) dz}{\int_{-\infty}^{\infty} 1_{(-\infty, y-\delta)} e^z \lambda_{\sigma(y-z-\delta)}(z) dz} \end{aligned} \tag{3.35}$$

Again, we observe that Lebesgue's dominated convergence theorem can be applied. The reasoning is analogous as before and will not be repeated. Thus, taking the limit with respect to $u \uparrow 1$ yields:

$$\lim_{u \uparrow 1} P\{X > x | Y = y\} = \frac{\int_{-\infty}^{\ln(E(e^K))} e^z \lambda_{\sigma_1}(z) dz}{\int_{-\infty}^{\infty} e^z \lambda_{\sigma_1}(z) dz} \tag{3.36}$$

Adding results (3.34) and (3.36), from Proposition 11 we obtain:

$$\lambda^U = \int_{\frac{\ln(E(e^K))}{\sigma_1}}^{\infty} \lambda_1(z) dz + \frac{\int_{-\infty}^{\ln(E(e^K))} e^z \lambda_{\sigma_1}(z) dz}{\int_{-\infty}^{\infty} e^z \lambda_{\sigma_1}(z) dz} \tag{3.37}$$

Since λ_σ is a normal density function with mean zero and standard deviation $\sigma > 0$, the upper tail dependence coefficient λ^U simplifies to:

$$\lambda^U = 2(1 - \Phi(\sigma_1/2)) = 2\Phi(-\sigma_1/2)$$

where we have used the result (3.3). Thus, the proof is accomplished.

Remark 21 *In the proof of (3.30) and (3.37) we assumed λ_σ to be a normal density. It seems likely that these results will also hold for more general densities. Such an extension has not been studied yet. Moreover, we intuitively expect that if $\lim_{x\to\infty} \sigma(x) = \infty$, then the upper tail dependence coefficient is 0 (for λ_σ normal as well as more general).*

3.5 Models and the related Copulas

The purpose of this section is to derive the copula functions that correspond to the discussed models. Moreover, we are interested whether the resulting copulas belong to the Archimedean class.

3.5.1 The Constant Spread Model and the related Copula

In this subsection, the copula function that corresponds to the Constant Spread Model will be calculated. Using the definition of the copula function for this model we obtain:

$$\begin{aligned} C(u,v) &= P\{X \leq F_X^{-1}(u), Y \leq F_Y^{-1}(v)\} \\ &= \int_0^{F_X^{-1}(u)} \left\{ \int_{-\infty}^{F_Y^{-1}(v)} e^{-s}\lambda_\sigma(t-s-\delta)dt \right\} ds \\ &= \int_0^{F_X^{-1}(u)} e^{-s}\Lambda_\sigma(F_Y^{-1}(v)-s-\delta)ds \end{aligned}$$

For the density $c(u,v)$ we obtain in a standard way the formula:

$$c(u,v) = \frac{f_{X,Y}(x,y)}{f_X(x)f_Y(y)}\Big|_{x=F_X^{-1}(u),y=F_Y^{-1}(u)}$$

Hence, denoting $x = F_X^{-1}(u)$ and $y = F_Y^{-1}(v)$, we obtain:

$$\begin{aligned} c(u,v) &= \frac{f_{X,Y}(x,y)}{f_X(x)f_Y(y)} = \frac{e^{-x}\lambda_\sigma(y-x-\delta)}{e^{-x}\int_0^\infty e^{-t}\lambda_\sigma(y-t-\delta)dt} \\ &= \frac{\lambda_\sigma(y-x-\delta)}{\int_0^\infty e^{-t}\lambda_\sigma(y-t-\delta)dt} \end{aligned}$$

Figure 3.4 pictures the density c and its contour plot under the assumption that λ_σ is a normal density function with mean zero and standard deviation $\sigma = 2$. Moreover, we assume that $\delta = 0$.

3.5.2 The Constant Symmetric Spread Model and the related Copula

In this subsection the copula function for the Constant Symmetric Spread Model will be derived. Using the formulas of section 3.3, we obtain:

$$\begin{aligned} C(u,v) &= P\{X \leq F_X^{-1}(u), Y \leq F_Y^{-1}(v)\} \\ &= \int_{s_0}^{F_X^{-1}(u)} \left\{ \int_{s_0}^{F_Y^{-1}(v)} e^{-s}\frac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{(t-\mu(s))^2}{2\sigma^2}\right\} dt \right\} ds \\ &= \int_{s_0}^{F_X^{-1}(u)} e^{-s}\left\{\Lambda_\sigma[F_Y^{-1}(v)-\mu(s)] - \Lambda_\sigma[s_0-\mu(s)]\right\} ds \end{aligned}$$

whereas the density function c is:

$$\begin{aligned} c(u,v) &= \frac{f_{X,Y}(x,y)}{f_X(x)f_Y(y)} = \frac{A(x)e^{-x}\frac{1}{A(x)}\frac{1}{\sigma\sqrt{2\pi}}\exp\{-[y-\mu(x)]^2/\sigma^2\}}{A(x)e^{-x}A(y)e^{-y}} \\ &= \frac{1}{A(x)A(y)\sigma\sqrt{2\pi}}\exp\left\{-[y-\mu(x)]^2/\sigma^2+y\right\} \end{aligned}$$

where $x = F_X^{-1}(u)$ and $y = F_Y^{-1}(v)$, for $u, v \in [0,1]$.
Figure 3.5 presents the density c, for the case $\sigma = 2$.

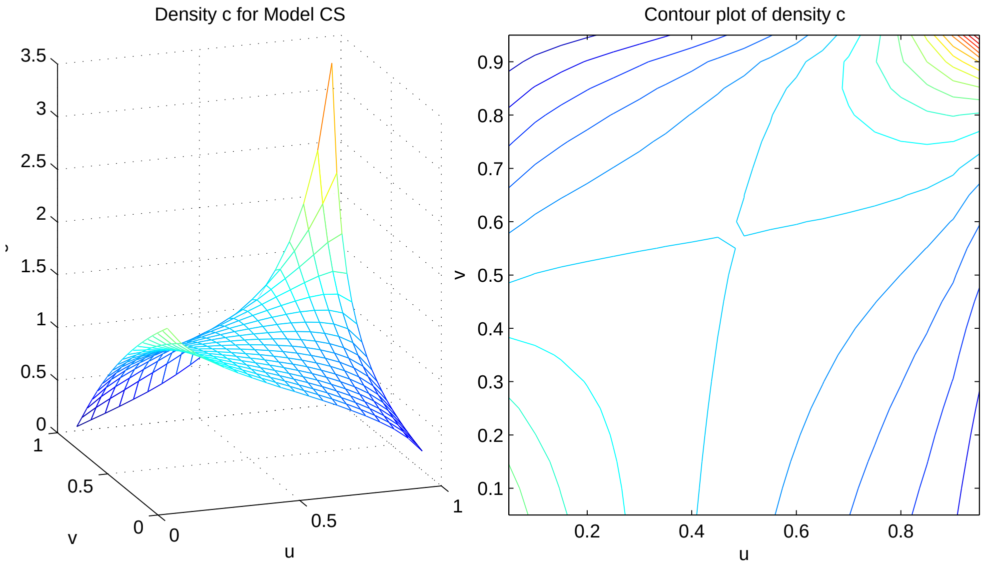


Figure 3.4: The density function c and its contour plot, where λ_σ is a normal density function with mean zero and standard deviation $\sigma = 2$, $\delta = 0$. The contour plot shows that the density c is not a symmetric function.

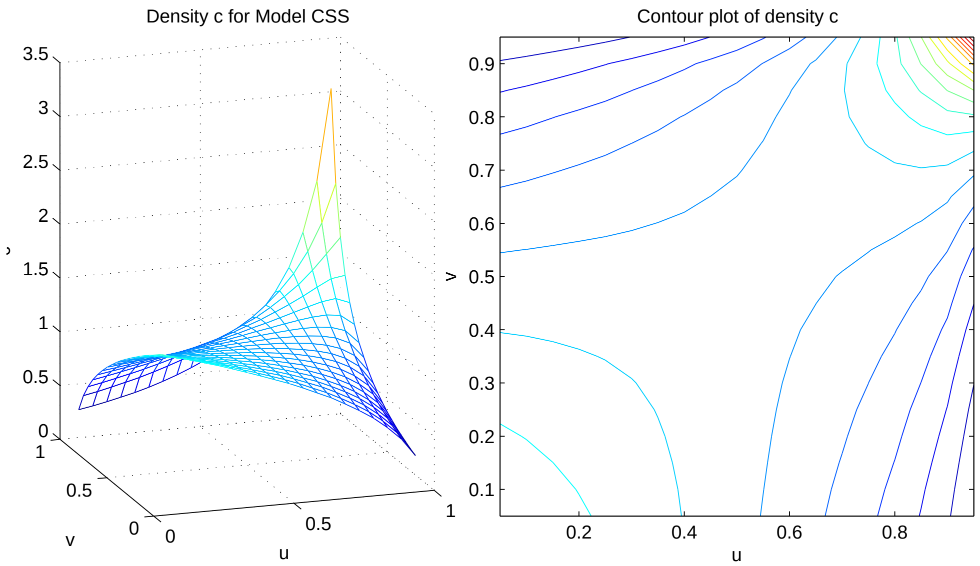


Figure 3.5: The density function c and its contour plot for $\sigma = 2$.

3.5.3 The Variable Spread Model and the related Copula

In this subsection the copula function related to the Variable Spread Model will be computed. We obtain:

$$\begin{aligned} C(u,v) &= P\{X \leq F_X^{-1}(u), Y \leq F_Y^{-1}(v)\} \\ &= \int_0^{F_X^{-1}(u)} \left\{ \int_{-\infty}^{F_Y^{-1}(v)} e^{-s} \lambda_{\sigma(s)}(t-s-\delta) dt \right\} ds \end{aligned}$$

$$= \int_0^{F_X^{-1}(u)} e^{-s}\Lambda_{\sigma(s)}(F_Y^{-1}(v) - s - \delta)ds$$

whereas the density function c is:

$$\begin{aligned} c(u,v) &= \frac{f_{X,Y}(x,y)}{f_X(x)f_Y(y)} = \frac{e^{-x}\lambda_{\sigma(x)}(y-x-\delta)}{e^{-x}\int_0^\infty e^{-t}\lambda_{\sigma(t)}(y-t-\delta)dt} \\ &= \frac{\lambda_{\sigma(x)}(y-x-\delta)}{\int_0^\infty e^{-t}\lambda_{\sigma(t)}(y-t-\delta)dt} \end{aligned}$$

where $x = F_X^{-1}(u)$ and $y = F_Y^{-1}(v)$, for $u, v \in [0,1]$.
The density function c and its contour plot with $\sigma(x) = 2 + x/2$ and $\delta = 0$ is presented in Figure 3.6.

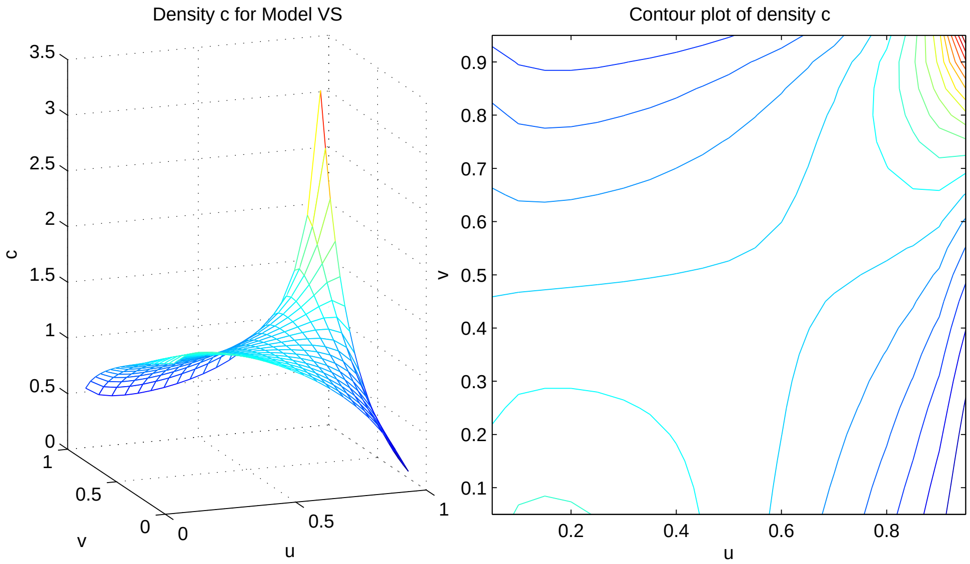


Figure 3.6: The density function c and its contour plot with $\sigma(x) = 2 + x/2$ and $\delta = 0$. Observe, that the contour plot of the density c suggests that the corresponding copula function is not symmetric.

3.5.4 Do copulas of the considered models belong to the Archimedean class?

The purpose of this subsection if a verification whether the copula functions derived in subsections 3.5.1, 3.5.2 and 3.5.3 are Archimedean. Let us recall the definition of the bivariate Archimedean copula:
The copula C is an Archimedean copula if:

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

where ϕ is a generator, (for further details we refer to Chapter 2).
One of the properties of this class is symmetry. Indeed, it follows straightforwardly

from the above definition:

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

Moreover, if the joint cumulative distribution is symmetric, the corresponding joint density function is also symmetric. The symmetry property is crucial in finding the counter examples for the copulas related to the Constant Spread Model and Variable Spread Model.
Thus, consider the copula that corresponds to the Constant Spread Model with λ_σ as a normal density function with mean zero and standard deviation $\sigma > 0$. If we substitute $u = 0.95$, $v = 0.85$, $\sigma = 2$ and $\delta = 0$, then we obtain:

$$c(0.95, 0.85) = 1.976 \neq 1.7456 = c(0.85, 0.95)$$

Since both quantities are different, we conclude that this copula function is not Archimedean. Note that the contour plot of the Figure 3.4 also shows that $c(u,v)$ is not symmetric.
Next, consider the copula function related to the Variable Spread Model with $\sigma(x) = 2 + x/2$. If we substitute $u = 0.3$, $v = 0.7$ and $\delta = 0$, then:

$$C(0.3, 0.7) = 0.2505 \neq 0.2376 = C(0.7, 0.3)$$

Both quantities are different. Thus, we conclude that this copula function is not Archimedean. Figure 3.6 also shows that $c(u,v)$ is not symmetric.
We observe that in case of the copula related to the Constant Symmetric Spread Model the symmetry property holds. In order to show this, the following proposition is required:

Proposition 14 *If a bivariate distribution function H is symmetric then the related copula function C is symmetric.*

Proof Let F and G be the marginal distribution functions related to the distribution H, then:

$$F(x) = \lim_{y\to\infty} H(x,y) = \lim_{y\to\infty} H(y,x) = G(x)$$

Hence, the marginal distribution functions are equal. Then we have:

$$\begin{aligned} C(u,v) &= H(F^{-1}(u), G^{-1}(v)) = H(F^{-1}(u), F^{-1}(v)) \\ &= H(F^{-1}(v), F^{-1}(u)) = H(F^{-1}(v), G^{-1}(u)) = C(v,u) \end{aligned}$$

Thus, the copula C is a symmetric function and the proof is complete.

The bivariate distribution function that corresponds to the Constant Symmetric Spread Model is symmetric. Hence, the above proposition entails that the copula related to this model is also symmetric. Figure 3.5 also shows this symmetry. To investigate whether the Constant Symmetric Spread Model is Archimedean we use the following theorem:

Theorem 3 *Let C be an associative (i.e. $C(C(u,v),w) = C(u,C(v,w))$ for $u,v,w \in [0,1]$) copula such that $C(u,u) < u$ for all $u \in (0,1)$. Then C is Archimedean.*

For details we refer to [13].
We cannot prove that the Constant Symmetric Spread Model has an associative copula. Therefore, we investigate associativity numerically, for the case $\sigma = 2$. Figure 3.7 illustrates the following error:

$$E = |C(C(u,v),w) - C(u,C(v,w))| \quad \text{for } u,v,w \in [0.1, 0.2, 0.3, ..., 0.9]$$

The second condition, namely $C(u,u) < u$ for $u \in (0,1)$, is satisfied in a numerical sense, as shown in Figure 3.8.

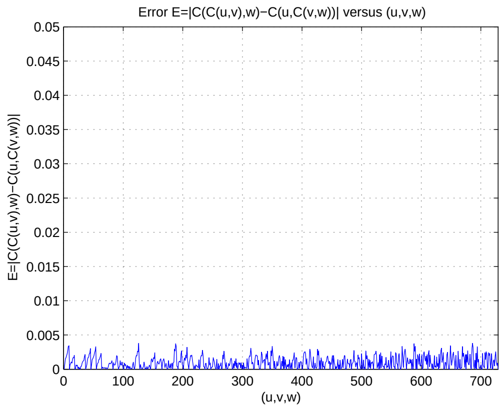


Figure 3.7: The error $E = |C(C(u,v),w) - C(u,C(v,w))|$ for $u,v,w \in [0.1, 0.2, 0.3, ..., 0.9]$ and $\sigma = 2$. The notation (u,v,w) on the horizontal line of the plot is used for simplicity. In fact the horizontal line corresponds to the following "ordered" set of points: $\{(u,v,w) : w = 0.1, 0.2, ..., 0.9, v = 0.1, 0.2, ..., 0.9, u = 0.1, 0.2, ..., 0.9\}$, where we first fix the values of u and v and go across the values of w. Then we change the value of v and we repeat the procedure. At the end we change the value of u and again we repeat everything.

