Dependence Structures and Risk Aggregation Using Copulas

Thesis submitted for the degree of

Doctor of Philosophy

at the

University of Leicester

by

Isaudin Ismail

Department of Mathematics

University of Leicester

September 2018

“The expert at anything was once a beginner.”

Helen Hayes

Abstract

Insurance and reinsurance companies have to calculate solvency capital requirements in order to ensure that they can meet their future obligations to policyholders and beneficiaries. The solvency capital requirement is a risk management tool essential when extreme catastrophic events happen, resulting in high number of possibly interdependent claims. In this thesis, we study the problem of aggregating the risks coming from several insurance lines of business and analyse the effect of reinsurance in the level of risk. Our starting point is to use a Hierarchical Risk Aggregation method, which was initially based on 2-dimensional elliptical copulas. We use copulas from the Archimedean family and a mixture of different copulas. The results show that a mixture of copulas can provide a better fit to the data than the plain (single) copulas and consequently avoid overestimation or underestimation of the capital requirement of an insurance company. We also investigate the significance of reinsurance in reducing the insurance company's business risk and its effect on diversification. The results show that reinsurance does not always reduce the level of risk but can reduce the effect of diversification for insurance companies with multiple business lines. To extend the literature on modelling multivariate distributions, we investigate the dependence structure of multiple insurance business lines risks using C-vine copulas. In particular, we use bivariate copulas, and aggregate the insurance risks. We employ three C-vine models such as mixed C-vine, C-vine Gaussian and C-vine t-copula to develop a new capital requirement model for insurance companies. Our findings suggest that the mixed C-vine copula is the best model which allows a variety of dependence structure estimated by its respective copula families.

Acknowledgements

First and foremost I would like to express my sincere gratitude to my Ph.D supervisors, Dr. Aihua Zhang (principal supervisor) and Dr. Alexandra Dias (co-supervisor), for their unparalleled support and thoughtful guidance. They gave me the freedom to do my research during my Ph.D studies while continuously provide constructive feedbacks and helpful suggestions through our fruitful discussions.

I would also like to extend my gratitude to my wonderful wife, Siti Mariam for her countless prayers and unlimited support. She has been my inspiration and motivation for me to keep positive and finish my Ph.D studies. I also thank my wonderful children, Irfan and Kayyisa, for being stress relievers to me and for understanding on those days especially school holidays when I was writing this thesis instead of spending time with them. I hope that one day they can understand why I spent so much time in my study room.

Special thanks to family in Malaysia especially my dad, Ismail and mum, Kintan as well as my mother in law, Siti Saleha for their prayers, love and unconditional support. To my brothers and sisters, thank you for your support and encouragement.

To all my Ph.D colleagues at Department of Mathematics, thank you for making my time spent at the department stress free through our casual talks and discussions. To my lunch buddies group, Simran, Num, Ka Kin, Aon and Eliyas, I enjoyed every lunch especially the lunch picnic we had at Victoria park.

It is a pleasure to thank all my Malaysian friends in Leicester for organizing a long list of great activities through Malaysian-Leicester Postgraduate Society (MASLEICS). I appreciated all moments we shared in the past.

I also gratefully acknowledge the funding of the Department of Mathematics, University of Leicester to purchase data for research during my Ph.D studies.

Finally, I would like to express my appreciation to the Ministry of Education Malaysia and University Tun Hussein Onn Malaysia (UTHM) for providing me the full scholarship to pursue my Ph.D studies at University of Leicester.

Contents

List of Figures

List of Tables

Abbreviations

AIC = Akaike Information Criterion

A-D = Anderson Darling

ABI = Association of British Insurers

APRA = Australian Prudential Regulatory Authority

BIC = Bayesian Information Criterion

C-vine = Canonical-vine

CDO = Collateralized Debt Obligation

CTP = Compulsory Third Party

CDF = Cumulative Density Function

D-vine = Drawable-vine

DFA = Dynamic Financial Analysis

EU = European Union

i.i.d = Independent and Identically Distributed

LAGIC = Life and General Insurance Capital

LR = Loss Ratio

ML = Maximum Likelihood

MLE = Maximum Likelihood Estimator

MME = Method-of-Moments Estimator

MCR = Minimum Capital Requirement

NAC = Nested Archimedean Copula

PCC = Pair-Copula Construction

PDF = Probability Density Function

RBC = Risk-Based Capital

SE = Sequential Estimation

SCR = Solvency Capital Requirement

TVaR = Tail-Value-at-Risk

UK = United Kingdom

VaR = Value-at-Risk

To my wife, children and parents

Chapter 1

Introduction

1.1 Background

Aggregation of risks from insurance companies perspective is a process of combining risks from all insurance business lines. In particular, it is important to understand the dependence structure between different insurance business lines. In this regard, dependence structure for insurance companies is interpreted as the behaviour or interaction of one business line to another business lines beyond linear dependence. Generally, insurance companies are divided into two main categories: general insurance and life insurance. These insurance companies are different from one another in terms of the protection they provide. General insurance protects against damages or losses to an asset while life insurance promises a lump sum or regular payments to beneficiary of a life policy upon the death of the policyholder. In this thesis, we focus on general insurance (hereafter referred to as insurance) whose losses are highly exposed to extreme events. The focal point is loss ratio as a proxy for insurance business risks. Risk managers and actuaries aggregate the risk of losses from insurance business lines to determine the capital requirement. Through a proper level of capital requirement, an insurance company is able to reduce the risk of insolvency as a part of the company's business continuity plans.

This thesis has a triple goal. First, we develop a new model to aggregate risks of an insurance company. In particular, we use copulas as a tool to model the dependence structure between insurance business lines risks. Second, by considering reinsurance business in an insurance company, we investigate the effects of reinsurance on the level of risks and analyse the influence of reinsurance on the dependence structure between different business lines. The first and second research goals are addressed by the first aggregation approach, hierarchical risk aggregation. Finally, we develop a new capital requirement estimation methodology for general insurance companies. This model focuses on the second aggregation approach using vine copula.

According to Embrechts et al. (2003), since it was introduced by Sklar (1959), copula is widely accepted and covered in the finance and insurance literature. (Darsow et al., 1992; Joe, 1996; Wang, 1998; Frees and Valdez, 1998; and Klugman and Parsa, 1999) are among the first using copulas. In finance, Breymann et al. (2003) investigate the dependence structure of two dimensional high-frequency FX spots data between US Dollar/Deutsch Mark[1] and US Dollar/Japanese Yen. Dias and Embrechts (2009) and Fortin and Kuzmics (2002) introduce a new method to model asymmetric dependence in asset returns. Jondeau and Rockinger (2006) model dependence between two stock returns while Hofert and Scherer (2011) use copula to model Collateralized Debt Obligation (CDO) prices.

Copulas are also frequently used in insurance. In the 1990s, the concept of copula was relatively unknown to insurance and finance. Frees and Valdez (1998) introduce basic properties of copulas and provide resources for future research. They also investigate the relationship of multivariate outcomes from financial systems and estimate joint life mortality and multi-decrement model. Wang (1998) pioneers the literature on risk aggregation of insurance business lines using Gaussian copula. According to Wang (1998), insurance risks are determined by the loss distribution of claims data. Then, these data are combined and the dependence structure of loss distributions is modelled using a Gaussian copula to derive the insurance company's risk aggregation model. However, Embrechts et al. (2003) highlight that Gaussian copulas are symmetric copulas and have a limited ability to model insurance losses especially in extreme events where potential contemporaneous high losses from insurance claims are expected. To address this issue, we propose in Chapter 4 to model the risk aggregation of an insurance company using copulas from the Archimedean family, in addition to Gaussian and t-copula. More precisely, we introduce hierarchical risk aggregation model to combine all insurance business lines' risks in general insurance companies using Archimedean copulas. This study focuses on the risk of general insurance companies. In addition, previous literature on copula in insurance predominately focus on claim data to model insurance risks (Frees and Valdez, 1998, Wang, 1998, Klugman and Parsa,

[1]Deutsch Mark was replaced with the Euro since 1 January 1999.

1999, among others). In this thesis, we explore a new method to model insurance risk. Instead of using claim data directly, we use loss ratio which is derived by claims per unit premiums. We discuss loss ratios in Section 2.3.

In general, insurance risks arising from large claims can be passed to reinsurance companies. In this regard, it is crucial to measure the dependence structure between insurance and reinsurance business. Using Dynamic Financial Analysis (DFA), Eling and Toplek (2009) analyse the risk and return profile of reinsurance contracts using copula models. They evaluate reinsurance contracts and use ruin probability with expected policyholder's deficit as benchmark for risk assessments. In contrast to Eling and Toplek (2009), in our second objective, we investigate the reinsurance effects on each insurance business line risk using loss ratio distributions and measure the risk reduction effects from reinsurance to the insurance company total risk in Chapter 4.

Modelling risk of multiple insurance business lines involves high dimensional distributions. This can be challenging and requires complex numerical computation. Bivariate distributions are proposed in the literature to simplify the modelling process. Due to its simpler implementation than high dimensional distributions, bivariate distributions are widely used in the literature to model dependence structure (Aas et al., 2009; Kurowicka and Joe, 2010; Arbenz et al., 2012; Brechmann and Schepsmeier, 2013; Côté and Genest, 2015; Mai and Scherer, 2012; and Cossette et al., 2017). Based on bivariate distributions, Arbenz et al. (2012), Côté and Genest (2015), Mai and Scherer (2012), and Cossette et al. (2017) decompose high dimensional distributions into bivariate distributions using Hierarchical copula models to model dependence. Similarly, Aas et al. (2009), Kurowicka and Joe (2010), Brechmann and Schepsmeier (2013) use bivariate copulas through vine copula models to model high dimensional distributions. Vine copulas reduce modelling complexity by pairing two datasets (bivariate) at the time. We use hierarchical copula models in Chapter 4 to model high dimensional insurance loss distributions for risks aggregation. In Chapter 5 we use vine copula models to model the capital requirement for insurance companies.

The research in this study focuses on two datasets. We use insurance data from Australia and the United Kingdom (UK) in Chapter 4 and 5 respectively. We investigate the dependence structure of insurance business lines from the risk perspective for both countries. We develop new models to aggregate risks and determine the capital requirement. In the UK and other European countries, Solvency

II was introduced on 1st January 2016 to regulate insurance companies by three different pillars. Pillar 1 concerns quantitative measures and relates to the capital requirement which is the central interest of this study. This pillar is the technical provision and includes guidelines for insurance companies to determine two level of capital requirements. The first one is Solvency Capital Requirement (SCR), a safe level for insurance companies, and the second one is Minimum Capital Requirement (MCR), a critical level that need immediate attention from insurance companies. Insurance companies with capital requirement below the MCR level are reported as insolvents. On the other hand, Pillar 2 focuses on the qualitative measures useful for risk management and governance while Pillar 3 details the disclosure requirements for annual published solvency and financial condition reports. Similarly, in the counterpart, Australia, the insurance companies are regulated by the Australian Prudential Regulatory Authority (APRA). Implemented 3 years earlier than Solvency II in the UK, the Life and General Insurance Capital (LAGIC) was introduced with similar objectives as Solvency II from its three pillars.

1.2 Thesis outline

In Chapter 2, we introduce some definitions for risk measures and loss ratio of an insurance company which are used in this thesis. We begin in Section 2.1 by giving the definitions of different risk measures such as Value-at-Risk (VaR) and Tail-Value-at-Risk (TVaR). We also discuss the importance of choosing a good (coherent) risk measure for modelling insurance risks and provide the properties of a coherent risk measure. In Section 2.2, we present methods for computing risk measures, and introduce loss ratios as a proxy for insurance risk in Section 2.3. In addition, we review the probability distributions, fitting distribution methods and standard test of randomness that are useful to model the aggregation of risks in this thesis.

Chapter 3 focuses on copulas and dependence measures. We first recall the definition of a copula in Section 3.1, and introduce the most commonly used copulas from Elliptical and Archimedean family in Section 3.2. Then, we introduce the statistical inference for copula in Section 3.3 including copula density, conditional copula and its statistical inference. In Section 3.4, we give two types of dependence measures such as linear dependence and non-linear dependence. We also provide a simple motivating example for choosing the best dependence measure. We end this chapter by giving an overview of graph theory which is important for modelling the aggregation of risk of an insurance company in this thesis.

Hierarchical risk aggregation model, one of the key methodologies in this thesis is introduced in Chapter 4. The hierarchical risk aggregation model is developed based on the application of graph theory as discussed in Section 3.5. This chapter seeks to address the problem of modelling the dependence structure of losses in general insurance companies. In Section 4.1, we highlights the importance of aggregation of risks and discuss relevant literature on risks aggregation using copulas. We introduce the methods for aggregating risk using hierarchical aggregation model in Section 4.2. This includes the conditions for constructing the aggregation model, simulation of observations from the aggregation model, numerical approximation algorithm and risk measures to estimate the risks of the aggregate losses. In Section 4.3, we provide the data we use for analysis in this chapter. We use data on general insurance from the Australia insurance industry. We use gross (from reinsurance) incurred claims, gross earned premiums, net incurred claims and net earned premiums variables to estimate the loss ratio distribution as a proxy for insurance risk. In this thesis, our second goal is to investigate the effect of reinsurance on capital requirements. We use both the gross and the net variables to investigate this effect. The results from the copula aggregation model and the effects of reinsurance in the level of risk and diversification of the portfolio of different business lines are presented in Section 4.4. In Section 4.5, we provide analysis of the results from focusing on the effect of reinsurance business to insurance companies from risk perspective. Conclusion is provided in Section 4.6. This chapter is based on a working paper by Dias et al. (2018) and a similar version of this chapter has been presented at the following international conferences: Institute & Faculty of Actuaries (IFoA) Asia Conference 2018, Bangkok, Thailand, 10-11 May 2018; 21st International Congress on Insurance: Mathematics and Economics, Vienna, July 6-7, 2017; 3rd Symposium on Quantitative Finance & Risk Analysis 2017, Corfu, Greece, 15-16 June 2017.

Chapter 5 introduces pair-copula constructions or vine copula to model high dimension vectors. In particular, we apply a vine copula method in modelling aggregation of multiple insurance business lines risks, to address the third research goal of this thesis. We review significant contributions in the literatures on vine copula in Section 5.2. Then, we develop theoretical foundations for vine copula constructions in Section 5.3. This includes, the simplifying assumptions to construct the vine copula and h-function for solving high dimensional conditional distribution functions. We particularly focus on C-vine copulas and discuss the procedures to build its tree structure in Section 5.4. In addition, also in Section 5.4, we provide the vine copula inference, a numerical example for C-vine model estimation, copula selection and simulation of a C-vine. It is important, from a statistical point of view, to investigate the impact of dependence structures to insurance companies. To this end, in Section 5.5, we introduce an empirical analysis on aggregation of multiple insurance business lines risks in the UK using data sourced from the Association of British Insurers (ABI). The types of data used in this chapter to derive loss ratios distributions is similar to Chapter 4. However, due to different reporting format in the UK, we use written premiums instead of earned premiums as one of the variables in this chapter. Statistical results including the C-vine structure and simulations to obtain VaR and TVaR are also presented. Section 5.6 concludes the chapter. A similar version of this chapter has been presented at Institute & Faculty of Actuaries (IFoA) Asia Conference 2018, Bangkok, Thailand, 10-11 May 2018.

