LISBOA
SCHOOL OF
ECONOMICS &
MANAGEMENT

MASTER OF SCIENCE IN

ACTUARIAL SCIENCE

MASTERS FINAL WORK

DISSERTATION

ANALYSIS OF NEW TECHNIQUES FOR RISK AGGREGATION AND DEPENDENCE MODELLING

MOSTAFA SHAMS ESFAND ABADI

JULY - 2015

LISBOA SCHOOL OF ECONOMICS & MANAGEMENT

SINCE 1911

MASTER OF SCIENCE IN

ACTUARIAL SCIENCE

MASTERS FINAL WORK

DISSERTATION

ANALYSIS OF NEW TECHNIQUES FOR RISK AGGREGATION AND DEPENDENCE MODELLING

MOSTAFA SHAMS ESFAND ABADI

SUPERVISOR:

PROF. DR MARIA DE LOURDES CENTENO

JULY - 2015

Abstract

In risk aggregation we are interested in the distribution of the sum of dependent risks. The objective of risk aggregation and dependence modeling is to model adequately dependent insurance portfolios in order to evaluate the overall risk exposure. This master thesis investigates some practical aspects of modeling risk aggregation and dependency. We give an introduction to *copula-based hierarchical aggregation model through reordering algorithm*. This approach can be easily applicable in high dimensions and consists of a tree structure, bivariate copulas, and marginal distributions. This method is empirically illustrated using data set of *Danish Fire Insurance Data*. These data were collected at Copenhagen Reinsurance over the period 1980 to 1990 and every total claim has been divided into three risks consisting of a building loss, a loss of contents and a loss of profits caused by the same fire.

Acknowledgements

I would like to express my gratitude to my supervisor Professor Maria de Lourdes Centeno for introducing me to the topic, for her helpful advice, comments, and motivation. It has been an honor to work with her. I would like to thank my father, my mother and my sister for supporting me throughout my life, for their encouragement. Words cannot express how grateful I am to them. Without their support and encouragement, it was impossible for me to successfully complete my master's degree. I would like to thank my best friend Cristina who was always my support throughout this dissertation journey. In addition, a thank you to Dr. Philipp Arbenz for putting the reordering algorithm **R** code program in his web page, for his papers and his PhD thesis. My master's thesis is inspired by his work.

Contents

1. Introduction **1**

2. Multivariate Models **2**
2.1. Variance-Covariance method . . . 3
2.2. Risk factor models . . . 4

3. Copula Models **6**
3.1. Sklar's Theorem [1959] and Copulas . . . 6
3.2. Frechet bounds for copulas . . . 8
3.3. Dependency and Association Measures . . . 9
3.3.1. Spearman's rho . . . 9
3.3.2. Kendall's tau . . . 10
3.4. Tail Dependence . . . 11
3.5. Elliptical Copulas . . . 12
3.5.1. Gaussian (Normal) Copula . . . 12
3.6. Archimedean Copulas . . . 14
3.6.1. Clayton Copula . . . 15
3.6.2. Gumbel Copula . . . 17

4. Copula-based hierarchical model for risk aggregation **20**
4.1. Hierarchical aggregation . . . 21
4.2. Reordering algorithm for the numerical approximation . . . 23

5. Application to Danish Fire Insurance Data **26**
5.1. R Packages fitdistrplus and copula . . . 27
5.2. Determination of the tree structure, hierarchical clustering . . . 27
5.3. Choice of marginal distributions for the risks . . . 28
5.4. Choice of bivariate copulas . . . 32
5.5. Hierarchical aggregation through reordering algorithm . . . 33
5.6. Conclusions and results . . . 34

A. Appendix A **36**

B. Appendix B **37**

1. Introduction

Dependence modelling can not be ignored. This motivated me to write my master's thesis in this field. The 2008 financial crisis have shown that the modelling dependence between risks is necessary for prudent risk aggregation. Most regulatory frameworks such as Solvency II need such a dependence modelling. I am interested in new risk aggregation techniques in this thesis and I will focus on practical aspects of modelling risk aggregation and dependency. From a practical point of view, *copula-based hierarchical model through reordering algorithm* is a flexible approach proposed for risk aggregation. My master's thesis is inspired by the work of Dr. Philipp Arbenz, especially his paper on *Copula based hierarchical risk aggregation through sample reordering* (see Arbenz et al. [2012]) and his PhD thesis (see Arbenz [2012]). To illustrate hierarchical aggregation method empirically, we apply this approach to *Danish Fire Insurance Data*. I use the statistical software package **R** many times in my thesis.

This thesis is organized as follows. We give a brief introduction to multivariate models and their applications in Chapter 2 and we present three multivariate models popular in insurance. Copula models are the most popular method in modelling dependence among the risks. Chapter 3 deals with the concept of copula and various forms of copulas and practical aspects of copulas. In Chapter 4, we explore copula-based hierarchical approach for risk aggregation and dependence modeling and we introduce reordering algorithm that allows us to numerically approximate the hierarchical risk aggregation structure. In Chapter 5, we apply this approach to the danish fire insurance data and present some conclusions.

2. Multivariate Models

Life is full of risks. Probability Theory as a special language is being applied to the area of risk modeling. Risks represented by random variables mapping unpredictable future events into actual amounts standing for profits and losses. In this thesis we are interested in aggregate risks as the overall risk or the risk of a portfolio. Risk aggregation is the aggregation of individual risks using a model for aggregation. Different methods for risk aggregation can be considered in practice. A good aggregation model is providing at the same time both probabilistic descriptions of individual risks and of their dependence or correlation structure.

A multivariate model has many applications in insurance mathematics and one of its most important application is risk aggregation model. A multivariate model for risks in the form of a joint cdf, survival function or density allows us to understand and approximate the unknown true joint distribution and *dependence structure* among the risks. There are many methods for modeling multivariate risks and their dependencies. Depending on the available input and the required output and constraints such as regulatory rules, different multivariate models can be selected. In this chapter and next chapter, briefly we talk about multivariate models and we explain basically three popular multivariate models for risk aggregation (see Arbenz [2012]);

- *variance-covariance method*
- *risk factor models*
- *copula models*

Let $X_i \colon \Omega \to \mathbb{R}$ for $i \in \mathbb{N}$ denote individual risks describing values of losses or gains. Assume we are exposed to d risks $X = (X_1, ..., X_d)$. In this case, our interest is the distribution of the random vector;

$$(X_1, ..., X_d) \colon \Omega \to \mathbb{R}^d$$

The *marginal* cdf of individual risk X_i, $F_i(x) = P(X_i \leq x)$, does not have any information on dependence structure and it is only characterizing the stochasticity of the single risk X_i. We call $F_X(x)$ the *joint* cdf of the random vector of risks $X = (X_1, ..., X_d)$. When we fix the joint cdf $F_X(x)$, we state a multivariate model containing the all marginal behaviour distributions and the dependence structure of the risks. This joint cdf contains all information on the distribution of random vector $X = (X_1, ..., X_d)$:

$$F_X(x) = F_X(x_1, ..., x_d) = P(X_1 \leq x_1, ..., X_d \leq x_d), \qquad (x_1, ..., x_d) \in \mathbb{R}^d$$

This distribution captures the important properties of $(X_1, ..., X_d)$ and a multivariate model is a model that allows us to approximate it.
In some cases it is useful to work with the survival function of random vector $X = (X_1, ..., X_d)$ defined by

$$\overline{F}_X(x) = \overline{F}_X(x_1, ..., x_d) = P(X_1 > x_1, ..., X_d > x_d)$$

The marginal cdf of the risk factor X_i, written F_i, is easily calculated from the joint cdf. For all $i \in \mathbb{N}$ we have

$$F_i(x) = P(X_i \leq x_i) = F_X(\infty, ..., \infty, x_i, \infty, ..., \infty)$$

If $F_i(x)$ is absolutely continuous, so then we refer to its derivative $f_i(x)$ as the marginal density of X_i . It is possible to define the cdf of a random vector $X = (X_1, ..., X_d)$ based on its joint density f_X;

$$F_X(x_1, ..., x_d) = \int_{-\infty}^{x_1} ... \int_{-\infty}^{x_d} f_X(u_1, ..., u_d)\ du_1 ...\ du_d$$

Note that existence of a joint density f_X implies existence of marginal densities $f_1, ..., f_d$ (but not vice versa).
We can make conditional probability statements. For instance, the conditional distribution of X_2 given $X_1 = x_1$ has density

$$f_{X_2|X_1}(x_2|x_1) = \frac{f_X(x_1, x_2)}{f_{X_1}(x_1)}$$

and its corresponding cdf is $F_{X_2|X_1}(x_2|x_1)$.
For the bivariate random vector $X = (X_1, X_2)$, in the case of existing a joint density, if the joint density of X factorizes into $f_X(x) = f_{X_1}(x_1)f_{X_2}(x_2)$, then we say X_1 and X_2 are independent. We recall that X_1 and X_2 are independent if and only if

$$F(x) = F_{X_1}(x_1)F_{X_2}(x_2), \quad \forall x$$

In general $X_1, ..., X_d$ are said to be mutually independent if and only if $F_X(x) = \prod_{i=1}^{d} F_i(x_i)$ for all $x \in \mathbb{R}^d$ or, in the case of existing a joint density $f_X(x) = \prod_{i=1}^{d} f_i(x_i)$. In many situations the risk factors X_i cannot be assumed to be independent. In these cases, the joint cdf can not decouple into marginal cdfs:

$$P(X_1 \leq x_1, ..., X_d \leq x_d) \neq \prod_{i=1}^{d} P(X_i \leq x_i)$$

In the following sections, we give a brief introduction to the *variance-covariance method* and to *risk factor models*, and in the next chapter we explore *copula models*. Each of these three popular models in finance and insurance has its own advantages and disadvantages. I discuss their own weaknesses and I give an alternative approach in Chapter 4.

2.1. Variance-Covariance method

Value at Risk (VaR) is a widely used risk measure of the risk of loss. What is Value-at-Risk(VaR)? We start this section with this question. The best answer is the following: VaR is the maximum loss value (measured in monetary units) that an institution is likely to face on a portfolio, not exceeded with a specified probability level of confidence over a given time period. There are three main methods that are used to estimate the Value-at-Risk (see Skoglund [2010]). It can be estimated by Monte Carlo simulations or by running a Historical simulation, or finally by Variance-Covariance method that is the most straightforward method of estimating the Value-at-Risk. This approach is widely used and is also known as Linear VaR or Delta normal VaR. Variance-Covariance method observes historical data over time, makes assumptions about

theoretical distribution of asset returns (usually normal distribution), and uses the variances and covariances to compute the likely maximum loss.

In *Variance-Covariance method*, random variables X_i and their correlations are characterised by their mean $\mu_i = E(X_i)$ and variance $\sigma_i^2 = var(X_i)$ and covariance $\sigma_{ij} = \sigma_i\sigma_j\rho_{ij} = cov(X_i, X_j)$. The *dependence structure* is characterised through the (Pearson) correlation coefficients, defined by

$$\rho_{ij} = \rho(X_i, X_j) = \frac{cov(X_i, X_j)}{\sqrt{var(X_i)var(X_j)}} = \frac{\sigma_{ij}}{\sigma_i\sigma_j} \quad i, j = 1, \ldots, d.$$

Therefore, the model is specified by a *mean vector* and a *covariance matrix*. The *mean vector* of $X = (X_1, ..., X_d)$ is given by

$$E(X) = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_d \end{pmatrix} \in \mathbb{R}^d$$

The *covariance matrix* is the matrix $cov(X)$ defined by

$$cov(X) = \begin{pmatrix} \sigma_{11} & \ldots & \sigma_{1d} \\ \vdots & \ddots & \vdots \\ \sigma_{d1} & \ldots & \sigma_{dd} \end{pmatrix} \in \mathbb{R}^{d\times d}$$

Note that the diagonal elements are the variances of the components of X.
The *correlation matrix* is given by

$$C = \begin{pmatrix} 1 & \ldots & \rho_{1d} \\ \vdots & \ddots & \vdots \\ \rho_{d1} & \ldots & 1 \end{pmatrix}$$

We now can see the application of Variance-Covariance method in risk aggregation. In this method, the characteristics of the distribution of the sum of all risks $S = X_1 + \ldots + X_d$ are given by

$$E[S] = \mu_1 + \ldots + \mu_d$$

and

$$Var[S] = \sum_{i=1}^{d}\sum_{j=1}^{d} cov(X_i, X_j) = \sum_{i=1}^{d}\sum_{j=1}^{d} \sigma_i\sigma_j\rho_{ij}$$

As can be seen above, this model is restricted to the first and second moments, i.e. mean, variance and covariances of the risks $X_1, \ldots, X_d$. Thus further properties of the distribution of sum of risks, S, cannot be understood.