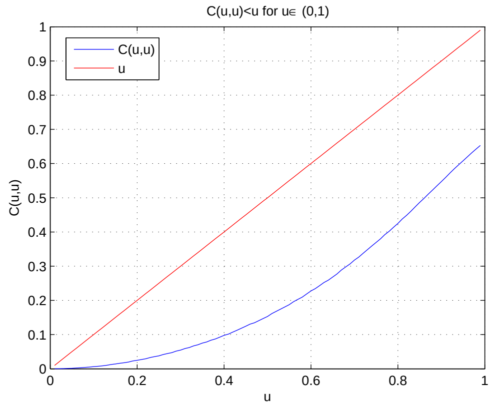


Figure 3.8: The condition $C(u,u) < u$ for $u \in (0,1)$ and $\sigma = 2$. This condition is satisfied in a numerical sense, where we have used the grid $u = [0.01, 0.02, ..., 0.99]$.

We believe that the small fluctuations, which are visible in Figure 3.7, are due to numerical errors. Therefore, we suppose that this copula for $\sigma = 2$ belongs to the Archimedean class of copulas or at least it is very "close" to the Archimedean class. We expect a similar behavior for other choices of the parameter σ.

3.6 Maximum Likelihood Method and evaluation of the fit

The purpose of this section is to provide a theoretical background that will be applied in the following case study. Recall that so far we discussed three bivariate parametric models (sections 3.1, 3.2 and 3.3). For a given dataset we would like to estimate the parameters of the underlying probability distributions and then evaluate the fit. The latter will be done using the so-called percentile lines method - explained in subsection 2.5.6. In the next two subsections, we describe the Maximum Likelihood Method and recall the percentile lines method. We denote the probability density corresponding to (V, W) by $f_{V,W}(v, w; \theta)$, where θ is a vector of model parameters. Moreover, we assume that $f_{V,W}(v, w; \theta)$ is related to a model in terms of a random vector (X, Y) as in (3.1), with probability density $f_{X,Y}(x, y; \theta)$ and marginals $F_X(x; \theta)$ and $F_Y(y; \theta)$.

3.6.1 Maximum Likelihood Method

Let $(V_i, W_i), i = 1, ..., n$, be a sample of n bivariate observations originating from a random vector (V, W) with known marginal distributions $F_V(v)$ and $F_W(w)$. Thus, the dataset is transformed to the model space as follows:

$$X_i = F_X^{-1}(F_V(V_i); \theta) \quad \text{and} \quad Y_i = F_Y^{-1}(F_W(W_i); \theta)$$

with $f_{V,W}(v, w; \theta)$ obtained from $f_{X,Y}(x, y; \theta)$ by using (3.1):

$$f_{V,W}(v, w; \theta) = \frac{f_{X,Y}(x, y; \theta) f_V(v) f_W(w)}{f_X(x; \theta) f_Y(y; \theta)}$$

where $x = F_X^{-1}(F_V(v); \theta), y = F_Y^{-1}(F_W(w); \theta)$ and $f_X(x; \theta)$ and $f_Y(y; \theta)$ are the margins of $f_{X,Y}(x, y; \theta)$. The quantities $(X_i, Y_i), i = 1, ..., n$, constitute the observations in the model space. The Maximum Likelihood Method now yields the estimator $\hat{\theta}$ of the parameter θ (we stress that θ can be a vector of parameters):

$$\hat{\theta} = \arg\max_\theta \prod_{i=1}^{n} \frac{f_{X,Y}(X_i, Y_i; \theta) f_V(V_i) f_W(W_i)}{f_X(X_i; \theta) f_Y(Y_i; \theta)}$$

Let us now consider the likelihood function:

$$L(\theta) = \prod_{i=1}^{n} \frac{f_{X,Y}(X_i, Y_i; \theta) f_V(V_i) f_W(W_i)}{f_X(X_i; \theta) f_Y(Y_i; \theta)}$$

and the corresponding log-likelihood function given by:

$$\begin{aligned} l(\theta) &= \ln L(\theta) = \ln \left(\prod_{i=1}^{n} \frac{f_{X,Y}(X_i, Y_i; \theta) f_V(V_i) f_W(W_i)}{f_X(X_i; \theta) f_Y(Y_i; \theta)} \right) \\ &= \ln \left(\prod_{i=1}^{n} \frac{f_{X,Y}(X_i, Y_i; \theta)}{f_X(X_i; \theta) f_Y(Y_i; \theta)} \right) + \ln \left(\prod_{i=1}^{n} f_V(V_i) f_W(W_i) \right) \end{aligned}$$

Remark 22 *Since we assume that the cumulative distribution functions $F_V(v)$ and $F_W(w)$ are known, the second term in the above expression is irrelevant in the maximization process of $l(\theta)$.*

3.6.2 The percentile lines

Consider the bivariate random vector (V, W) with marginal distribution functions $F_V(v)$ and $F_W(w)$. Let us transform (V, W) into the model (X, Y) with the margins $F_X(x)$ and $F_Y(y)$ as in the previous subsection. We call (V, W) *the original or physical space*, whereas the space (X, Y) is called *the transformed or model space.*

Definition 15 *The p-percentile line in transformed space is given by the following function:*

$$y = f(x; p), \quad x \in D$$

where p is a percentage and $f(x; p)$ is obtained by solving the equation $P\{Y \leq y | X = x\} = p$ with respect to variable y. The set D is the support of the distribution $F_X(x)$.

The transformation of the p-percentile line from the model space into the physical space is given as follows:

$$\{(x, f(x; p)) : x \in D\} \quad \rightarrow \quad \{(F_V^{-1}(F_X(x)), F_W^{-1}(F_Y(f(x; p)))) : x \in D\}$$

Observe that this is exactly the inverse of the transformation (3.1).
As it will become clear below, percentile lines can be used as a visual tool for determining the best model. They indicate the trend of the considered model. Moreover, if the considered distribution fits well, then approximately $r\%$ of the dataset should fall below the percentile lines corresponding to $p = r\%$ (both in the transformed and in the original space). Moreover, this should also be satisfied in the subregions of the physical and model space.
We stress that the method of the percentile lines is a bit subjective and requires experience in interpreting the results.

Remark 23 *Recall that we have derived the corresponding copula functions for each considered model. It will be interesting to see how the percentile lines in the copula space will look like (the percentile lines in the copula space are obtained by applying the proper transformations, which we will not explicitly describe here).*

3.7 Case study

In this section, the bivariate models CS, CSS and VS will be applied to a given dataset.

3.7.1 Assumptions

In this subsection, we list the assumptions used to carry out the case study:

1. *We consider the dataset of $n = 89$ bivariate observations of the water levels and wind speeds already introduced in Chapter 2. This dataset constitutes the observations $(V_i, W_i), i = 1, ..., n$.*

2. *The cumulative distribution functions $F_V(v)$ and $F_W(w)$ are prescribed (for further details we refer to Chapter 2).*

3. *The Constant Spread Model (Model CS) - We assume that $\delta = 0$ and λ_σ is a normal density function with mean zero and standard deviation $\sigma > 0$. The unknown parameter is σ.*

4. *The Constant Symmetric Spread Model (Model CSS) - The unknown parameter is $\sigma > 0$, which is asymptotically equal to the standard deviation of the normal distribution in the transformed space.*

5. *The Variable Spread Model (Model VS) - We assume that $\delta = 0$ and function $\sigma_s(x)$ is given as follows:*

$$\sigma_s(x) = \begin{cases} \frac{\sigma}{5}x + \sigma & \text{for } x \in \ [0,5] \\ 2\sigma & \text{for } x > 5 \end{cases}$$

where $\sigma > 0$. Clearly, σ is our parameter of interest.
This choice of function $\sigma_s(x)$ is very subjective and was made for simplicity, in order to illustrate a simple application of this model. It could be modified in various ways. For instance, $\sigma_s(x)$ could be changed into a two-parametric function.

3.7.2 Estimation

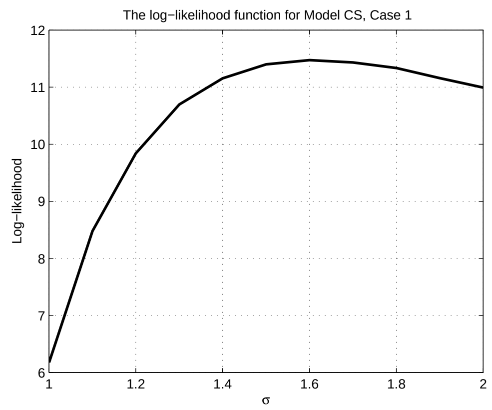
In this subsection we present the estimations of the unknown parameters by using the Maximum Likelihood Method, they are given in the following table:

$\hat{\sigma}$ Model CS	$\hat{\sigma}$ Model CSS	$\hat{\sigma}$ Model VS
1.6	1.4	1.1

In order to confirm the above results, we present the Figures 3.9, 3.10, 3.11 and 3.12, which picture the log-likelihood functions of the corresponding models.

3.7.3 Evaluation of the fit by using the percentile lines

In this subsection the transformed spaces derived under the estimated parameters are presented. In addition, the percentile lines in the original, transformed and copula space are pictured. The Figures 3.13, 3.14 and 3.15 illustrate this. We stress that in our application the interest is in the extreme region. Suppose we only interpret the outcomes in the copula space (the rightmost pictures). It is difficult to judge the fit in the extreme region. A copula is focused on the whole transformed dataset, whereas the picture of the fit in the tail is a bit blurry - it is easier to evaluate the fit in the extreme area in the physical space (the leftmost pictures).
In case of the percentile lines in the model space (middle pictures), we can see precisely



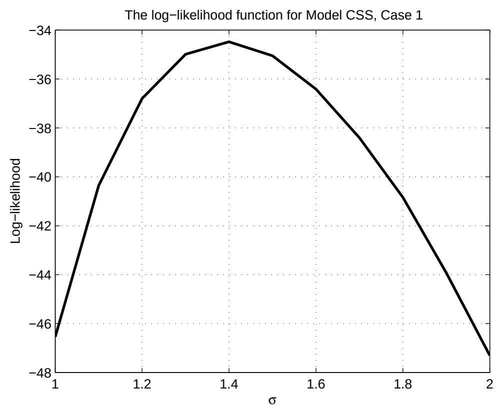
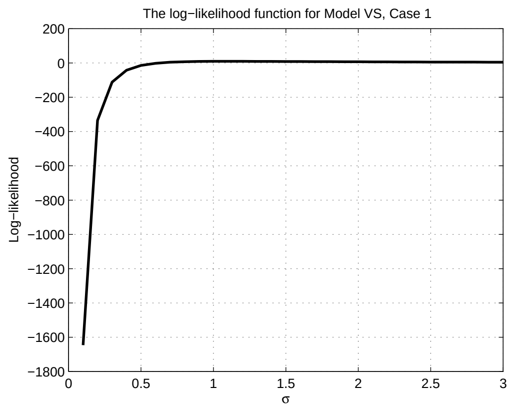


Figure 3.9: The log-likelihood function for Model CS, $\delta = 0$ - maximum is reached for $\hat{\sigma} = 1.6$

Figure 3.10: The log-likelihood function for Model CSS - maximum is reached for $\hat{\sigma} = 1.4$



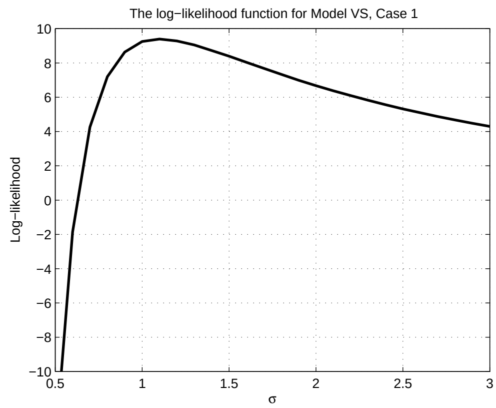


Figure 3.11: The log-likelihood function for Model VS, $\delta = 0$ - maximum is reached for $\hat{\sigma} = 1.1$

Figure 3.12: The log-likelihood function for Model VS, $\delta = 0$ - zoom of the area where the maximum is reached

what the model "says" about the tail. For instance, the Model CS "predicts" that the variables will increase jointly and that there will be a constant spread of variable Y conditioning on the variable $X = x$ (even in the tail) - as a direct consequence of its construction. Moreover, we can judge the fit in the extreme region, because we have the picture of the percentile lines in this area.
Generally, the three pictures of the original, model and copula space together with the percentile lines provide the complete information. However, in this case, the picture of the copula space could be omitted, since it does not contribute much into the information about the extreme fit. The percentile lines in the model space give more information about the fit in the extreme region than the percentile lines in the copula space.
Let us now interpret the results looking at the percentile lines in the physical and model space. The percentile lines describe well the increasing tendency of the dataset.

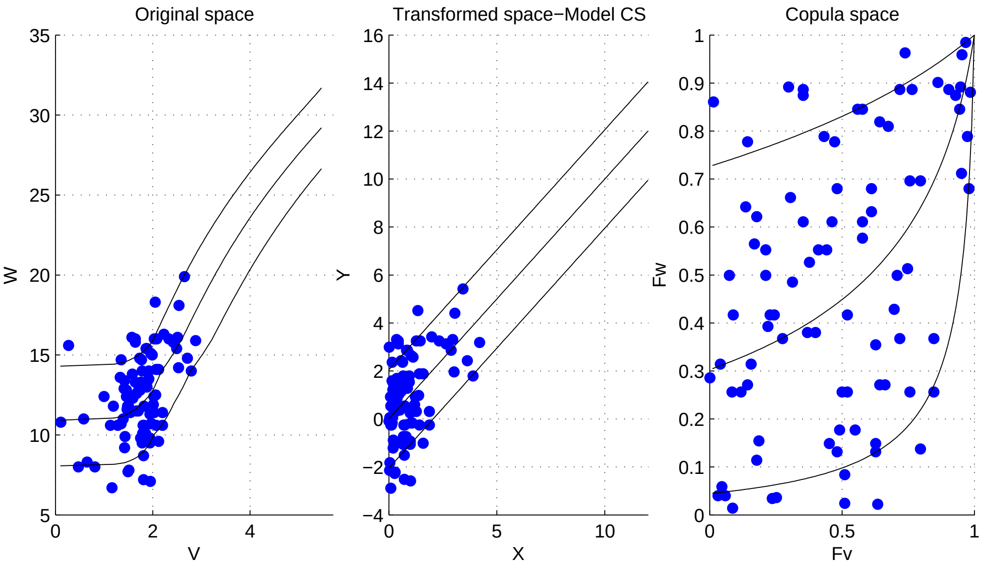


Figure 3.13: The original, transformed and copula space with the percentile lines $(10\%, 50\%, 90\%)$ for Model CS, $\hat{\sigma} = 1.6$, $\delta = 0$. We observe that approximately 8 points fall below the 10%-percentile line and approximately 7 points fall above the 90%-percentile line when we consider the whole dataset. Recall that it should be around 9 data points below the 10%-line and above the 90%-line when we consider the whole dataset.

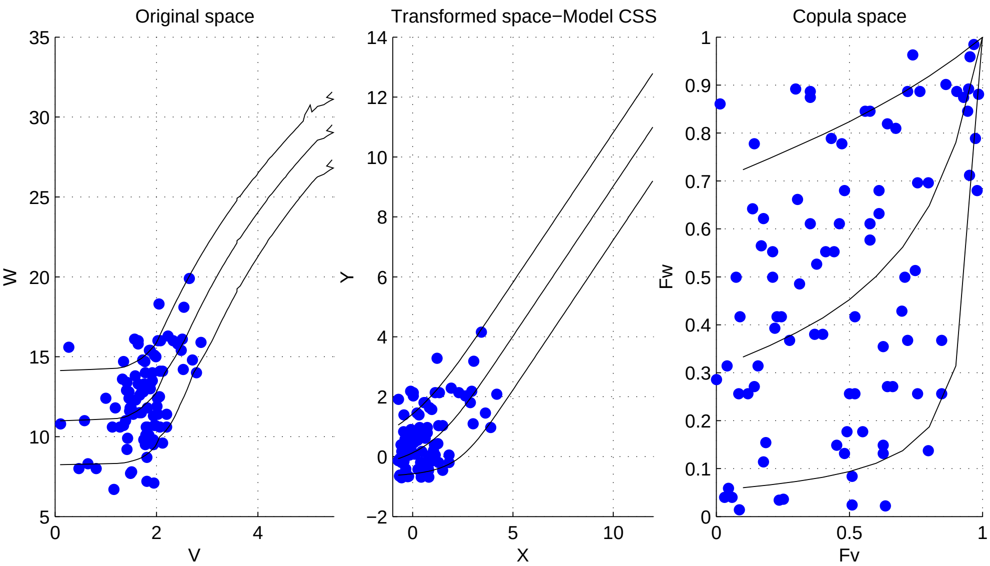


Figure 3.14: The original, transformed and copula space with the percentile lines $(10\%, 50\%, 90\%)$ for Model CSS, $\hat{\sigma} = 1.4$. We observe that approximately 9 points fall below the 10%-percentile line and around 7 points fall above the 90%-percentile line when we consider the whole dataset.