A summary of the thesis and further work are given in Chapter 6.

Chapter 2

Risk measures for general insurance

2.1 Introduction

In this chapter we discuss risk measures that are particularly important and frequently used in insurance industry. Although some of the risk measures can be applied to life insurance business, we primarily focus on the application in the general insurance businesses. We start in Section 2.1.1 by giving an overview of Value-at-Risk (VaR) and its application in determining the capital requirement for insurance companies. In addition, we also highlight the key challenges for VaR as a risk measure. Properties for a good risk measure (coherent) are introduced by Artzner et al. (1999) and details of these properties are explained in Section 2.1.2. We introduce in Section 2.1.3, Tail Value-at-Risk (TVaR), a coherent risk measure to estimate the capital requirement. In Section 2.2 we discuss the two different methods to compute a risk measure and define loss ratio from insurance perspective. Moreover, we also introduce six different types of probability distributions that are candidate distributions to be fitted to loss data in this thesis such as Log-normal, Gamma, Weibull, Log-logistic, Pareto and Burr distribution. In Section 2.3.2, we describe the procedure for selecting the best fitted distribution for a dataset. Finally, in Section 2.3.3, we explain the standard tests of randomness for testing the stationary and serial dependence of data sets.

2.1.1 Value-at-Risk (VaR)

In practice, Value-at-Risk (VaR) is the most widely used risk measure in financial and insurance industry. The regulator for insurance industry in the European Union (EU) countries implements Solvency II whereby VaR is used as a risk measure to determine the level of capital requirement. Generally, VaR can be defined as the maximum loss that should not be exceeded during a specified time horizon with a given confidence level. The loss or loss ratio in the context of insurance risk is defined in Section 2.3. Rosenberg and Schuermann (2006), Dowd and Blake (2006) and Ye and Li (2012) interpret VaR from the statistical perspective as a quantile of the distribution of portfolio returns or losses. The quantile is subsequently used to determine the level of risk. Based on 95% confident level, VaR is interpreted as the maximum loss that a company could receive on the "best" 95 days from the total of 100 days. Mathematically, VaR_α of the loss X can be written as

$$
\begin{aligned}
VaR_\alpha(X) &= \inf\{x \in \mathbb{R} : P(X > x) \leq 1 - \alpha)\}, \\
&= \inf\{x \in \mathbb{R} : P(X \leq x) \geq \alpha\}
\end{aligned}
$$

where α is the confidence level in (0,1).

Although VaR is considered a better way to measure risk or specifically downside risk, it is also has its disadvantages as explained by Ye and Li (2012). VaR is only useful in good states where a tail event is not present. In this case we may be able to know the maximum of possible losses. However, no guidance is provided on possible losses that may occur during bad states where a tail event does occur (Dowd and Blake, 2006). VaR lacks of sub-additivity and as alternative risk measure, TVaR has been introduced to overcome this issue. Numerous studies on VaR and TVaR can be found in the literature. Acerbi and Tasche (2002) introduce risk management strategy using both VaR and TVaR as risk measures. Emmer et al. (2001) and Alexander and Baptista (2004) analyse on the optimal portfolio choices with VaR and TVaR constraint in a single period. Tsai et al. (2010) extend the application of TVaR by introducing Tail Value-at-Risk Minimization (TVaRM) model to generate the optimal mix of insurance products.

We discuss the properties of a coherent risk measure and alternative risk measure to VaR in the following sections.

2.1.2 Coherent risk measures

It is important to distinguish and determine a good risk measure or in other words a coherent risk measure. We refer to Artzner et al. (1999) which provide properties of a coherent risk measure. Consider $\rho(X_i)$ to denote as risk measure for a set of possible outcomes X_i then the following are the properties of a coherent risk measure:

- **Sub-additivity**: $\rho(X_1 + X_2) \leq \rho(X_1) + \rho(X_2)$.
  Based on this property, the total risk of two added business lines is less than the total risk of two business lines added separately. This property is suitable for diversification. An insurance company with two business lines needs less capital requirement than the sum of capital requirement for the two business lines considered as separate business.

- **Monotonicity**: If $X_2 \leq X_1$, then $\rho(X_1) \leq \rho(X_2)$.
  If a business line has a better value than another business lines under all scenarios, then its risk will be less than the other business lines as well.

- **Positive homogeneity**: For all $\lambda \geq 0, \rho(\lambda X) = \lambda\rho(X)$.
  This property says that if we double the size of the position, then we are actually doubling the risk. From a business perspective, if we increase the size of a business by 50%, then we are actually increasing the business risk by 50%.

- **Translation invariant**: For all constant c, $\rho(X + c) = \rho(X) - c$.
  Translation invariant property says that adding a certain amount constant c will reduce risk by the same amount.

VaR is not sub-additive. Therefore, VaR is not a coherent risk measure. In the next section, we consider TVaR as an alternative risk measure to VaR.

2.1.3 Tail Value-at-Risk (TVaR)

We estimate the risk of the aggregate loss ratio based on the distribution of the aggregate loss X, obtained with the copula hierarchical model and C-vine copula model in Chapters 4 and 5 respectively. We use a coherent risk measure, tail value at risk (TVaR) as introduced in Acerbi and Tasche (2002). The TVaR of the loss represented by X at confidence level α, for $\alpha \in (0,1)$, is defined by

$$TVaR_\alpha(X) = \frac{1}{1-\alpha}\int_\alpha^1 VaR_u(X)\, du,$$

where the VaR_α of the loss X is given by

$$VaR_\alpha = \inf\left\{x \in \mathbb{R} : P(X \le x) \ge \alpha\right\}.$$

In risk management α typically takes values between 90% and 99%. In order to estimate TVaR we use the nonparametric estimator (see Adam et al. (2008) for details). Given n observations $\{x_1, x_2, \ldots, x_n\}$ of the variable X, the TVaR estimator is given by

$$\widehat{\mathrm{TVaR}}_\alpha = \frac{1}{n(1-\alpha)}\left(\sum_{i=1}^{\lfloor n(1-\alpha)\rfloor} x_{(n-i+1)} + (n(1-\alpha) - \lfloor n(1-\alpha)\rfloor)x_{(n-\lfloor n(1-\alpha)\rfloor)}\right),$$

where $\{x_{(1)}, x_{(2)}, \ldots, x_{(n)}\}$ is the ordered sample, $\lfloor v \rfloor$ denotes the largest integer not greater than v, and in our case $\alpha \in \{0.9, 0.95, 0.99\}$.

2.2 Bootstrap methods for computing risk measures

We estimate the risk measure from the aggregate loss ratios in our empirical research in Chapters 4 and 5 using bootstrapping method. In this section, we discuss two possible bootstrapping methods. The first method is parametric bootstrap method and the second one is non-parametric bootstrapping method. We summarize these methods in the following sections.

2.2.1 Parametric bootstrap

In the parametric bootstrap method, we use the fitted distributions to respective loss ratios and simulate according N simulations to obtain simulated losses for the risk measure.

Consider $\hat{F}$ is the empirical cdf of loss ratios, $X_1, ..., X_d$. The parametric bootstrap to estimate the risk measures, VaR and TVaR in Chapters 4 and 5 is given by the following steps

- Simulate N independent samples from $\hat{F}$,
- Estimate the risk measure such as VaR_α denoted as $\widehat{\mathrm{VaR}}_\alpha^{(i)}, i = 1, 2, ..., N$ and TVaR_α denoted as $\widehat{\mathrm{TVaR}}_\alpha^{(i)}, i = 1, 2, ..., N$
- Determine the median and confident intervals of $\widehat{\mathrm{VaR}}_\alpha^{(i)}, i = 1, 2, ..., N$ and $\widehat{\mathrm{TVaR}}_\alpha^{(i)}, i = 1, 2, ..., N$ respectively.

2.2.2 Non-parametric bootstrap

In non-parametric bootstrapping, the original loss ratios are treated as a complete dataset representing the whole general insurance industry. In this regard, this method simulates a new loss ratios data set, $X_1^*, ..., X_d^*$ by re-sampling with replacement from the original loss ratios, $X_1, ..., X_d$ according to the bootstrap distribution, given by

$$P(X^* = x_i) = \frac{1}{d}, \tag{2.1}$$

where $i = 1, 2, ..., d$ and $x_1, x_2, ..., x_d$ are elements of the drawn samples.

According to Pekasiewicz (2016), this method however does not allow to obtain a good estimation of distribution tail and therefore is unable to provide a good approximation for the quantile distribution.

2.3 Loss ratio and its distribution

In insurance industry, loss ratio has been widely used particularly by actuaries and insurance managers to measure a company profitability, products' pricing, business strategies and also capital management. A loss ratio usually expresses the relationship between insurance losses and premiums.

Generally, the loss ratio (LR) is defined as

$$LR = \frac{losses}{premiums}, \tag{2.2}$$

where the numerator is the losses from insurance claims and the denominator is the premiums which can be from earned premium or written premium. The selection of premium depends on the types of exposure. Earned premiums is a correct measure if the data is recorded based on accident year exposure while written premiums are used for policy year exposure data (Taylor, 1997).

We base our study on the variable loss ratios. For business line i and time period t, we define the *loss ratio* as

$$LR_{i,t} = \frac{IC_{i,t}}{EP_{i,t}},$$

where $IC_{i,t}$ denotes the incurred claims corresponding to the earned premium $EP_{i,t}$ based on accident year insurance company accounting principal; see Taylor (1997) for details on the loss ratio variable. The loss ratio can be seen as a measure of claims standardized by the risk exposure (given by the earned premium). Using loss ratios eliminates temporal effects of business growth and inflation, and it allows to make comparisons between business lines with different risk exposures. The loss ratios are subsequently added up to form the aggregate loss ratio used for capital requirement estimation.

The aggregate loss ratio at time t, LR_t, can then be written as a weighted sum of the loss ratios for each of the d lines of business as

$$\begin{aligned} LR_t &= \frac{IC_t}{EP_t} \\ &= \frac{\sum_{i=1}^{d} IC_{i,t}}{\sum_{i=1}^{d} EP_{i,t}} \\ &= \frac{\sum_{i=1}^{d} \left(\frac{IC_{i,t}}{EP_{i,t}} \times EP_{i,t} \right)}{\sum_{i=1}^{d} EP_{i,t}} \\ &= \sum_{i=1}^{d} \left(LR_{i,t} \times \frac{EP_{i,t}}{\sum_{i=1}^{d} EP_{i,t}} \right) \\ &= \sum_{i=1}^{d} \Big(LR_{i,t} \times w_{i,t} \Big), \end{aligned} \tag{2.3}$$

where IC_t and EP_t are the incurred claims and earned premium aggregated across all lines of business, and $w_{i,t}$ is the weight of business line i in period t. In our work we consider gross loss ratios and net (of reinsurance) loss ratios. The gross loss ratios are computed with gross claims and gross premiums. The net loss ratios are based on net claims and net premiums.

2.3.1 Probability distribution

We only consider the families of heavy-tailed distributions since this research is interested on analysing the tail part of distribution and according to the literature, insurance loss ratios are typically heavy-tailed. Heavy-tailed distributions have been widely used in the literature (Dropkin, 1964, Bickerstaff, 1972, Kleiber, 2003, Mikosch, 2009, Klugman et al., 2012). To this end, we fit the following families of distributions: Lognormal, Gamma, Weibull, Loglogistic, Pareto and Burr distribution.

Lognormal distribution

If a random variable Y follows a Log-normal distribution, then $X = \ln(Y)$ is normally distributed. Similarly, if X has a Normal distribution, then $Y = e^X$ is Log-normally distributed. The cumulative distribution function (cdf) of a Log-normal is given by

$$\begin{aligned} F(y;\mu,\sigma) &= \Phi\left(\frac{\ln(y)-\mu}{\sigma}\right) \\ &= \int_0^y \frac{1}{\sqrt{2\pi}\sigma x} e^{-\frac{1}{2}\left(\frac{\ln(x)-\mu}{\sigma}\right)^2} dx, \end{aligned} \tag{2.4}$$

where μ is the mean, σ is the standard deviation and Φ is the cdf of a standard Normal distribution with $y, \sigma > 0$ and $\mu \in \mathbb{R}$.

Gamma distribution

If a random variable X has a Gamma distribution then its cdf is given by

$$F(x;\alpha,\beta) = \int_0^x \frac{1}{\Gamma(\alpha)\beta^\alpha} y^{\alpha-1} e^{-y/\beta} dy, \tag{2.5}$$

where α is the shape parameter and β is the scale parameter with α, β and $x > 0$.

Weibull distribution

A random variable X has Weibull distribution if $X \in [0, \infty)$ has the following cdf

$$F(x;\alpha,\beta) = \begin{cases} 1 - e^{-(x/\beta)^\alpha}, & \text{if } x \geqslant 0 \\ 0 & \text{otherwise,} \end{cases} \tag{2.6}$$

where the shape parameter $\alpha > 0$ and the scale parameter $\beta > 0$. This distribution is heavy-tailed if $\alpha < 1$ since

$$\lim_{x\to\infty} \Big(\lambda x - (x/\beta)^\alpha\Big) = \infty$$

holds for every $\lambda > 0$.

Log-logistic distribution

If X has a Log-logistic distribution, we write $X \sim LL(\alpha,\beta)$, then its cdf is

$$\begin{aligned} F(x;\alpha,\beta) &= \frac{1}{1+(x/\beta)^{-\alpha}} \\ &= \frac{(x/\beta)^\alpha}{1+(x/\beta)^\alpha} \\ &= \frac{x^\alpha}{\beta^\alpha + x^\alpha}, \end{aligned} \tag{2.7}$$

where $\alpha > 0$ is the shape parameter, $\beta > 0$ is the scale parameter and $x \geq 0$.

Pareto distribution

A random variable X is said to have Pareto distribution if its cdf is given by

$$F(x;\alpha,\beta) = 1 - \left(\frac{\beta}{\beta + x}\right)^\alpha, \alpha, \beta > 0, x \geq 0, \tag{2.8}$$

where α is the shape parameter and β is the scale parameter. The parameter α controls the thickness of the tail of the distribution. A smaller α represents a heavier tail.

Burr distribution

$$F(x;\alpha,\tau,\beta) = 1 - \left(\frac{\beta^\tau}{\beta^\tau + x^\tau}\right)^\alpha, \tag{2.9}$$

where α is the shape parameters, β is the scale parameter and x, α, β, τ are positive. The Burr distribution is a generalization of Pareto distribution and does have greater flexibility than the Pareto distribution.

2.3.2 Fitting distribution

Univariate distributions are fitted by estimating its parameters using maximum likelihood. We decide between distributions based on the Anderson and Darling (1954) goodness of fit test.

Let $x = (x_1, x_2, ..., x_n)$ be a vector with n independent observations and probability density function (pdf) $f(x;\theta')$, where $\theta' = (\theta_1, \theta_2, ..., \theta_q)$ is a vector of q unknown parameters. The likelihood function, $L(\theta';x)$ is defined by

$$L(\theta';x) = \prod_{j=1}^{n} f(x_j;\theta'). \tag{2.10}$$

The maximum likelihood estimates, $\hat{\theta} = \hat{\theta}(x)$, are such that $\hat{\theta}$ is the value that maximizes the likelihood function $L(\theta';x)$.