2.2. Risk factor models

Risk factors are different types of systematic factors affecting portfolio losses. The assets and liabilities are exposed to many risk factors. Risk factors are important in understanding the sources of risk in a portfolio. *Risk factor models* are now often used in finance and insurance. For example, Swiss Re is using a risk factor model to estimate the distribution of the Profit-and-Loss distribution in its Solvency II internal model (see Arbenz [2012]).

Here we focus on the simplest form of risk factor models that is the linear factor model. In this model with k risk factors, each risk X_i can be written in the form of

$$X_i = \mu_i + r_{i,1}Y_1 + \ldots + r_{i,k}Y_k + \epsilon_i$$

Where;

- μ_i is the unconditional expectation, $\mu_i = E(X_i)$.
- The Y_j are the risk factors, that can be assumed to be distributed according to the standard normal distribution, $Y_j \sim \mathcal{N}(0, 1)$.
- The $r_{i,j}$ represent the sensitivity of risks X_i with respect to the risk factors Y_j.
- The ϵ_i denote the residuals, which are commonly assumed to be independent.

If the distribution of the risk factors Y_j, the distribution of residuals ϵ_i and the parameters μ_i, $r_{i,j}$ and σ_i^2 are known, then a sample of the random vector $(X_1, \ldots, X_d)$ can easily be obtained by simulating the risk factors Y_j and residuals ϵ_i using Monte Carlo simulations.

A risk factor model is a dimension reduction tool, as it allows us to model high number d of risks with just few risk factors ($k \ll d$). The risks' variability can be explained through a potentially low number of risk factors and the dependence between risks is understood through dependence structure between the risk factors.

3. Copula Models

In this chapter we concentrate on the concept of a *Copula* and we will be interested in how *dependence* among the risks, X_i, can be understood and modeled using Copula. Dependencies happen in many other fields, not only in finance and insurance. Copula is a popular tool in multivariate modeling for understanding relationships among random variables. In actuarial science, copulas are frequently used in modeling dependence structure.

We begin our discussion about Copulas with *Sklar's Theorem [1959]* that allows us to separate the dependence structure from the marginal distributions (see Panjer [2006]).

3.1. Sklar's Theorem [1959] and Copulas

A copula $C : [0,1]^k \to [0,1]$ is a function with the following properties (see McNeil et al. [2010]):

(1) $C(u_1, u_2, \ldots, u_k)$ is increasing in each component u_i.

(2) $C(1, \ldots, 1, u_i, 1, \ldots, 1) = u_i$ for all $i \in 1, ..., k$, and $u_i \in [0,1]$.

(3) For all $(a_1, \ldots, a_k), (b_1, \ldots, b_k) \in [0,1]^k$ with $a_i \leqslant b_i$ we have

$$\sum_{i_1=1}^{2} \ldots \sum_{i_k=1}^{2} (-1)^{i_1+\ldots+i_k} C(u_{1i_1}, \ldots, u_{ki_k}) \geqslant 0,$$

where $u_{j1} = a_j$ and $u_{j2} = b_j$ for all $j \in \{1, \ldots, k\}$ ♦

If a function C fulfills the above properties, then it is a copula. A copula function is a multivariate distribution whose marginal distributions are all $Uniform(0,1)$. We define a copula function C as the joint distribution function of standard $Uniform(0,1)$ random variables;

$$C(u_1, u_2, \ldots, u_k) = Pr(U_1 \leq u_1, U_2 \leq u_2, \ldots, U_k \leq u_k).$$

Copula functions are helpful to simulate a dependence structure independently from the marginal distributions. The idea behind a copula is to translate any multivariate distribution $F(x_1, x_2, \ldots, x_k)$ into its marginal distributions $F_1(x_1), F_2(x_2), \ldots, F_k(x_k)$ and its copula $C(F_1(x_1), F_2(x_2), \ldots, F_k(x_k))$ describing the dependence among the random variables $X_1, \ldots, X_k$.

We know from basic probability that the probability integral transforms $F_1(X_1), F_2(X_2)$ $, \ldots, F_k(X_k)$ are each distributed as Uniform $(0,1)$, hence the copula at $F_1(x_1), F_2(x_2), \ldots, F_k(x_k)$ can be written as

$$\begin{aligned} C(F_1(x_1), F_2(x_2), \ldots, F_k(x_k)) &= Pr(U_1 \leq F_1(x_1), U_2 \leq F_2(x_2), \ldots, U_k \leq F_k(x_k)) \\ &= Pr(F_1^{-1}(U_1) \leqslant x_1, \ldots, F_d^{-1}(U_d) \leqslant x_d) \\ &= Pr(X_1 \leqslant x_1, \ldots, X_d \leqslant x_d) \\ &= F(x_1, \ldots, x_d) \end{aligned}$$

where we define the quantile function as $F_j^{-1}(u) = inf\{x : F_j(x) \geqslant u\}$

Sklar's theorem states the above result that there is a unique copula C for any joint distribution function F, that satisfies

$$F(x_1, \ldots, x_d) = C(F_1(x_1), F_2(x_2), \ldots, F_k(x_k)) \tag{3.1}$$

Sklar's theorem shows that we can form a multivariate joint distribution F from a group of marginal distributions $F_1(x_1), F_2(x_2), \ldots, F_d(x_d)$, and a selected copula C. The function $C(F_1(x_l), \ldots, F_d(x_d))$ is the multivariate joint distribution function F. In practice, usually distributions of different risk types are modeled separately. The dependence structure in the copula function is independent of the form of the marginal distributions. Sklar's theorem allows us to use different copulas while keeping identical marginal distributions.

In the rest of this chapter, we focus on dependency structures among pairs of random variables - in other words, on bivariate copulas. Here I cite two examples mentioned in Frees and Valdez [1998]. Although Sklar's theorem proves that a copula function always exists, Example (1.1) shows that it is not always easy to discover the copula. Example (1.2) considers the important question of how to build a copula function for a existing problem and provides a useful way of building a copula, using the method of compounding.

Example 1.1 Marshall-Olkin (1967) Exponential Shock Model Suppose that we wish to model $p = 2$ lifetimes that we suspect are subject to some common disaster that may cause a dependency between the lives. For simplicity, let us assume that Y_1 and Y_2 are two independent lifetimes with distribution functions H_1 and H_2. We further assume there exists an independent exponential random variable Z with parameter λ that represents the time until common disaster. Both lives are subject to the same disaster, so that actual ages-at-death are represented by $X_1 = min(Y_1, Z)$ and $X_2 = min(Y_2, Z)$. Thus, the marginal distributions are

$$\begin{aligned} Pr(X_j \leq x_j) &= F_j(x_j) \\ &= 1 - exp(-\lambda x_j)(1 - H_j(x_j)), \qquad j = 1, 2 \end{aligned}$$

Basic calculations show that the joint distribution is

$$F(x_1, x_2) = F_1(x_1) + F_2(x_2) - 1 + exp(\lambda min(x_1, x_2))(1 - F_1(x_1))(1 - F_2(x_2)).$$

This expression is not in the form of the copula construction (Equation (3.1)) because the joint distribution function F is not *just* a function of the marginals $F_1(x_1)$ and $F_2(x_2)$. ♦

Example 1.2 Bivariate Pareto Model Consider a claims random variable X that, given a risk classification parameter γ, can be modeled as an exponential distribution; that is,

$$Pr(X \leq x \mid \gamma) = 1 - e^{-\gamma x}$$

As is well known in probability theory, if γ has a gamma(α, λ) distribution, then the marginal distribution (over all risk classes) of X is Pareto;

$$F(x) = 1 - (1 + x/\lambda)^{-\alpha}$$

Suppose, conditional on the risk class γ, that X_1 and X_2 are independent and identically distributed. Assuming that they come from the same risk class γ induces a dependency. The joint distribution is

$$F(x_1, x_2) = F_1(x_1) + F_2(x_2) - 1 + [(1 - F_1(x_1))^{-1/\alpha} + (1 - F_2(x_2))^{-1/\alpha} - 1]^{-\alpha}$$

This yields the copula function

$$C(u_1, u_2) = u_1 + u_2 - 1 + [(1 - u_1)^{-1/\alpha} + (1 - u_2)^{-1/\alpha} - 1]^{-\alpha}$$

With this function, we can express the bivariate distribution function as $F(x_1, x_2) = C(F_1(x_1), F_2(x_2))$. ♦

The Sklar's theorem and copula approach allows us to separate the marginal distributions from the copula. The marginals contain the information of the different individual risks. The copula contains the information on the dependence structure.

3.2. Frechet bounds for copulas

Here we establish the important *Frechet bounds* for copulas (see Panjer [2006]). In the bivariate case, it is interesting to note from basic probability that

$$Pr(U_i > u_i, U_j > u_j) = 1 - u_i - u_j + C(u_i, u_j)$$

Then we have

$$\begin{aligned} C(u_i, u_j) &= u_i + u_j - 1 + Pr(U_i > u_i, U_j > u_j) \\ &\geqslant u_i + u_j - 1 \end{aligned}$$

Therefore, we have got a *Frechet* lower bound on the copula cdf;

$$C(u_i, u_j) \geq max\{0, u_i + u_j - 1\}$$

We can obtain *Frechet* upper bound on the copula cdf from the simple fact that both

$$Pr(U_i \leqslant u_i, U_j \leqslant u_j) \leqslant Pr(U_i \leqslant u_i) = u_i$$

and

$$Pr(U_i \leqslant u_i, U_j \leqslant u_j) \leqslant Pr(U_j \leqslant u_j) = u_j$$

so that

$$Pr(U_i \leqslant u_i, U_j \leqslant u_j) \leqslant min\{u_i, u_j\}$$

Thus we have got *Frechet* bounds on the copula cdf;

$$max\{0, u_i + u_j - 1\} \leq C(u_i, u_j) \leq min\{u_i, u_j\}$$

3.3. Dependency and Association Measures

The linear correlation coefficient ρ, also called Pearson's coefficient, of random variables (X_1, X_2) is defined as;

$$\rho(X_1, X_2) = \frac{cov(X_1, X_2)}{\sqrt{var(X_1)var(X_2)}} = \frac{E(X_1X_2) - E(X_1)E(X_2)}{\sqrt{var(X_1)var(X_2)}}$$

Correlation plays a central role in financial theory. It is important to remember that dependence and correlation are different concepts. We have X_1 and X_2 are independent $\Rightarrow$ X_1 and X_2 are uncorrelated or $\rho(X_1, X_2) = 0$ but the converse is in general false. The classical measure of dependence is the correlation coefficient. The correlation coefficient is a measure of the linearity between random variables. For two random variables X_1 and X_2, the correlation coefficient is exactly equal to 1 or -1 if there is a perfect linear relationship between X_1 and X_2, that is, if $X_2 = aX_1 + b$. If a is positive, the correlation coefficient is equal to 1; if a is negative, the correlation coefficient is equal to -1. This explains why the correlation described here is often called linear correlation.

$$\rho(X_1, X_2) = 1 \Longleftrightarrow X_2 = aX_1 + b, \qquad a > 0$$

$$\rho(X_1, X_2) = -1 \Longleftrightarrow X_2 = aX_1 + b, \qquad a < 0$$

We know that the linear correlation coefficient is related to the marginal distributions that its value will change according to the form of the marginals. But the copula does not depend on the form of the marginals. Therefore, it would be much more natural to have dependency measures using copulas that depend only on the copula and not on the marginal distributions F_1 and F_2.

As mentioned in Frees and Valdez [1998], Schweizer and Wolff (1981) showed that two dependence measures could be expressed only in terms of the copula. These two measures of association among random variables are Spearman's rho (explained in subsection 3.3.1) and Kendall's tau (explained in subsection 3.3.2). Similar to the linear correlation coefficient, these measures of dependence take on values of 1 for perfect positive dependence and -1 for perfect negative dependence.

3.3.1. Spearman's rho

Here we introduce *Spearman's rho* as given in Panjer [2006]. The measure of association Spearman's rho $\rho_S(X_1, X_2)$, also sometimes called rank correlation, of a couple of random variable (X_1, X_2) with marginal distributions $F_1(x_1)$ and $F_2(x_2)$ is given by

$$\rho_S(X_1, X_2) = \rho(F_1(X_1), F_2(X_2))$$

where ρ denotes linear correlation coefficient.

Because $U_1 = F_l(X_1)$ and $U_2 = F_2(X_2)$ are both $Uniform(0, 1)$ random variables with mean $1/2$ and variance $1/12$, we can easily rewrite Spearman's rho as

$$\begin{aligned}\rho_S(X_1, X_2) &= \frac{E[F_1(X_1)F_2(X_2)] - E[F_1(X_1)]E[F_2(X_2)]}{\sqrt{Var(F_1(X_1))Var(F_2(X_2))}} \\ &= 12E[F_1(X_1)F_2(X_2)] - 3\end{aligned}$$

It is straightforward to write the Spearman's rho $\rho_S(X_1, X_2)$ in terms of copula function C and we get;

$$\begin{aligned}\rho_S(X_1, X_2) &= 12E[U_1U_2] - 3 \\ &= 12\int_0^1\int_0^1 u_1u_2 \; dC(u_1, u_2) - 3 \\ &= 12\int_0^1\int_0^1 C(u_1, u_2) \; du_1 \; du_2 - 3\end{aligned}$$

Spearman's rho ρ_S replaces the bivariate random variable (X_1, X_2) by the bivariate $(F_1(X_1), F_2(X_2))$ before considering linear correlation. Hence, in the same way as Sklar's theorem, Spearman's rho ρ_S only depends on the copula function and not on the marginal distributions F_1 and F_2.