The number of data points that fall below the percentile lines agree with the theory for Model CS and CSS (even in the upper subregion). Thus, we may conclude that these models fit well.
We emphasize that the Model VS with the spread function $\sigma_s(x) > 0$, defined in subsection (3.7.1), was used as an example - generally, it should be verified if this

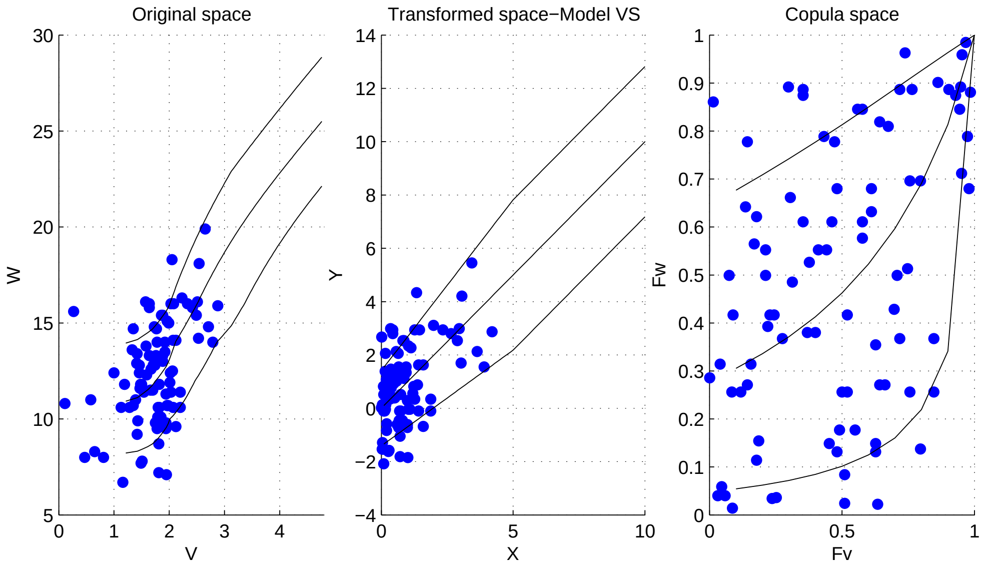


Figure 3.15: The original, transformed and copula space with the percentile lines $(10\%, 50\%, 90\%)$ for Model VS, $\hat{\sigma} = 1.1$, $\delta = 0$. We observe that approximately 7 points fall below the 10%-percentile line and above the 90%-percentile line when we consider the whole dataset.

model is proper for modeling the considered dataset. In case of this model and under assumption of function $\sigma_s(x)$, the fit is a bit doubtful in the upper subregion. In the model space for $x \in [2, 5]$ there are no points that fall above the 90%-percentile line and below the 10%-percentile line. This is of course subjective judgment.
Moreover, the percentile lines in the copula space (for each model) are very similar to the picture of the percentile lines for the Gumbel copula presented in the previous chapter. Our conclusion from Chapter 2 was that the Gumbel copula could be used as a possible representation of our dataset. This fact is consistent with the pictures of the copulas in this paragraph.

3.8 Conclusions

In this chapter three models were considered: the Constant Spread Model, the Variable Spread Model and the Constant Symmetric Spread Model. The Constant Spread Model contains the (constant) parameter σ, with which the dependence in the model can be varied. In the Variable Spread Model, which is a generalization of the Constant Spread Model, the parameter σ no longer needs to be constant, but can be modeled using the function $\sigma(x)$. The Constant Symmetric Spread Model uses, just as in the Constant Spread Model, a constant parameter σ with which the dependence in the model can be varied.
The Models CS, VS and CSS are constructed in such a way that the behavior of these models in the extreme region can more or less be "predicted" from the very construction of the models. So, the extrapolation from the data to the extreme region occurs more or less in a "controlled way", which is believed to be an advantage of these

models.
In the first part of this chapter, we focused on the mathematical description of these models. Later on, we provided some theoretical results concerning the tail dependence coefficients. For every model, we proved lower tail independence and upper dependence (under some special assumptions in case of the Variable Spread Model). Also, explicit formulas were derived for the upper tail coefficients. Moreover, we tackled the question whether the copulas corresponding to these models are Archimedean. For the Constant Spread and the Variable Spread Model (with $\sigma(x) = 2 + x/2 > 0$), we have provided counter examples, using the symmetry property of the Archimedean class. For the Symmetric Spread Model, we performed some computer experiments, which seem to confirm that the related copula is Archimedean or at least very "close" to this class.
At the end, we provided a bivariate case study using observations of water levels and wind speeds with prescribed marginals. For each model the unknown parameters were estimated using the Maximum Likelihood Method. The goodness-of-fit of the models was judged, both in the model space and the original space, using the visual method of the percentile lines. These lines (in the original and model space) not only allow for a proper judgment in the data region, but also in the extreme region (beyond the data). Moreover, we observed that the related copula models are "similar" (judging on the basis of the percentile lines) to the Gumbel copula - presented in Chapter 2 - which was concluded as the best copula fit.

Chapter 4

Rotation Model

This chapter is devoted to the development of the so-called Rotation Model in the bivariate and 3-variate case. The principle of this model was proposed by Chris Geerse. The model arises from the rotation of the system of coordinates; that is where the name of this model originates from.
The Rotation Model is related to the conditional models discussed in Chapter 3, because it is constructed as a modification of the product of the conditional distribution function and the distribution of the conditioning variable. This model is relatively easily extended to the 3-variate case. Moreover, the behavior in the tail of the resulting distribution in the original space can be approximately predicted from the very construction of the model.
It is important to mention that, given the bivariate random vector (V, W), the model will arise from the transformation - introduced in the previous chapter - of random variables (V, W) into (X, Y) (the analogous situation occurs in the 3-variate case).
We will prove some useful properties of this model like the exponentiality of the marginal distributions and the values of the associated tail dependence coefficients (under special assumptions). At the end, we will provide a bivariate case study in which the same statistical tools as for the dependency models introduced in Chapter 3 will be used. However, the case study will reveal a serious drawback of the model, which implies that practical applications of the model seem to be very limited. Notwithstanding its lack of practical use, it seems worthwhile to present the mathematics of the model.

4.1 The 2-dimensional Basic Model

In this section, we introduce the so-called 2-dimensional Basic Model, which gives the foundation to the main goal of the chapter, namely to the Rotation Model. Let us assume that S and T are two random variables, such that the density function $f_S(s)$ of random variable S is:

$$f_S(s) = e^{-s} \quad \text{for } s \geq 0 \tag{4.1}$$

and the conditional density function $f_{T|S=s}(t|s)$ of random variable T given $S = s$, for $s \geq 0$, is represented by:

$$f_{T|S=s}(t|s) = \begin{cases} \lambda_\sigma(t) & \text{for } t \in D \\ 0 & \text{otherwise} \end{cases} \tag{4.2}$$

where λ_σ is a density function with corresponding mean zero and standard deviation $\sigma > 0$, and $D = \{t : \lambda_\sigma(t) > 0\} = [y_b, y_e]$ and $-\infty < y_b < 0 < y_e < \infty$.
We note that S is a standard exponentially distributed random variable and that the joint density function $f_{S,T}(s,t)$ of variables S and T is given by:

$$f_{S,T}(s,t) = f_S(s)f_{T|S=s}(t|s) = \begin{cases} e^{-s}\lambda_\sigma(t) & \text{for } (s,t) \in [0,\infty) \times D \\ 0 & \text{otherwise} \end{cases} \tag{4.3}$$

Then, we can derive the formula for the density function $f_T(t)$ of random variable T as follows:

$$f_T(t) = \int_0^\infty e^{-s}\lambda_\sigma(t)ds = \lambda_\sigma(t) \tag{4.4}$$

Since $f_{S,T}(s,t) = f_S(s)f_T(t)$, we observe that this model yields independence of random variables S and T. A sketch of this model is presented in Figure 4.1.

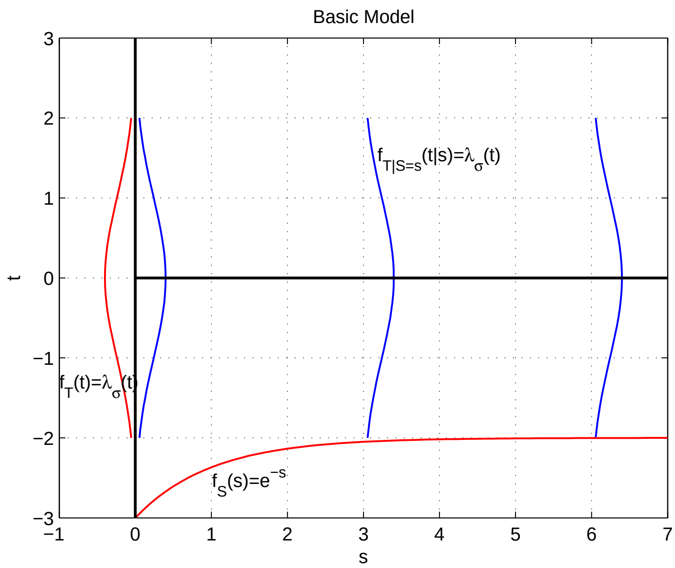


Figure 4.1: Basic model

4.2 Rotation in 2 dimensions

In this section the rotation of a system of two coordinates will be discussed. Suppose we have a point A in a plane with perpendicular x and y coordinate axes subscribed

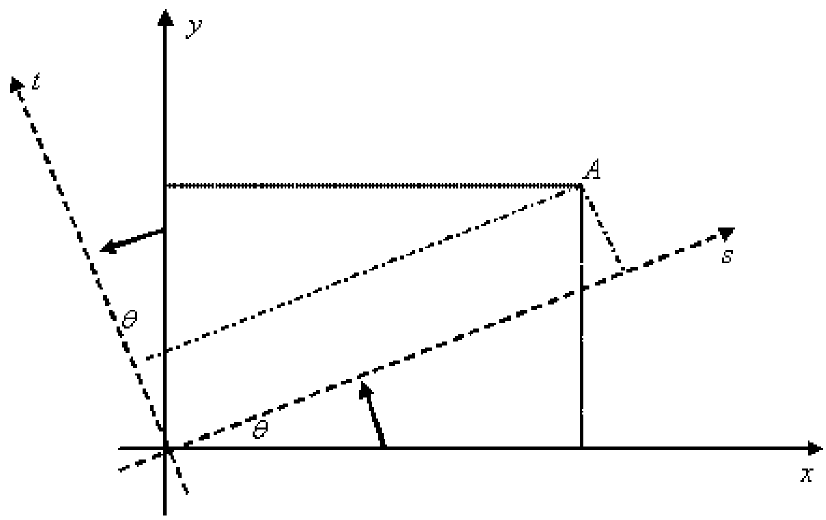


Figure 4.2: Rotation

on it, as pictured in Figure 4.2. We can introduce a different pair of perpendicular coordinate axes s and t on the same plane surface, using dashed lines, by simply rotating the original coordinate axes by an angle θ about their common origin $(0, 0)$. Let us assume that we have the coordinates (x_A, y_A) of point A in the original coordinate system (x, y). The basic trigonometry rules provide the transformation of coordinates (x_A, y_A) into the coordinates (s_A, t_A) in the system (s, t) as follows:

$$\begin{cases} s_A = x_A \cos\theta + y_A \sin\theta \\ t_A = -x_A \sin\theta + y_A \cos\theta \end{cases} \tag{4.5}$$

The above system of equations yields the relation between space (x, y) *and* (s, t).

4.3 Transformation of 2-dimensional probability densities

Let:

$$u = u(x, y) \quad \text{and} \quad v = v(x, y) \tag{4.6}$$

be continuous functions defined on set B with continuous first order partial derivatives. Assume that (4.6) - treated as a system of equations - has exactly one solution, that is given by:

$$x = x(u, v) \quad \text{and} \quad y = y(u, v) \tag{4.7}$$

with continuous first order partial derivatives and that:

$$\frac{D(x, y)}{D(u, v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} \neq 0 \quad \text{for } (u, v) \in \Delta \tag{4.8}$$

where Δ is obtained by applying the relations (4.6) on set B.
If $f_{X,Y}$ is a density function of the random vector (X, Y), then the density $f_{U,V}$ of vector (U, V), such that $U = u(X, Y)$ and $V = v(X, Y)$ is given by:

$$f_{U,V}(u, v) = \begin{cases} f_{X,Y}(x(u, v), y(u, v)) \left|\frac{D(x,y)}{D(u,v)}\right| & \text{for } (u, v) \in \Delta \\ 0 & \text{otherwise} \end{cases} \tag{4.9}$$

4.4 The 2-dimensional Rotation Model

In this section the 2-dimensional Rotation Model will be introduced. We will start with the simplest situation, from which the essence of the model becomes clear. Later on, it will be generalized to more complicated cases; although all calculations are analogous.

4.4.1 The 2-dimensional Rotation Model with finite endpoints and $\theta = \frac{\pi}{4}$

Consider a space (x, y), where x and y constitute the perpendicular axis. We rotate the axis about an angle $\theta = \frac{\pi}{4}$ to the new coordinate system (s, t). We assume that the random vector (S, T), defined on the space (s, t), follow the assumptions of the Basic Model discussed in details in section 4.1. Moreover, we assume that the domain of the function λ_σ, namely the set D, is an interval $[y_b, y_e]$, where $-\infty < y_b < 0 < y_e < \infty$ (finite endpoints). We say that the two random variables X and Y defined on the space (x, y) state the Rotation Model. A visualization of this model is presented in Figure 4.3.

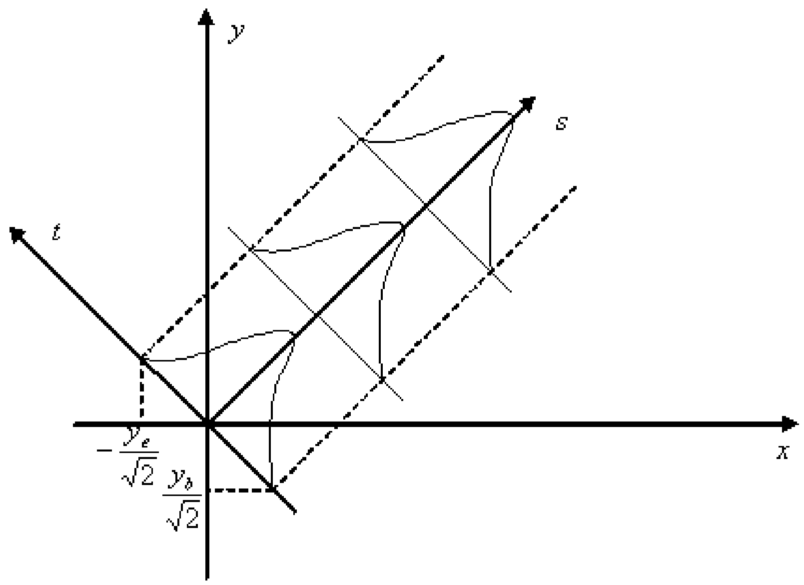


Figure 4.3: The 2-dimensional Rotation Model with finite endpoints and $\theta = \frac{\pi}{4}$.

Clearly, the relation between random vectors (S, T) and (X, Y) is of the form (4.6), with $u = s$ and $r = t$.
Since the angle θ is equal to $\frac{\pi}{4}$, we obtain:

$$\begin{cases} S = \frac{X+Y}{\sqrt{2}} \\ T = \frac{-X+Y}{\sqrt{2}} \end{cases}$$

Invoking expressions (4.3) and (4.9), we obtain the formula for the joint density function $f_{X,Y}(x,y)$ of variables X and Y in terms of the joint density function $f_{S,T}(s,t)$ defined on the set $D = \{(s,t) \in [0,\infty) \times [y_b, y_e]\}$. Since it is easily checked that $D(x,y)/D(s,t) = 1$, we can write:

$$f_{X,Y}(x,y) = \begin{cases} e^{-\left(\frac{x+y}{\sqrt{2}}\right)} \lambda_\sigma\left(\frac{-x+y}{\sqrt{2}}\right) & \text{for } (x,y) \in \Delta \\ 0 & \text{otherwise} \end{cases} \tag{4.10}$$

where Δ is given as follows:

$$\Delta = \{(x,y) : 0 \leq \frac{x+y}{\sqrt{2}}, y_b \leq \frac{-x+y}{\sqrt{2}} \leq y_e\}$$

The integration of the joint density function $f_{X,Y}(x,y)$ over all y values yields the density function $f_X(x)$ of variable X. For $x \geq -\frac{y_e}{\sqrt{2}}$, we have:

$$\begin{aligned} f_X(x) &= \int_{\max\{-x,\sqrt{2}y_b+x\}}^{\sqrt{2}y_e+x} f_{X,Y}(x,y)dy \\ &= \int_{\max\{-x,\sqrt{2}y_b+x\}}^{\sqrt{2}y_e+x} e^{-\left(\frac{x+y}{\sqrt{2}}\right)} \lambda_\sigma\left(\frac{-x+y}{\sqrt{2}}\right) dy \\ &= \sqrt{2}e^{-\sqrt{2}x} \int_{\max\{-\sqrt{2}x,y_b\}}^{y_e} e^{-k}\lambda_\sigma(k)dk \end{aligned} \tag{4.11}$$

We observe that for $x \geq -\frac{y_b}{\sqrt{2}}$ *the density function* $f_X(x)$ *takes the following form:*

$$f_X(x) = \sqrt{2}e^{-\sqrt{2}x}E(e^{-K}) \tag{4.12}$$

where E *denotes the expected value with respect to the density function* λ_σ*.*
Thus, for $x \geq -\frac{y_b}{\sqrt{2}}$ *the function* $f_X(x)$ *becomes a shifted exponential density with a scale parameter* $1/\sqrt{2}$ *and location parameter* $\ln(E(e^{-K}))/\sqrt{2}$*.*

Remark 24 *For* $x \geq -\frac{y_b}{2}$ *the distribution function* $F_X(x)$ *of variable* X *takes the following form:*

$$F_X(x) = 1 - E(e^{-K})e^{-\sqrt{2}x}$$

Hence for $u \in [0,1]$*, the inversion* $F_X^{-1}(u)$ *of function* $F_X(x)$ *becomes:*

$$x = F_X^{-1}(u) = -\frac{1}{\sqrt{2}} \ln\left(\frac{1-u}{E(e^{-K})}\right)$$

The integration of the joint density function $f_{X,Y}(x,y)$ over all x values yields the density function $f_Y(y)$ of variable Y. For $y \geq \frac{y_b}{\sqrt{2}}$, we have:

$$\begin{aligned} f_Y(y) &= \int_{\max\{-y,y-\sqrt{2}y_e\}}^{y-\sqrt{2}y_b} f_{X,Y}(x,y)dx \\ &= \int_{\max\{-y,y-\sqrt{2}y_e\}}^{y-\sqrt{2}y_b} e^{-\left(\frac{x+y}{\sqrt{2}}\right)}\lambda_\sigma\left(\frac{-x+y}{\sqrt{2}}\right)dx \\ &= \sqrt{2}e^{-\sqrt{2}y}\int_{y_b}^{\min\{\sqrt{2}y,y_e\}} e^k\lambda_\sigma(k)dk \end{aligned} \tag{4.13}$$

We observe that for $y \geq \frac{y_e}{\sqrt{2}}$ the density function $f_Y(y)$ takes the following form:

$$f_Y(y) = \sqrt{2}e^{-\sqrt{2}y}E(e^K) \tag{4.14}$$

Thus, for $y \geq \frac{y_e}{\sqrt{2}}$ the function $f_Y(y)$ becomes a shifted exponential density with a scale parameter $1/\sqrt{2}$ and location parameter $\ln(E(e^K))/\sqrt{2}$.