Hypothesis testing

Consider $\boldsymbol{X}$ independent and identically distributed (i.i.d) random variables with distribution function F. Then, the hypothesis test for Kolmogorov-Smirnov (Kolmogorov, 1933 and Smirnoff, 1939) and Anderson and Darling (Anderson and Darling, 1954) is given by

$$H_0\colon F = F_0,$$

$$H_1\colon F \neq F_0,$$

where F_0 is some specified distribution function.

The Kolmogorov-Smirnov test (K-S test) measures dissimilarity between empirical cumulative distribution $F_n(x)$ of the data and the fitted cumulative distribution $F_0(x)$. The test statistics is given by

$$d_n = \sup_x |F_n(x) - F_0(x)|, \tag{2.11}$$

where n is the sample size.

The empirical cdf, $F_n(x)$ is written as

$$F_n(x) = \begin{cases} 0 & \text{for} \quad x < x_1 \\ \frac{i}{n} & \text{for} \quad x_i \leq x < x_{i+1}, \quad i = 1, 2, ..., n-1 \\ 1 & \text{for} \quad x \geq x_n, \end{cases} \tag{2.12}$$

where $x_1 \leq x_2 \leq ... \leq x_n$ are the n-sample X values arranged based on ascending order. If d_n is greater than the critical value (from K-S test table), the null hypothesis is rejected.

The Anderson-Darling test is a modified K-S test. The test statistic is given by

$$AD^2 = -n - \frac{\sum_{j=1}^{n} \Big[(2j-1)\ln(z_j) + (2n+1-2j)\ln(1-z_j)\Big]}{n}, \tag{2.13}$$

where n is the sample size and $z_j = F_n(x_j)$ is the cdf for the specified distribution. The null hypothesis is rejected if AD^2 is greater than the critical value.

2.3.3 Standard tests of randomness

We test the data in this study using standard tests of randomness to check if the data present trend, seasonality or serial dependence.

The Ljung-Box test was introduced by Ljung and Box (1978) to examine if residuals derived from a time series model exhibit white noise. We regard a time series with no serial dependence as i.i.d. The test statistic is given by

$$\hat{Q} = n(n+2)\sum_{k=1}^{m} \frac{\hat{r}_k^2}{n-k}, \tag{2.14}$$

where n is the length of the time series, m is the number of lags to test and $\hat{r}_k$ is the autocorrelation coefficient of residuals at lag k. Ljung-Box test is developed by examining the following hypotheses

$$H_0 : \hat{r}_1 = \hat{r}_2 =, ..., = \hat{r}_k = 0, \qquad H_a : \text{at least one} \quad \hat{r}_k \neq 0. \tag{2.15}$$

The hypothesis of zero dependence is rejected if

$$\hat{Q} > \chi^2_{1-\alpha,m}, \tag{2.16}$$

where $\chi^2_{1-\alpha,m}$ is the value of the chi-square distribution table with m degree of freedom and significant level α.

Chapter 3

Copulas and dependence measures

3.1 Copulas

As mentioned in the previous section, copula is very useful in modelling dependence structures. The basic idea of copula is that every joint distribution from a random vector can be translated into its marginal distribution and the dependence structure can be described by its copula. By now, the theory of copula is well established. In fact, for reference to copula, Joe (1997), Nelsen (2006) and McNeil et al. (2015) provide excellent textbook references to understand the theoretical foundations of copula. Generally, a n-dimensional copula is a multivariate distribution function on $[0,1]^n$ with standard uniform univariate marginal distribution.

Following copula definition by McNeil et al. (2015), we define n-dimensional copula according to its properties in Definition 3.1.

Definition 3.1. *A copula with n-dimensional is a function* $C : [0,1]^n \rightarrow [0,1]$ *that satisfies the following properties*

- ***Groundedness:*** $C(u_1, ..., u_n) = 0$ *if* $u_i = 0$ *for any* i
- ***Normalize margin:*** $C(1, ..., 1, u_i, 1, ..., 1) = u_i$ *for all* $i \in 1, ..., n, u_i \in [0,1]$. *This means the marginal distributions need to be in uniform scale.*

- ***n-increasingness:*** *For all* $(a_1, ..., a_n), (b_1, ..., b_n) \in [0, 1]^n$ *with* $a_i \leq b_i$ *we have*
$$\sum_{i_1=1}^{2} ... \sum_{i_n=1}^{2} (-1)^{i_1+...+i_n} C(u_{1i_1}, ..., u_{ni_n}) \leq 0,$$
where $u_{j1} = a_j$ *and* $u_{j2} = b_j$ *for all* $j \in 1, ..., n$.

Copula is useful to measure dependence and its role is explained in theorem of Sklar (1959). It explains the link between multivariate distribution function and its univariate margins.

Theorem 3.2. (Sklar's theorem). *Let* F *be a joint distribution function with univariate margins* $F_1, ..., F_n$. *Then there exists a function* $C : [0, 1]^n \to [0, 1]$ *such that*

$$F(x_1, ..., x_n) = \mathbb{P}(X_1 \leq x_1, ..., X_n \leq x_n) = C(F_1(x_1), ..., F_n(x_n)). \tag{3.1}$$

If the margins $F_1, ..., F_n$ *are continuous, then* C *is unique. Conversely, if* C *is a copula with univariate distribution functions* $F_1, ..., F_n$, *then* $F(x_1, ..., x_n)$ *defined in (3.1) is a joint distribution function with univariate margins* $F_1, ..., F_n$.

Theorem 3.2 provides method for constructing a copula based on inverse distribution function. Consider a multivariate distribution function F and inverse distribution functions $F_1^{-1}, ..., F_n^{-1}$, a copula C is given by

$$C(u_1, ..., u_n) = F\Big(F_1^{-1}(u_1), ..., F_n^{-1}(u_n)\Big). \tag{3.2}$$

The construction of Gaussian and t-copula are primarily based on Sklar's Theorem 3.2 and (3.2).

In this thesis, we use copula to model the dependence structure of insurance loss ratios. In particular, copula function is used to determine the marginal distributions of each loss ratio and the dependence structure between loss ratios. To obtain the joint distribution function of two loss ratios X and Y, these loss ratios are transformed into random variables U_1 and U_2 respectively that is uniformly distributed on $[0, 1]$ by using the corresponding marginal distribution F_X and F_Y

as given by the following

$$
\begin{aligned}
U_1 &= F_X(X) \\
U_2 &= F_Y(Y).
\end{aligned} \tag{3.3}
$$

Next, the multivariate distribution function can be derived by replacing (3.3) into the copula function:

$$
\begin{aligned}
F_{XY}(x, y) &= P(X \leq x, Y \leq y) \\
&= P(F_X^{-1}(U_1) \leq x, F_Y^{-1}(U_2) \leq y) \\
&= P(U_1 \leq F_X(x), U_2 \leq F_Y(y)) \\
&= C(F_X(x), F_Y(y)),
\end{aligned} \tag{3.4}
$$

where F_X^{-1} and F_Y^{-1} are the inverse functions of the univariate marginals F_X and F_Y respectively and this is hold when function F is continuous and strictly increasing.

Clearly (3.4) proves that from Sklar theorem, a bivariate join distribution can be modelled by marginal distributions of each loss ratio F_X, F_Y and a copula C. The Sklar's theorem in (3.1) can easily be extended to general case and its proof can be derived following (3.3) and (3.4).

3.1.1 Copula density

If the joint cumulative distribution function $F(x_1, ..., x_n)$ is absolutely continuous and its marginal distribution functions $F_1, ..., F_n$ are strictly increasing and continuous, then using (3.1), the copula density can be written as

$$
\begin{aligned}
c(F_1(x_1), ..., F_n(x_n)) &= \frac{\partial^n C\big(F_1(x_1), ..., F_n(x_n)\big)}{\partial F_1(x_1), ..., \partial F_n(x_n)} \\
&= \frac{\partial^n F(x_1, ..., x_n)}{\partial F_1(x_1), ..., \partial F_n(x_n)} \\
&= \frac{f(x_1, ..., x_n)}{\prod_{j=1}^{n} f_j(x_j)}.
\end{aligned} \tag{3.5}
$$

Finally, we can re-write (3.5) as

$$f(x_1, ..., x_n) = \prod_{j=1}^{n} f_j(x_j) c\big(F_1(x_1), ..., F_n(x_n)\big). \tag{3.6}$$

3.1.2 Invariance properties

Another important property of copula is invariance under strictly monotone transformation. We state this property in the following proposition.

Proposition 3.3. *Let $\boldsymbol{X} = (X_1, ..., X_d)$ be a random vector with continuous univariate marginal distributions and copula C. Then C is invariant under strictly, monotone increasing transformations of the component $\boldsymbol{X}$.*

3.2 Relevant families of copulas

In this section, we review five copulas from the family of Elliptical copulas and Archimedean copulas. More precisely, Gaussian and t-copula from the family of Elliptical copulas and Clyton, Gumbel and Frank copulas from the family of Archimedean copulas. For other types of copulas see Joe (1997). In this thesis, we primarily focus on bivariate copulas to address issues on modelling high dimensional copulas. Hence, the following explanations are based on bivariate copulas.

3.2.1 Elliptical copulas

Elliptical copulas have been widely studied in finance and insurance. As discussed in Section 3.1, this copula is derived from multivariate distribution function using Skalar's theorem. Although elliptical copula has been widely used in finance and insurance, however Embrechts et al. (2003) highlighted that it has certain drawbacks such as its inability to capture different dependence structures and are restricted to radial symmetry. Especially in extreme events, the dependence of big losses from different lines of business can not be modelled by elliptical copula. The examples of Elliptical copulas are Gaussian and t-copula.

Example 3.1. (Gaussian Copula)

Gaussian copula is derived from Sklar theorem. We get bivariate Gaussian copula with parameter $\rho \in (-1, 1)$, by applying Sklar theorem to the bivarate standard normal cdf. The Gaussian copula is given by

$$C(u_1, u_2) = \Phi_2\big(\Phi^{-1}(u_1), \Phi^{-1}(u_2), \rho\big), \tag{3.7}$$

where Φ_2 is the bivariate distribution with two random variables normally (standard) distributed, ρ is the dependence and Φ^{-1} is the inverse function of Φ of standard normal distribution. The density of bivariate Gaussian copula is written as

$$c(u_1, u_2) = \frac{1}{\sqrt{1-\rho^2}} \exp\left[-\frac{\rho^2(x_1^2 + x_2^2) - 2\rho x_1 x_2}{2(1-\rho^2)} \right], \tag{3.8}$$

where $x_1 = \Phi^{-1}(u_1)$ and $x_2 = \Phi^{-1}(u_2)$. Φ^{-1} is the inverse function of the standard normal distribution.

Example 3.2. (t- copula)

t-copula is also constructed by Sklar theorem. The distribution function is given by

$$C(u_1, u_2) = t_{\theta,v}(t_v^{-1}(u_1), t_v^{-1}(u_2)), \tag{3.9}$$

where $t_{\theta,v}$ denotes the distribution function of the bivariate t-distribution with parameters $\theta \in (-1, 1)$ and degree of freedom, $v > 0$. t_v^{-1} represents the inverse distribution function of a standard univariate t-distribution with v degree of freedom. Then, the density is defined as

$$c(u_1, u_2) = \frac{\Gamma(\frac{v+2}{2})/\Gamma(\frac{v}{2})}{v\pi dt_v(x_1) dt_v(x_2)\sqrt{1-\theta^2}} \left[1 + \frac{x_1^2 + x_2^2 - 2\theta x_1 x_2}{v(1-\theta^2)} \right]^{-\frac{v+2}{2}}, \tag{3.10}$$

where $x_1 = t_v^{-1}(u_1)$ and $x_2 = t_v^{-1}(u_2)$. dt_v is the density of a standard univariate t-distribution with v degree of freedom.

Figure 3.1 presents the scatter plots (left) and density plots (right) for Gaussian and t-copula.

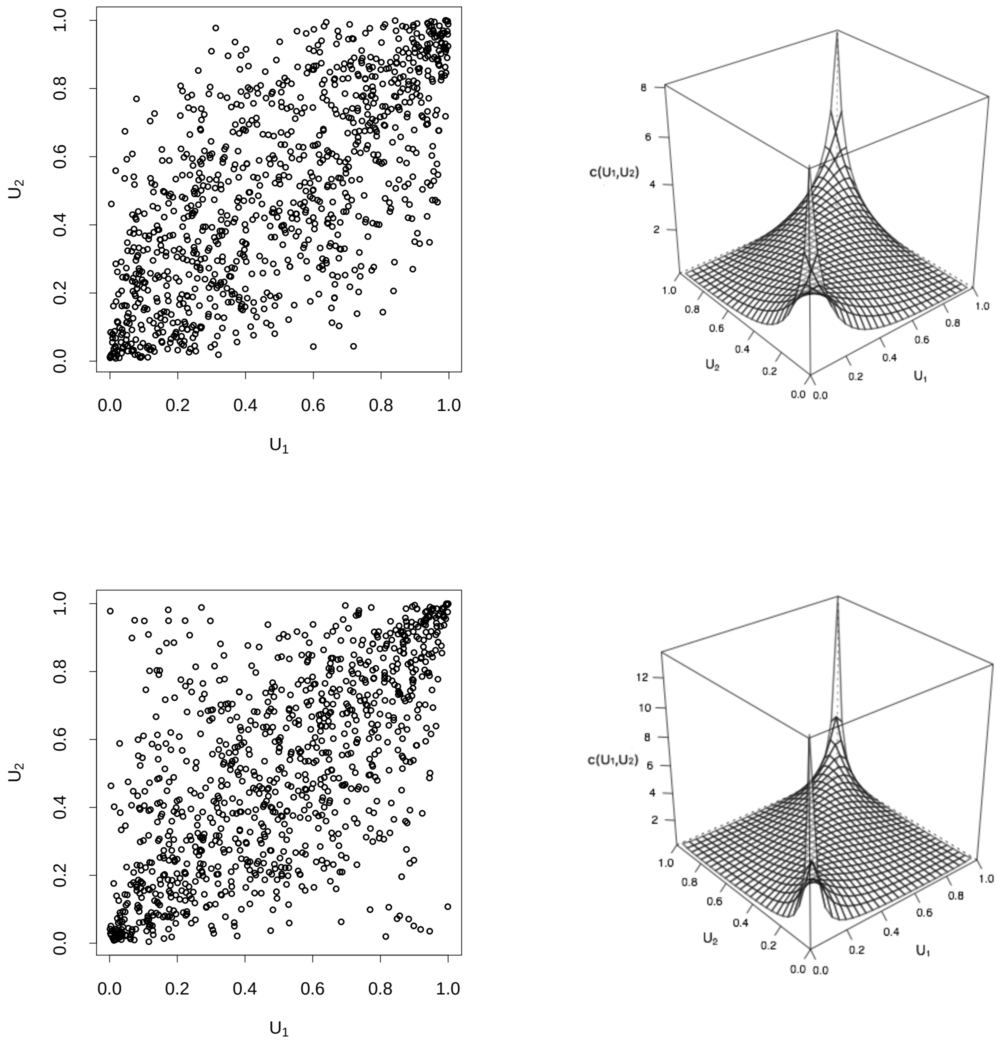


FIGURE 3.1: Scatter plots of Gaussian copula (top-left) and t-copula (bottom-left). The top-right is the density plot for Gaussian copula and at the bottom-right is the density plot for t-copula.