3.3.2. Kendall's tau

The measure of association Kendall's tau, $\tau_K(X_1, X_2)$, of bivariate random variables (X_1, X_2) is given by;

$$\tau_K(X_1, X_2) = Pr[(X_1 - X_1^*)(X_2 - X_2^*) > 0] - Pr[(X_1 - X_1^*)(X_2 - X_2^*) < 0]$$

where (X_1^*, X_2^*) is a bivariate random variable distributed as (X_1, X_2) and independent from (X_1, X_2), i.e. the same marginal distribution $F_1(x_1)$ for X_1 and X_1^* and the same marginal distribution $F_2(x_2)$ for X_2 and X_2^* .
The first term is the probability of concordance, the differences between the random variables have the same signs. The second term then is the probability of discordance, the differences between the random variables have opposite signs. It is easy now to obtain an expression for Kendall's tau in terms of the copula function as follows;

$$\begin{aligned}\tau_K(X_1, X_2) &= Pr[(X_1 - X_1^*)(X_2 - X_2^*) > 0] - Pr[(X_1 - X_1^*)(X_2 - X_2^*) < 0] \\ &= Pr[(X_1 - X_1^*)(X_2 - X_2^*) > 0] - \{1 - Pr[(X_1 - X_1^*)(X_2 - X_2^*) > 0]\} \\ &= 2Pr[(X_1 - X_1^*)(X_2 - X_2^*) > 0] - 1\end{aligned}$$

Because the random variables are interchangeable;

$$\begin{aligned}\tau_K(X_1, X_2) &= 4Pr[(X_1 < X_1^*, X_2 < X_2^*)] - 1 \\ &= 4E\{Pr[(X_1 < X_1^*, X_2 < X_2^*)|X_1^*, X_2^*]\} - 1 \\ &= 4\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} Pr[X_1 < x_1, X_2 < x_2] \; dF(x_1, x_2) - 1 \\ &= 4\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(x_1, x_2) \; dF(x_1, x_2) - 1 \\ &= 4\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} C(F_1(x_1), F_2(x_2)) \; dC(F_1(x_1), F_2(x_2)) - 1 \\ &= 4\int_{-1}^{1}\int_{-1}^{1} C(u_1, u_2) \; dC(u_1, u_2) - 1\end{aligned}$$

Thus, Kendall's tau in terms of the copula function $C(u_1, u_2)$ with $Uniform(0, 1)$ marginals is given by

$$\begin{aligned}\tau_K(X_1, X_2) &= 4\int_{-1}^{1}\int_{-1}^{1} C(u_1, u_2)\ dC(u_1, u_2) - 1 \\ &= 4E[C(U_1, U_2)] - 1\end{aligned}$$

which proves that Kendall's tau measure only depends on the dependence structure among the two variables and not on their marginal distributions.

3.4. Tail Dependence

The deviation from normality in actuarial science is called fat-tails or tail dependence. We represent here the dependence measure of tail dependence that is a very useful measure in describing a copula and it is especially important in risk management. It has been observed in risk management that especially in bad times there may be significant correlation between risks. Thus it is important to focus on the extreme events, in other words how strong the correlation is in the upper or lower tails.

As given in Denuit et al. [2006], for the bivariate random variables (X_1, X_2), the tail dependence measures the concordance between the extreme events of X_1 and X_2. We are concerned with the probability of observing an unusually large loss for X_1 given that an unusually large loss has occurred for X_2. The *upper tail dependence* λ_U measures the probability of X_1 is very large if it is known that X_2 is very large, thus is defined as;

$$\lambda_U(X_1, X_2) = \lim_{u\to 1} P(F_1(X_1) \geq u \mid F_2(X_2) \geq u).$$

This tail dependence measure does not depend on F_1 and F_2 and only depends on the dependence between X_1 and X_2. Because $F_1(X_1) = U_1$ and $F_2(X_2) = U_2$ are both $Uniform(0, 1)$ random variables, thus for the random variable couple (X_1, X_2) with copula C we can rewrite:

$$\begin{aligned}\lambda_U(X_1, X_2) &= \lim_{u\to 1} P(F_1(X_1) \geq u \mid F_2(X_2) \geq u) \\ &= \lim_{u\to 1} P(U_1 \geq u \mid U_2 \geq u) \\ &= \lim_{u\to 1} \frac{1 - P(U_1 < u) - P(U_2 < u) + P(U_1 < u, U_2 < u)}{1 - P(U_2 < u)} \\ &= \lim_{u\to 1} \frac{1 - 2u + C(u, u)}{1 - u}\end{aligned}$$

In a similar way, the *lower tail dependence* λ_L is defined as;

$$\begin{aligned}\lambda_L(X_1, X_2) &= \lim_{u\to 0} P(F_1(X_1) \leq u \mid F_2(X_2) \leq u) \\ &= \lim_{u\to 0} P(U_1 \leq u \mid U_2 \leq u) \\ &= \lim_{u\to 0} \frac{P(U_1 \leq u, U_2 \leq u)}{P(U_2 \leq u)} \\ &= \lim_{u\to 0} \frac{C(u, u)}{u}\end{aligned}$$

According to these formulas, the tail dependency of X_1 and X_2 can be measured only by looking at copula C. These measures help us to understand the tail events. We say that the random

variable couple (X_1, X_2) has upper (lower) tail dependence when $\lambda_U > 0$ ($\lambda_L > 0$). In practice they are used to in order to choose suitable copula to describe the dependence structure among the risks. For more details on tail dependence, see Embrechts et al. [2003].

Among many different copula families, only a few are used in practice. The most important families are the Archimedean Copulas and Elliptical Copulas. There are many books and papers on Copulas families, Archimedean and Elliptical copulas, see for instance Embrechts et al. [2003], Frees and Valdez [1998], Bürgi et al. [2008], Nelsen [2007], Panjer [2006].

3.5. Elliptical Copulas

Elliptical copulas are the copulas associated with elliptical distributions, which have a elliptical form and symmetry in the tails. An elliptical distribution is an extension of multivariate normal distribution. An elliptical distribution is uniquely determined by its mean, correlation matrix and the type of its margins, hence the copula of an elliptical distribution is uniquely determined by its correlation matrix and knowledge of its type. For further details on elliptical distributions we refer to Embrechts et al. [2003]. Elliptical copulas are generally defined as copulas of elliptical distributions. Important copulas in this copula family are the Gaussian copula associated with the multivariate normal distribution and the Students T copula associated with the multivariate t distribution. We describe Gaussian (Normal) copula in the following subsection.

3.5.1. Gaussian (Normal) Copula

The Gaussian copula is based on the multivariate normal distribution. The Gaussian copula generated by a multivariate normal distribution with linear correlation matrix Σ is given by:

$$C(u_1, \ldots, u_k) = \Phi_\Sigma(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_k)).$$

where Φ is the cdf of the standard univariate normal distribution, Φ^{-1} denotes the inverse of a standard normal distribution and Φ_Σ denotes the joint cdf of the standard multivariate normal distribution (with zero mean and variance of 1 for each component). The correlation matrix Σ is defined as:

$$\Sigma = \begin{pmatrix} 1 & \rho_{12} & \ldots & \rho_{1k} \\ \rho_{21} & 1 & \ldots & \rho_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{k1} & \rho_{k2} & \ldots & 1 \end{pmatrix}$$

where ρ_{ij} denotes the correlation coefficients.

Here we refer to a part of abstract of paper on "The Role of Copulas in the Housing Crisis" (see Zimmer [2012]):

> *"Due to its simplicity and familiarity, the Gaussian copula is popular in calculating risk in collaterized debt obligations, but it imposes asymptotic independence such that extreme events appear to be unrelated. This restriction might be innocuous in normal times, but during extreme events, such as the housing crisis, the Gaussian copula might be inappropriate."*

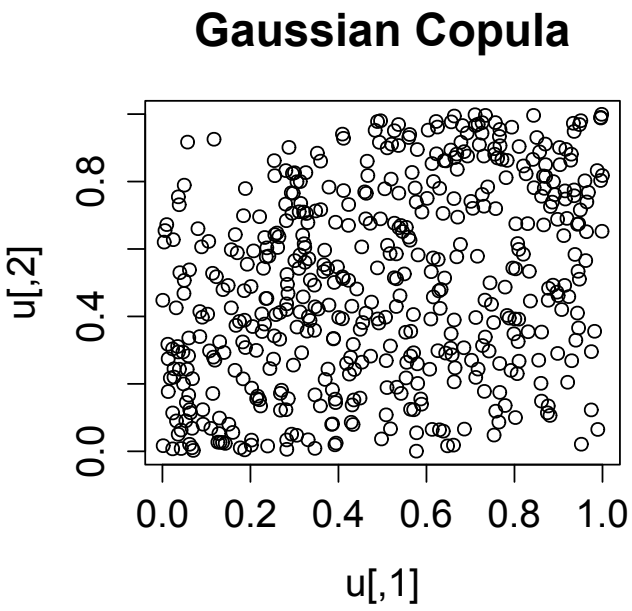

Figure 3.1.: Scatter plot of the Gaussian copula for $\rho = 0.3$

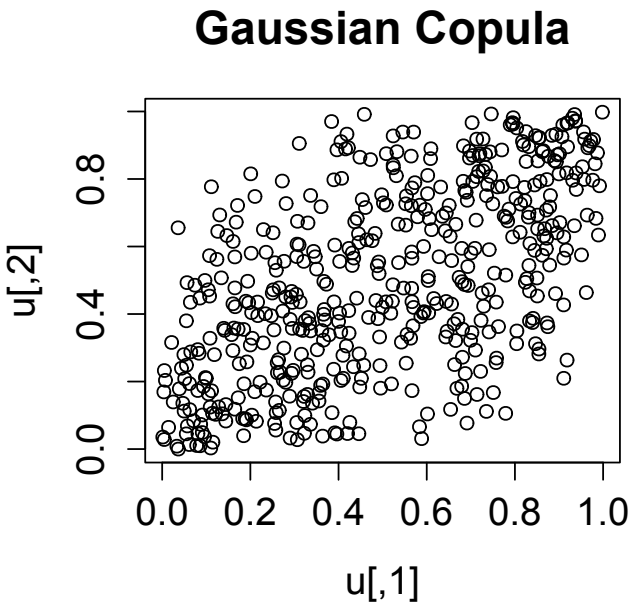

Figure 3.2.: Scatter plot of the Gaussian copula for $\rho = 0.6$

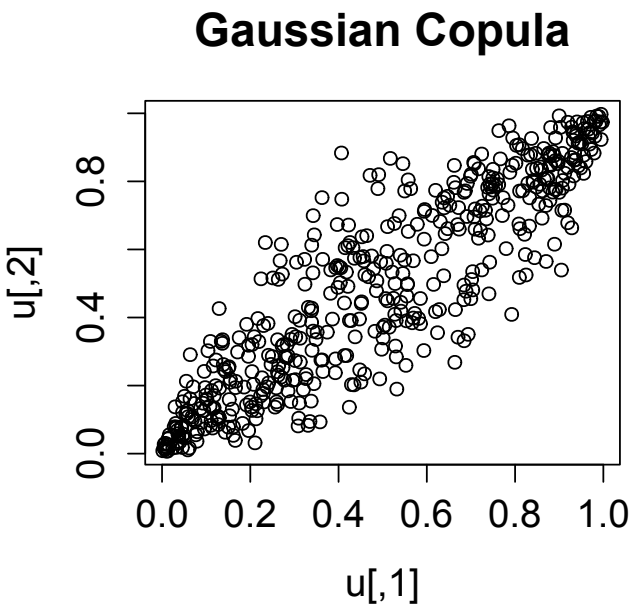

Figure 3.3.: Scatter plot of the Gaussian copula for $\rho = 0.9$

It is interesting to note that Gaussian (Normal) copulas have zero tail dependence ($\lambda_U = \lambda_L = 0$). See Embrechts et al. [2003] for a proof of this result. Using **R**, scatter plots of the bivariate Gaussian copula for various values of the correlation parameter $\rho = 0.3$, $\rho = 0.6$ and $\rho = 0.9$ are shown in Figures 3.1, 3.2 and 3.3 respectively. The symmetry of the Gaussian copula can also be seen in these scatter plots.