Remark 25 *For $y \geq \frac{y_e}{\sqrt{2}}$ the distribution function $F_Y(y)$ of variable Y takes the following form:*

$$F_Y(y) = 1 - E(e^K)e^{-\sqrt{2}y}$$

Hence for $u \in [0,1]$, the inversion $F_Y^{-1}(u)$ of function $F_Y(y)$ becomes:

$$y = F_Y^{-1}(u) = -\frac{1}{\sqrt{2}}\ln\left(\frac{1-u}{E(e^K)}\right)$$

Example 4 *Let us assume that λ_σ, for $\sigma > 0$, is a density function of the uniform $(-\sqrt{3}\sigma, \sqrt{3}\sigma)$ random variable, hence:*

$$\lambda_\sigma(x) = \begin{cases} \frac{1}{2\sqrt{3}\sigma} & \textit{for } -\sqrt{3}\sigma < x < \sqrt{3}\sigma \\ 0 & \textit{otherwise} \end{cases}$$

the corresponding distribution function Λ_σ is:

$$\Lambda_\sigma(x) = \begin{cases} 0 & \textit{for } x \leq -\sqrt{3}\sigma \\ \frac{x+\sqrt{3}\sigma}{2\sqrt{3}\sigma} & \textit{for } -\sqrt{3}\sigma < x < \sqrt{3}\sigma \\ 1 & \textit{for } x \geq \sqrt{3}\sigma \end{cases}$$

It is straightforward to check that the expected value and standard deviation are 0 *and σ, respectively. Under the above assumptions we obtain the exact expressions for $f_X(x)$ and $F_X(x)$, which are given as follows:*

$$f_X(x) = \begin{cases} 0 & \textit{for } x \leq -\frac{\sqrt{3}\sigma}{\sqrt{2}} \\ \frac{\sqrt{2}e^{-\sqrt{2}x}}{2\sqrt{3}\sigma}(e^{\sqrt{2}x} - e^{-\sqrt{3}\sigma}) & \textit{for } -\frac{\sqrt{3}\sigma}{\sqrt{2}} < x < \frac{\sqrt{3}\sigma}{\sqrt{2}} \\ \frac{\sqrt{2}e^{-\sqrt{2}x}}{2\sqrt{3}\sigma}(e^{\sqrt{3}\sigma} - e^{-\sqrt{3}\sigma}) & \textit{for } x \geq -\frac{\sqrt{3}\sigma}{\sqrt{2}} \end{cases}$$

$$F_X(x) = \begin{cases} 0 & \textit{for } x \leq -\frac{\sqrt{3}\sigma}{\sqrt{2}} \\ \frac{(e^{-\sqrt{2}x-\sqrt{3}\sigma}-1)}{2\sqrt{3}\sigma} + \frac{\sqrt{2}(x+\frac{\sqrt{3}\sigma}{\sqrt{2}})}{2\sqrt{3}\sigma} & \textit{for } -\frac{\sqrt{3}\sigma}{\sqrt{2}} < x < \frac{\sqrt{3}\sigma}{\sqrt{2}} \\ 1 - \frac{(e^{\sqrt{3}\sigma-\sqrt{2}x}-e^{-\sqrt{3}\sigma-\sqrt{2}x})}{2\sqrt{3}\sigma} & \textit{for } x \geq -\frac{\sqrt{3}\sigma}{\sqrt{2}} \end{cases}$$

The formulas for $f_Y(y)$ and $F_Y(y)$ are analogous. Figure 4.4 pictures the density function $f_X(x)$ for $\sigma = 2$.

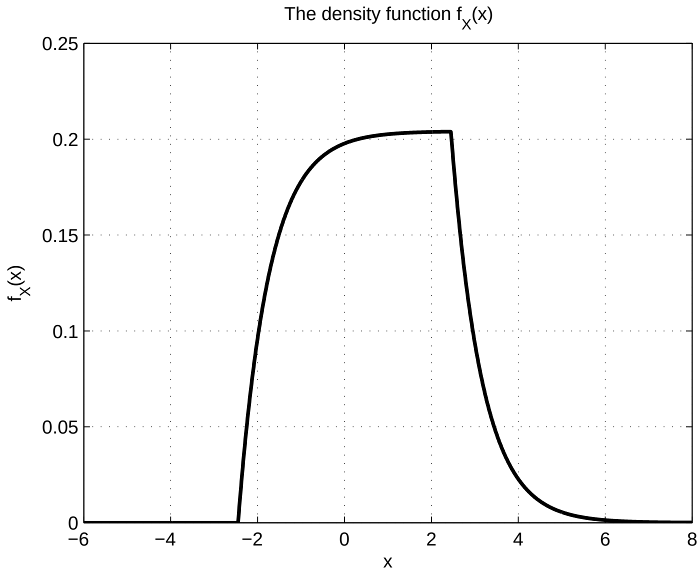


Figure 4.4: The density function $f_X(x)$ when λ_σ is a uniform $(-2\sqrt{3}, 2\sqrt{3})$ density.

4.4.2 The 2-dimensional Rotation model with finite endpoints and $\theta \in [0, \pi/2]$

In this subsection the 2-dimensional Rotation Model with an arbitrary angle $\theta \in [0, \pi/2]$ will be discussed. We again assume that the domain of the function λ_σ, namely the set D, is an interval $[y_b, y_e]$, where $-\infty < y_b < 0$ and $0 < y_e < \infty$. Clearly, the system of equations (4.5) describes the dependency between spaces (X, Y) and (S, T). Then, recalling (4.3) and (4.9) results in:

$$f_{X,Y}(x,y) = \begin{cases} e^{-(x\cos\theta + y\sin\theta)}\lambda_\sigma(-x\sin\theta + y\cos\theta) & \text{for } (x,y) \in \Delta \\ 0 & \text{otherwise} \end{cases} \tag{4.15}$$

where the set Δ is given by:

$$\Delta = \{(x,y) : 0 \leq x\cos\theta + y\sin\theta \text{ , } y_b \leq -x\sin\theta + y\cos\theta \leq y_e\}$$

The integration of the joint density function $f_{X,Y}(x,y)$ over all y values yields the density function $f_X(x)$ of variable X. For $x \geq -y_e \sin\theta$ we have:

$$f_X(x) \;\; = \;\; \int_{\max\{-x\frac{\cos\theta}{\sin\theta}, \frac{y_b + x\sin\theta}{\cos\theta}\}}^{\frac{y_e + x\sin\theta}{\cos\theta}} f_{X,Y}(x,y)dy$$

$$
\begin{aligned}
&= \int_{\max\{-x\frac{\cos\theta}{\sin\theta},\frac{y_b+x\sin\theta}{\cos\theta}\}}^{\frac{y_e+x\sin\theta}{\cos\theta}} e^{-(x\cos\theta+y\sin\theta)}\lambda_\sigma\left(-x\sin\theta+y\cos\theta\right)dy \\
&= \frac{1}{\cos\theta}e^{-\frac{x}{\cos\theta}}\int_{\max\{-\frac{x}{\sin\theta},y_b\}}^{y_e} e^{-k\frac{\sin\theta}{\cos\theta}}\lambda_\sigma(k)dk \qquad (4.16)
\end{aligned}
$$

We observe that for $x \geq -y_b\sin\theta$ the density function $f_X(x)$ takes the following form:

$$
\frac{1}{\cos\theta}e^{-\frac{x}{\cos\theta}}E\left(e^{-K\frac{\sin\theta}{\cos\theta}}\right) \qquad (4.17)
$$

Thus, for $x \geq -y_b\sin\theta$ the function $f_X(x)$ becomes a shifted exponential density with scale parameter $\cos\theta$ and location parameter $\ln(E(e^{-K\frac{\sin\theta}{\cos\theta}}))\cos\theta$.
The integration of the joint density function $f_{X,Y}(x,y)$ over all y values yields the density function $f_X(x)$ of variable X. For $y \geq y_b\cos\theta$ we have:

$$
\begin{aligned}
f_Y(y) &= \int_{\max\{-y\frac{\sin\theta}{\cos\theta},\frac{-y_e+y\cos\theta}{\sin\theta}\}}^{\frac{-y_b+y\cos\theta}{\sin\theta}} f_{X,Y}(x,y)dx \\
&= \int_{\max\{-y\frac{\sin\theta}{\cos\theta},\frac{-y_e+y\cos\theta}{\sin\theta}\}}^{\frac{-y_b+y\cos\theta}{\sin\theta}} e^{-(x\cos\theta+y\sin\theta)}\lambda_\sigma\left(-x\sin\theta+y\cos\theta\right)dx \\
&= \frac{1}{\sin\theta}e^{-\frac{y}{\sin\theta}}\int_{y_b}^{\min\{\frac{y}{\cos\theta},y_e\}} e^{k\frac{\cos\theta}{\sin\theta}}\lambda_\sigma(k)dk \qquad (4.18)
\end{aligned}
$$

We observe that for $y \geq y_e\cos\theta$ the density function $f_Y(y)$ takes the following form:

$$
\frac{1}{\sin\theta}e^{-\frac{y}{\sin\theta}}E\left(e^{K\frac{\cos\theta}{\sin\theta}}\right) \qquad (4.19)
$$

Thus, for $y \geq y_e\cos\theta$ the function $f_Y(y)$ becomes a shifted exponential density with scale parameter $\sin\theta$ and location parameter $\ln(E(e^{K\frac{\cos\theta}{\sin\theta}}))\sin\theta$.

Remark 26 *Let us assume that random variables X and Y follow the Rotation Model with finite endpoints and $\theta = 0$, then this model becomes equal to the Basic Model. Indeed, using (4.3), (4.5) and (4.9) we obtain:*

$$
f_{X,Y}(x,y) = \begin{cases} e^{-x}\lambda_\sigma(y) & \textit{for } (x,y) \in \Delta \\ 0 & \textit{otherwise} \end{cases}
$$

where Δ is:

$$
\Delta = \{(x,y) : x \geq 0, y_b \leq y \leq y_e\}
$$

Hence, X becomes exactly standard exponential random variable. Moreover, X and Y are independent.
If we assume that random variables X and Y satisfy the assumptions of the Rotation Model with finite endpoints and $\theta = \frac{\pi}{2}$, then using (4.3), (4.5) and (4.9) we obtain:

$$
f_{X,Y}(x,y) = \begin{cases} e^{-y}\lambda_\sigma(-x) & \textit{for } (x,y) \in \Delta \\ 0 & \textit{otherwise} \end{cases}
$$

where Δ *is:*
$$\Delta = \{(x,y) : y \geq 0, y_b \leq -x \leq y_e\}$$
Therefore, the random variable Y *becomes exactly standard exponential, and* X *and* Y *are independent.*

4.4.3 The 2-dimensional Rotation Model with infinite endpoints

In this subsection the Rotation Model with infinite endpoints will be considered. The term infinite endpoints refers to the function λ_σ and its domain D. We assume that the set D becomes $(-\infty, \infty)$.
For purpose of application we assume that the angle θ is equal to $\frac{\pi}{4}$ and that λ_σ is a normal density function with corresponding mean zero and standard deviation $\sigma > 0$. We expect the model to be symmetric, due to the symmetry of the normal density around its mean. The above assumptions and relations (4.3) and (4.9) entail that the joint density function $f_{X,Y}(x,y)$ of variables X and Y is given by the following expression:

$$f_{X,Y}(x,y) = \begin{cases} e^{-\left(\frac{x+y}{\sqrt{2}}\right)} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(-x/\sqrt{2}+y/\sqrt{2})^2}{2\sigma^2}} & \text{for } (x,y) \in \Delta \\ 0 & \text{otherwise} \end{cases} \tag{4.20}$$

where Δ is given as follows:

$$\Delta = \{(x,y) : x + y \geq 0\}$$

Because $f_{X,Y}(x,y) = f_{X,Y}(y,x)$ then we conclude that the model is symmetric.
The calculations below give the density function $f_X(x)$ of variable X. For $x \in \Re$ we have:

$$\begin{aligned} f_X(x) &= \int_{-x}^{\infty} e^{-(x/\sqrt{2}+y/\sqrt{2})} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(-x/\sqrt{2}+y/\sqrt{2})^2}{2\sigma^2}} dy \\ &= \sqrt{2} \int_{-\sqrt{2}x}^{\infty} e^{-(k+\sqrt{2}x)} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{k^2}{2\sigma^2}} dk \\ &= \sqrt{2} e^{-\sqrt{2}x} e^{\frac{\sigma^2}{2}} \int_{-\sqrt{2}x}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(k+\sigma^2)^2}{2\sigma^2}} dk \\ &= \sqrt{2} e^{-\sqrt{2}x} e^{\frac{\sigma^2}{2}} \int_{\frac{-\sqrt{2}x+\sigma^2}{\sigma}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \\ &= \sqrt{2} e^{-\sqrt{2}x} e^{\frac{\sigma^2}{2}} [1 - \Phi((-\sqrt{2}x + \sigma^2)/\sigma)] \end{aligned} \tag{4.21}$$

Then, the distribution function $F_X(x)$ of variable X arises form the integration of expression (4.21):

$$F_X(x) = \sqrt{2} e^{\frac{\sigma^2}{2}} \int_{-\infty}^{x} e^{-\sqrt{2}t} [1 - \Phi((-\sqrt{2}t + \sigma^2)/\sigma)] dt$$

Due to the symmetry of the model:

$$F_Y(y) = \sqrt{2}e^{\frac{\sigma^2}{2}} \int_{-\infty}^{y} e^{-\sqrt{2}t}[1 - \Phi((-\sqrt{2}t + \sigma^2)/\sigma)]dt$$

Figure 4.5 presents the density function $f_X(x)$ with $\sigma = 2$. Now, we would like to

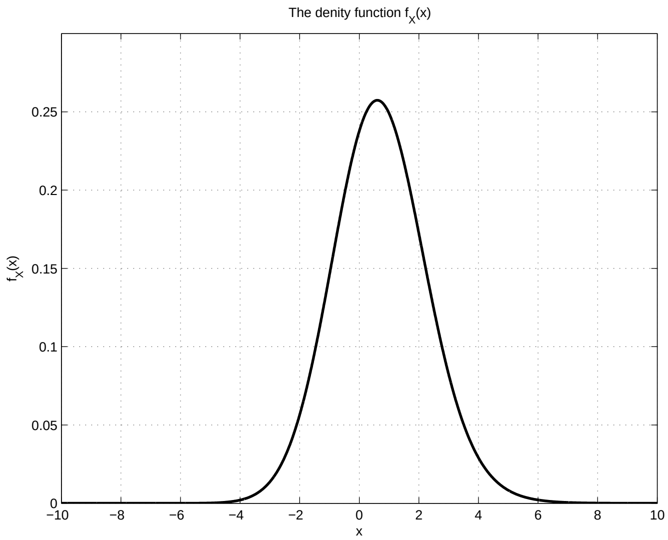


Figure 4.5: The density function $f_X(x)$ when λ_σ is a normal density function with mean zero and standard deviation $\sigma = 2$.

focus on the conditional density function $f_{Y|X=x}(y|x)$ of variable $Y|X = x$, for $x \in \Re$. It is derived as follows. For $y \geq -x$:

$$f_{Y|X=x}(y|x) \quad = \quad \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{e^{-(x/\sqrt{2}+y/\sqrt{2})}\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(-x/\sqrt{2}+y/\sqrt{2})^2}{2\sigma^2}}}{\int_{-x}^{\infty} e^{-(x/\sqrt{2}+y/\sqrt{2})}\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(-x/\sqrt{2}+y/\sqrt{2})^2}{2\sigma^2}}dy}$$

After reduction of common terms and rearrangement of variables we obtain:

$$f_{Y|X=x}(y|x) \quad = \quad \frac{e^{-\frac{(y/\sqrt{2}-x/\sqrt{2}+\sigma^2)^2}{2\sigma^2}}}{\int_{-x}^{\infty} e^{-\frac{(y/\sqrt{2}-x/\sqrt{2}+\sigma^2)^2}{2\sigma^2}}dy} \quad \text{for } y \geq -x$$

If we put $\sigma_1 = \sqrt{2}\sigma$, this can also be written as:

$$f_{Y|X=x}(y|x) = \frac{\frac{1}{\sigma_1\sqrt{2\pi}}e^{-[y-(x-\sigma_1^2/\sqrt{2})]^2/(2\sigma_1^2)}}{\int_{-x}^{\infty}\frac{1}{\sigma_1\sqrt{2\pi}}e^{-[y-(x-\sigma_1^2/\sqrt{2})]^2/(2\sigma_1^2)}dy} \quad \text{for } y \geq -x$$

Thus, we conclude that function $f_{Y|X=x}(y|x)$, for $x \in \Re$ and $y \geq -x$, is a left-truncated normal density function.