3.2.2 Archimedean copulas

Another popular copula family is called Archimedean copulas. A bivariate Archimedean copula is given by

$$C(u_1, u_2) = \varphi^{-1}\big(\varphi(u_1) + \varphi(u_2)\big), \tag{3.11}$$

where function $\varphi : [0,1] \to [0,\infty]$ is the generator and has to be continuous and strictly monotonic decreasing convex with $\varphi(0) = 1$ and $\lim_{d\to\infty} \varphi(d) = 0$. The pseudo-inverse φ^{-1} is defined as

$$\varphi^{-1}(u) = \begin{cases} \varphi^{-1}(u) & \text{for} \quad 0 \leq u \leq \varphi(0) \\ 0 & \text{for} \quad \varphi(0) < u < \infty. \end{cases} \tag{3.12}$$

Archimedean copula family is asymmetric copula. It is capable to capture variety of different dependence structures. In other words, this copula is suitable for modelling risk of extreme events such as catastrophe and earth quakes which incur big losses for insurance companies (Embrechts et al. (2003)).

The following are three examples of the most commonly used copulas from the Archimedean family.

Example 3.3. (Clayton copula)

Clayton copula is also known as Cook-Johnson copula (Cook and Johnson, 1981) and the generator can be defined by

$$\varphi(u) = \frac{1}{\theta}(u^{-\theta} - 1). \tag{3.13}$$

Then the Clayton copula can be described by the following

$$C(u_1, ..., u_n) = \left[\sum_{i=1}^{n} u_i^{-\theta} - n + 1\right]^{-1/\theta}, (u_1, ..., u_n) \in [0,1]^n, \tag{3.14}$$

and by considering bivariate copula, the above can be re-defined as

$$C(u_1, u_2) = \left[u_1^{-\theta} + u_2^{-\theta} - 1\right]^{-1/\theta}. \tag{3.15}$$

Another important element of Clayton copula is the dependence which is present heavier in the lower tail than the upper tail and suitable for modelling negative dependence. The dependence structure can be visualised from scatter plot and density plot in Figure 3.2.

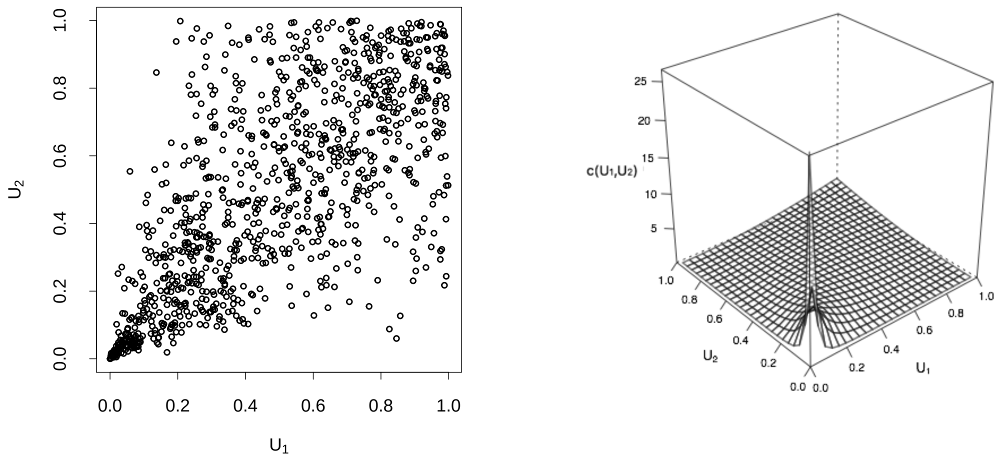


FIGURE 3.2: Scatter plot (left) and the density plot (right) of bivariate Clayton copula with Kendall's tau =0.5 and number of simulations, N=1000.

Example 3.4. (Gumbel Copula)

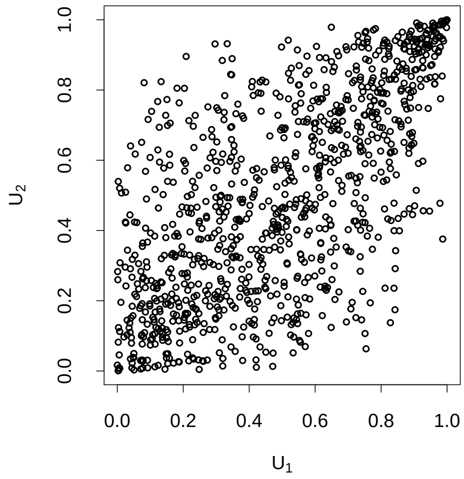
Gumbel copula or also known as Gumbel-Hougaard copula (Hutchinson and Lai, 1990) is another type in Archimedean copula family. Unlike Clayton copula, Gumbel copula has greater dependence in the upper tail than the lower tail. The generator for Gumbel copula is given by

$$\varphi(u) = (-\ln u)^{\theta}, \tag{3.16}$$

for $\theta \geq 1$ Then, Gumbel copula can be described as

$$\begin{aligned} C(u_1, u_2) &= \varphi^{-1}\Big(\varphi(u_1) + \varphi(u_2)\Big) \\ &= \exp\Bigg(- \Big[(-\ln u_1)^{\theta} + (-\ln u_2)^{\theta}\Big]^{1/\theta}\Bigg). \end{aligned} \tag{3.17}$$

Figure 3.3 shows the scatter plot of Gumbel copula on the left and its density plot on the right. We can also observe heavy upper tail.



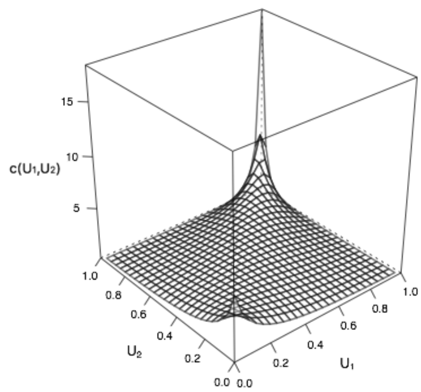


FIGURE 3.3: Scatter plot (left) and the density plot (right) of bivariate Gumbel copula with Kendall's tau =0.5 and number of simulations, N=1000.

Example 3.5. (Frank Copula)

The third type of Archimedean copula family is Frank copula. This copula exhibits symmetric dependence structure and can be seen in Figure 3.4. Its generator is given by:

$$\varphi(u) = -\ln\left(\frac{e^{-\theta u} - 1}{e^{-\theta} - 1}\right). \tag{3.18}$$

Hence, the Frank copula with n-dimension can be expressed as

$$C(u_1, ..., u_n) = -\frac{1}{\theta}\ln\left[1 + \frac{\prod_{i=1}^{n}(e^{-\theta u_i} - 1)}{(e^{-\theta} - 1)^{n-1}}\right], \tag{3.19}$$

and for bivariate Frank copula

$$C(u_1, u_2) = -\frac{1}{\theta}\ln\left[1 + \frac{(e^{-\theta u_1} - 1)(e^{-\theta u_2} - 1)}{e^{-\theta} - 1}\right], \tag{3.20}$$

where $\theta \in (-\infty, \infty)\setminus\{0\}$. If $\theta \to \infty$, the Frank copula presents completely positive dependence and if $\theta \to$-∞, it exhibits negative dependence.

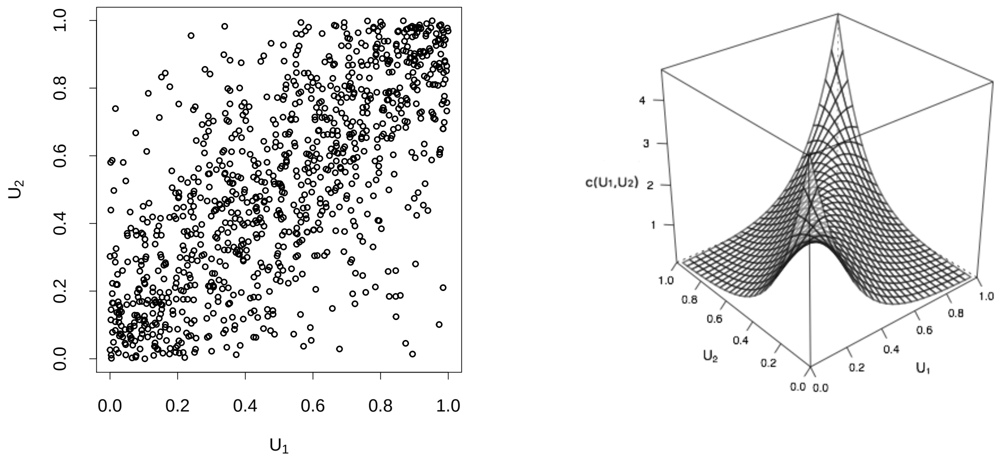


FIGURE 3.4: Scatter plot (left) and density plot (right) of bivariate Frank copula with Kendall's tau =0.5 and number of simulations, N=1000.

Copula	Parameter range	λ_L	λ_U	Kendall's tau
Clayton	$\theta \leqslant -1$	$2^{\frac{-1}{\theta}}$	0	$\frac{\theta}{\theta+2}$
Gumbel	$\theta \leqslant 1$	0	2-$2^{\frac{1}{\theta}}$	1-$\frac{1}{\theta}$
Frank	$\theta \in (-\infty, \infty) \setminus \{0\}$	0	0	1-4 θ+4D$_1(\theta)$ θ

TABLE 3.1: Copula properties for Clayton, Gumbel and Frank copula. λ_L and λ_U are the lower and upper tail dependence respectively. The Debey function is used to estimate the Kendall's tau for Frank copula. Debey function $D_1(\theta) = \int_0^\theta \frac{x/\theta}{exp(x)-1}dx$

.

The summary of copula properties include the tail dependence and Kendall's tau for Clayton, Gumbel and Frank copula are given in Table 3.1.

3.2.3 Survival copula

Another important type of copula is survival copula. We discuss the properties of survival copula in the following proposition.

Proposition 3.4. *Given $\boldsymbol{X} = (X_1, ..., X_d)$ random vector with distribution function F and marginal distribution function $F_1, ..., F_d$. Then, there exist a survival copula $\widehat{C}$ such that*

$$\begin{aligned}
\bar{F}(x_1, ..., x_d) &= \widehat{C}(\bar{F}(x_1), ..., \bar{F}_d(x_d)), \\
&= P(X_1 > x_1, ..., X_d > x_d) \\
&= P(1 - F_1(X_1) \le \bar{F}_1(x_1), ..., 1 - F_d(X_d) \le \bar{F}_d(x_d),
\end{aligned}$$

where $\bar{F}_i = 1 - F_i, i = 1, ..., d$. Generally, survival copula $\widehat{C}$ of a copula C with random vector $\boldsymbol{U} = (U_1, ..., U_d)$ is used to denotes the distribution function of $\mathbf{1} - \boldsymbol{U}$ where $\boldsymbol{U} := (F_1(X_1), ..., F_d(x_d))$.

Note that copulas are simply multivariate distribution functions. In particular, copulas have survival functions $\bar{C}$. If $\boldsymbol{U}$ has distribution function C and the survival copula of C is $\widehat{C}$, then the relationship between a survival copula $\widehat{C}$ and a survival function $\bar{C}$ is given by,

$$\begin{aligned}
\bar{C}(u_1, ..., u_d) &= P(U_1 > u_1, ..., U_d > u_d) \\
&= P(1 - U_1 \le 1 - u_1, ..., 1 - U_d \le 1 - u_d) \\
&= \widehat{C}(1 - u_1, ..., 1 - u_d).
\end{aligned}$$

Similarly, the relationship between a copula and its survival copula in the bivariate case can be expressed as

$$\widehat{C}(1 - u_1, 1 - u_2) = 1 - u_1 - u_2 + C(u_1, u_2)$$

3.3 Copula inference

3.3.1 Method-of-moments estimator (MME)

The estimation of copulas can be based on an empirical rank correlation coefficient, which are scalar dependence measures. This method is similar to the well-known method of moments in the sense that the method of moments equates the empirical moments to the parameters of the distribution while the rank correlation estimator equates the empirical rank correlation, Kendall's τ or Spearman's ρ, to the parameter of the copula. Given that these two rank correlation measures do not depend on the distribution of the margins but only on the distribution of the copula, this estimation method does not require modelling the margins in order to fit the copula model. This method is considered a simpler method than the maximum likelihood estimator (MLE).

3.3.2 Maximum likelihood estimator (MLE)

Consider a random sample $(x_1, y_1), (x_2, y_2), ..., (x_d, y_d)$ from a bivariate distribution F with marginal distributions F_X and F_Y and copula C. Then, assuming f_X, f_Y, and c are the density for marginal distributions and copula respectively, we derive the joint density $(x_t, y_t), t = 1, ..., d$ by taking the partial derivatives with respect to x_t and y_t and is given by

$$\begin{aligned} \frac{\partial^2}{\partial x_t \partial y_t} F(x_t, y_t; \theta_x, \theta_y, \delta) &= f(x_t, y_t; \theta_x, \theta_y, \delta) \\ &= \frac{\partial^2}{\partial x_t \partial y_t} C(F_X(x_t; \theta_x), F_Y(y_t; \theta_y); \delta) \\ &= c\Big(F_X(x_t; \theta_x), F_Y(y_t; \theta_y); \delta\Big) f_X(x_t; \theta_x) f_Y(y_t; \theta_y), \end{aligned} \tag{3.21}$$

where δ is the parameter of the copula and θ_x and θ_y are the parameters of the marginal distributions F_X and F_Y respectively. The log-likelihood function is given by

$$L(\delta; x, y) = \sum_{t=1}^{d} \left(\ln c\Big(F_X(x_t; \theta_x), F_Y(y_t; \theta_y); \delta\Big) + \ln f_X(x_t; \theta_x) + \ln f_Y(y_t; \theta_y) \right). \tag{3.22}$$

Then, the MLE is written as

$$\hat{\delta}_{MLE} = \arg\max_{\delta} L(\delta; x, y). \tag{3.23}$$

Based on MLE, the parameters for marginal distributions θ_x and θ_y are estimated jointly with the copula parameter δ. In particular, the MLE estimation is performed by one-stage. However, this method is computationally challenging for high dimensional copula. To solve this issue, one may consider the inference for margin (IFM). We discuss this method in the following section.