The Gaussian copula plots for $\rho = 0.3$ and $\rho = 0.6$ (Figures 3.1 and 3.2) show that Gaussian copulas do not have upper tail dependence and lower tail dependence ($\lambda_U = \lambda_L = 0$) except in the special case with ρ close to 1 (Figure 3.3), where there is perfect correlation. However, this approach does not show tail dependence.

As mentioned in Economist [2009] and Bogard [2011], the Gaussian copula was used widely before the housing crisis to simulate the dependence between housing prices in various geographic areas. Looking at the scatter plots for the Gaussian copulas above, it can be seen that extreme events (very high values of U_1 and U_2 or very low values of U_1 and U_2) seem very weakly correlated. Archimedean copulas, e.g. Clayton and Gumbel copulas are good alternatives to Gaussian copula.

3.6. Archimedean Copulas

Archimedean copulas are widely used because they can be easily constructed and many copula functions belong to this copula family. In comparison to Elliptical copulas, Archimedean copulas have only one dependency parameter θ instead of a correlation dependency matrix parameter. As mentioned in Embrechts et al. [2003], Elliptical copulas are restricted to have radial symmetry. In many finance and insurance applications it seems reasonable that there is a stronger dependence between big losses (e.g. a stock market crash) than between big gains. Such asymmetries cannot be modeled with Elliptical copulas. For simplicity, we consider bivariate copulas. Let φ be a continuous, strictly decreasing convex function with domain $(0, 1]$

and range $[0, \infty)$ satisfying $\varphi(0) = \infty$ and $\varphi(1) = 0$. Use φ^{-1} for the inverse function of φ. Then Archimedean copulas are defined as;

$$C_\varphi(u_1, u_2) = \varphi^{-1}(\varphi(u_1) + \varphi(u_2)) \qquad u_1, u_2 \in (0, 1]$$

where φ is called a generator function of te Copula C_φ. The Archimedean representation allows us to reduce the study of a multivariate copula to a single univariate function. Archimedean copulas have a commutative property, i.e. $C_\varphi(u_1, u_2) = C_\varphi(u_2, u_1)$. This class family of copulas have many different forms and you can see three often used Archimedean copulas in the following subsections.

3.6.1. Clayton Copula

The Clayton copula is also known as Cook-Johnston copula, whose generator φ_θ is defined by

$$\varphi_\theta(u) = \frac{1}{\theta}(u^{-\theta} - 1)$$

Hence, the Clayton copula is in the form of:

$$C_\theta(u_1, u_2) = \left(u_1^{-\theta} + u_2^{-\theta} - 1\right)^{-1/\theta}$$

The upper tail dependence measure for Clayton copula is zero ($\lambda_U = 0$) and the lower tail dependence measure is positive ($\lambda_L = 2^{-1/\theta} > 0$). See Panjer [2006] for a proof of these results. The Clayton copula has a single parameter θ that can be estimated from data using a statistical methods, such as maximum likelihood. Using **R**, scatter plots of the bivariate Clayton copula for various values of $\theta = 0.5$, $\theta = 2$ and $\theta = 10$ are shown in Figures 3.4 , 3.5 and 3.6 respectively.

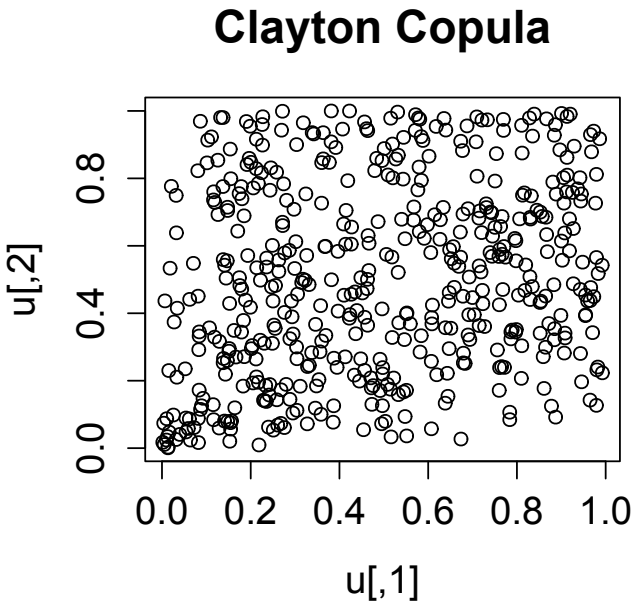

Figure 3.4.: Scatter plot of the Clayton copula for $\theta = 0.5$

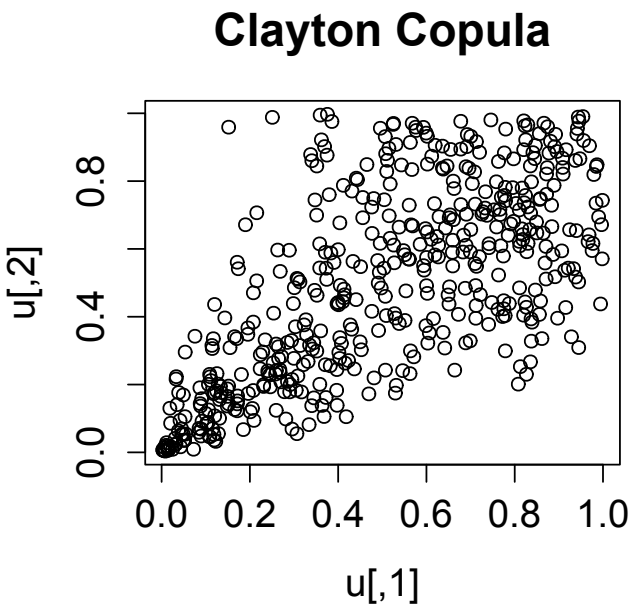

Figure 3.5.: Scatter plot of the Clayton copula for $\theta = 2$

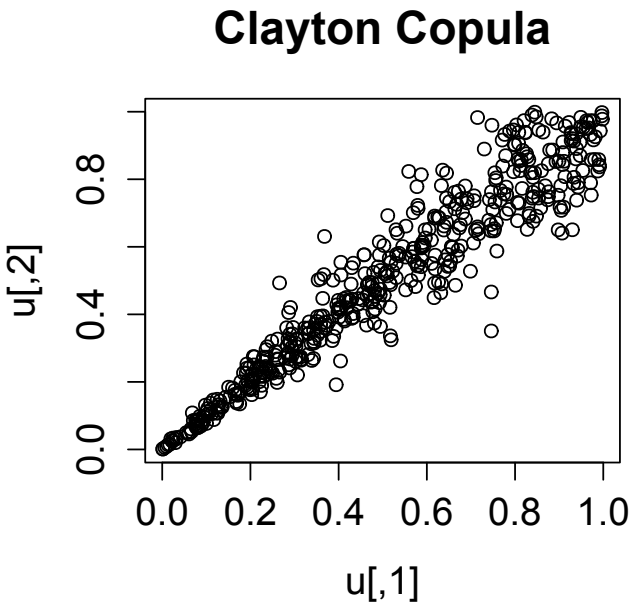

Figure 3.6.: Scatter plot of the Clayton copula for $\theta = 10$

Note that there is a strong correlation in the lower left corner of each plot and it represents positive lower tail dependence. In the upper right corner, there is no evidence of dependence and it indicates zero upper tail dependence. It can be seen in rank scatter plots of the bivariate

Clayton copula 3.4, 3.5 and 3.6 above that the higher the value of θ, the more the two random variables depend on each other in lower tail. As mentioned in Bürgi et al. [2008], the Clayton copula is not symmetric and acts on the lower tail of the distribution, whereas for upper tail the random variables are hardly dependent on each other. In insurance, however, the dependence should be modeled for the upper tails. This can easily be obtained by mirroring the copula by a transformation $(u_1, u_2) \longrightarrow (1 - u_1, 1 - u_2)$.

3.6.2. Gumbel Copula

The Gumbel copula is also known as the Gumbel-Hougaard copula and has the generator:

$$\varphi_\theta(u) = (-\ln u)^\theta \qquad \theta \geq 1$$

Hence, the Gumbel copula has the form:

$$C_\theta(u_1, u_2) = \exp\left(-\left[(-\ln u_1)^\theta + (-\ln u_2)^\theta\right]^{1/\theta}\right)$$

The measure of upper tail dependence is $\lambda_U = 2 - 2^{1/\theta}$. See Panjer [2006] for a proof of this result. The Gumbel copula also has a single parameter θ. Using **R**, scatter plots of the bivariate Gumbel copula for various values of $\theta = 1.5$, $\theta = 2$ and $\theta = 5$ are shown in Figures 3.7 , 3.8 and 3.9 respectively.

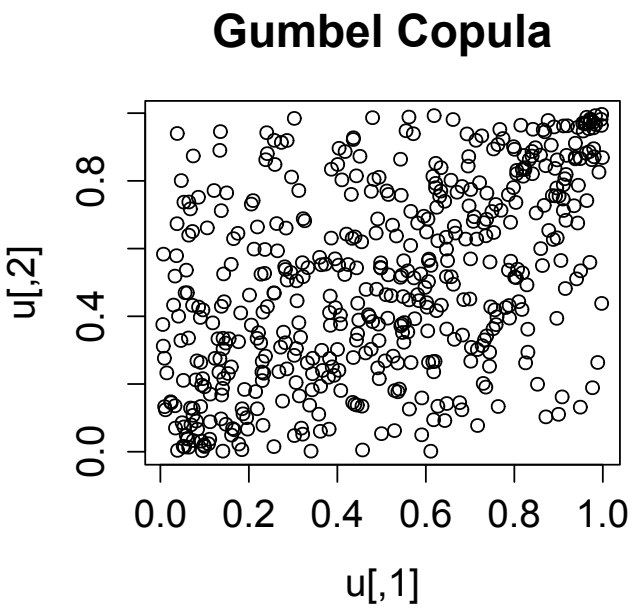

Figure 3.7.: Scatter plot of the Gumbel copula for $\theta = 1.5$

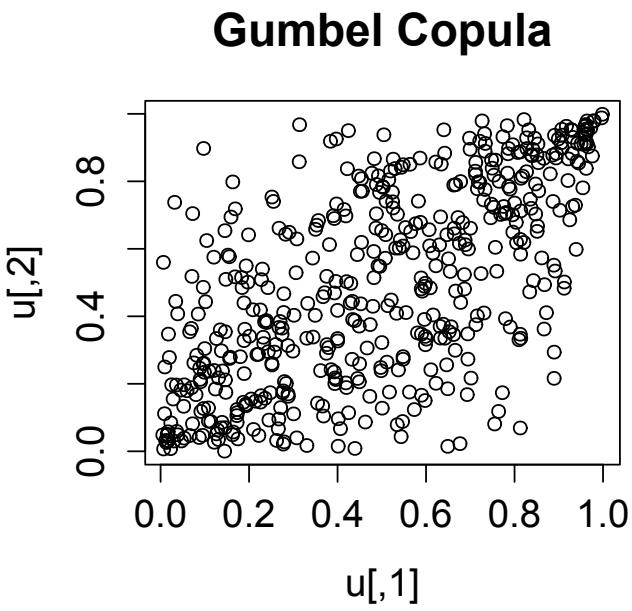

Figure 3.8.: Scatter plot of the Gumbel copula for $\theta = 2$

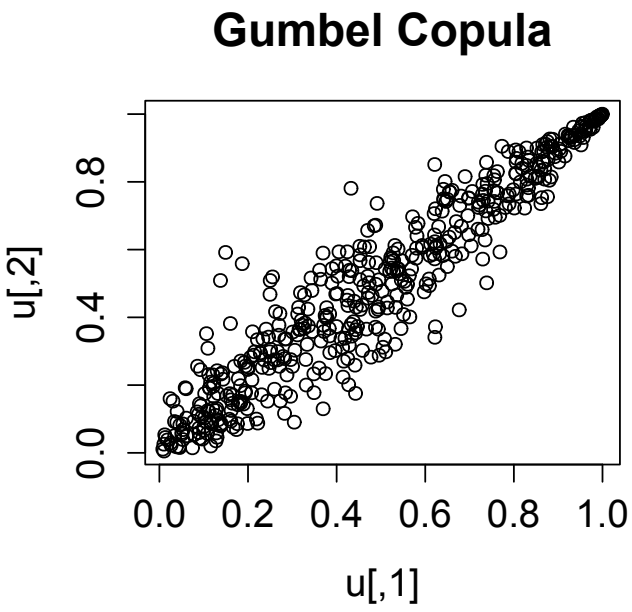

Figure 3.9.: Scatter plot of the Gumbel copula for $\theta = 5$

The upper tail dependence is evident in the upper right corner of each plot in above figures. Note that the dependence of the random variables grows with increasing of θ and also there is no upper tail dependence when $\theta = 1$ because we have $\lambda_U = 2 - 2^{1/\theta}$. As mentioned in the paper

Bürgi et al. [2008], in comparison to the Clayton copula, the Gumbel dependence acts on both upper and lower tails. Nevertheless, the Gumbel copula is not symmetric, i.e. the dependence of high Quantiles is stronger than the one of low Quantiles.