4.4.4 Extension of the Rotation Model with infinite endpoints and $\theta = \frac{\pi}{4}$

One of the possible extensions of the Rotation Model with infinite endpoints and $\theta = \pi/4$ is the introduction of the function $\sigma(x)$ for $x \in [0, \infty)$, which is "responsible" for the spread of the variables. Hence, we assume that the 2-dimensional Basic Model is given by the random variables S and T, such that:

$$f_{T|S=s}(s,t) = \lambda_{\sigma(s)}(t) \quad \text{for } s \geq 0 \text{ and } t \in \Re$$

where λ_σ is a normal probability density function with the corresponding mean zero and standard deviation $\sigma > 0$. Moreover, we assume that the density function $f_S(s)$ of variable S is given by:

$$f_S(s) = e^{-s} \quad \text{for } s \geq 0$$

We observe that variables S and T are no longer independent, as it was in the previous case. This is due to the fact that the conditional density function $f_{T|S=s}(t|s)$ contains the term $\sigma(s)$, which is a function of s.
Under the assumption of the angle $\theta = \pi/4$, the Rotation Model in terms of the random variables X and Y yields the density:

$$f_{X,Y}(x,y) = \begin{cases} e^{-(x/\sqrt{2}+y/\sqrt{2})}\lambda_{\sigma(x/\sqrt{2}+y/\sqrt{2})}(-x/\sqrt{2}+y/\sqrt{2}) & \text{for } x+y \geq 0 \\ 0 \quad \text{otherwise} \end{cases}$$

Figure 4.6 pictures the joint density function $f_{X,Y}(x,y)$ and its contour plot with the assumption $\sigma(x) = 2 + x/2$.
The assumption of the normal density function λ_σ entails that $f_{X,Y}(x,y)$ is symmetric as confirmed in the contour plot presented in Figure 4.6. Moreover, since the density function is symmetric, it immediately follows that the marginal distributions $F_X(x)$ and $F_Y(y)$ take the same form.

4.5 Tail dependence for the 2-dimensional Rotation Model

In this section the upper tail dependence coefficient for the 2-dimensional Rotation Model will be derived. The tail dependence coefficients will be calculated for the models discussed in subsections 4.4.1, 4.4.3 and 4.4.4. We refer to Chapter 2, where the definition and necessary propositions concerning the tail dependence coefficients are introduced.

4.5.1 Tail dependence for the 2-dimensional Rotation Model with finite endpoints and $\theta = \frac{\pi}{4}$

We assume that vector (X, Y) is related to the Rotation Model with finite endpoints and $\theta = \frac{\pi}{4}$. Our interest is to calculate the lower and upper tail dependence coefficients. We begin with the definition of the lower tail dependence coefficient λ^L,

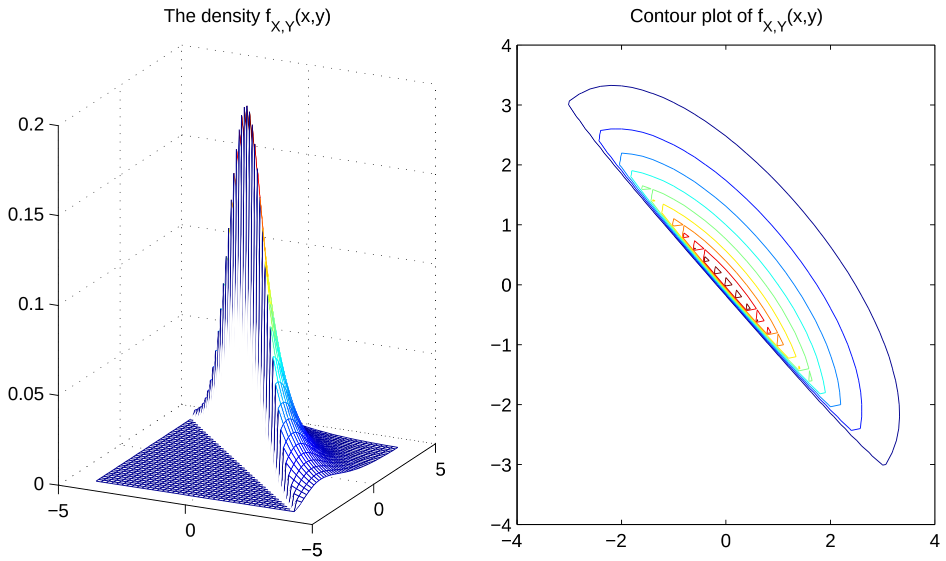


Figure 4.6: The joint density function $f_{X,Y}(x,y)$ and its contour plot under the assumption $\sigma(x) = 2 + x/2$.

namely:

$$\lambda^L = \lim_{u \downarrow 0} P\{X \leq F_X^{-1}(u) | Y \leq F_Y^{-1}(u)\}$$

then, using the definition of the conditional probability and the fact that $P\{Y \leq F_Y^{-1}(u)\} = u$ we obtain:

$$\lambda^L = \lim_{u \downarrow 0} \frac{P\{X \leq F_Y^{-1}(u), Y \leq F_Y^{-1}(u)\}}{u} \tag{4.22}$$

We observe that $F_X^{-1}(u) \downarrow -y_e/\sqrt{2} < 0$ and $F_Y^{-1}(u) \downarrow y_b/\sqrt{2} < 0$ as $u \downarrow 0$. Moreover, by the definition of the joint density function $f_{X,Y}(x,y)$, $P\{X \leq x, Y \leq y\} = 0$ for $x < 0$ and $y < 0$. This implies that for some $\delta > 0$, the numerator of (4.22) becomes 0 for all $u < \delta$, see Figure 4.7 as an illustration of this fact. Therefore, we conclude that $\lambda^L = 0$.
Let us recall that the upper tail dependence coefficient λ^U can be computed from the following formula:

$$\lambda^U = \lim_{u \uparrow 1} (P\{X > F_X^{-1}(u) | Y = F_Y^{-1}(u)\} + P\{Y > F_Y^{-1}(u) | X = F_X^{-1}(u)\})$$

Let us denote $x = F_X^{-1}(u)$ and $y = F_Y^{-1}(u)$. Using relations (4.10) and (4.13) we first compute the probability:

$$\begin{aligned} P\{X > x | Y = y\} &= \int_x^\infty f_{X|Y=y}(k|y) dk = \int_x^\infty \frac{f_{X,Y}(k,y)}{f_Y(y)} dk \\ &= \frac{1}{f_Y(y)} \int_x^\infty e^{-(k/\sqrt{2} + y/\sqrt{2})} \lambda_\sigma(-k/\sqrt{2} + y/\sqrt{2}) dk \end{aligned}$$

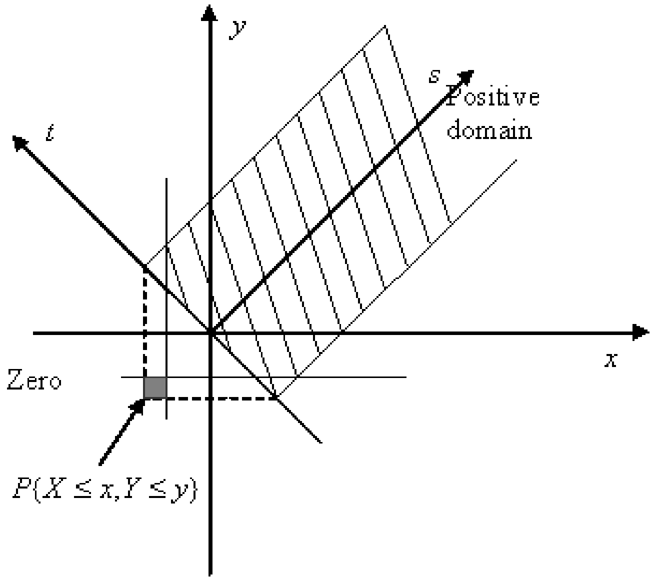


Figure 4.7: Lower tail dependence for the 2-dimensional Rotation Model with finite endpoints and $\theta = \frac{\pi}{4}$.

$$
\begin{aligned}
&= \frac{1}{f_Y(y)}\sqrt{2}e^{-\sqrt{2}y}\int_{-\infty}^{-x/\sqrt{2}+y/\sqrt{2}} e^z\lambda_\sigma(z)dz \\
&= \frac{1}{f_Y(y)}\sqrt{2}e^{-\sqrt{2}y}\int_{y_b}^{-x/\sqrt{2}+y/\sqrt{2}} e^z\lambda_\sigma(z)dz \\
&= \frac{\int_{y_b}^{-x/\sqrt{2}+y/\sqrt{2}} e^z\lambda_\sigma(z)dz}{\int_{y_b}^{\min\{\sqrt{2}y,y_e\}} e^z\lambda_\sigma(z)dz} \\
&= \frac{\int_{y_b}^{-x/\sqrt{2}+y/\sqrt{2}} e^z\lambda_\sigma(z)dz}{E(e^Z)} \qquad (4.23)
\end{aligned}
$$

where in the last step we assumed $\sqrt{2}y > y_e$, which will be fulfilled for u sufficiently close to 1.
Now, we focus on the limit:

$$\lim_{u\uparrow 1}\frac{\sqrt{2}}{2}(y-x) \qquad (4.24)$$

In order to deal with (4.24), we consider the distributions $F_X(x)$ and $F_Y(y)$ for $x \geq -y_b/\sqrt{2}$ and $y \geq y_e/\sqrt{2}$, because we are interested in large values of x and y. The small values are "forgotten" while the limit is taken. Thus, Remarks 24 and 25 entail:

$$
\begin{aligned}
\lim_{u\uparrow 1}\frac{\sqrt{2}}{2}(y-x) &= \lim_{u\uparrow 1}\frac{\sqrt{2}}{2}(F_Y^{-1}(u) - F_X^{-1}(u)) \\
&= \lim_{u\uparrow 1}\frac{\sqrt{2}}{2}\left(-\frac{1}{\sqrt{2}}\ln\left(\frac{1-u}{E(e^K)}\right) + \frac{1}{\sqrt{2}}\ln\left(\frac{1-u}{E(e^{-K})}\right)\right)
\end{aligned}
$$

$$= \frac{1}{2}\ln\left(\frac{E(e^K)}{E(e^{-K})}\right) \tag{4.25}$$

Let us define the following quantity:

$$y_0 = \frac{1}{2}\ln\left(\frac{E(e^K)}{E(e^{-K})}\right)$$

The result (4.25) allows us to compute the following limit:

$$\lim_{u\uparrow 1} P\{X > x|Y = y\} = \frac{\int_{y_b}^{y_0} e^z \lambda_\sigma(z)dz}{E(e^K)} \tag{4.26}$$

Using relations (4.10) and (4.11) we compute the probability:

$$\begin{aligned}
P\{Y > y|X = x\} &= \int_y^\infty f_{Y|X=x}(k|x)dk = \frac{1}{f_X(x)}\int_y^\infty f(x,k)dk \\
&= \frac{1}{f_X(x)}\int_y^\infty e^{-(x/\sqrt{2}+k/\sqrt{2})}\lambda_\sigma(-x/\sqrt{2}+k/\sqrt{2})dk \\
&= \frac{1}{f_X(x)}\sqrt{2}e^{-\sqrt{2}x}\int_{-x/\sqrt{2}+y/\sqrt{2}}^{y_e} e^{-z}\lambda_\sigma(z)dz \\
&= \frac{\int_{-x/\sqrt{2}+y/\sqrt{2}}^{y_e} e^{-z}\lambda_\sigma(z)dz}{\int_{\max\{-\sqrt{2}x,y_b\}}^{y_e} e^{-k}\lambda_\sigma(k)dk} \\
&= \frac{\int_{(-x/\sqrt{2}+y/\sqrt{2})}^{y_e} e^{-z}\lambda_\sigma(z)dz}{E(e^{-K})}
\end{aligned}$$

where in the last step we assumed $-\sqrt{2}x < y_b$, which will be fulfilled for u sufficiently close to 1. Now, using this and result (4.25) we obtain:

$$\lim_{u\uparrow 1} P\{Y > y|X = x\} = \frac{\int_{y_0}^{y_e} e^{-z}\lambda_\sigma(z)dz}{E(e^{-K})} \tag{4.27}$$

The combination of (4.26) and (4.27) yields the upper tail dependence coefficient λ^U:

$$\lambda^U = \frac{\int_{y_b}^{y_0} e^z\lambda_\sigma(z)dz}{E(e^K)} + \frac{\int_{y_0}^{y_e} e^{-z}\lambda_\sigma(z)dz}{E(e^{-K})}$$

Proposition 15 *If K is a random variable with a symmetric density function λ_σ defined on a domain $[y_b, y_e]$ with finite endpoints and $y_b = -y_e$ then:*

1. $E(e^{-K}) = E(e^K)$
2. $\int_{y_b}^0 e^k\lambda_\sigma(k)dk = \int_0^{y_e} e^{-k}\lambda_\sigma(k)dk$

Proof Using the definition of the expected value and the above assumptions we prove the first part of the proposition:

$$\begin{aligned} E(e^{-K}) &= \int_{y_b}^{y_e} e^{-k}\lambda_\sigma(k)dk = \int_{y_b}^{y_e} e^{-k}\lambda_\sigma(-k)dk \\ &= \int_{-y_e}^{-y_b} e^{z}\lambda_\sigma(z)dz = \int_{y_b}^{y_e} e^{z}\lambda_\sigma(z)dz \\ &= E(e^{K}) \end{aligned}$$

The assumptions entail the proof of the second part of the proposition:

$$\begin{aligned} \int_{y_b}^{0} e^{k}\lambda_\sigma(k)dk &= \int_{y_b}^{0} e^{k}\lambda_\sigma(-k)dk \\ &= -\int_{-y_b}^{0} e^{-z}\lambda_\sigma(z)dz = \int_{0}^{y_e} e^{-k}\lambda_\sigma(k)dk \end{aligned}$$

Hence, both parts of the proposition have been proved.