3.3.3 Inference function for margin (IFM)

Based on this method, the estimation of parametric copulas is derived in two steps. The first step is performed by estimating marginal distribution $F_1, ..., F_d$ and denote as $\hat{F}_1, ..., \hat{F}_d$. Then, given estimates $\hat{F}_1, ..., \hat{F}_d$, we construct the pseudo-observation $\hat{U}_t$ consist of the vectors $\hat{U}_1, ..., \hat{U}_n$ from the original data vectors $X_1, ..., X_n$ where

$$\hat{U}_t = (U_{t,1}, ..., U_{t,d})' = \left(\hat{F}_1(X_{t,1}), ..., \hat{F}_d(X_{t,d})\right)', t = 1, ..., n. \tag{3.24}$$

In the second step, consider C_θ a parametric copula where θ is the parameter vector to be estimated by the MLE, that is the maximizer of the log-likelihood function

$$L(\boldsymbol{\theta}; \hat{U}_1, ..., \hat{U}_n) = \sum_{t=1}^{n} \ln c_\theta(\hat{U}_t; \boldsymbol{\theta}), \tag{3.25}$$

with respect to $\boldsymbol{\theta}$ where c_θ is the copula density and $\hat{U}_t$ is the pseudo-observation of the copula.

3.4 Dependence measures

Dependence measures are used in financial industry to quantitatively evaluate the dependence structure across risk distributions or loss distributions for insurance case. There are three most commonly used dependence measures: Linear dependence or also known as Pearson correlation, rank dependence, and the coefficient of tail dependence. Both rank dependence and the coefficient of tail dependence are non-linear dependence. Rank dependence includes Spearman's ρ and Kendall's τ. We discuss the concept of these dependence measures in the following subsection.

3.4.1 Linear dependence

Linear dependence is a well known theory and widely used to measure dependences. In fact, the Solvency II propose to use linear dependence to determine the capital requirement. However, it has certain pitfalls and only true for multivariate normal or elliptical model. In other words, this measure can be applied to bivariate normal random variables X_1 and X_2. Linear dependence can be defined as

Definition 3.5. Given two random variables X_1 and X_2 with $\mathbb{E}(X_j^2) < \infty$ and $j \in \{1, 2\}$, the Pearson's linear correlation is defined by

$$\begin{aligned} \rho_{12} &= \frac{Cov(X_1, X_2)}{\sqrt{Var(X_1)}\sqrt{Var(X_2)}} \\ &= \frac{\mathbb{E}\big[\big(X_1 - \mathbb{E}(X_1)\big)\big(X_2 - \mathbb{E}(X_2)\big)\big]}{\sqrt{\mathbb{E}\big(X_1 - \mathbb{E}(X_1)\big)^2}\sqrt{\mathbb{E}\big(X_2 - \mathbb{E}(X_2)\big)^2}}. \end{aligned} \tag{3.26}$$

3.4.2 Non-linear dependence

Non-linear dependence is suitable to measure non-linear dependence structure. We discuss three types of non-linear dependence measures in the following definitions:

Spearman's ρ

Definition 3.6. Spearman's ρ measures the association of the ranks between two random variables X_1 and X_2 and its corresponding marginals F_1 and F_2 respectively. It is defined as

$$\rho_s = \rho\big(F_1(X_1), F_2(X_2)\big), \tag{3.27}$$

where ρ is the linear correlation. In the multivariate random vector X case, the Spearman's ρ matrix is given by

$$\rho_s(X) = \rho\big(F_1(X_1), ..., F_n(X_n)\big), \tag{3.28}$$

where ρ refers to the correlation matrix.

Kendall's τ

Definition 3.7. Kendall's τ is also a measure of dependence of the ranks between two random variables X_1 and X_2. Now, let consider two additional random variables $\widetilde{X}_1$ and $\widetilde{X}_2$ for comparison. Both of the additional variables have the same joint distribution but independent from the previous variables. Thus, using expectation, we define Kendall's τ by

$$\rho_\tau(X_1, X_2) = E\Big(sign\big((X_1 - \widetilde{X}_1)(X_2 - \widetilde{X}_2)\big)\Big), \tag{3.29}$$

where sign(x)=$\mathbf{1}_{(0,\infty)}(x) - \mathbf{1}_{(-\infty,0)}(x)$.

Tail dependence

Tail dependence is also known as a fat-tail which is very useful in measuring and describing a copula. It is important to distinguish between upper (right upper corner) and lower (lower left corner) tail dependence.

According to Denuit et al. (2006), considering two random variables X_i and X_j the concordance between the two extreme events of the random variables is measured by tail dependence. More precisely, the concern is on the probability that loss for X_i exceeds a limit d given that loss for X_j has exceeded the same level. On the other hand, if the level recorded is lower than the limit given, then tail dependence does not present. The coefficient of upper tail dependence can be defined by

$$\lambda_u := \lambda_u(X_i, X_j) = \lim_{d\to 1} P(X_j > F_j^{-1}(d) \mid X_i > F_i^{-1}(d)), \tag{3.30}$$

provided that the limit $\lambda_u \in [0, 1]$ exists. Analogously, the coefficient of lower tail dependence can be defined by

$$\lambda_l := \lambda_l(X_i, X_j) = \lim_{d\to 0} P(X_j \leq F_j^{-1}(d) \mid X_i \leq F_i^{-1}(d)), \tag{3.31}$$

provided that the limit $\lambda_l \in [0, 1]$ exists.

If both F_i and F_j are continuous dfs, simple expressions are derived for the coefficient in terms of copula C using Bayes' rule and we have

$$\begin{aligned} \lambda_l &= \lim_{d\to 0} \frac{P\big(X_j \leq F_j^{-1}(d), X_i \leq F_i^{-1}(d)\big)}{P(X_i \leq F_i^{-1}(d))} \\ &= \lim_{d\to 0} \frac{C(d,d)}{d}, \end{aligned} \tag{3.32}$$

and similarly for upper tail dependence, we have

$$\lambda_u = \lim_{d \to 1} \frac{C(d,d) - 2d + 1}{1 - d}. \tag{3.33}$$

3.4.3 A simple motivating example

Having the concept and theory of dependence as explained in the Section 3.4.1 and 3.4.2 in mind, we now consider the following examples as introduced by Hofert. et al. (2017) to further understand the importance of dependence.

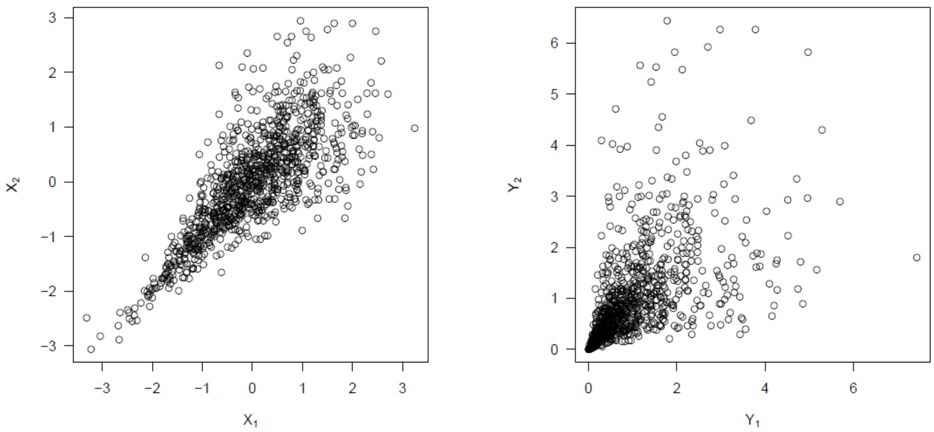


FIGURE 3.5: Scatter plots of random variables X_1 and X_2 on the left side and Y_1 and Y_2 on the right side.

In Figure 3.5, the left panel shows the scatter plot of dependence between bivariate random vectors (X_1, X_2), while the right figure shows the dependence between bivariate random vectors (Y_1, Y_2). We can clearly distinguish that these scatter plots show different dependence structures and values. However, this might not be the case for this example. To further analyse the dependence between these scatter plots, consider to analyse the empirical marginal distribution for each dataset.

Figure 3.6 illustrates the estimated marginal distribution functions for bivariate random vectors (X_1, X_2) on the left and (Y_1, Y_2) on the right. Bivariate random vectors (X_1, X_2) is best fitted by standard normal distribution $N(0, 1)$ while (Y_1, Y_2) is best fitted by standard exponential distribution $Exp(1)$. At this stage, we conclude that the two datasets in Figure 3.5 differ by the marginal distribution.

However, the comparison in terms of dependence between the two datasets can be done on fairer grounds. Given the marginal distributions of the underlying vectors (X_1, X_2) and (Y_1, Y_2), we transform them to a similar underlying marginals known as standard uniform distribution using probability integral transformation (PIT). We define the PIT in the following.

Definition 3.8. (Probability integral transformation)

Let F be a cumulative distribution function and $X \sim F$. Then $F(X)$ is a standard uniform random variable, satisfying $F(X) \sim U(0,1)$.

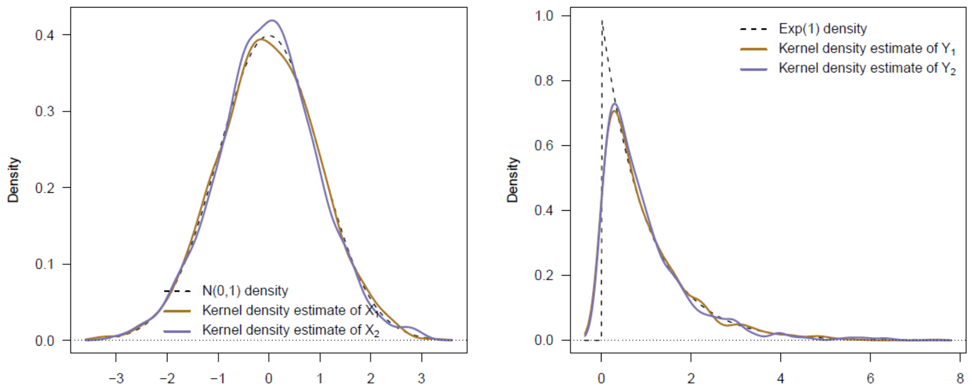


FIGURE 3.6: (Left) Empirical standard normal distribution, $N(0,1)$ with Kernel density estimates of X_1 and X_2. (Right) Standard exponential distribution, $Exp(1)$ with Kernel density estimates of Y_1 and Y_2.

In Figure 3.7, similar dependence structure can be observed in both datasets. To this end, we confirm that the dependence structures for both datasets are identical and only differ by its corresponding marginal.

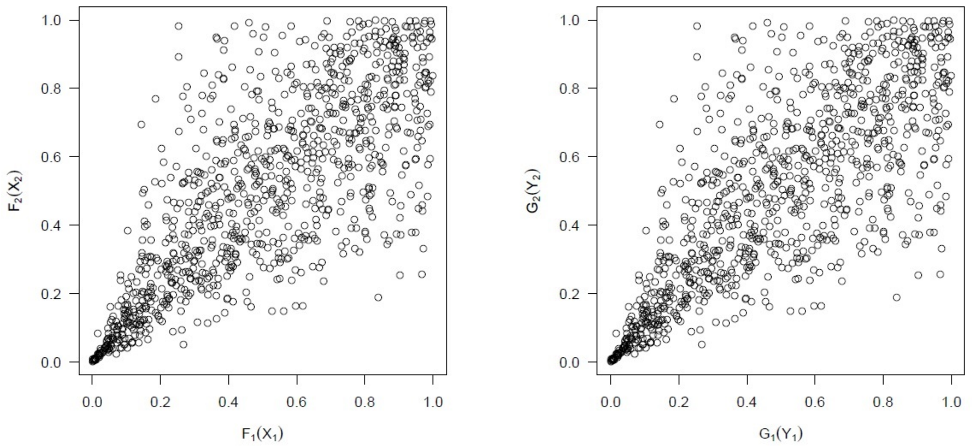


FIGURE 3.7: Scatter plots of bivariate random vectors $(F_1(X_1), F_2(X_2))$ on the left and $(G_1(Y_1), G_2(Y_2))$ on the right.

3.5 Application of graph theory

In this section, we introduce the fundamentals of graph theory which is essential for constructing hierarchical aggregation models in Chapter 4 and C-vine copula models in Chapter 5 of this thesis. In particular, we use notation from graph theory such as nodes, edges and trees to represent random variables, bivariate copulas and aggregation levels respectively.

Arbenz et al. (2012) and Côté and Genest (2015) use graph theory to construct hierarchical aggregation tree structures. Similarly, Dißmann et al. (2013) and Kraus and Czado (2017) use graph theory to develop structures of vine copula by representing random variables and copulas with nodes and edges respectively.

The aggregation modelling using copula utilises the application of graph theory. In this section, following Diestel (2017), we explain the concepts of graph theory important for modelling aggregation of risks.

3.5.1 Graph, nodes and edges

Definition 3.9. A graph G is a pair of non-empty set $N(G)$ of vertices (or nodes) and a set $E(G)$ of unordered pairs of edges. For example, a graph can be written as $G = (N, E)$ where

$N = \{x_1, x_2, x_3, x_4, x_5\}$,
$E = \{(x_1, x_2), (x_1, x_3), (x_3, x_4), (x_3, x_5), (x_2, x_5)\}$.

To futher understand the above definition, we now consider the diagram in Figure 3.8.

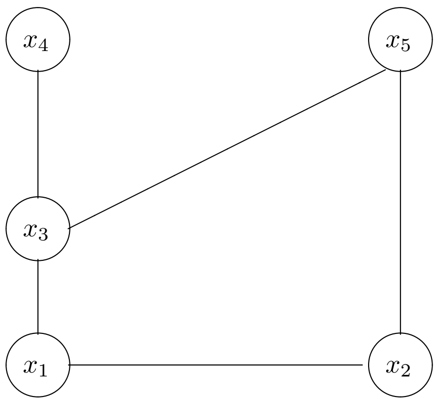


FIGURE 3.8: Graph G=(N(G),E(G)) consists of 5 nodes $N = \{x_1, x_2, x_3, x_4, x_5\}$ and 5 edges $(x_1, x_2), (x_2, x_5), (x_1, x_3), (x_3, x_4)$ and (x_3, x_5).

Each point is represented by x_1, x_2, x_3, x_4, x_5 and is called a node and each line linking between 2 nodes is called an edge. Figure 3.8 consist of 5 nodes with 5 edges.

3.5.2 Cycle, trees and forest

Definition 3.10. A cycle in a Graph, $G = \big(N(G), E(G)\big)$ is a sequence of nodes

$$x_1, x_2, x_3 ..., x_k,$$

and edges

$$\{x_1, x_2\}, \{x_2, x_3\}, ..., \{x_{k-1}, x_k\},$$

where x_1 is the start node and the same as the end node, $x_k (x_1 = x_k)$ and $\{x_i, x_{i+1}\}$ is an edge of G for all i where $1 \leq i \leq k$.

Definition 3.11. A graph without any cycles is called acyclic graph and a forest is an acyclic graph. A connected forest is called a tree. In other words, a forest can be defined as a graph whose components are trees.

Definition 3.12. If each of the connected components of a graph G is a tree, then G is a forest.