Simulations using copulas can be implemented in **R** package copula (see Yan et al. [2007] *"Enjoy the joy of copulas: with a package copula"*). The **R** codes that I have used to create the scatter plots of the bivariate Gaussian, Clayton and Gumbel copulas above, can be found in Appendix A.

4. Copula-based hierarchical model for risk aggregation

We discuss copula-based hierarchical approach for risk aggregation and dependence modeling here. In the next chapter, to empirically illustrate this method we will apply the hierarchical aggregation model to the real data, *Danish fire insurance data*, in detail and present some conclusions. This chapter is inspired by the paper on "*Copula based hierarchical risk aggregation through sample reordering*" (Arbenz et al. [2012]).

The objective of risk aggregation and dependence modeling is to model adequately dependent insurance portfolios in order to evaluate the overall risk exposure. In risk aggregation we are interested in the distribution of sum of insurance claims:

$$S = \sum_{i=1}^{d} X_i$$

where X_i are the value of losses due to some risks that are correlated. Dependence between risks cannot be ignored and their dependence must be modelled appropriately. As outlined in the previous chapters, there are three popular risk aggregation methods and modeling dependence structure in insurance:

- *variance-covariance method*
- *risk factor models*
- *copula models*

But each of these three models has its own weaknesses and in high dimensions, they become problematic (see Arbenz [2012]):

- In *variance-covariance method*, the conclusions that can be drawn from this model are only limited to the first and second moments, i.e. mean, variance and covariance of the risks $X_1, \ldots, X_d$, and thus further properties of the distribution of sum of the risks, S, cannot be understood. Number of correlation parameters in this method ($= d(d-1)/2$) become confusingly large in high dimensions.

- In *risk factor method*, modelling risk factors and estimating risk factor sensitivities for all risks can be difficult in high dimensions.

- In *copula models*, fitting a copula in high dimensions is problematic. The number of parameters is not appropriate in high dimensions, e.g. for large d number of margins, the number of parameters is too many for Elliptical copulas ($= d(d-1)/2$) and too few for Archimedean copulas ($= 1$). Copula model simulation is numerically slow in high dimensions.

Hierarchical risk aggregation model can avoid the above problems. In this method we do not need to specify the whole multivariate dependence structure. It is enough to specify lower dimensional dependence structures for each of the aggregation steps. We can obtain the distribution of sum of the risks using partial sums. This method is appropriate for high dimensional problems, but for making it clearer we will consider a simple three-dimensional problem example.

4.1. Hierarchical aggregation

This model consists of an aggregation tree structure, marginal distributions for each risk and bivariate copulas for each aggregation step. Suppose we have three risks X, Y, Z and we want to compute the distribution of sum of these three risks, i.e. the total aggregate S;

$$S = X + Y + Z$$

The classical copula approach consists of modelling the joint distribution of (X, Y, Z) using one trivariate copula $C_{X,Y,Z}(F_X(x), F_Y(y), F_Z(z))$ and directly computing the distribution of sum, S. Instead, we do the *"hierarchical aggregation approach"*, in the following steps;

(1) We set the aggregation tree structure. Number of tree structures to aggregate 3 risks is equal to 3, to aggregate 4 risks is equal to 15 and so on. To set the aggregation tree structure, first we select the greatest dependencies either positive or negative, i.e. we select the two risks X_i, X_j that are most dependent (based on Kendall's tau τ) and combine them. The aggregation tree structure of this example is shown in Figure (4.1). In this specific case, a joint model for the pair (X, Y) would first be constructed. For further details on existence and uniqueness of aggregation tree structure we refer to Arbenz et al. [2012].

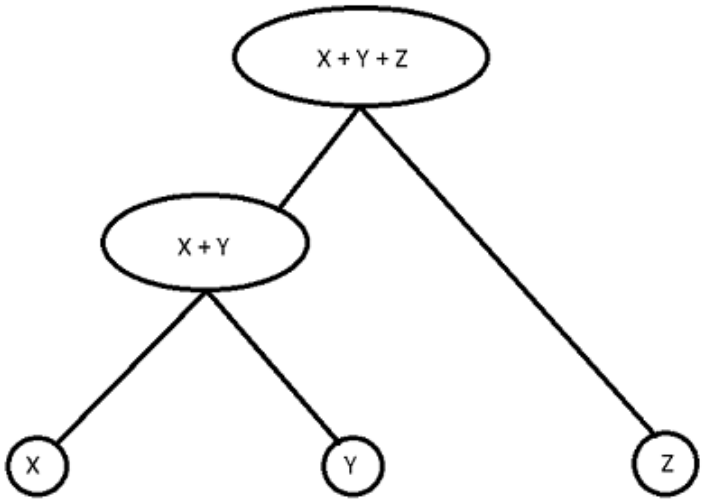

Figure 4.1.: An aggregation tree structure involving three risks.

(2) We have to specify the marginal distributions for each three risks X, Y, Z and estimate their parameters:

$$F_X(x) = P[X \leq x] \qquad F_Y(y) = P[Y \leq y] \qquad F_Z(z) = P[Z \leq z].$$

(3) We model the dependence structure among X and Y with a bivariate copula $C_{X,Y}$:

$$P[X \leq x, Y \leq y] = C_{X,Y}(F_X(x), F_Y(y)).$$

This determines the distribution of the bivariate random vector (X, Y), and as a result, it determines the distribution of the sub-aggregate T:

$$T = X + Y.$$

The cumulative distribution function (F_T) of sub-aggregate T is given by:

$$F_T(t) = P[X + Y \leq t] = \int_{\mathbb{R}^2} \mathbb{1}\{x + y \leq t\} \, dC_{X,Y}(F_X(x), F_Y(y)).$$

where $\mathbb{1}\{x + y \leq t\}$ is an indicator function defined as:

$$\mathbb{1}\{x + y \leq t\} = \begin{cases} 1 & \text{if } x + y \leq t \\ 0 & \text{if } x + y \nleq t \end{cases}$$

(4) Having fixed the distribution of $T = X + Y$ in the previous step, we would build a joint model for the pair (T, Z) by combining F_T and F_Z using the bivariate copula $C_{T,Z}$:

$$P[T \leq t, Z \leq z] = C_{T,Z}(F_T(t), F_Z(z)).$$

This determines the distribution of the bivariate random vector (T, Z), and as a result, it determines the distribution of the total aggregate S:

$$S = T + Z = X + Y + Z.$$

The cumulative distribution function F_S of total aggregate S is given by:

$$F_S(s) = P[T + Z \leq s] = \int_{\mathbb{R}^2} \mathbb{1}\{t + z \leq s\} \, dC_{T,Z}(F_T(t), F_Z(z)).$$

♦

This approach has many advantages. As mentioned above, in the classical copula approach for the calculation of the distribution of the sum of three risks, $S = X + Y + Z$, we would determine the trivariate copula $C_{X,Y,Z}$. As opposed to the classical approach, the hierarchical aggregation model involves only two bivariate copulas $C_{X,Y}$ and $C_{T,Z}$. This is an advantage when the number of risks is large, i.e. in higher dimensions. In general, the hierarchical aggregation for d risks $(X_1, X_2, \ldots, X_d)$ can be modelled as follows. First, select and combine the two risks (X_i, X_j) that are most dependent through a bivariate copula model C_{X_i,X_j}. Kendall's tau dependence measure can used to determine the order in which risks are aggregated. Next, replace the individual risks X_i and X_j by their sum $S_{ij} = X_i + X_j$ and repeat these steps for the new combined risk and the remaining risks, i.e. it leaves one new aggregation tree with $d - 1$ risks, that the procedure can be repeated for the new tree. Continue this procedure in a iterative way until all the risks have been aggregated in a single sum $S = X_1 + \ldots + X_d$. This procedure involves d marginal distributions and $d - 1$ bivariate copulas. (see Côté and Genest [2015])

4.2. Reordering algorithm for the numerical approximation

After the above steps in the hierarchical aggregation approach, i.e. selecting the aggregation tree structure, finding the marginal distributions for each risk, and fitting the bivariate copulas for each aggregation step, we need to numerically approximate the hierarchical risk aggregation structure. The "*reordering algorithm*" allows us to do it. Classical multivariate models use approximations through i.i.d. sampling (Monte Carlo simulations), but generating i.i.d. samples from this aggregation tree is not possible because it is not easy to deal numerically with the joint density (if existing) of all risks (X, Y, Z) and the copula function between all risks $C_{X,Y,Z}$. Instead of the classical i.i.d. sampling, we suggest the reordering algorithm for approximation that is inspired by the the Iman-Conover method (see Arbenz et al. [2012]). In this section we discuss the reordering algorithm that obtains numerical approximations through a bottom-up approach. For illustrative purposes, we again suppose we have three risks X, Y, Z example as described in the previous section. Therefore, our purpose is to approximate the distribution function of total aggregate:

$$S = T + Z = (X + Y) + Z.$$

Assume the marginal distributions of X, Y, Z are estimated and given by:

$$F_X(x) = P[X \leq x] \qquad F_Y(y) = P[Y \leq y] \qquad F_Z(z) = P[Z \leq z].$$

And two bivariate copulas $C_{X,Y}$ and $C_{T,Z}$ are fitted and given by;

$$P[X \leq x, Y \leq y] = C_{X,Y}(F_X(x), F_Y(y))$$

$$P[T \leq t, Z \leq z] = C_{T,Z}(F_T(t), F_Z(z)).$$

We do the reordering algorithm using the following steps;

(1) Fix number n.

(2) Simulate independently marginal samples of size n from X, Y and Z;

- $X_i \sim F_X$
- $Y_i \sim F_Y$
- $Z_i \sim F_Z$

for $i = 1, \ldots, n$.

(3) Simulate independently copula samples of size n from $C_{X,Y}$ and $C_{T,Z}$;

- $U_i \sim C_{X,Y}$
- $V_i \sim C_{T,Z}$

for $i = 1, \ldots, n$.

(4) For the first aggregation step, construct the bivariate reordered samples of (X, Y) by reordering the marginal samples X_i and Y_i based on the joint ranks of the copula sample U_i. Thus we get a sample of T by summing up $T = X + Y$.

(5) For the second aggregation step, construct the bivariate reordered samples of (T, Z) by reordering the marginal samples Z_i and T_i based on the joint ranks of the copula sample V_i. Thus we get a sample of total aggregate S by summing up $S = T + Z = (X + Y) + Z$.

(6) Define the empirical distribution function for S.

Thanks to Dr. Philipp Arbenz for putting the reordering algorithm **R** code program in his web page `https://sites.google.com/site/philipparbenz/`. For illustrating hierarchical aggregation through reordering algorithm, we give a trivariate example here:

We choose equal marginal distributions for all three risks. As mentioned in (Bürgi et al. [2008]), the biggest effect of dependence can be seen when aggregating *equal risks*. A canonical type of aggregate loss model used in insurance is the *lognormal distribution*. Therefore, we choose lognormal marginal distributions with parameters meanlog $\mu = 1$ and sdlog $\sigma = 1$ for each risk in our example. For generating the copula samples, we use Clayton copulas that is the most asymmetric one of the copulas (Clayton copula with parameter $\theta = 2$ for first aggregation step and with parameter $\theta = 1$ for second aggregation step). Recall that the Clayton copula is in the form of $C_\theta(u_1, u_2) = \left(u_1^{-\theta} + u_2^{-\theta} - 1\right)^{-1/\theta}$. We do the reordering algorithm example using the following steps:

- Fix number of simulations $n = 4$.
- Generate lognormal marginal samples i.i.d. $X_i \sim Lognormal(\mu = 1, \sigma = 1)$, for $i = 1, 2, 3, 4$, and then sort them. We denote the sorted X_i with notation $X_{(i)}$. This gives:

$$X_{(1)} = 0.56 \qquad X_{(2)} = 0.64 \qquad X_{(3)} = 1.60 \qquad X_{(4)} = 5.83$$

- Generate lognormal marginal samples i.i.d. $Y_i \sim Lognormal(\mu = 1, \sigma = 1)$, independent of the X_i. Then sort them. We denote the sorted Y_i with notation $Y_{(i)}$. This gives:

$$Y_{(1)} = 0.67 \qquad Y_{(2)} = 2.40 \qquad Y_{(3)} = 4.53 \qquad Y_{(4)} = 19.37$$

- Generate Clayton copula i.i.d. samples $U_i = (U_i^1, U_i^2)$, independent of the X_i and Y_i. the copula simulation yielded the following samples:

$$(U_1^1, U_1^2) = (0.74, 0.54) \qquad (U_2^1, U_2^2) = (0.46, 0.92)$$
$$(U_3^1, U_3^2) = (0.92, 0.70) \qquad (U_4^1, U_4^2) = (0.35, 0.38)$$