Thus, under the assumptions of symmetry of function λ_σ and $y_b = -y_e$ the upper tail dependence coefficient λ^U becomes:

$$\lambda^U = \frac{2\int_0^{y_e} e^{-k}\lambda_\sigma(k)dk}{E(e^K)}$$

4.5.2 Tail dependence for the Rotation Model with infinite endpoints and $\theta = \frac{\pi}{4}$

We consider the 2-dimensional Rotation Model with infinite endpoints, where λ_σ is a normal density function with mean zero and standard deviation $\sigma > 0$, and $\theta = \frac{\pi}{4}$.
The argument in the derivation of the lower tail dependence coefficient λ^L is analogous as for the Rotation Model with finite endpoints and $\theta = \frac{\pi}{4}$. Thus, we observe that $F_X^{-1}(u) \to \infty$ and $F_Y^{-1}(u) \to \infty$ as $u \to 0$. By definition, the "positive" domain of the random vector (X, Y) is $\{(x, y) : x + y \geq 0\}$. Hence, $P\{X \leq x, Y \leq y\} = 0$ for $x < 0$ and $y < 0$. This entails that $\lambda^L = 0$. Figure 4.8 depicts this situation.
In order to calculate the upper tail dependence coefficient we recall that the considered model is symmetric, what entails the usage of Corollary 1. Let us denote $x = F_X^{-1}(u)$ for $u \in [0, 1]$, then we observe that $x \to \infty$ as $u \uparrow 1$. Hence, computation of the upper tail dependence coefficient λ^U diminishes to:

$$\begin{aligned} \lambda^U &= \lim_{u\uparrow 1} 2P\{X \geq x | Y = x\} = \lim_{u\uparrow 1} 2\int_x^\infty f_{X|Y=x}(k|x)dk = \lim_{u\uparrow 1} \frac{2}{f_Y(x)}\int_x^\infty f_{X,Y}(k, x)dk \\ &= \lim_{u\uparrow 1} \frac{2}{f_Y(x)}\int_x^\infty e^{-(k/\sqrt{2}+x/\sqrt{2})}\lambda_\sigma(-k/\sqrt{2} + x/\sqrt{2})dk \\ &= \lim_{u\uparrow 1} \frac{2}{f_Y(x)}\sqrt{2}e^{-\sqrt{2}x}\int_{-\infty}^{0} e^{z}\lambda_\sigma(z)dz \end{aligned}$$

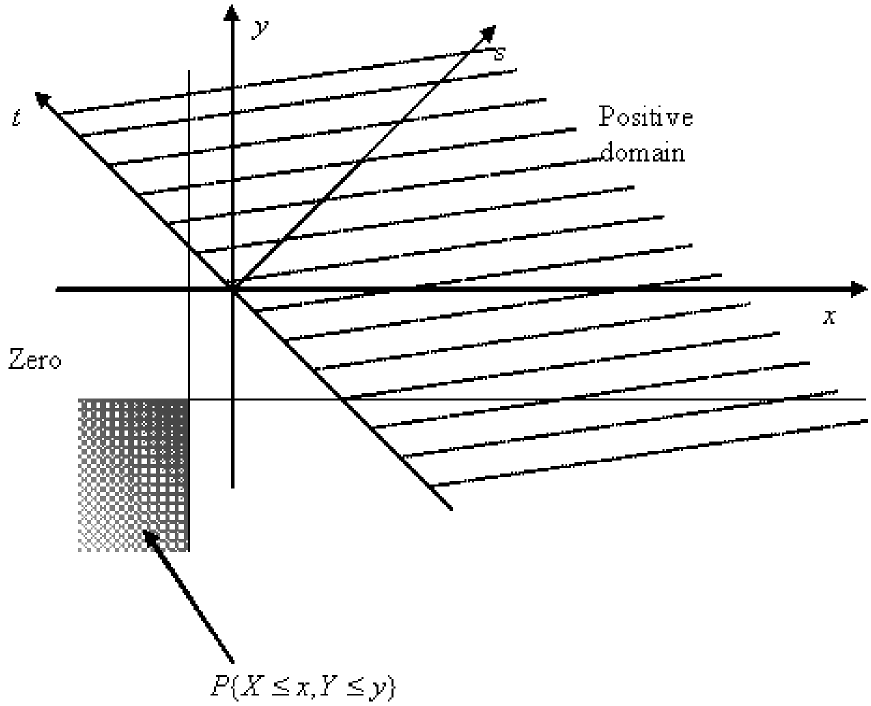


Figure 4.8: Lower tail dependence for the 2-dimensional Rotation Model with infinite endpoints and $\theta = \frac{\pi}{4}$.

$$
\begin{aligned}
&= \lim_{u \uparrow 1} \frac{2e^{-\frac{\sigma^2}{2}}}{\Phi((\sqrt{2}x - \sigma^2)/\sigma)} \int_{-\infty}^{0} e^z \lambda_\sigma(z) dz \\
&= 2e^{-\frac{\sigma^2}{2}} \int_{-\infty}^{0} e^z \lambda_\sigma(z) dz = 2e^{-\frac{\sigma^2}{2}} \int_{-\infty}^{0} e^z \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{z^2}{2\sigma^2}} dz \\
&= 2e^{-\frac{\sigma^2}{2}} \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(z-\sigma^2)^2}{2\sigma^2}} e^{\frac{\sigma^2}{2}} dz = 2 \int_{-\infty}^{-\sigma} \frac{1}{\sqrt{2\pi}} e^{-\frac{w^2}{2}} dw \\
&= 2\Phi(-\sigma) \qquad (4.28)
\end{aligned}
$$

We observe that the resulting upper tail dependence coefficient is less than 1. Since $\Phi(-\sigma) < \Phi(0) = 1/2$, we have:

$$\lambda^U = 2\Phi(-\sigma) < 2\Phi(0) = 1$$

Thus, if σ is equal to $\epsilon > 0$, which is assumed to be a very small number, then λ^U becomes close to 1. It means that in the tail the variables X and Y then become "almost" completely dependent.

4.5.3 The tail dependence for the extension of the Rotation Model with infinite endpoints and $\frac{\pi}{4}$

In this subsection the tail dependence coefficients for the generalization of the Rotation Model with infinite endpoints and $\theta = \pi/4$ will be derived. In addition to the assumptions of subsection 4.4.4, we assume that function $\sigma(x)$ satisfies the following conditions:

1. $\sigma(x)$ is an increasing function and $\sigma(0) = \sigma_0 > 0$
2. $\lim_{x\to\infty} \sigma(x) = \sigma_1 < \infty$

Clearly, the lower tail dependence coefficient λ^L is 0. The explanation is analogous to that for the lower tail dependence coefficient for the Rotation Model discussed in subsection 4.4.3, and follows from the definition of the joint density function $f_{X,Y}(x,y)$. In order to derive the upper tail dependence coefficient λ^U, we recall that the considered model is symmetric, what entails the usage of the Corollary 1. Let us denote $x = F_X^{-1}(u)$, then we observe that $x \to \infty$ as $u \uparrow 1$. Consider first the following limit:

$$\begin{aligned}
\lim_{u\uparrow 1} P\{Y \le x | X = x\} &= \lim_{u\uparrow 1} \int_{-x}^{x} f_{Y|X=x}(k|x)dk \\
&= \lim_{u\uparrow 1} \frac{1}{f_X(x)} \int_{-x}^{x} f_{X,Y}(x,k)dk \\
&= \lim_{u\uparrow 1} \frac{\int_{-x}^{x} e^{-(x/\sqrt{2}+k/\sqrt{2})} \lambda_{\sigma(x/\sqrt{2}+k/\sqrt{2})}(-x/\sqrt{2}+k/\sqrt{2})dk}{\int_{-x}^{\infty} e^{-(x/\sqrt{2}+y/\sqrt{2})} \lambda_{\sigma(x/\sqrt{2}+y/\sqrt{2})}(-x/\sqrt{2}+y/\sqrt{2})dy} \\
&= \lim_{u\uparrow 1} \frac{\int_{-\sqrt{2}x}^{0} \sqrt{2} e^{-(\sqrt{2}x+z)} \lambda_{\sigma(\sqrt{2}x+z)}(z)dz}{\int_{-\sqrt{2}x}^{\infty} \sqrt{2} e^{-(\sqrt{2}x+z)} \lambda_{\sigma(\sqrt{2}x+z)}(z)dz} \\
&= \lim_{u\uparrow 1} \frac{\int_{-\sqrt{2}x}^{0} e^{-z} \lambda_{\sigma(\sqrt{2}x+z)}(z)dz}{\int_{-\sqrt{2}x}^{\infty} e^{-z} \lambda_{\sigma(\sqrt{2}x+z)}(z)dz} \\
&= \frac{\int_{-\infty}^{0} e^{-z} \lambda_{\sigma_1}(z)dz}{\int_{-\infty}^{\infty} e^{-z} \lambda_{\sigma_1}(z)dz} = \Phi(\sigma_1)
\end{aligned}$$

In the above calculations we have used Lebesgue's dominated convergence theorem, which was introduced in subsection 3.4.4. The procedure is the same as in case of the Variable Spread Model and will not be repeated. Then, the upper tail dependence coefficient λ^U is:

$$\lambda^U = 2 \lim_{u\uparrow 1} (1 - P\{Y \le x | X = x\}) = 2(1 - \Phi(\sigma_1)) = 2\Phi(-\sigma_1)$$

Remark 27 *Suppose that:*

$$\lim_{x\to\infty} \sigma(x) = \infty$$

Then we may intuitively expect that $\lambda^U = 0$, although we have not given a mathematically rigorous proof. The lower tail dependence coefficient would remain equal to zero, by the same reasoning as before.

4.6 The 3-dimensional Basic Model

The aim of this section is to provide the theoretical background which will help us to understand the 3-dimensional Rotation Model. We start with the description of the 3-dimensional Basic Model, which is the generalization of the 2-dimensional Basic Model.
We assume that S, T and W are three random variables, such that the density function $f_S(s)$ of random variable S is given as follows:

$$f_S(s) = e^{-s} \quad \text{for } s \geq 0 \tag{4.29}$$

whereas, the conditional density function $f_{T,W|S=s}(t, w|s)$ of variable $(T, W)|S = s$ is:

$$f_{T,W|S=s}(t, w|s) = \begin{cases} \lambda_\sigma(t, w) & \text{for } (t, w) \text{ such that } t^2 + w^2 \leq \sigma^2 \\ 0 & \text{otherwise} \end{cases} \tag{4.30}$$

where λ_σ is a density function and $\sigma > 0$. (We note that σ is no longer the standard deviation of λ_σ.)
Hence, the joint density function $f_{S,T,W}(s, t, w)$ of random vector (S, T, W) arises from (4.29) and (4.30):

$$f_{S,T,W}(s, t, w) = f_S(s) f_{T,W|S=s}(t, w|s) = \begin{cases} e^{-s}\lambda_\sigma(t, w) & \text{for } s \geq 0,\ t^2 + w^2 \leq \sigma^2 \\ 0 & \text{otherwise} \end{cases} \tag{4.31}$$

Figure 4.9 illustrates the construction of this model. This model is the natural generalization of the 2-dimensional Basic Model. *The above assumptions state that S is an exponential random variable and that vector (T, W) does not depend on the random variable S. These facts are analogous to the 2-dimensional Basic model.*

4.7 Rotation in 3 dimensions

In this section a rotation in the 3-dimensional space will be described - that will be used to formulate the 3-dimensional Rotation Model.
Roughly speaking, the rotation of the 3-dimensional system of coordinates is based on fixing one axis and performing a 2-dimensional rotation of the remaining two axis. We assume that the space (x, y, z) is our basic system of coordinates. We start with a rotation about the z axis through an angle $\alpha = \frac{\pi}{4}$ counterclockwise relative to (x, y, z) to give the new system of coordinates (x', y', z'). The rotation matrix R_1 that describes this operation is given by:

$$R_1 = \begin{pmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and we have} \quad \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = R_1 \begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{4.32}$$

Next, the system (x', y', z') is rotated about the y' axis through an angle β counterclockwise to generate the new coordinate system (s, t, w). Analogous to the first

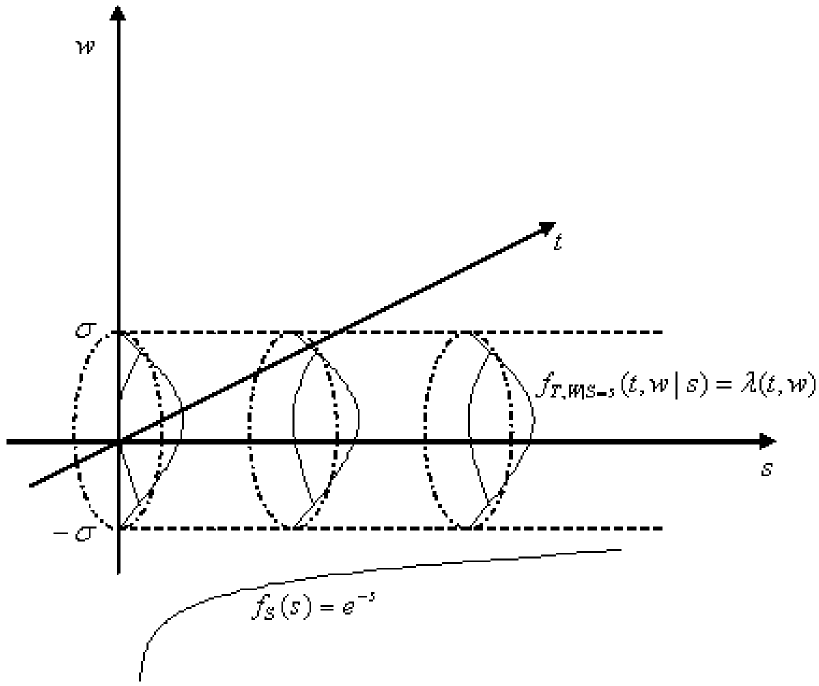


Figure 4.9: The 3-dimensional Basic Model.

rotation, this mixes the coordinates along x' and z', while the coordinate along y' remains unaffected. The angle β is chosen such that:

$$\cos\beta = \sqrt{\frac{2}{3}}$$

The rotation matrix R_2 that corresponds to the above operation is given as follows:

$$R_2 = \begin{pmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{pmatrix} \quad \text{and we have} \quad \begin{pmatrix} s \\ t \\ w \end{pmatrix} = R_2 \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} \tag{4.33}$$

The rotation matrix R, that straightforwardly connects system (x, y, z) with (s, t, w), arises as a product of matrices given by formulas (4.33) and (4.32). The resulting matrix R is then (with $\alpha = \frac{\pi}{4}$ and $\cos\beta = \sqrt{2/3}, \sin\beta = \sqrt{1/3}$) given by:

$$R = R_2R_1 = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & \sqrt{\frac{2}{3}} \end{pmatrix} \quad \text{and we have} \quad \begin{pmatrix} s \\ t \\ w \end{pmatrix} = R \begin{pmatrix} x \\ y \\ z \end{pmatrix} \tag{4.34}$$

Figure 4.10 illustrates the 3-dimensional rotation. *The above relation is indispensable in the next sections*

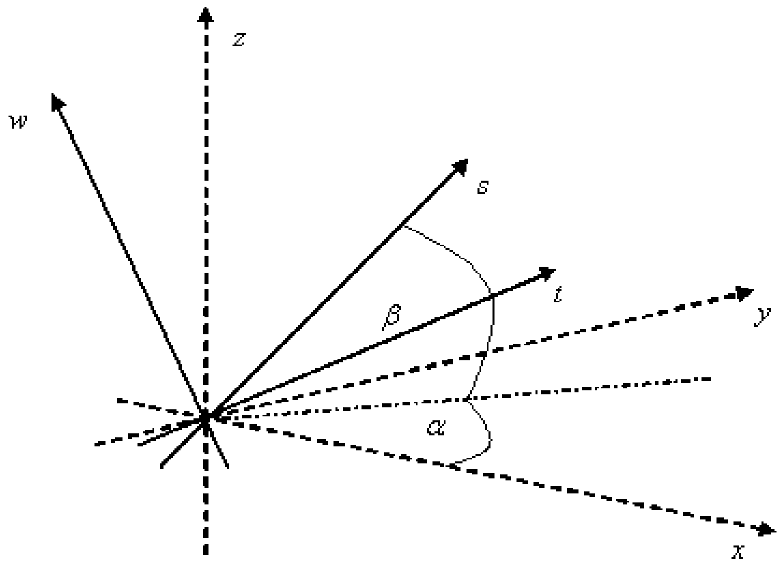


Figure 4.10: The rotation in 3 dimensions.

Remark 28 *Note that* $(s,t,w) = (1,0,0)$ *corresponds to* $(x,y,z) = (1,1,1)/\sqrt{3}$*, from which easily follows* $\cos(x,s) = \cos(y,s) = \cos(z,s) = 1/\sqrt{3}$*, where* $\cos(a,b)$ *denotes the cosine of the angle between axis* a *and* b*. So the* s *axis makes the same angle with each of the axis* x, y *and* z*. The axis* t *and* w *are lying in a plane perpendicular to the* s *axis.*

4.8 Functions of three random variables

Let:

$$u = u(x,y,z), \quad v = v(x,y,z) \quad \text{and} \quad r = r(x,y,z) \tag{4.35}$$

be continuous functions defined on a set D with continuous first order partial derivatives. Assume, that (4.35) - treated as a system of equations - has exactly one solution:

$$x = x(u,v,r) \quad , \quad y = y(u,v,r) \quad z = z(u,v,r) \tag{4.36}$$

with continuous first order partial derivatives and that:

$$\frac{D(x,y,z)}{D(u,v,r)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial r} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial r} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & \frac{\partial z}{\partial r} \end{vmatrix} \neq 0 \quad \text{for } (u,v,r) \in \Delta \tag{4.37}$$

where Δ is obtained by applying the relations (4.35) on set D.
If $f_{X,Y,Z}$ is a density function of random vector (X,Y,Z), then the density $f_{U,V,R}$ of vector the (U,V,R), such that $U = u(X,Y,Z)$, $V = v(X,Y,Z)$ and $R = r(X,Y,Z)$,

is given by:

$$f_{U,V,R}(u,v,r) = \begin{cases} f_{X,Y,Z}(x(u,v,r), y(u,v,r), z(u,v,r)) \left|\frac{D(x,y,z)}{D(u,v,r)}\right| & \text{for } (u,v,r) \in \Delta \\ 0 & \text{otherwise} \end{cases} \tag{4.38}$$

4.9 The 3-dimensional Rotation Model with the restricted support

In this section, we will generalize the 2-dimensional Rotation Model to the 3-variate case. Moreover, we will prove that the resulting marginal distributions are exponential. This fact is consistent with the results of the 2-dimensional Rotation Model.
Let us assume that random vector (S,T,W) follows the 3-dimensional Basic Model described in subsection 4.6. The Rotation Model is represented by random vector (X,Y,Z), such that the relation (4.34) connects it with vector (S,T,W). This fact together with (4.38) yield that the density function $f_{X,Y,Z}(x,y,z)$ has the following form:

$$f_{X,Y,Z}(x,y,z) = e^{-(x/\sqrt{3}+y/\sqrt{3}+z/\sqrt{3})}\lambda_{\sigma}(-\sqrt{2}x/2+\sqrt{2}y/2, -x/\sqrt{6}-y/\sqrt{6}+\sqrt{2/3}z) \tag{4.39}$$

This form is valid for $(x,y,z) \in \Delta$, where set Δ is given by:

$$\Delta = \{(x,y,z) : x/\sqrt{3}+y/\sqrt{3}+z/\sqrt{3} \geq 0, (-\sqrt{2}x/2+\sqrt{2}y/2)^2+(-x/\sqrt{6}-y/\sqrt{6}+\sqrt{2/3}z)^2 \leq \sigma^2\}$$

The other choice of (x,y,z) yields $f_{X,Y,Z}(x,y,z) = 0$. Thus, the model is defined. Figure 4.11 visualizes the construction of this model. From the picture we can suspect the behavior

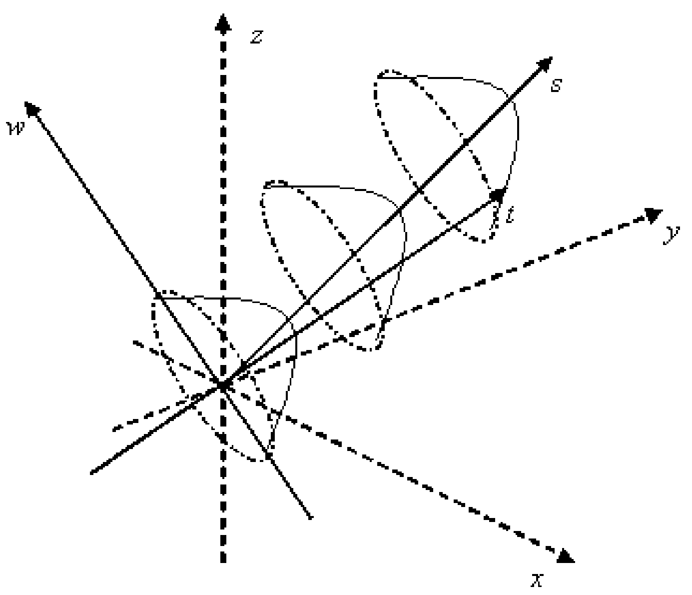


Figure 4.11: The 3-dimensional Rotation Model.

of the distribution in the tail.