Example

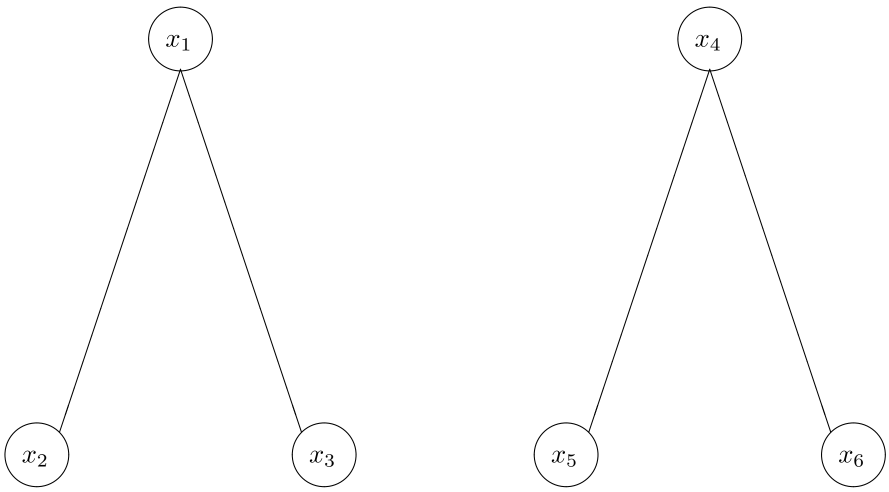


FIGURE 3.9: A forest consists of 2 component trees with 6 nodes $N = \{x_1, x_2, x_3, x_4, x_5, x_6\}$. This graph is not connected and therefore do not considered as a tree.

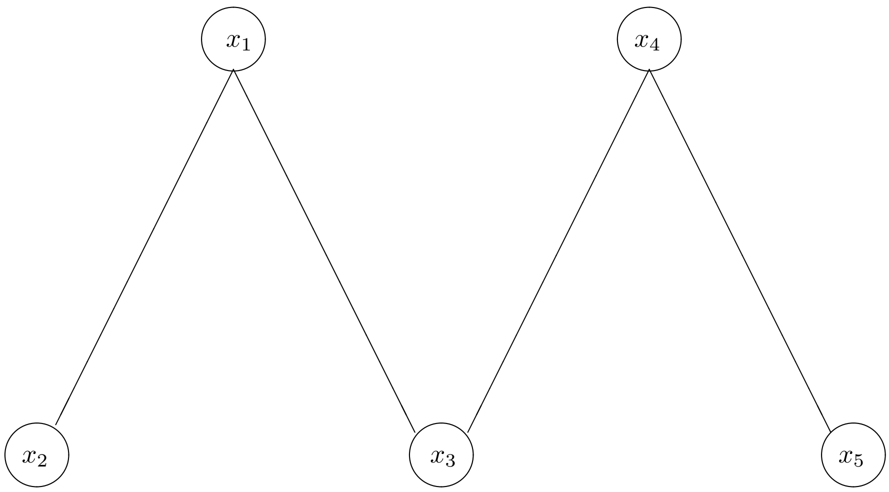


FIGURE 3.10: A tree connecting 5 nodes, $N = \{x_1, x_2, x_3, x_4, x_5\}$.

In Figure 3.9 we can observe two components of a forest. Each component is also known as a tree but the 6-nodes graph (whole graph) is not a tree since the two components (trees) are not connected. To further understand the concept of a tree, consider another 6-nodes graph presented in Figure 3.10. All 6 nodes are connected by edges and therefore the graph itself is a tree.

Chapter 4

Copula-based hierarchical aggregation

4.1 Introduction

Determining the level of capital required for business continuity is essential for insurance companies. This capital requirement can help an insurance company to minimize the risk of insolvency and to serve its obligation to the policyholders. In extreme events, such as floods, earthquakes, hurricanes and other catastrophic events, the number of claims to be paid by an insurance company can be extremely high even though part of the claims can be passed to reinsurance companies. In some cases these large number of claims can originate not just from one business line but involve other products as well. In other words, some insurance business lines are dependent on each other, in the sense that an increase on the number of claims being filled in one business line is accompanied by a higher number of claims in other business lines too. Hence, there is a need to properly model the aggregate risk of losses across a broad range of insurance products.

Aggregating the risk of losses for insurance companies is challenging where the most crucial aspect of the aggregation process is modelling the dependence structure between the risks of losses across different business lines. Linear correlation is the classic approach to model risk dependence but fails to incorporate all possible dependence structures. The appropriate method to model the dependence structure is using copulas, which have received increasing interest from researchers and practitioners in recent years.

This chapter is twofold. First we focus on modelling the aggregation of risks of an insurance company. Second, we explore the effect of reinsurance on the level of risk and how this relates with the dependence structure between the different business lines. As a technique for aggregating the risks of an insurance company, this chapter focusses on an hierarchical risk aggregation method which is based on two dimensional copulas.

The hierarchical risk aggregation approach has been recently adopted by Côté and Genest (2015) and was previously developed by Arbenz et al. (2012) and references therein. The hierarchical aggregation procedure, developed by Arbenz et al. (2012), is based on rooted trees that include branching and leaf nodes, and uses the elliptical copula family for each aggregation step. However, as highlighted, for instance, in Embrechts et al. (2003), this copula family has certain drawbacks, such as its inability to capture dependence structures, which are not radially symmetric. Especially in the case of extreme events, the dependence of large losses from different business lines cannot be modelled by the elliptical copula family (see Nguyen and Molinari, 2011). To overcome this problem, we propose to use copulas from the Achimedean family in the construction of the hierarchical model. Archimedean copulas can be asymmetric and capture a variety of dependence structures. We also include the mixture of rotated Archimedean copulas, which are the most appropriate copulas in some cases, based on the goodness of fit tests.

In this chapter, we use data from the Australian Prudential Regulation Authoruty (APRA) as in Tang and Valdez (2006). Tang and Valdez (2006) analyse 19 semi-annual gross incurred claims and earned premiums data from December 1992 to June 2002. In contrast, we choose a more recent time horizon and quarterly frequency in order to increase the sample size and improve the estimation of the risk aggregation model. As a result, a total of 28 observations, consisting of quarterly premium earned and incurred claims, gross and net of reinsurance, for five business lines, were selected for the period between September 2010 and June 2017. The quarterly incurred claims and premium earned are then used to build loss ratios for five different business lines, which will be used to investigate the risk aggregation. The gross and net of reinsurance loss ratios are used to examine the change in the level of risk for each business line and for the aggregate risk.

Research on risk aggregation with copulas applied to insurance was pioneered by Wang (1998). This research introduces the concept of copula and chooses Gaussian copula as one of the useful tools in determining the risk aggregation of an insurance company by combining correlated loss distributions. More precisely, the aggregate loss distribution is determined by the combination between the effect from claim frequency and claim severity distribution. By contrast, Tang and Valdez (2006) use copula models to aggregate risks in order to determine the economic capital as well as the diversification benefits focusing on the insurance industry. Using multiple insurance business lines data, they analyse the importance of selecting an appropriate copula model to avoid underestimation or overestimation of capital required, which consequently may affect the level of capital for insurance products.

Further, Bürgi et al. (2008) highlight that modelling the dependence between risks is important as it is a form of rule for risk aggregation. Their research also considers various methods to model dependencies, which subsequently affect the diversification benefits and show that overestimation of diversification may cause inaccurate computation of risk-based capital (RBC). Also, Nguyen and Molinari (2011) use copula to cover the loopholes of Solvency II which among others relies on linear correlation to measure the dependence structure of correlated risks. However, linear correlation may not be suitable for modelling dependence structure and may not be able to capture all information of a tail distribution. To overcome this problem, the authors proposed method of risk aggregation via copula, which allows to determine completely the dependence structure between risks. Nevertheless, they mainly focus on analysing the concept of the Solvency II rather than modelling copula using real data.

Modelling risk aggregation for high dimensional copula can be very challenging and requires more parameters to be estimated than the traditional two dimensional copulas or bivariate copula models (Bürgi et al., 2008). With this in mind, we consider hierarchical aggregation as an alternative modelling technique based on two dimensional copula. This model, introduced by Arbenz et al. (2012), does not requires specification of copula for all business lines. Instead, a copula and the joint dependence between the aggregated sub-business lines will be determined in each aggregation step. The aggregation model is represented by a rooted tree, which consists of branching nodes and leafs based on graph theory.

In addition, we also investigate the significance of reinsurance from the risk management perspective. Previous research by Cummins et al. (2008) proves that insurance companies purchase reinsurance for the benefits of reducing the loss ratio measured by its volatility. It also provides protection against catastrophes by limiting the liability on specific risks. The drawback of reinsurance is that insurers' cost for production is increased. Furthermore, reinsurance also provides other benefits, such as for capital relief as well as flexible financing. Insurance companies are able to transfer risks to reinsurance and as a result capital is saved from being allocated to these risks (Baur et al., 2004).

The remaining of this chapter is organized as follows. Section 4.2 discusses the methods for aggregating risk using hierarchical copula aggregation model, copula simulations and determination of capital requirement. Section 4.3 describes the estimation of the hierarchical aggregation copula model. In Section 4.4 we analyse the results of the copula aggregation model and the effects of reinsurance in the level of risk and diversification of the portfolio of different business lines. Section 4.5 concludes the chapter.

4.2 Copula-based hierarchical aggregation models

In finance and insurance popular models for problems involving a large number of random variables have been based on copula functions. Different models have been proposed. These include Archimedean and elliptical copula models (Genest and Nešlehová, 2012), vine copula models (Aas et al., 2009; Kurowicka and Joe, 2010), and hierarchical copula models (Mai and Scherer, 2012; Côté and Genest, 2015; Cossette et al., 2017). Some of these models impose a more restrictive dependence structure than others more flexible which in turn imply more difficult inference. This chapter adopts the hierarchical copula model with the goal of achieving a good compromise between flexibility and ease of estimation.

4.2.1 Copula functions

Bivariate copulas are the main building block of hierarchical aggregation copula models. Here we only give the definition in order to introduce the notation and we refer the readers to Nelsen (2006) and McNeil et al. (2015) for an introduction to copulas and the definition of specific copula families. Given a d-dimensional random vector $(X_1, X_2, \ldots, X_d)'$, from Sklar (1959) there exists a function C :

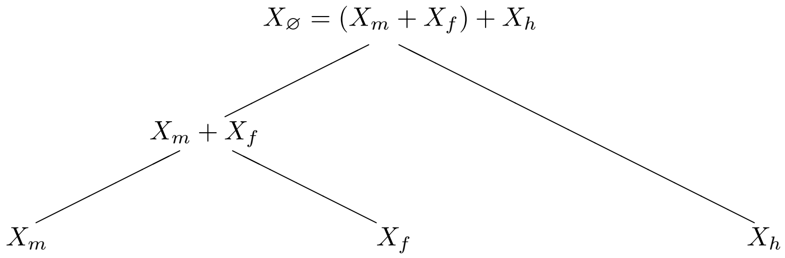


FIGURE 4.1: Illustration of an hierarchical loss aggregation copula model built by allocating each of three individual business lines, represented by X_m, X_f and X_h, to a leaf node of a rooted tree. The structure of the tree in this example is determined by the assumption that the pair (X_m, X_f) have the strongest dependence among the three possible pairs of individual business lines.

$[0,1]^d \to [0,1]$ such that

$$\mathbb{P}(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_n \leq x_d) = C(F_1(x_1), F_2(x_2), \ldots, F_d(x_d)),$$

where $F_i(x_i) = \mathbb{P}(X_i \leq x_i)$ for $i = 1, 2, \ldots, d$ and C is a copula function. In fact, a copula is a multivariate joint cumulative distribution function (cdf) with uniform margins. If the univariate cdf's F_i are continuous then the copula function C is unique.

4.2.2 Hierarchical aggregation copula models

Hierarchical copula models draw on results from graph theory on rooted trees (Diestel, 2017). Following the notation used in Arbenz et al. (2012), a rooted tree τ is composed by leaf nodes and branching nodes where one of the branching nodes is the root. The subset of branching nodes is denoted by $\mathscr{B}(\tau)$, the subset of leaf nodes is denoted by $\mathscr{L}(\tau)$, and the root node by $\varnothing$. Naturally, $\mathscr{B}(\tau) \cup \mathscr{L}(\tau) = \tau$ and $\mathscr{B}(\tau) \cap \mathscr{L}(\tau) = \emptyset$. In order to use rooted trees to aggregate the losses of several business lines we assume the following:

- Each leaf node in the rooted tree is associated with the loss of business line i, represented by a random variable X_i.
- Each branching node is associated with the sum of the business lines mapped to that node's children.

In Figure 4.1 we illustrate the mapping to a rooted tree of three insurance loss variables X_m, X_f and X_h representing the business lines Motor, Fire and Household respectively. Each leaf node corresponds to a business line and each branching node corresponds to the sum of the variables associated with its children nodes. We assume that each branching node has two children, although the results on rooted trees used in this chapter are valid for branching nodes with any number of children (see Arbenz et al., 2012). Assuming that each branching node has only two children keeps the construction and estimation of the model simpler as only bivariate copulas are necessary. In order to define the aggregation model we denote by $(X_i)'_{i\in\tau} = (X_1, X_2, \ldots, X_d)'$ the vector of random variables, where each X_i represents the loss in the business line i. The rooted tree aggregation model for the random vector $(X_i)_{i\in\tau}$ is determined by

- a rooted tree structure τ,
- univariate cdf's $F_i : \mathbb{R} \to [0,1]$ for all leaf nodes i in $\mathscr{L}(\tau)$, and
- bivariate copula functions $C_j : [0,1]^2 \to [0,1]$ for the two children of each branching node j in $\mathscr{B}(\tau)$.

We denote the tree aggregation model by $(\tau, (F_i)_{i\in\mathscr{L}(\tau)}, (C_j)_{j\in\mathscr{B}(\tau)})$. Using this modelling approach we obtain the distribution of the root node which represents the aggregate total loss

$$X_\varnothing = X_1 + X_2 + \ldots + X_d = \sum_{i\in\mathscr{L}(\tau)} X_i$$

based on the univariate cdf's for the business lines associated with the leaf nodes, and the bivariate copulas associated with the branching nodes.

Existence and uniqueness of a joint distribution

The existence and uniqueness of the joint distribution of the hierarchical aggregation copula model for the vector $(X_1, X_2, \ldots, X_d)'$ has been studied in Arbenz et al. (2012). Here we summarise the conditions and the main results necessary in this chapter. Given a rooted tree aggregation model $(\tau, (F_i)_{i\in\mathscr{L}(\tau)}, (C_j)_{j\in\mathscr{B}(\tau)})$ where each branching node $j \in \mathscr{B}(\tau)$ is the sum of its children, the random vector $(X_i)_{i\in\tau}$ is called a *mildly tree dependent*. A mildly tree dependent random vector $(X_i)_{i\in\tau}$ is called *tree dependent* if for each branching node $i \in \mathscr{B}(\tau)$, given X_i, its descendants $(X_j)_{j\in\mathscr{D}(i)}$, where $\mathscr{D}(i)$ is the set of descendent nodes, are conditionally independent of the remaining nodes $(X_j)_{j\in\tau\setminus\mathscr{D}(i)}$,

$$(X_j)_{j\in\mathscr{D}(i)} \perp (X_j)_{j\in\tau\setminus\mathscr{D}(i)} \mid X_i \qquad \text{for all } i \in \mathscr{B}(\tau).$$

This conditional independence condition does not imply that $(X_j)_{j\in\mathscr{D}(j)}$ is independent of $(X_j)_{j\in\tau\setminus\mathscr{D}(i)}$ rather that their dependence comes from X_i.