- Obtain the joint ranks of the copula samples. We denote the joint ranks with notation (R_i^1, R_i^2). This gives:

$$(R_1^1, R_1^2) = (3, 2) \qquad (R_2^1, R_2^2) = (2, 4)$$
$$(R_3^1, R_3^2) = (4, 3) \qquad (R_4^1, R_4^2) = (1, 1)$$

- Reorder and couple the samples $X_{(i)}$ and $Y_{(i)}$ such that the new bivariate sample has the same joint ranks as the copula samples. This gives the bivariate sample $(X_{(R_i^1)}, Y_{(R_i^2)})$ of:

$$(X_{(R_1^1)}, Y_{(R_1^2)}) = (X_{(3)}, Y_{(2)}) = (1.60, 2.40) \qquad (X_{(R_2^1)}, Y_{(R_2^2)}) = (X_{(2)}, Y_{(4)}) = (0.64, 19.37)$$
$$(X_{(R_3^1)}, Y_{(R_3^2)}) = (X_{(4)}, Y_{(3)}) = (5.83, 4.53) \qquad (X_{(R_4^1)}, Y_{(R_4^2)}) = (X_{(1)}, Y_{(1)}) = (0.56, 0.67)$$

- Calculate samples of the sub-aggregate $T = X + Y$ by taking the component wise sum of the bivariate reordered samples $(X_{(R_i^1)}, Y_{(R_i^2)})$ above. This gives:

$$T_1 = X_{(3)} + Y_{(2)} = 1.60 + 2.40 = 4.00 \qquad T_2 = X_{(2)} + Y_{(4)} = 0.64 + 19.37 = 20.01$$
$$T_3 = X_{(4)} + Y_{(3)} = 5.83 + 4.53 = 10.36 \qquad T_4 = X_{(1)} + Y_{(1)} = 0.56 + 0.67 = 1.23$$

- Repeat the previous procedure for samples T_i and $Z_i \sim Lognormal(\mu = 1, \sigma = 1)$ and reorder the samples T_i and Z_i such that their linked ranks are equal to the joint ranks of the copula samples $V = (V_i^1, V_i^2)$. This gives the following results:
  Sorted sub-aggregate samples $T_{(i)}$:

$$T_{(1)} = 1.23 \qquad T_{(2)} = 4.00 \qquad T_{(3)} = 10.36 \qquad T_{(4)} = 20.01$$

  Sorted lognormal marginal samples $Z_{(i)}$:

$$Z_{(1)} = 1.15 \qquad Z_{(2)} = 1.53 \qquad Z_{(3)} = 2.01 \qquad Z_{(4)} = 21.97$$

  Clayton copula simulation $V = (V_i^1, V_i^2)$ yielded:

$$(V_1^1, V_1^2) = (0.02, 0.28) \qquad (V_2^1, V_2^2) = (0.34, 0.06)$$

$$(V_3^1, V_3^2) = (0.33, 0.40) \qquad (V_4^1, V_4^2) = (0.17, 0.09)$$

  Joint ranks of the copula samples:

$$(R_1^1, R_1^2) = (1, 3) \qquad (R_2^1, R_2^2) = (4, 1)$$

$$(R_3^1, R_3^2) = (3, 4) \qquad (R_4^1, R_4^2) = (2, 2)$$

  Reordered bivariate sample $(T_{(R_i^1)}, Z_{(R_i^2)})$ according to the joint ranks of the copula V samples:

$$(T_{(R_1^1)}, Z_{(R_1^2)}) = (T_{(1)}, Z_{(3)}) = (1.23, 2.01) \qquad (T_{(R_2^1)}, Z_{(R_2^2)}) = (T_{(4)}, Z_{(1)}) = (20.01, 1.15)$$

$$(T_{(R_3^1)}, Z_{(R_3^2)}) = (T_{(3)}, Z_{(4)}) = (10.36, 21.97) \qquad (T_{(R_4^1)}, Z_{(R_4^2)}) = (T_{(2)}, Z_{(2)}) = (4.00, 1.53)$$

- Calculate samples of the total aggregate $S = T + Z = X + Y + Z$ by taking the component wise sum of the bivariate reordered samples $(T_{(R_i^1)}, Z_{(R_i^2)})$ above. This gives:

$$S_1 = T_{(1)} + Z_{(3)} = 1.23 + 2.01 = 3.24 \qquad S_2 = T_{(4)} + Z_{(1)} = 20.01 + 1.15 = 21.16$$

$$S_3 = T_{(3)} + Z_{(4)} = 10.36 + 21.97 = 32.33 \qquad S_4 = T_{(2)} + Z_{(2)} = 4.00 + 1.53 = 5.53$$

  This defines the empirical distribution function for the total aggregate S as:

$$F_S^n(s) = F_S^4(s) = \frac{1}{4}\mathbb{1}\{3.24 \leq s\} + \frac{1}{4}\mathbb{1}\{21.16 \leq s\} + \frac{1}{4}\mathbb{1}\{32.33 \leq s\} + \frac{1}{4}\mathbb{1}\{5.53 \leq s\}$$

  where $\mathbb{1}\{.\}$ is an indicator function. ♦

The reordering algorithm produces the approximations of the distributions of (X, Y), (T, Z), and S. The final result, F_S^4, is an empirical distribution function for the total aggregate S. Suppose the F_S is the cumulative distribution function of S, then F_S^4 is an approximation of F_S. In general, for fixed number of simulations n, the F_S^n is an approximation of F_S. A theorem given in (Arbenz et al. [2012], subsection 3.2, page 7) proves that when $n \to \infty$, we obtain convergence $F_S^n \to F_S$.

The R programming code that used for the implementation of the reordering algorithm for the described trivariate example, can be found in Appendix B.

5. Application to Danish Fire Insurance Data

These Danish Fire Insurance Data were collected at Copenhagen Reinsurance and comprise 2167 fire losses over the period 1980 to 1990 and are expressed in millions of Danish Krone (DKK). Every total claim has been divided into three risks consisting of a building loss, a loss of contents and a loss of profits caused by the same fire. These three risks in our study are denoted by:

$$X = \text{loss of building} \qquad Y = \text{loss of content} \qquad Z = \text{loss of profit}$$

As mentioned in Esmaeili and Klüppelberg [2010], Haug et al. [2011] and Dreesa and Müllerb [2007], the claims are recorded only if the sum of all three risks is greater or equal to 1 million Danish Kroner (DKK). Because of this, there is an artificial negative dependence between the risks components X, Y and Z, i.e. if one risk component is smaller than 1 million DKK, the sum of the others must be accordingly larger. Therefore, **we assume the values in each risk component had been truncated from below at** 1 **million DKK**. The risks X , Y and Z are clearly dependent, as can be seen from Figure 5.1. Note that the number of observations is not the same each time I consider a pair of risks.

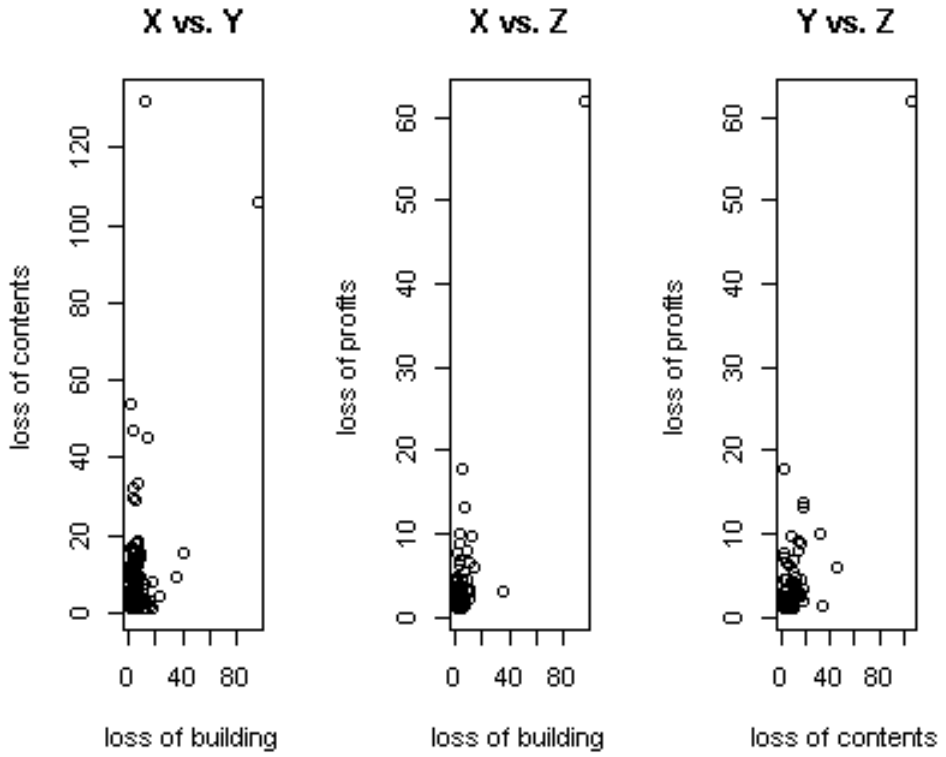

Figure 5.1.: Bivariate scatterplots of Danish Fire Insurance Data

5.1. R Packages fitdistrplus and copula

Here we briefly introduce two important **R** packages. In the following subsections we are frequently using these **R** packages to statistically analyze this data set. The *copula* package provides a platform for multivariate modeling with copulas in **R** (see Yan et al. [2007] and Kojadinovic et al. [2010]). We are using *copula* package for constructing Archimedean copula and Elliptical copula class objects with its corresponding parameters and dimension, goodness-of-fit tests and parametric estimation for copulas, and multivariate independence test of continuous random variables based on the empirical copula process.

In *fitdistrplus* package, the multivariate Danish Fire Insurance Data set is stored in *"danishmulti"*. This data set has been divided into a building loss, a loss of contents and a loss of profits (see Delignette-Muller and Dutang [2014]). In this package, *"danishmulti"* contains five columns:

- *Date*: The day of claim occurence.
- *Building*: The loss amount (mDKK) of the building coverage.
- *Contents*: The loss amount (mDKK) of the contents coverage.
- *Profits*: The loss amount (mDKK) of the profit coverage.
- *Total*: The total loss amount (mDKK).

where all columns are numeric except *Date* columns of class Date. We can import the multivariate Danish Fire Insurance Data set in **R** by:

```
> library(fitdistrplus)
> data(danishmulti)
```

5.2. Determination of the tree structure, hierarchical clustering

As mentioned in section 4.1 in the previous chapter, to determine the aggregation tree structure, first we select the two risks that are most dependent and combine them. *Kendall's tau* dependence measure can be used to determine the order in which risks are aggregated. A procedure for selecting the tree structure based on Kendalls tau τ is called *"hierarchical clustering technique"*. Consider two arbitrary risks X_i and X_j. The measure of distance $D(X_i, X_j)$ between these two risks based on Kendall's tau τ is defined as:

$$D(X_i, X_j) = \sqrt{1 - \tau^2(X_i, X_j)} \tag{5.1}$$

where $\tau(X_i, X_j)$ denotes *Kendall's tau* dependence measure between pair X_i and X_j. The principle of *hierarchical clustering method* is to identify the two risks that are *the closest*, i.e. the measure of distance D is minimal for them, and then to combine them into a group. Then we repeat this procedure until only one risk is left (see Côté and Genest [2015]). As can be seen from equation (5.1), the minimal measure of distance D is related to the largest Kendall's tau τ dependence measure. Therefore, first we select and model the two risks with the largest Kendall's tau τ.

To test if Kendall's tau τ is greater than zero at 0.05 significance level for components of our data set, we could test the null hypothesis that $\tau = 0$ by applying the function "*cor.test*" with the "*kendall*" option in **R**. For instance:

```
> cor.test(contents, profits, method = "kendall", alternative = "greater")
```

We found *p-values* of $2.495e-08$ for (X, Y), 0.001346 for (X, Z) and $4.629e-06$ for (Y, Z) that they are all less than 0.05 significance level. Hence we reject the null hypothesis and Kendall's tau τ is significantly greater than zero. It can be seen this result from Figure 5.1 also, that the risks are clearly dependent. Recall that X = loss of building, Y = loss of content, Z = loss of profit.

We apply the function "*cor*" with the "*kendall*" option in **R** for comparison of Kendall's tau τ between three different pairs of components of danish fire insurance data. For instance:

```
> cor(contents, profits, method="kendall")
```

After applying function "*cor*", we got these results:

$$\tau(X, Y) = 0.211032, \qquad \tau(X, Z) = 0.261722, \qquad \tau(Y, Z) = 0.328090$$

From equation (5.1) we could calculate the measure of distance D:

$$D(X, Y) = 0.977480, \qquad D(X, Z) = 0.965143, \qquad D(Y, Z) = 0.944646$$

The measure of distance $D(Y, Z)$ is the smallest one, i.e. Y and Z are *the closest*. Thus we first construct the joint model for the pair (Y ="*loss of contents*" , Z ="*loss of profits*"). The aggregation tree structure of this data set is shown in Figure 5.2.