The aim of this section is the calculation of the density functions $f_X(x)$, $f_Y(y)$ and $f_Z(z)$ of variables X, Y and Z, respectively. As for the 2-dimensional model with finite endpoints and $\theta = \frac{\pi}{4}$, we will show that the mentioned density functions become exponential from some point.
The density function $f_X(x)$ of random variable X can be derived as follows:

$$f_X(x) = \int\int_{\Delta_x} f_{X,Y,Z}(x,y,z)dydz$$

where $f_{X,Y,Z}(x,y,z)$ is given by equation 4.39 and Δ_x is:

$$\Delta_x = \{(y,z) : x/\sqrt{3}+y/\sqrt{3}+z/\sqrt{3} \geq 0, (-\sqrt{2}x/2+\sqrt{2}y/2)^2+(-x/\sqrt{6}-y/\sqrt{6}+\sqrt{2/3}z)^2 \leq \sigma^2\}$$

Substituting $y = \sqrt{2}k + x$ and $z = \sqrt{3/2}u + \sqrt{2}k/2 + x$ with the Jacobian $J = \sqrt{3}$ results in:

$$f_X(x) = \sqrt{3}e^{-\sqrt{3}x} \int\int_{\Delta'_x} e^{-(\sqrt{3/2}k+u/\sqrt{2})}\lambda_\sigma(k,u)dkdu \tag{4.40}$$

where Δ'_x is:

$$\Delta'_x = \{(k,u) : \sqrt{3/2}k + u/\sqrt{2} + \sqrt{3}x \geq 0, k^2 + u^2 \leq \sigma^2\} \tag{4.41}$$

We observe that for $x \geq \sqrt{2/3}\sigma$, the set Δ'_x becomes equal to:

$$\Delta'_x = \{(k,u) : k^2 + u^2 \leq \sigma^2\}$$

Hence, function $f_X(x)$ appears to be:

$$f_X(x) = \sqrt{3}e^{-\sqrt{3}x}E\left(e^{-\sqrt{3/2}K-U/\sqrt{2}}\right)$$

what entails that X is shifted exponentially distributed for $x \geq \sqrt{2/3}\sigma$ with scale parameter $1/\sqrt{3}$ and location parameter $\ln(E(e^{-\sqrt{3/2}K-U/\sqrt{2}}))/\sqrt{3}$.
Let us consider the density function $f_Y(y)$ of random variable Y. It is computed as follows:

$$f_Y(y) = \int\int_{\Delta_y} f_{X,Y,Z}(x,y,z)dxdz$$

where the set Δ_y is:

$$\Delta_y = \{(x,z) : x/\sqrt{3}+y/\sqrt{3}+z/\sqrt{3} \geq 0, (-\sqrt{2}x/2+\sqrt{2}y/2)^2+(-x/\sqrt{6}-y/\sqrt{6}+\sqrt{2/3}z)^2 \leq \sigma^2\}$$

Substituting $x = y - \sqrt{2}k$ and $z = \sqrt{3/2}u - \sqrt{2}k/2 + y$ with the Jacobian $J = \sqrt{3}$ entails:

$$f_Y(y) = \sqrt{3}e^{-\sqrt{3}y} \int\int_{\Delta'_y} e^{-(-\sqrt{3/2}k+u/\sqrt{2})}\lambda_\sigma(k,u)dkdu \tag{4.42}$$

where the set Δ'_y is given as follows:

$$\Delta'_y = \{(k,u) : -\sqrt{3/2}k + u/\sqrt{2} \geq 0, k^2 + u^2 \leq \sigma^2\} \tag{4.43}$$

We note that for $y \geq \sqrt{2/3}\sigma$, the set Δ'_y becomes equal to:

$$\Delta'_y = \{(k,u) : k^2 + u^2 \leq \sigma^2\}$$

Hence, function $f_Y(y)$ appears to be:

$$f_Y(y) = \sqrt{3}e^{-\sqrt{3}y} E\left(e^{\sqrt{3/2}K - U/\sqrt{2}}\right)$$

what entails that variable Y is shifted exponentially distributed for $y \geq \sqrt{2/3}\sigma$ with scale parameter $1/\sqrt{3}$ and location parameter $\ln(E(e^{\sqrt{3/2}K-U/\sqrt{2}}))/\sqrt{3}$.

Consider the density function $f_Z(z)$ of random variable Z. It is computed as follows:

$$f_Z(z) = \int\int_{\Delta_z} f_{X,Y,Z}(x,y,z)dxdy$$

where the set Δ_z is:

$$\Delta_z = \{(x,y) : x/\sqrt{3}+y/\sqrt{3}+z/\sqrt{3} \geq 0, (-\sqrt{2}x/2+\sqrt{2}y/2)^2+(-x/\sqrt{6}-y/\sqrt{6}+\sqrt{2/3}z)^2 \leq \sigma^2\}$$

Substituting $x = -\sqrt{2}k/2 - \sqrt{3/2}u + z$ and $y = -\sqrt{3/2}u + \sqrt{2}k/2 + z$ with the Jacobian $J = \sqrt{3}$ gives:

$$f_Z(z) = \sqrt{3}e^{-\sqrt{3}z} \int\int_{\Delta'_z} e^{\sqrt{2}u}\lambda_\sigma(k,u)dkdu \tag{4.44}$$

where Δ'_z is:

$$\Delta'_z = \{(k,u) : \sqrt{3}z - \sqrt{2}u \geq 0, k^2 + u^2 \leq \sigma^2\} \tag{4.45}$$

We remark that for $z \geq \sqrt{2/3}\sigma$, the set Δ'_z becomes equal to:

$$\Delta'_z = \{(k,u) : k^2 + u^2 \leq \sigma^2\}$$

Hence, function $f_Z(z)$ appears to be:

$$f_Z(z) = \sqrt{3}e^{-\sqrt{3}z} E\left(e^{\sqrt{2}U}\right)$$

what entails that variable Z is shifted exponentially distributed for $z \geq \sqrt{2/3}\sigma$ with scale parameter $1/\sqrt{3}$ and location parameter $\ln(E(e^{\sqrt{2}U}))/\sqrt{3}$.

4.10 Case study - the 2-dimensional Rotation Model with infinite endpoints

In this section a bivariate case study will be performed using the same dataset as in section 2.6. Thus, given the dataset we will try to find the appropriate Rotation Model. First, by estimating the unknown parameters and later on by applying the percentile lines method. The theoretical background is provided in the previous chapter. As before, we will use the transformation of random variables given by (3.1), which will allow us to model the bivariate density function of the considered dataset.

4.10.1 Assumptions

In this subsection we list the assumptions, that have been made to carry out the case study:

1. *We consider the dataset introduced in subsection 2.6.1,* $(V_i, W_i), i = 1, ..., n,$ *where* $n = 89$*. Observations* $V_i, i = 1, ..., n,$ *denote the water levels, and observations* $W_i, i = 1, ..., n$ *denote the wind speeds.*

2. *We assume that the distribution functions* $F_V(v)$ *and* $F_W(w)$ *of variables* V *and* W *are given and of the form described in subsection 2.6.3.*

3. *For the 2-dimensional Rotation Model with infinite endpoints and constant standard deviation (Model RC) we assume that* λ_σ *is a normal density function with mean zero and standard deviation* $\sigma > 0$*. Moreover, we assume that* σ *is constant in this model. Clearly,* σ *is our parameter of interest and has to be estimated from the data.*

4. *For the 2-dimensional Rotation Model with infinite endpoints and varying standard deviation (Model RV) we assume that* λ_σ *is a normal density function mean zero and standard deviation* $\sigma > 0$*. Moreover, we assume that function* $\sigma_s(x)$ *takes the following form:*

$$\sigma_s(x) = \begin{cases} \frac{\sigma}{5}x + \sigma & \textit{for } 0 \leq x \leq 5 \\ 2\sigma & \textit{for } x > 5 \end{cases} \tag{4.46}$$

The parameter $\sigma > 0$ *is unknown and has to be estimated from the data.*

4.10.2 Estimation of the unknown parameters - Part 1

In this subsection we present results of the Maximum Likelihood Method of estimation of the unknown parameters, when the whole dataset is taken into consideration. The Figures 4.12 and 4.13 illustrate the likelihood functions.

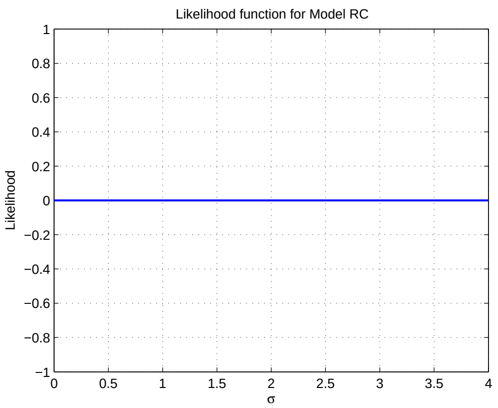


Figure 4.12: The likelihood function for Model RC

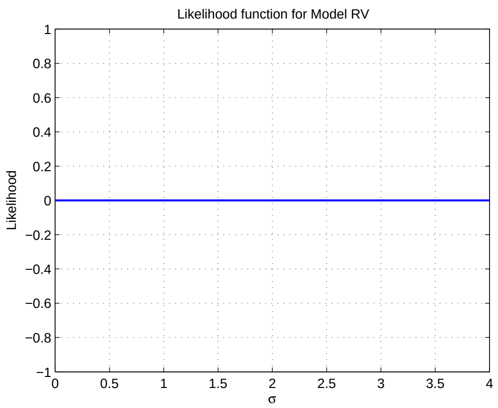


Figure 4.13: The likelihood function for Model RV

Clearly, we cannot observe a unique maximum point, since the likelihood functions are constant and take the value zero. The reason of such a "behavior" is associated with the fact that after the transformation some points have fallen outside the region $D = \{(x, y) : x + y \geq 0\}$, where the joint density function $f_{X,Y}(x, y)$ takes value zero. Moreover, the number of points that fall in the region D vary with respect to the value of parameter σ. This is shown in the Figures 4.14 and 4.15.

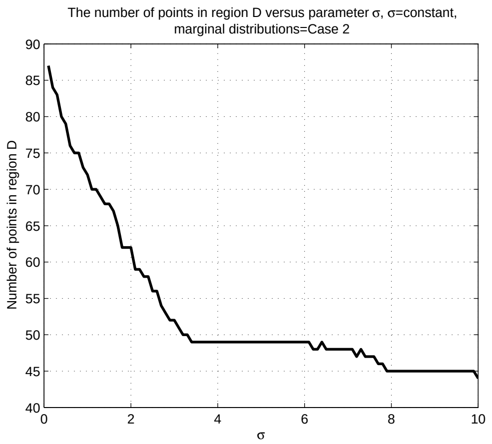


Figure 4.14: The number of the data points in the region D versus the value of parameter σ - Model RC

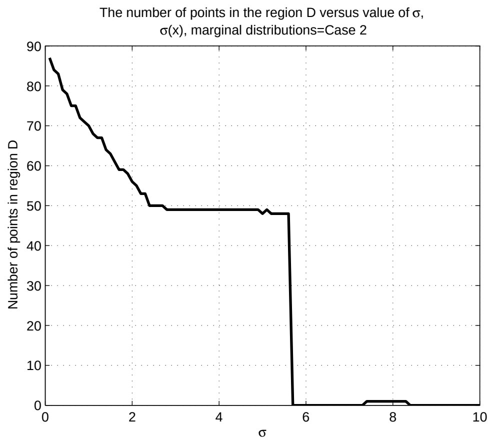


Figure 4.15: The number of the data points in the region D versus the value of parameter σ - Model RV

We observe that the functions decrease. Hence, the larger parameter σ, the more points will fall outside region D. We note that a point (X_i, Y_i) for which $f_{X,Y}(X_i, Y_i) = 0$ corresponds to a point (V_i, W_i) for which $f_{V,W}(V_i, W_i) = 0$. This shows that the Rotation Model does not provide a good fit to the data. In fact, it provides a bad fit, at least to the data lying in the lower left corner of the (V, W)-plane. In the next section, we investigate the fit to the data in the upper right corner of the plane.

4.10.3 Removing the data points

We decide to remove some data points, that are "responsible" for the zero-constant likelihood functions. Hence, we remove the data points that obtained $f_{V,W}(V_i, W_i) = 0$ after the transformation of the dataset to the model space. For Model RC, we decide to remove the points that fall outside the region D for $\sigma = 3.5$ and for Model RV, we decide to remove the points that fall outside the region D for $\sigma = 2.8$. As a result, the size of the dataset diminishes to 49. The new dataset used for Model RC is the same as for Model RV, therefore it will be possible to compare the results for both models. Removing the points is presented in Figures 4.16 and 4.17.

4.10.4 Estimation of the unknown parameters - Part 2

Having the new dataset, we can derive the estimation of the unknown parameter σ for Model RS and Model RV by using the Maximum Likelihood Method. The resulting estimations are presented in the following table:

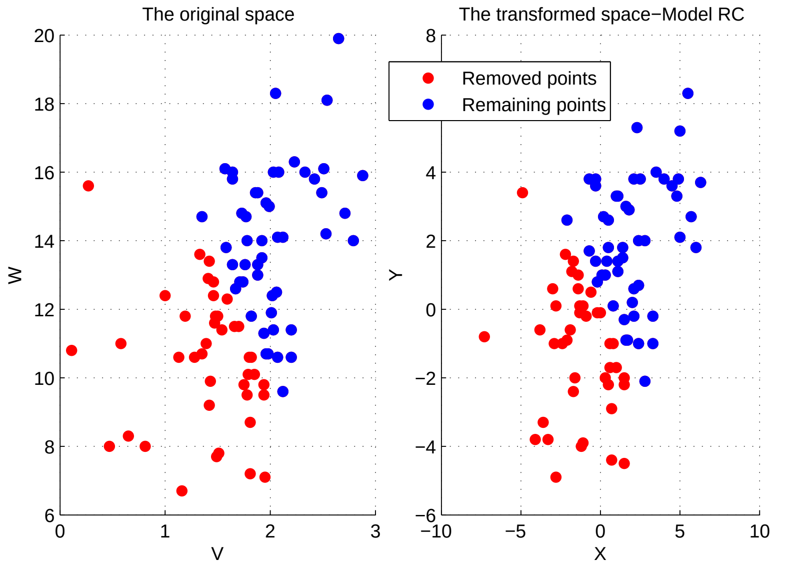


Figure 4.16: The removing of points, $\sigma = 3.5$ - Model RC

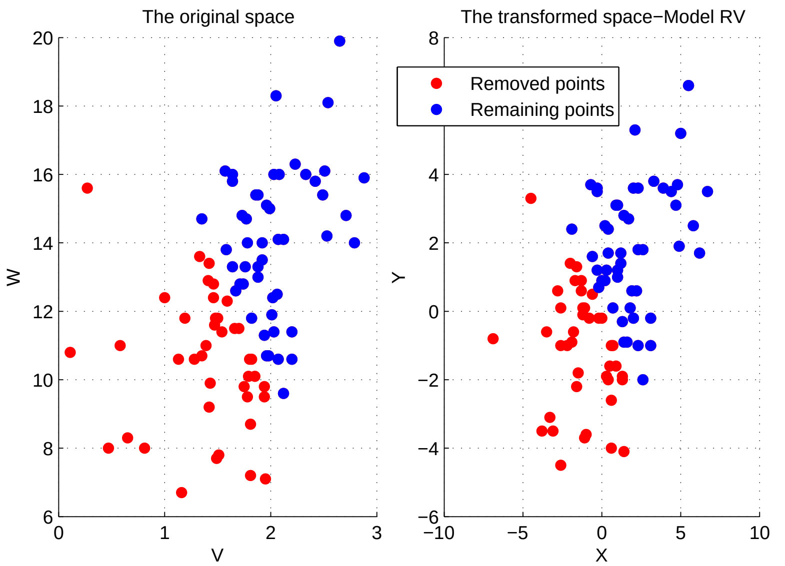


Figure 4.17: The removing of points, $\sigma = 2.8$ - Model RV

Model RS $\hat{\sigma}$	Model RV $\hat{\sigma}$
0.6	0.5

Figures 4.18 and 4.19 illustrate the log-likelihood functions.