Theorem 4.1. *Given a rooted tree aggregation model $\left(\tau, (F_i)_{i\in\mathscr{L}(\tau)}, (C_j)_{j\in\mathscr{B}(\tau)}\right)$, a tree dependent random vector exists and its joint distribution is unique.*

For the proof of this result see Arbenz et al. (2012). For the example illustrated in Figure 4.1 the joint distribution of the hierarchical aggregation copula model for the vector $(X_m, X_f, X_h)'$ exists and is unique if and only if

$$(X_m, X_f) \ \perp \ (X_\varnothing, X_h) \mid X_m + X_f,$$

where $X_\varnothing = X_m + X_f + X_h$. This means that all the information in X_m and X_f that influences X_h is contained in $X_m + X_f$.

Under the above theorem, if all the univariate and copula distributions are absolutely continuous then the joint density function is given by the following proposition showed in Côté and Genest (2015).

Proposition. *Given a rooted tree aggregation model $(\tau, (F_i)_{i\in\mathscr{L}(\tau)}, (C_j)_{j\in\mathscr{B}(\tau)})$ with d leaf nodes associated with the vector $\mathbf{X} = (X_1, X_2, \ldots, X_d)'$, the joint density function of the vector $\mathbf{X}$ is given by*

$$f_{\mathbf{X}}(x_1,\ldots,x_d) = \prod_{j=1}^{d-1} c_j \left[F_{\mathscr{LD}(j1)}\left(\sum_{i\in\mathscr{LD}(j1)} x_i\right), F_{\mathscr{LD}(j2)}\left(\sum_{i\in\mathscr{LD}(j2)} x_i\right)\right] \prod_{i=1}^{d} f_i(x_i),$$

for all $x_1, \ldots, x_d \in \mathbb{R}$, where $\mathscr{LD}(ji)$ represents the leaf nodes in the set of descendants of child node i of the branching node j, $F_{\mathscr{LD}(jk)}$ is the cdf of the sum of the leaf nodes in $\mathscr{LD}(jk)$, $f_1, \ldots, f_d$ are the univariate density functions of $X_1, X_2, \ldots, X_d$ respectively, and c_j is the copula density function of the children of X_j for $j \in \mathscr{B}(\tau)$.

As an example, for the business lines represented by the random vector $(X_m, X_f, X_h)'$ associated with the rooted tree τ illustrated in Figure 4.1 the joint density function is given by

$$f_{\mathbf{X}}(x_m, x_f, x_h) = c_{m,f}\left(F_m(x_m), F_f(x_f)\right) c_{m+f,h}\left(F_{m+f}(x_m + x_f), F_h(x_h)\right) \\ .f_m(x_m) f_f(x_f) f_h(x_h),$$

for all (x_m, x_f, x_h), where F_i is the cdf of the univariate random variable X_i with density function f_i, F_{m+f} is the cdf of $X_m + X_f$, $c_{m,f}$ is the copula density function of (X_m, X_f) and $c_{m+f,h}$ is the bivariate copula density function of $((X_m + X_f), X_h)$.

Simulation of joint distribution

Given the set of d business lines represented by the random variables $X_1, X_2, \ldots, X_d$ we determine the tree structure by aggregating iteratively the pair of variables with the strongest dependence. We use Kendall's tau to measure the dependence between pairs of random variables in the hierarchical aggregation procedure. We refer to Côté and Genest (2015) for the motivation and justification for using Kendall's tau in this setting.

After defining the structure of the tree we proceed with selecting the probability distribution for each random variable allocated to a leaf node and the copula family for the two children of each branching node in order to specify the hierarchical aggregation model. We use maximum likelihood to estimate the parameters, and Anderson and Darling (1954) and Genest et al. (2009) goodness-of-fit methods to select the best distributions.

The hierarchical aggregation model allows to estimate measures of risk on the sum of the individual variables considered. We estimate these risk measures based on the simulation of observations from the aggregation model. We follow the algorithm introduced in Arbenz et al. (2012) that consists of a numerical approximation procedure where sample reordering induces the dependence structure, a technique that goes back to the work of Iman and Conover (1982).

We present below the algorithm for the case when all branching nodes have two children and the functional that produces the aggregation is a weighted sum of the branching nodes. This later aspect is a straightforward generalization of the case of a sum presented by Arbenz et al. (2012). A generalization for the case when the aggregation functionals are Kendall functions can be found in Brechmann (2014).

Sample reordering numerical approximation algorithm:

1. Define the number of simulations $N \in \mathbb{N}$.

2. Simulate N independent samples from the univariate variables X_i ($i \in \mathscr{L}(\tau)$) associated with the d leaf nodes: $X_i^k \sim F_i$ for $k = 1, \ldots, N$ and $i = 1, \ldots, d$. F_i is the pre-determined univariate cdf for i.
3. Simulate N independent samples from the bivariate copula C_j ($j \in \mathscr{B}(\tau)$) associated with each of the $d-1$ branching nodes: $\mathbf{U}_j^k = (U_{j1}^k, U_{j2}^k) \sim C_j$ for $k = 1, \ldots, N$ and $j = 1, \ldots, d-1$.
4. Following a bottom-up approach, recursively (beginning at the branching nodes closer to the leaf nodes and ending at the root node) define the approximation for the cdf of each branching node $j \in \mathscr{B}(\tau)$ as

$$F_j^N(x) = \frac{1}{N} \sum_{k=1}^{N} \mathbb{1} \left\{ w_{j1} \, x_{j1}^{(r_{j1}^k)} + w_{j2} \, x_{j2}^{(r_{j2}^k)} \leq x \right\},$$

where $\mathbb{1}$ is the indicator function given by

$$\mathbb{1}\{s_j \leq x\} = \begin{cases} 1, & \text{if } s_j \leq x \\ 0, & \text{otherwise,} \end{cases}$$

x_{j1}^k and x_{j2}^k are (simulated) sample values of the variables associated with the two nodes children of the branching node j, w_{ji} is the weight given to variable X_{ji}, r_{ji}^k is the (componentwise) rank of u_{ji}^k, and $\{x_{ji}^{(1)}, x_{ji}^{(2)}, \ldots, x_{ji}^{(N)}\}$ is the ordered sample for $i = 1, 2$.

To better understand the above algorithm, in the following we present a reworked example from Arbenz et al. (2012) illustrating the estimation of the distribution of a variable representing the aggregate loss using the sample reordering algorithm.

Example. Consider the hierarchical copula model associated with the rooted tree depicted in Figure 4.1. Suppose that the results from simulating $N = 3$ samples of X_i^k and $\mathbf{U}_j^k$ are

$$\begin{aligned}
&x_m^1 = 0.2, \qquad x_m^2 = 0, \qquad x_m^3 = 0.1, \\
&x_f^1 = 1, \qquad x_f^2 = 0, \qquad x_f^3 = 2, \\
&x_h^1 = 20, \qquad x_h^2 = 10, \qquad x_h^3 = 0, \\
&\mathbf{u}_{m+f}^1 = (0.5, 0.2), \qquad \mathbf{u}_{m+f}^2 = (0.3, 0.9), \qquad \mathbf{u}_{m+f}^3 = (0.7, 0.4), \\
&\mathbf{u}_{\varnothing}^1 = (0.9, 0.5) \qquad \mathbf{u}_{\varnothing}^2 = (0.6, 0.8), \qquad \mathbf{u}_{\varnothing}^3 = (0.1, 0.4).
\end{aligned}$$

The componentwise ranks (ordered from smallest to largest) of the copula samples are

$$\mathbf{r}^{(1)}_{m+f} = (2,1), \qquad \mathbf{r}^{(2)}_{m+f} = (1,3), \qquad \mathbf{r}^{(3)}_{m+f} = (3,2),$$
$$\mathbf{r}^{1}_{\varnothing} = (3,2), \qquad \mathbf{r}^{2}_{\varnothing} = (2,3), \qquad \mathbf{r}^{3}_{\varnothing} = (1,1).$$

For simplicity we assume unit weights, $\omega_{j1} = \omega_{j2} = 1$ for all branching nodes j. The approximation given by the sample reordering algorithm for the cfd of the variable X_{m+f} is

$$\begin{aligned} F^3_{m+f}(x) &= \frac{1}{3}\left(\mathbb{I}\{x^{(2)}_m + x^{(1)}_f \le x\} + \mathbb{I}\{x^{(1)}_m + x^{(3)}_f \le x\} + \mathbb{I}\{x^{(3)}_m + x^{(2)}_f \le x\}\right) \\ &= \frac{1}{3}\left(\mathbb{I}\{0.1 + 0 \le x\} + \mathbb{I}\{0 + 2 \le x\} + \mathbb{I}\{0.2 + 1 \le x\}\right) \\ &= \frac{1}{3}\left(\mathbb{I}\{0.1 \le x\} + \mathbb{I}\{2 \le x\} + \mathbb{I}\{1.2 \le x\}\right). \end{aligned}$$

Hence, the atoms $\{x^1_{m+f}, x^2_{m+f}, x^3_{m+f}\}$ of the distribution F^3_{m+f} are

$$x^1_{m+f} = 0.1, \qquad x^2_{m+f} = 2 \qquad \text{and} \qquad x^3_{m+f} = 1.2.$$

Now we can derive the approximation for the cdf of the total aggregate loss represented by the variable $X_{\varnothing}$ which is

$$\begin{aligned} F^3_{\varnothing}(x) &= \frac{1}{3}\left(\mathbb{I}\{x^{(3)}_{m+f} + x^{(2)}_h \le x\} + \mathbb{I}\{x^{(2)}_{m+f} + x^{(3)}_h \le x\} + \mathbb{I}\{x^{(1)}_{m+f} + x^{(1)}_h \le x\}\right) \\ &= \frac{1}{3}\left(\mathbb{I}\{2 + 10 \le x\} + \mathbb{I}\{1.2 + 20 \le x\} + \mathbb{I}\{0.1 + 0 \le x\}\right) \\ &= \frac{1}{3}\left(\mathbb{I}\{12 \le x\} + \mathbb{I}\{21.2 \le x\} + \mathbb{I}\{0.1 \le x\}\right). \end{aligned}$$

Once we have the estimate for the cdf of the total aggregate loss we can estimate the risk based on the estimated cdf. We will now turn to the risk estimation.

4.2.3 Risk estimation of the aggregate loss

After building the model for the aggregate loss, using the hierarchical copula model we can estimate the risk of the aggregate loss. As a coherent measure of risk we use the tail value at risk (TVaR) introduced by Acerbi and Tasche (2002). The TVaR of the loss represented by the variable X at the confidence level α, for $\alpha \in (0,1)$, is defined as

$$\mathrm{TVaR}_\alpha(X) = \frac{1}{1-\alpha}\int_\alpha^1 \mathrm{VaR}_u(X)\, du,$$

where the VaR_α of the loss X is given by

$$\mathrm{VaR}_\alpha(X) = \inf\left\{x \in \mathbb{R} : P(X \leq x) \geq \alpha\right\}.$$

For risk measurement in finance and insurance α typically takes the values 90%, 95% or 99%. In order to estimate the TVaR we use the following nonparametric estimator that can be found in more detail in Adam et al. (2008). Given n observations $\{x_1, x_2, \ldots, x_n\}$ of the variable X the TVaR estimator is given by

$$\widehat{\mathrm{TVaR}}_\alpha = \frac{1}{n(1-\alpha)}\left(\sum_{i=1}^{\lfloor n(1-\alpha)\rfloor} x_{(n-i+1)} + (n(1-\alpha) - \lfloor n(1-\alpha)\rfloor)x_{(n-\lfloor n(1-\alpha)\rfloor)}\right), \tag{4.1}$$

where $\{x_{(1)}, x_{(2)}, \ldots, x_{(n)}\}$ is the ordered sample, $\lfloor v \rfloor$ denotes the largest integer not greater than v, and in our case $\alpha \in \{0.9, 0.95, 0.99\}$. In our setting, we estimate the TVaR by applying (4.1) to the N observations simulated by the sample reordering algorithm. Given its wide use, notably in Solvency II, we also report the VaR estimates at the three confidence levels given by the empirical quantile of the N simulations.

4.3 Empirical analysis on Australia insurance industry

4.3.1 The data

In this chapter we use data on general insurance obtained from the Australian Prudential Regulation Authority - APRA (`https://www.apra.gov.au/`) as in Tang and Valdez (2006). However, we use a more recent time period and quarterly data instead of annual data to increase the sample size. Our preference for the Australian insurance case is due to the data being freely available and its large insurance industry market share within developed countries. According to the data published by OECD (see OECD, 2017), Australia's general insurance is above the 70th percentile in terms of total gross premiums in 2016. This shows the relevance of our results. In September 2010 a change in the reporting format was introduced. Therefore we focus on the period after September 2010 until June 2017. This gives us 28 observations. We are interested in four variables: gross incurred claims (this includes movements in outstanding claims Liability during the period); gross earned premium; net incurred claims (net of reinsurance recoveries revenue); net earned premium (net of outwards reinsurance expense). We consider both the gross and the net variables as one of our goals is to evaluate the effect of reinsurance on the capital requirements. We source data for five insurance business lines, namely domestic Motor vehicle (hereafter referred to as Motor), houseowners/households (House from here onwards), Fire and ISR[1] (Fire), Liability, and compulsory third party Motor vehicle (CTP). According to the data collected from the APRA webpage these five business lines encompass more than 85% of the Australian general insurance market in terms of net earned premiums as at June 2017. In the process of cleaning the data we removed the observations from two quarters where there are two negative observations of gross incurred claims leading to counter intuitive negative loss ratios. Hence, our final data set has 26 observations per business line.

Loss ratios

We base our study on the variable loss ratio. For business line i and time period t, we define *loss ratio* as

$$LR_{i,t} = \frac{IC_{i,t}}{EP_{i,t}},$$

where $IC_{i,t}$ denotes the incurred claims corresponding to the earned premium $EP_{i,t}$ based on accident year insurance company accounting principal; see Taylor (1997) for details on the loss ratio variable. The loss ratio can be seen as a measure of claims standardized by the risk exposure (given by the earned premium). Using loss ratios eliminates temporal effects of business growth and inflation, and it allows to make comparisons between business lines with different risk exposures. The loss ratios are subsequently added up to form the aggregated loss ratio for capital requirement estimation.

[1]ISR stands for Industrial Special Risk

The aggregate loss ratio at time t, LR_t, can then be written as a weighted sum of the loss ratios for each of the d business lines as

$$\begin{aligned} LR_t &= \frac{IC_t}{EP_t} \\ &= \frac{\sum_{i=1}^d IC_{i,t}}{\sum_{i=1}^d EP_{i,t}} \\ &= \frac{\sum_{i=1}^d \left(\frac{IC_{i,t}}{EP_{i,t}} \times EP_{i,t}\right)}{\sum_{i=1}^d EP_{i,t}} \\ &= \sum_{i=1}^d LR_{i,t} \times \frac{EP_{i,t}}{\sum_{i=1}^d EP_{i,t}} \\ &= \sum_{i=1}^d LR_{i,t} \times w_{i,t}, \end{aligned} \tag{4.2}$$

where IC_t and EP_t are the incurred claims and earned premium aggregated across all business lines, and $w_{i,t}$ is the weight of business line i in period t. In our study we consider gross loss ratios and net (of reinsurance) loss ratios. The gross loss ratios are computed with gross claims and gross premiums. The net loss ratios are based on net claims and net premiums.