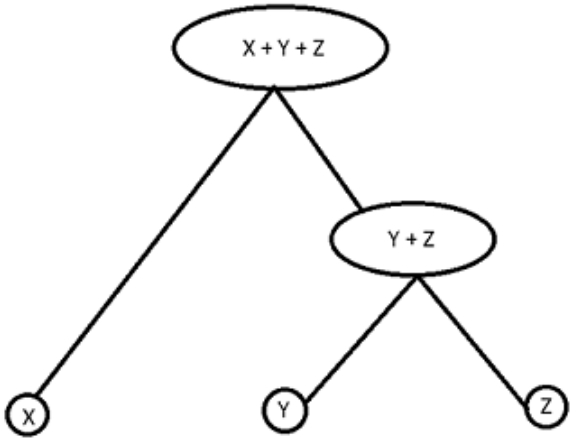

Figure 5.2.: An aggregation tree structure of Danish Fire Insurance Data.

5.3. Choice of marginal distributions for the risks

We have to specify the marginal distributions for the X = loss of building, Y = loss of contents, Z = loss of profits. Note that when estimating the marginals I consider the situation risk by risk and when I estimate the copulas I consider each pair of risks separately. As explained before about danish fire insurance data, **we assume the values in each risk component had been truncated from below at** 1 **million DKK**.

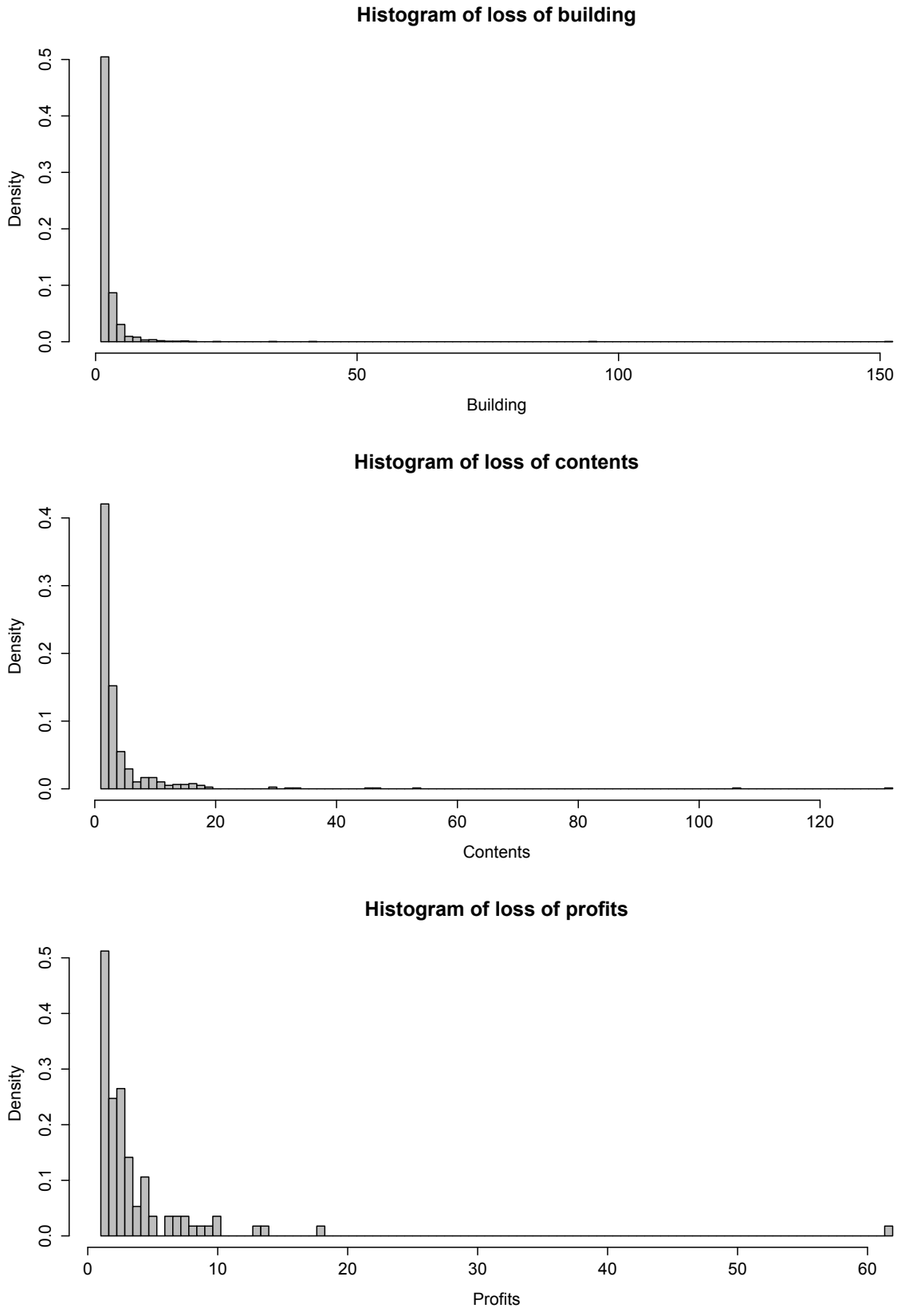

Figure 5.3.: Histograms of loss of building, loss of contents and loss of profits

The most direct way to see how well a distribution fits a data set is to plot the respective histogram, which can suggest a kind of distribution to use to fit the model. These plots are a very useful guide to heavy-tailedness. We present the histogram of each risk X, Y and Z in Figure 5.3. A first glance at Figure 5.3 clearly shows heavytailedness and skewness to the right for each data set. The danish fire insurance data show Pareto tail behaviour. The **Single-parameter Pareto** distribution is a classical skewed, heavy-tailed distribution. We estimate parameter shape (α) of Single-parameter Pareto distribution fitted to loss of building by maximum likelihood estimation method.

The Single-parameter Pareto distribution with parameter shape = α has PDF and CDF:

$$f(x) = \frac{\alpha\theta^{\alpha}}{x^{\alpha+1}} \qquad x > \theta.$$

$$F(x) = 1 - \left(\frac{\theta}{x}\right)^{\alpha} \qquad x > \theta.$$

Although there appears to be two parameters, only parameter shape = α is a true parameter. The value of lower bound = θ must be set in advance. In our case $\theta = 1$.

The Maximum Likelihood Estimation method (MLE) is the most popular method to estimate the distribution parameters from an empirical sample. It finds the model parameters that maximize the likelihood of the observed data with respect to the theoretical model. One of the attractive properties of the Single-parameter Pareto distribution is the ease of calculation of the maximum likelihood estimate of the parameter.

In general, the likelihood function for Single-parameter Pareto distribution with parameter shape = α, given data **truncated** from below at θ, is:

$$L(\alpha) = \prod_{j=1}^{n} \frac{f(x_j \mid \alpha)}{1 - F(\theta \mid \alpha)}$$

Note that in Single-parameter Pareto distribution the lower bound is equal to truncation point ($\theta = 1$) and the Survival function at lower bound $(1 - F(\theta \mid \alpha)) = 1$. Therefore the likelihood function is:

$$L(\alpha) = \prod_{j=1}^{n} \frac{\alpha}{x_j^{\alpha+1}}$$

The logarithmic likelihood function is:

$$\begin{aligned} l(\alpha) &= \sum_{j=1}^{n} (ln\alpha - (\alpha+1)ln(x_j)) \\ &= nln\alpha - (\alpha+1)\sum_{j=1}^{n} ln(x_j) \end{aligned}$$

To find the estimator for α, we compute the corresponding partial derivative and determine where it is zero:

$$\frac{\partial l}{\partial \alpha} = \frac{n}{\alpha} - \sum_{j=1}^{n} ln(x_j) = 0$$

Thus the maximum likelihood estimator for α is:

$$\hat{\alpha} = \frac{n}{\sum_{j=1}^{n} ln(x_j)}$$

As a result of above formula, we got parameter shape estimation $\alpha = 1.587$ and lower bound $\theta = 1$ for loss of building.

These real data are surely not exactly Single-parameter Pareto distributed, and for most practical applications the question would be "how good is the Single-parameter Pareto approximation?". The Q-Q plot is a good way to show the quality of such an approximation. A Q-Q plot represents the quantiles of the theoretical fitted distribution against the empirical quantiles of the data. The Q-Q plot of loss of building data against its estimated Single-parameter Pareto distribution is shown in Figure 5.4.

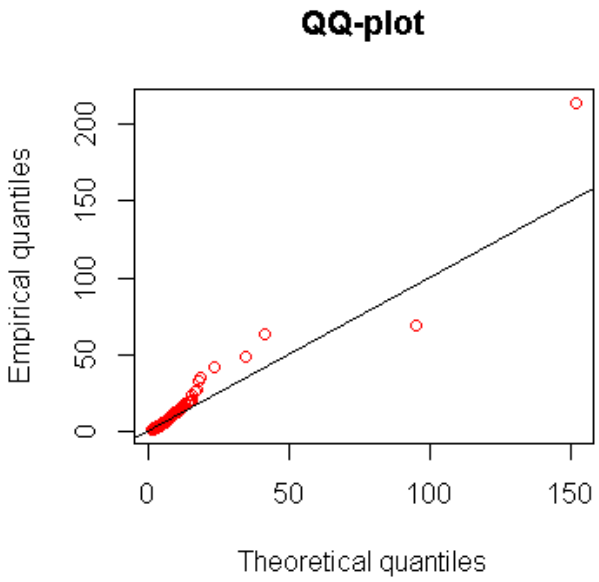

Figure 5.4.: Q-Q plot of loss of building data against its estimated Single-parameter Pareto distribution.

Any way we can do our test with the Kolmogorov-Smirnov statistic, after estimating the parameters by maximum likelihood. The KolmogorovSmirnov statistic quantifies a distance between the empirical distribution function of the sample and the cumulative distribution function of the reference distribution. We perform K-S test using *ks.test* function in **R**:

```
> ks.test(Building, "ppareto1", shape = 1.587, min = 1)
```

We got the following **R** output:

```
##################################################
One-sample Kolmogorov-Smirnov test

data:  Building
D = 0.0699, p-value = 1.324e-06
alternative hypothesis: two-sided
##################################################
```

which shows this data do not follows Single-parameter Pareto distribution.

Many articles have used the same danish fire insurance data and concluded that this data follows different distributions (see Esmaeili and Klüppelberg [2010], Haug et al. [2011], Dreesa and Müllerb [2007] and McNeil [1997]). Although these data are surely not exactly Single-parameter Pareto distributed, I assume they follow Single-parameter Pareto distribution in this thesis for the purposes of illustration of the hierarchical aggregation model.

Therefore, we select a marginal distribution Single-parameter Pareto for loss of building with parameter shape $\alpha = 1.587$ and lower bound $\theta = 1$. In a similar way, we select the marginal Single-parameter Pareto distribution for loss of contents with parameter shape $\alpha = 1.085$ and loss of profits with parameter shape $\alpha = 1.038$.

5.4. Choice of bivariate copulas

After determination of the aggregation tree structure, we select an appropriate bivariate copula at each aggregation step. I am analyzing the dependence only in the tails. As explained above, we first construct the joint model for the pair (Y ="*loss of contents*" , Z ="*loss of profits*") at first aggregation step because they are the two risks that are most dependent. To check the null hypothesis of independence of loss of contents and loss of profits, we implement an independence test using functions *indepTestSim* and *indepTest* of the *copula* package in **R**. This independence test consists of two steps: (**i**) *indepTestSim* function: a simulation step, which consists of simulating the distribution of the test statistics under independence for the sample size under consideration; (**ii**) *indepTest* function: the test itself, which consists of computing the approximate p-values of the test statistics with respect to the empirical distributions obtained in the first step (see Kojadinovic et al. [2010]). We apply the independence test to loss of contents and loss of profits:

```
> cp <- cbind(contents , profits)
> empsamp <- indepTestSim(nrow(cp), p = 2, N=1000)
> indepTest(cp,empsamp)
```

We obtain the **R** output:

```
##############################################################################
Global Cramer-von Mises statistic: 0.2200595 with p-value 0.0004995005
Combined p-values from the Mobius decomposition:
   0.0004995005  from Fisher's rule,
   0.0004995005  from Tippett's rule.
##############################################################################
```

These *p-values* give strong evidence against the null hypothesis of independence at the 0.05 significance level. The independence is rejected, and the next step is to fit an appropriate parametric copula $C_{Y,Z}$ function to the pair (Y ="*loss of contents*" , Z ="*loss of profits*"). As candidate copulas, we consider Gaussian (Normal), Clayton and Gumbel copula families to model the dependence among Y and Z. We perform several goodness-of-fit tests for these copula families using *gofCopula* function in *copula* package:

```
> normal.cop <- normalCopula(0.6, dim=2)
> gofCopula(normal.cop, cp)

> clayton.cop <- claytonCopula(2, dim=2)
```

```
> gofCopula(clayton.cop, cp)

> gumbel.cop <- gumbelCopula(2, dim=2)
> gofCopula(gumbel.cop, cp)
```

We found approximate *p-values* of 0.004496 for Normal copula, 0.0004995 for Clayton copula and 0.1683 for Gumbel copula respectively. Therefore, among all candidate copula families that we have tested, the Gumbel copula is the only one that is not rejected at the 0.05 significance level because its *p-value* is greater than 0.05. The parameter estimate for the copula fit is computed by the pseudo maximum likelihood method. The *gofCopula* function also returns the estimate of the parameters of the Gumbel copula, $\theta = 1.534$. In a similar way, we select an appropriate bivariate parametric copula $C_{X,T}$, Gumbel copula with parameter estimate $\theta = 1.282$, for the pair (X, T) at the second aggregation step where

$$X = \text{loss of building} \qquad T = (Y + Z) = (\text{loss of contents} + \text{loss of profits})$$

5.5. Hierarchical aggregation through reordering algorithm

After determining the aggregation tree structure, finding the marginal distributions for each risk, and fitting the bivariate copulas for each aggregation step, we need to numerically approximate the hierarchical risk aggregation structure. We do it through reordering algorithm. As explained in the previous chapter, we do the reordering algorithm using the following steps:

(1) Fix number of simulations $n = 1000$.