4.10.5 Results with percentile lines

In this subsection we present results for the original, transformed (model) and copula space using percentile lines, after removing the data points. See Figures 4.20 and 4.21. The

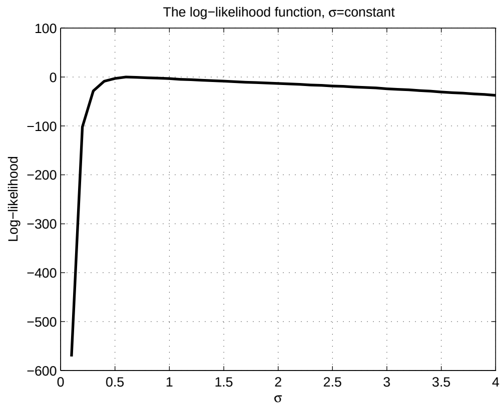


Figure 4.18: The log-likelihood function for Model RC - maximum is reached for $\hat{\sigma} = 0.6$

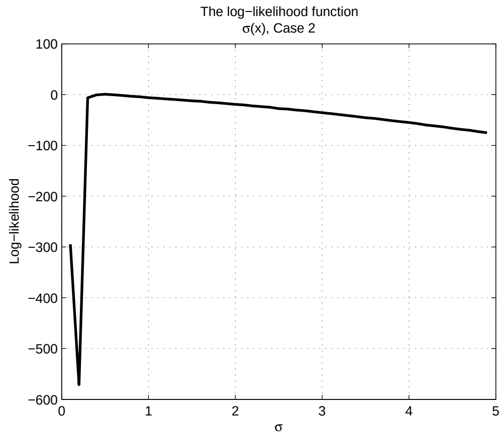


Figure 4.19: The log-likelihood function for Model RV - maximum is reached for $\hat{\sigma} = 0.5$ Observe that numerical errors occur at the beginning of this plot.

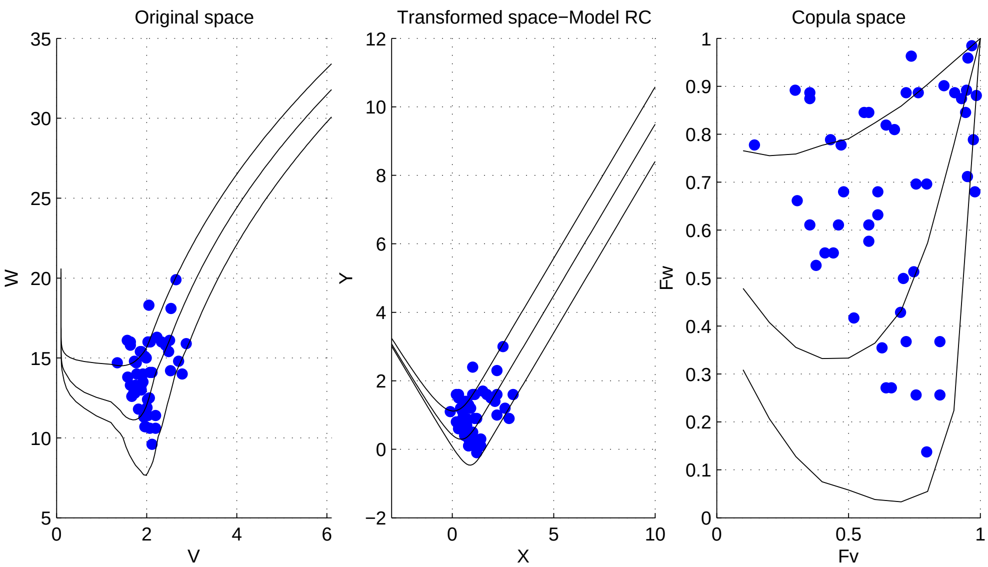


Figure 4.20: The original, transformed and copula space with the percentile lines $(10\%, 50\%, 90\%)$ for $\hat{\sigma} = 0.6$, Model RC There is 1 data point below the percentile line 10% and 5 points above the percentile line 90% when we consider the whole dataset. Recall that there should be approximately 7 points below and above the 10% and 90% percentile lines, respectively, when we consider the whole dataset.

resulting plots are not satisfactory, since the data points do not satisfy the criterion of the percentile lines. To be more precise, the number of data points below the percentile lines $p = 10\%$ and above the percentile lines $p = 90\%$ do not agree with the theory. We will not discuss the results any further. The overall conclusion is that the Rotation Model cannot be used in practice.

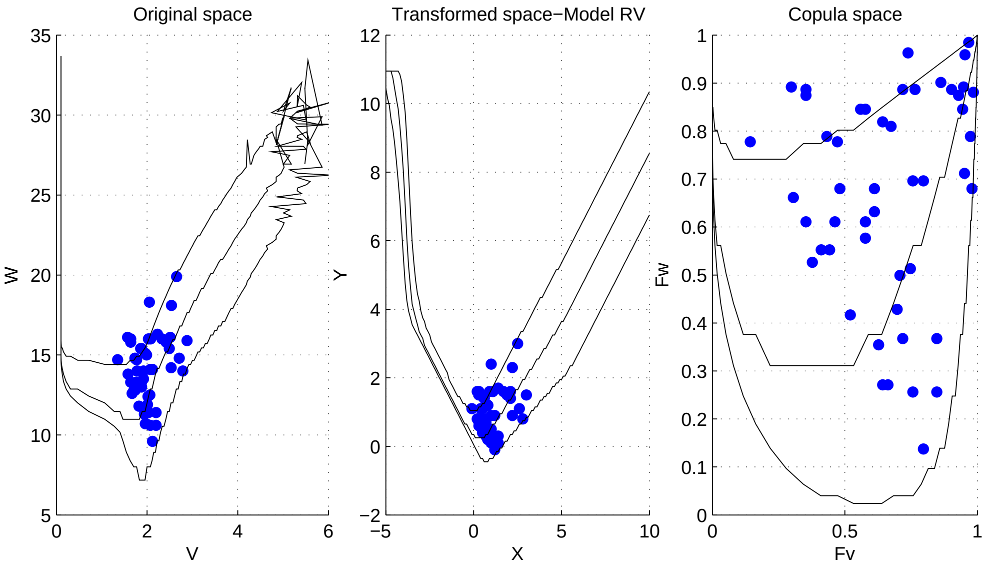


Figure 4.21: The original, transformed and copula space with the percentile lines $(10\%, 50\%, 90\%)$ for $\hat{\sigma} = 0.5$, Model RV There are no data points below the percentile line $p = 10\%$ and approximately 6 data points above the percentile line 90% when we consider the whole dataset. Moreover, we observe that numerical errors occur when the Model RV is applied, causing disturbance in the upper right corner of the plot in the original space.

4.11 Conclusions

This chapter was devoted to the construction of the Rotation Model. The construction of this model is very elegant in the 2-dimensional case. In the first part of this chapter we provided explicit formulas for this model and derived some useful properties. We showed that the margins of this model are shifted exponential (in case of restricted support) and we derived explicit formulas for the tail dependence coefficients. We concluded that lower tail independence and upper tail dependence occur.

At the end of this chapter we performed a case study, which revealed a serious drawback of this model. Because of the form of the marginal distributions, there is a possibility that some points after the transformation of the dataset to the model space will end up in the area where the probability density is zero. This is of course very unrealistic and we can exclude this model from further considerations.

Chapter 5

Summary and conclusions

In this report, we focused on the probabilistic modeling of two and three random variables (such as loads) and their dependencies. For this purpose the following three types of dependency models were studied: the bivariate copula function, bivariate conditional models and the Rotation Model. The purpose of these mathematical models is the modeling of bivariate distribution functions - if they are not known.
The notion of a copula function was intensively studied in Chapter 2. The copula model is a distribution function defined on the unit square with uniform margins and it describes the margin-free dependency of two random variables. We introduced the necessary definitions and propositions related to this concept. Moreover, we described the Archimedean class of copulas, which is known from its useful properties. We also provided statistical inference methods for fitting and testing the goodness-of-fit for copulas. In a case study we considered a dataset of 89 bivariate observations of water levels and wind speeds - the aim was to fit a copula model to this data. We applied two methods of estimation of the copula parameter and five methods of the evaluation of the fit. We concluded that the best model is provided by the Gumbel copula, although the results depend on the estimation methods and the assumptions concerning the marginal distributions. It turned out that visually judging the fit, by using the percentile lines, in the extreme region of the copula space (on unit square) is difficult, because the copula focuses on the entire dataset.
In Chapter 3, bivariate conditional models were discussed: the Constant Spread Model, the Variable Spread Model and the Constant Symmetric Spread Model. These models are constructed as a product of the conditional distribution and the distribution of the conditioning variable. The construction of these models is straightforward and makes it possible to predict the tail behavior of the distribution in the model space (the space in which the model is stated). The characteristic feature of this model is a constant spread parameter, which characterizes the spread of one random variable given the value of the other. An extension of this model is the Variable Spread Model, which allows for the spread to be dependent on the conditioning variable. The Constant Symmetric Spread Model entails symmetry with respect to the two random variables. We investigated the concept of the tail dependence coefficients and proved lower tail independence and upper tail dependence for each model. We also considered the concept of copula functions that arise for these models. It turned out that the copula functions related to the Constant Spread and Variable Spread Model are not Archimedean, since they do not satisfy the symmetry property. For the Constant Symmetric Spread Model, numerical experiments "suggest" that the corresponding copula belongs to the Archimedean class. Unfortunately, we were not able to find a mathematical

proof. At the end of this chapter, a case study was performed using the Maximum Likelihood Method and percentile lines method of visual evaluation of the fit. Moreover, we provided pictures of the percentile lines in the corresponding copula space, which simplified the comparison between the model space and copula space. We concluded that the visual evaluation of the fit, by using percentile lines, in the extreme region is easier in the model space than in the copula space.

In Chapter 4, we constructed the Rotation Model in the bivariate and 3-variate case. This model arises from the rotation of the coordinate system and can be seen as a modification of the conditional models. This models is newly proposed and has not been studied in the literature before. We considered some modifications of the 2-dimensional Rotation Model and calculated the tail dependence coefficients. In addition, we proved that the marginal distributions are shifted exponential in case of a restricted support. The performed case study revealed a serious disadvantage of this model. It may happen that after the transformation of the dataset to the model space some data points would fall in the region where the probability density is zero. More research should be performed to find out whether this drawback can be overcome.

Appendix A

Test statistic for Method 3

Consider the test statistic $S_n = n\int_0^1 |K_n(t) - K(t;\hat{\theta})|^2 k(\hat{\theta},t)dt$. It can be simplified as follows:

$$
\begin{aligned}
S_n &= \frac{n}{3} + n\sum_{j=1}^{n-1} K_n^2\left(\frac{j}{n}\right)\left\{K\left(\frac{j+1}{n};\hat{\theta}\right) - K\left(\frac{j}{n};\hat{\theta}\right)\right\} \\
&- n\sum_{j=1}^{n-1} K_n\left(\frac{j}{n}\right)\left\{K^2\left(\frac{j+1}{n};\hat{\theta}\right) - K^2\left(\frac{j}{n};\hat{\theta}\right)\right\}
\end{aligned}
$$

We will show this below.
Recall that $K_n(t)$ is a step function taking some nonnegative constant value c on the interval $[a,b] \subset [0,1]$, then:

$$
\begin{aligned}
& n\int_a^b |c - K(t;\hat{\theta})|^2 k(t;\hat{\theta})dt = n\int_a^b (c^2 - 2cK(t;\hat{\theta}) + K^2(t;\hat{\theta}))k(t;\hat{\theta})dt \\
&= nc^2\int_a^b k(t;\hat{\theta})dt - 2nc\int_a^b K(t;\hat{\theta})k(t;\hat{\theta})dt + n\int_a^b K^2(t;\hat{\theta})k(t;\hat{\theta})dt \\
&= nc^2(K(b;\hat{\theta}) - K(a;\hat{\theta})) - nc(K^2(b;\hat{\theta}) - K^2(a;\hat{\theta})) + \frac{n}{3}(K^3(b;\hat{\theta}) - K^3(a;\hat{\theta}))
\end{aligned}
$$

Let us now partition interval $[0,1]$ into the subintervals $[i/n,(i+1)/n]$, for $i = 0,...,n-1$. Additionally, we assume that $K_n(t)$ takes nonnegative constant values on each subinterval. Then, the test statistic S_n can be written as follows:

$$
\begin{aligned}
& S_n = n\sum_{i=0}^{n-1}\int_{i/n}^{(i+1)/n} |K_n(i/n) - K(t;\hat{\theta})|^2 k(t;\hat{\theta})dt \\
&= \sum_{i=0}^{n-1}\Big\{nK_n^2(i/n)[K((i+1)/n;\hat{\theta}) - K(i/n;\hat{\theta})] - nK_n(i/n)[K^2((i+1)/n;\hat{\theta}) - K^2(i/n;\hat{\theta})] \\
&+ \frac{n}{3}[K^3((i+1)/n;\hat{\theta}) - K^3(i/n;\hat{\theta})]\Big\} \\
&= n\sum_{i=1}^{n-1} K_n^2(i/n)[K((i+1)/n;\hat{\theta}) - K(i/n;\hat{\theta})] - n\sum_{i=1}^{n-1} K_n(i/n)[K^2((i+1)/n;\hat{\theta}) - K^2(i/n;\hat{\theta})] \\
&+ \frac{n}{3}\underbrace{(K^3(1;\hat{\theta}) - K^3(0;\hat{\theta}))}_{=1}
\end{aligned}
$$

where we used the fact that $K(0;\hat{\theta}) = K_n(0) = 0$ and $K(1;\hat{\theta}) = K_n(1) = 1$.

Bibliography

[1] P. Billingsley. *Probability and Measure.* John Wiley & Sons, first edition, 1978.

[2] R. de Matteis. Fitting copulas to data. Master's thesis, Institute of Mathematics of the University of Zurich, June 2001.

[3] J. Dobric and F. Schmidt. A Goodness of Fit Test for Copulas based on Rosenblatt's Transformation. *Computational Statistics & Data Analysis*, 51:4633–4642, 2007.

[4] M.T. Duits and J.M van Noortwijk. Parametrisch correlatiemodel voor extreme waarden. RO142 HKV Consultants, March 2006.

[5] C.P.M. Geerse. Bivariate correlatiemodellen met exponentiele en asymptotisch exponentiele marginale verdelingen. Technical report, HKV Consultants, Lelystad, May 2004.

[6] C.P.M Geerse. Hydraulische Randvoorwarden 2006 Vecht- en IJsseldelta. Statistiek IJsselmeerpeil, afvoeren en stormverlopoen voor Hydra-VIJ. Technical report, Rijkswaterstaat-RIZA, Lelystad, March 2006.

[7] C.P.M. Geerse and F. Diermanse. Correlaties en meerdimensionale statistiek. Technical report, HKV Consultants, Lelystad, November 2006.

[8] C. Genest and A.-C. Favre. Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of Hydrologic Engineering*, 12:347–368, 2007.

[9] C. Genest, J.-F. Quessy, and B. Rmillard. Goodness-of-fit procedures for copula models based on the probability integral transformation. *Scandinavian Journal of Statistics*, 33:337–366, 2006.

[10] H. Joe. Asymptotic efficiency of the two-stage estimation method for copula-based models. *Journal of Multivariate Analysis*, 94:401–419, 2005.

[11] D. Kurowicka and R. Cooke. *Uncertainty Analysis with High Dimensional Dependence Modeling.* John Wiley & Sons, first edition, 2006.

[12] R. Matus. Multivariate frequency analysis of hydrological extremes. Technical report, Slovak University of Technology, Radlinskeho 11, 813 68 Bratislava, Slovakia. Contact rastimatus@gmail.com.

[13] R.B. Nelsen. *An Introduction to Copulas.* Springer, first edition, 1999.

[14] A.J. McNeil P. Embrechts and D. Straumann. *Risk management: value at risk and beyond*, chapter Correlation and dependence in risk management: properties and pitfalls. Cambridge University Press.

[15] M. Rosenblatt. Remarks on a multivariate transformation. *Annals of Mathematical Statistic*, (27):470–472, 1952.

[16] G. van der Wulp. Using Copulas in Risk Management. Master's thesis, Tilburg University, August 2003.

[17] R.C. Weimer. *Applied elementary statistics*. Brooks/Cole Publishing Company, first edition, 1987.