(2) Generate Single-parameter Pareto marginal i.i.d. samples of size $n = 1000$ from X, Y and Z;

- $X_i \sim$ Single-parameter Pareto($\alpha = 1.587$)
- $Y_i \sim$ Single-parameter Pareto($\alpha = 1.085$)
- $Z_i \sim$ Single-parameter Pareto($\alpha = 1.038$)

for $i = 1, \ldots, 1000$.

(3) Generate copula i.i.d. samples of size $n = 1000$ from $C_{Y,Z}$, Gumbel copula with parameter $\theta = 1.534$, and $C_{X,T}$, Gumbel copula with parameter $\theta = 1.282$;

- $U_i \sim C_{Y,Z}$
- $V_i \sim C_{X,T}$

for $i = 1, \ldots, 1000$.

(4) For the first aggregation step, construct the bivariate reordered samples of (Y, Z) by reordering the marginal samples Y_i and Z_i based on the joint ranks of the copula sample U_i. Thus we get a sample of T by summing up $T = Y + Z$.

(5) For the second aggregation step, construct the bivariate reordered samples of (X, T) by reordering the marginal samples X_i and T_i based on the joint ranks of the copula sample V_i. Thus we get a sample of the total aggregate S by summing up $S = X+T = X+(Y+Z)$

(6) Calculate the samples of the total aggregate S. Denote by $S_1, \ldots, S_{1000}$ the simulated samples of S. The empirical distribution function for S is defined as:

$$F_S^n(s) = F_S^{1000}(s) = \frac{1}{1000} \sum_{i=1}^{1000} \mathbb{1}\{S_i \leq s\}$$

where $\mathbb{1}\{.\}$ is an indicator function. ♦

5.6. Conclusions and results

Suppose that F_S is the cumulative distribution function of S, then F_S^{1000} is a good approximation of F_S. **R** provides a very useful function *ecdf* for working with the empirical distribution function. It computes or plots an empirical cumulative distribution function. The *ecdf* function applied to a data sample returns a function representing the empirical cumulative distribution function. Let's use the *ecdf()* function to obtain some empirical CDF values of S. For example it is possible to see what the output looks like below:

```
> Fn <- ecdf(S)
> Fn(3.04)
[1] 0.001
> Fn(40)
[1] 0.943
> Fn(100)
[1] 0.976
```

The empirical cumulative distribution function of S is shown in Figure 5.5.

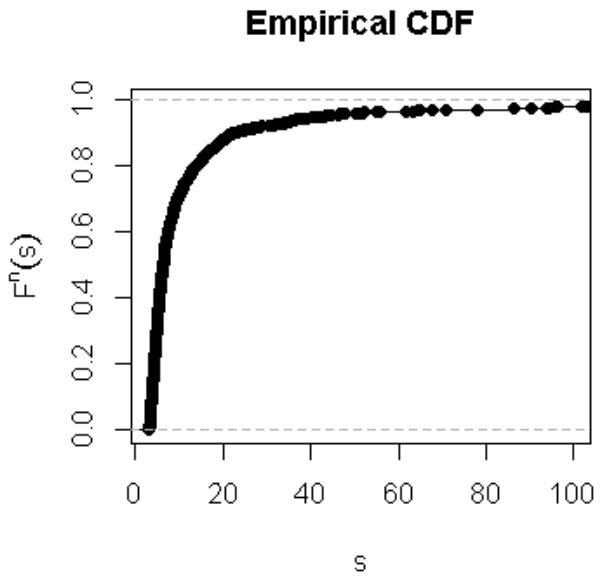

Figure 5.5.: The empirical cumulative distribution function of S

We present the histogram plot of the total aggregate samples S in Figure 5.6 to examine the distribution of S.

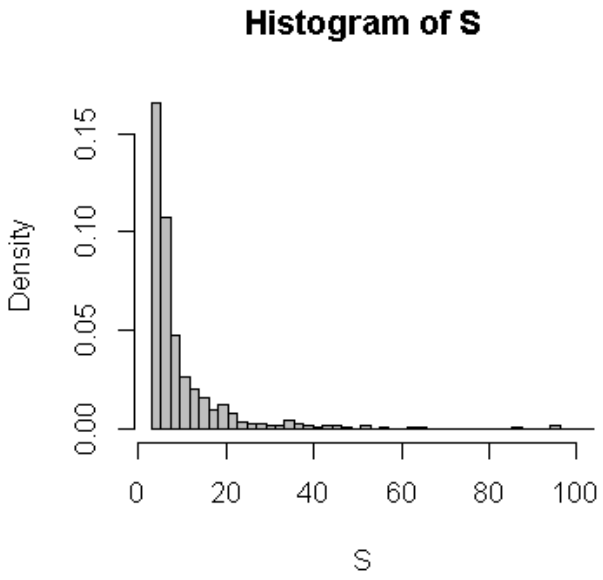

Figure 5.6.: Histogram and Density estimate of the total aggregate S

As can be seen in this master thesis, the *copula-based hierarchical aggregation model through reordering algorithm* provides a simple and practical framework to model the distribution of the sum of dependent risks. In general, a hierarchical aggregation model for d risks is defined with a tree structure, $d-1$ bivariate copulas and d marginal distributions. This approach can be easily applicable in high dimensions. The only restriction being that the marginal distributions for each risk, and the bivariate copulas for each aggregation step can be simulated. It would be interesting to try to find out more about this aggregation method in future research.

A. Appendix A

The R codes that I used to create the scatter plots of the bivariate Gaussian, Clayton and Gumbel copulas:

```
> library("copula")
> set.seed(1)

# Gaussian Copula
> norm.cop <- normalCopula(0.3)
> norm.cop
> u <- rcopula(norm.cop, 500)
> plot(u)
> title("Gaussian Copula")

# Clayton Copula
> clayton.cop <- claytonCopula(0.5)
> clayton.cop
> u <- rcopula(clayton.cop,500)
> plot(u)
> title("Clayton Copula")

# Gumbel Copula
> gumbel.cop <- gumbelCopula(1.5)
> gumbel.cop
> u <- rcopula(gumbel.cop,500)
> plot(u)
> title("Gumbel Copula")
```

B. Appendix B

The R programming code that used for the implementation of the reordering algorithm for the mentioned trivariate example:

```
# load package 'copula' library
> library(copula)

# fix number of simulations
> n

# generate marginal samples. Here, lognormal distributions is used.
> X = rlnorm(n, 1, 1)
> Y = rlnorm(n, 1, 1)
> Z = rlnorm(n, 1, 1)

# generate copula samples. Here, Clayton copulas is used.
> U = rcopula(claytonCopula(param = 2, dim = 2), n)
> V = rcopula(claytonCopula(param = 1, dim = 2), n)

# reordering according to U
> X[order(U[,1])] = sort(X)
> Y[order(U[,2])] = sort(Y)
> print(cbind(X,Y) , digits=3)

# calculate samples of the sub-aggregate T
> T = X+Y

# reordering according to copula V
> T[order(V[,1])] = sort(T)
> Z[order(V[,2])] = sort(Z)
> print(cbind(T,Z) , digits=3)

# calculate total aggregate S
> S = T+Z

# final result
> print(S , digits=3)
```

List of Figures

3.1. Gaussian Copula for $\rho = 0.3$ . . . 13
3.2. Gaussian Copula for $\rho = 0.6$ . . . 13
3.3. Gaussian Copula for $\rho = 0.9$ . . . 14
3.4. Clayton Copula for $\theta = 0.5$ . . . 15
3.5. Clayton Copula for $\theta = 2$ . . . 16
3.6. Clayton Copula for $\theta = 10$ . . . 16
3.7. Gumbel Copula for $\theta = 1.5$ . . . 17
3.8. Gumbel Copula for $\theta = 2$ . . . 18
3.9. Gumbel Copula for $\theta = 5$ . . . 18

4.1. An aggregation tree structure involving three risks . . . 21

5.1. Bivariate scatterplots of Danish Fire Insurance Data . . . 26
5.2. An aggregation tree structure of Danish Fire Insurance Data . . . 28
5.3. Histograms of loss of building, loss of contents and loss of profits . . . 29
5.4. Q-Q plot of loss of building data against its estimated Single-parameter Pareto distribution . . . 31
5.5. The empirical cumulative distribution function of S . . . 34
5.6. Histogram and Density estimate of the total aggregate S . . . 35

Bibliography

P. Arbenz. *Multivariate Modelling in Non-Life Insurance*. PhD thesis, ETH ZURICH, 2012.

P. Arbenz, C. Hummel, and G. Mainik. Copula based hierarchical risk aggregation through sample reordering. *Insurance: Mathematics and Economics*, 51(1):122–133, 2012.

M. Bogard. Copula functions, r, and the financial crisis. R-bloggers.com, March 2011. URL `http://www.r-bloggers.com/copula-functions-r-and-the-financial-crisis/`.

R. Bürgi, M. M. Dacorogna, and R. Iles. Risk aggregation, dependence structure and diversification benefit. *Stress Testing for Financial Institutions*, 2008.

M.-P. Côté and C. Genest. A copula-based risk aggregation model. *Canadian Journal of Statistics*, 43(1):60–81, 2015.

M. L. Delignette-Muller and C. Dutang. fitdistrplus: An r package for fitting distributions. 2014.

M. Denuit, J. Dhaene, M. Goovaerts, and R. Kaas. *Actuarial Theory for Dependent Risks: Measures, Orders and Models*. John Wiley & Sons, 2006.

H. Dreesa and P. Müllerb. Modeling dependencies in large claims. 2007.

T. Economist. In defense of the gaussian copula. The Economist online, Apr 2009. URL `http://www.economist.com/blogs/freeexchange/2009/04/in_defense_of_copula`.

P. Embrechts, F. Lindskog, and A. McNeil. Modelling dependence with copulas and applications to risk management. *Handbook of heavy tailed distributions in finance*, 8(1):329–384, 2003.

H. Esmaeili and C. Klüppelberg. Parameter estimation of a bivariate compound poisson process. *Insurance: Mathematics and Economics*, 47(2):224–233, 2010.

E. W. Frees and E. A. Valdez. Understanding relationships using copulas. *North American actuarial journal*, 2(1):1–25, 1998.

S. Haug, C. Klüppelberg, and L. Peng. Statistical models and methods for dependence in insurance data. *Journal of the Korean Statistical Society*, 40(2):125–139, 2011.

I. Kojadinovic, J. Yan, et al. Modeling multivariate distributions with continuous margins using the copula r package. *Journal of Statistical Software*, 34(9):1–20, 2010.

A. J. McNeil. Estimating the tails of loss severity distributions using extreme value theory. *Astin Bulletin*, 27(01):117–137, 1997.

A. J. McNeil, R. Frey, and P. Embrechts. *Quantitative risk management: concepts, techniques, and tools*. Princeton university press, 2010.

R. B. Nelsen. *An introduction to copulas*. Springer Science & Business Media, 2007.

H. H. Panjer. *Operational risk: modeling analytics*, volume 620. John Wiley & Sons, 2006.

J. Skoglund. Risk aggregation and economic capital. *Available at SSRN 2070695*, 2010.

J. Yan et al. Enjoy the joy of copulas: with a package copula. *Journal of Statistical Software*, 21(4):1–21, 2007.

D. M. Zimmer. The role of copulas in the housing crisis. *Review of Economics and Statistics*, 94(2):607–620, 2012.