	House	Fire	Motor	CTP	Liability	Aggregate loss
			Gross loss ratios			
Mean	0.5849	0.7820	0.7211	0.8172	0.7024	0.7005
Standard deviation	0.2981	0.8334	0.0682	0.3100	0.1566	0.1971
Skewness	2.6290	3.6449	0.9729	-0.7432	-0.2392	2.8759
Excess kurtosis	8.0694	13.819	0.0075	0.0036	0.0671	9.6254
Average weight, $\bar{w}_{i,t}$	0.25	0.14	0.33	0.11	0.18	1
Weight at June 2017, $w_{i,T}$	0.26	0.12	0.33	0.13	0.16	1
			Net loss ratios			
Mean	0.6272	0.6549	0.7394	0.8051	0.6499	0.7018
Standard deviation	0.2105	0.2639	0.0454	0.3333	0.1907	0.1659
Skewness	2.0440	1.4870	0.3835	-0.8458	-0.6556	1.3425
Excess kurtosis	5.6319	2.2074	-0.9542	0.0960	1.5980	2.4629
Average weight, $\bar{w}_{i,t}$	0.22	0.10	0.36	0.13	0.18	1
Weight at June 2017, $w_{i,T}$	0.24	0.09	0.36	0.13	0.17	1

TABLE 4.1: Summary statistics of the loss ratios. The five columns of values labelled with the individual business lines have the statistics for $LR_{i,t}$ defined in equation (4.2) for all business lines i and t from September 2010 to June 2017. The right column has the statistics for the aggregate loss ratio, LR, as defined in (4.2).

Table 4.1 lists descriptive statistics for the loss ratios obtained from the data on the five business lines. The column 'Aggregate loss' has the statistics for the aggregate loss ratio calculated as in equation (4.2). The first observation from

Table 4.1 concerns the mean. We observe that for all the business lines the average loss ratios gross and net of reinsurance are not statistically different. Although reinsurance is essentially a risk transfer (or sharing) tool, loss distributions tend to be positively skewed and hence we would expect the average loss ratio to reduce from gross to net of reinsurance. But reinsurance seems to have no strong effect on the average loss ratio. We explore later in the chapter how this may result from the interplay between the premium ceded to and claim recoveries from reinsurance. The standard deviation is higher for Fire. While for House, Motor an especially Fire standard deviation reduces, it actually increases for CTP and Liability when reinsurance is taken into account. The values estimated for the skewness show that the loss ratios for House and Fire do not have symmetric distributions. There is also significant excess kurtosis for House and Fire both reducing with reinsurance. In terms of the aggregate loss ratio, reinsurance has a larger effect on the skewness and kurtosis than on the mean and standard deviation of the loss ratio. Most notably, reinsurance reduces the excess kurtosis of the aggregate loss ratio by 74%.

4.4 Estimation of the hierarchical aggregation copula model

In this section we implement the estimation of the hierarchical copula model for the aggregate loss from the individual business lines as presented in Section 4.2.2.

4.4.1 Tree structure of the hierarchical copula model

The first element of the hierarchical copula model is the rooted tree τ associated with the variables representing the loss ratio for each business line. As explained in Section 4.2.2 to build the tree we start by allocating the loss ratio of each business line to one leaf node and then aggregate the two random variables, representing loss ratios, with the highest dependence measured by Kendall's tau. Table 4.2 shows the Kendall's tau estimates for each pair of business lines.

After allocating each business line to a leaf node, as in the bottom row of the tree depicted in Figure 4.2, we aggregate the two business lines with the strongest dependence. We first address the gross loss ratios and after building a new tree

Stage 1					
	House	Fire	Motor	CTP	Liability
House	1	–	–	–	–
Fire	**0.5262**	1	–	–	–
Motor	0.4338	0.2308	1	–	–
CTP	0.0154	-0.0523	-0.1815	1	–
Liability	0.0585	-0.1323	0.1446	0.3662	1

Stage 2				
	House + Fire	Motor	CTP	Liability
House + Fire	1	–	–	–
Motor	0.3169	1	–	–
CTP	-0.0400	-0.1815	1	–
Liability	-0.0338	0.1446	**0.3662**	1

Stage 3			
	House + Fire	Motor	CTP + Liability
House + Fire	1	–	–
Motor	**0.3169**	1	–
CTP+Liability	0.0154	-0.0523	1

Table 4.2: Sequential aggregation of the gross loss ratios for the five business lines. At each stage we aggregate the two loss ratio random variables with the strongest Kendall's tau estimate.

for the net loss ratios. From Table 4.2 we observe that House and Fire have the largest Kendall's tau. Hence, at this first stage, we aggregate these two business lines. In the second panel of Table 4.2, labelled stage 2, the largest Kendall's tau observed is between CTP and Liability. We then aggregate CTP and Liability defining the second row, from the bottom, of the rooted tree in Figure 4.2. In stage 3 the strongest dependence is between Motor and Fire plus House. Finally, aggregating Motor plus Fire plus House with CTP plus Liability defines the root node and completes the tree in Figure 4.2. The Kendall's tau between the gross loss ratios House plus Fire plus Motor and CTP plus Liability is 0.0154.

Table 4.3 contains the Kendall tau values for the case of the net (of reinsurance) loss ratio for the five business lines. The variables are more strongly dependent, at the different stages of the construction of the tree. The same scenario observed for the gross loss ratios. As a consequence the structure of the rooted tree for the net loss ratios hierarchical copula model is the same as for the gross loss ratios shown in Figure 4.2. In the last stage of the aggregation model the Kendall's tau between the net loss ratios for House, Fire, Motor, CTP and Liability together is -0.0892.

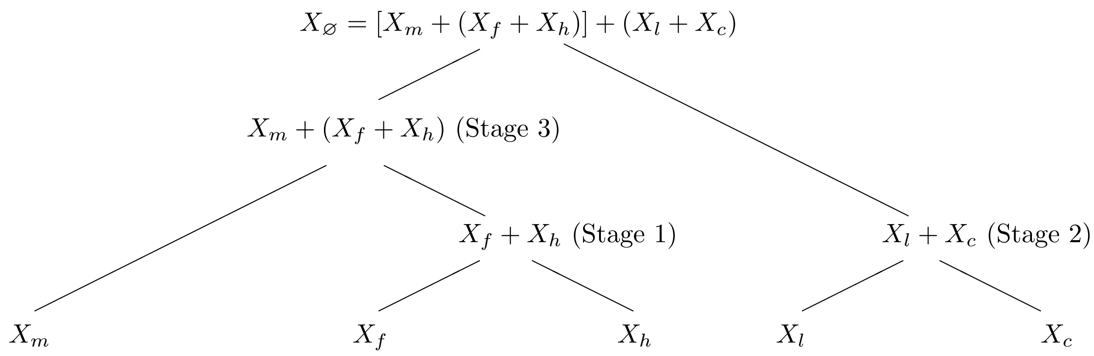


FIGURE 4.2: Hierarchical loss aggregation copula model for the gross (and net) loss ratio of the the five business lines Motor, Fire, House, Liability and CTP, represented by X_m, X_f, X_h, X_l and X_c, respectively. The structure of the tree is determined by aggregating iteratively the two nodes with the strongest dependence.

Stage 1					
	House	Fire	Motor	CTP	Liability
House	1	–	–	–	–
Fire	**0.5446**	1	–	–	–
Motor	0.4338	0.2492	1	–	–
CTP	-0.0154	-0.0031	-0.2369	1	–
Liability	0.0092	-0.0646	-0.0523	0.4954	1

Stage 2				
	House + Fire	Motor	CTP	Liability
House + Fire	1	–	–	–
Motor	0.3969	1	–	–
CTP	-0.0400	-0.2369	1	–
Liability	-0.0523	-0.0523	**0.4954**	1

Stage 3			
	House + Fire	Motor	CTP + Liability
House + Fire	1	–	–
Motor	**0.3969**	1	–
CTP+Liability	-0.0523	-0.2123	1

TABLE 4.3: Sequential aggregation of the net loss ratios for the five business lines.

4.4.2 Determination of univariate distributions

At this point we select a family of univariate distributions for the loss ratio of each business line. We use maximum likelihood to fit parametric families of distributions and decide between possible distributions according to the Anderson and Darling (1954) goodness of fit test. As we are primarily interested in estimating measures of risk, which are based on the tail of the distributions, it is important to use an appropriate test. It is known that the Anderson and Darling (A-D) test is more powerful and sensitive to the tails of the distribution (see Engmann and Cousineau, 2011) than other alternative tests such as the commonly used Kolmogorov-Smirnov[2] goodness-of-fit test. Hence, we choose the distribution giving the highest p-value according to the A-D test. For each business line we fit the following families of distributions: lognormal, gamma, Weibull, loglogistic, Pareto and Burr.

The results for the distribution with the highest A-D test p-value and corresponding parameter estimates in Table 4.4. The selection of families for the loss ratio of each business line falls into the set of log-logistic, Burr and Weibull distributions. These selections are also visualised in Figure 4.3 and Figure 4.4 for gross and net loss ratios respectively.

[2](Kolmogorov, 1933; Smirnov, 1948).

	House	Fire	Motor	CTP	Liability	Aggregate loss
			Gross loss ratios			
Distribution	Log-logistic	Burr	Burr	Weibull	Burr	Burr
Shape 1	4.76266	0.19159	0.04799	3.00527	7.70166	0.3732
(s.e.)	(0.776)	(0.122)	(0.042)	(0.505)	(22.63)	(0.199)
Shape 2	–	8.11427	189.928	–	5.64960	15.8580
(s.e.)	–	(4.012)	(155.0)	–	(1.555)	(5.441)
Scale*	0.52243	3.04747	1.55319	0.90936	0.92955	1.70254
(s.e.)	(0.037)	(0.415)	(0.014)	(0.061)	(0.604)	(0.095)
A-D statistic	0.294	0.147	0.335	1.417	0.270	0.230
A-D p-value	0.942	0.998	0.909	0.197	0.958	0.979
			Net loss ratios			
Distribution	Log-logistic	Log-logistic	Log-logistic	Weibull	Weibull	Burr
Shape 1	6.37499	4.96750	27.9840	2.53352	3.87399	0.50244
(s.e.)	(1.031)	(0.801)	(4.469)	(0.439)	(0.599)	(0.269)
Shape 2	–	–	–	–	–	18.4406
(s.e.)	–	–	–	–	–	(5.898)
Scale*	0.59180	0.59840	0.73616	0.89199	0.71298	1.55857
(s.e.)	(0.031)	(0.041)	(0.009)	(0.071)	(0.037)	(0.073)
A-D statistic	0.246	0.455	0.371	1.962	0.602	0.197
A-D p-value	0.971	0.791	0.875	0.097	0.643	0.991

TABLE 4.4: Family of distributions selected for each business line gross and net loss ratios. The parameter and corresponding standard errors estimates are listed for each business line together with the Anderson and Darling (A-D) statistic and p-value. For the purpose of comparison the table also has the estimates for the aggregate loss ratio with the weights fixed as at June 2017. *In the case of the Burr distribution the value listed in the table as being the scale is in fact the estimate for the rate which is 1/scale.

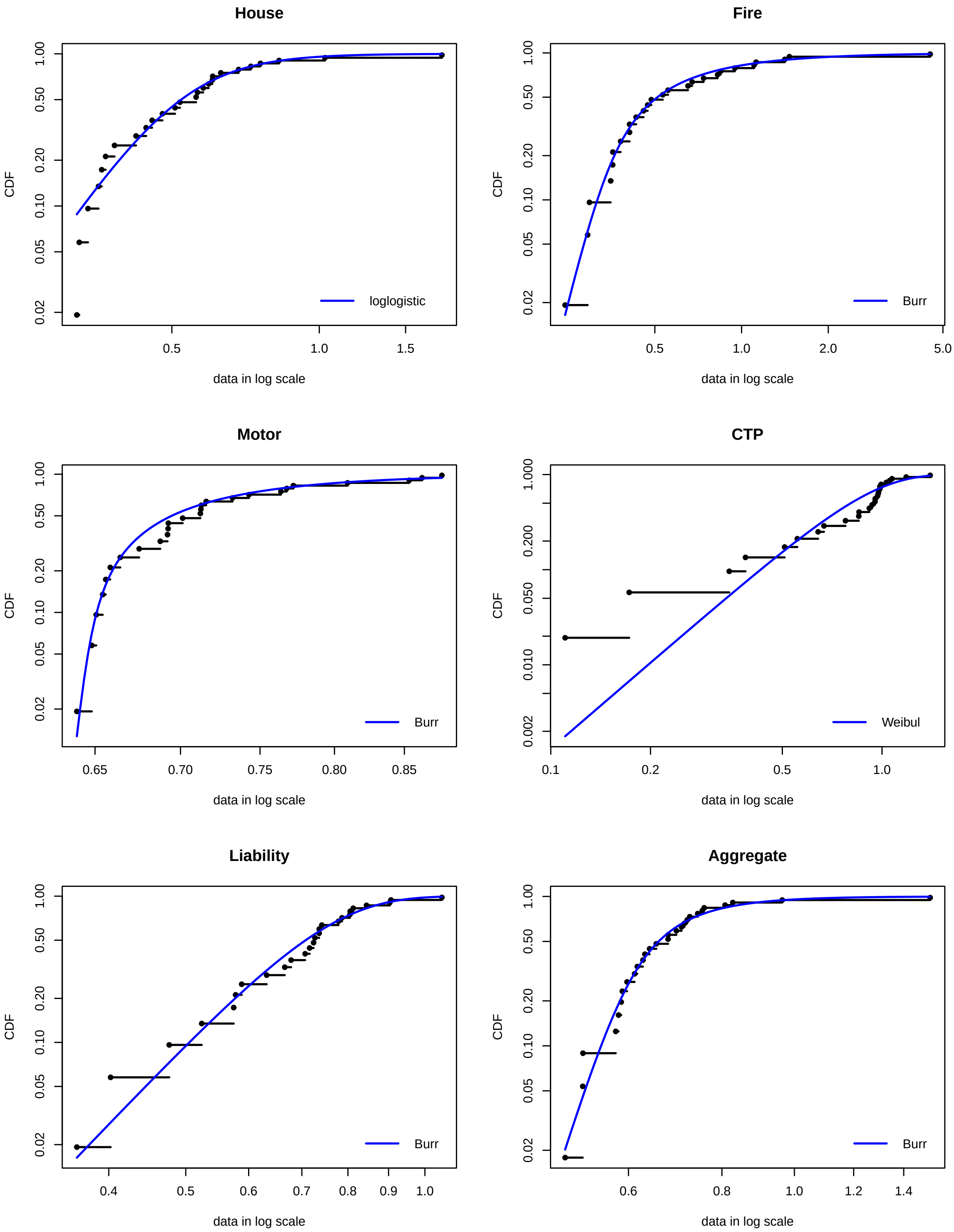


Figure 4.3: Subplots of empirical cumulative distribution with theoretical distributions corresponding to respective fitted distributions for gross loss ratios.

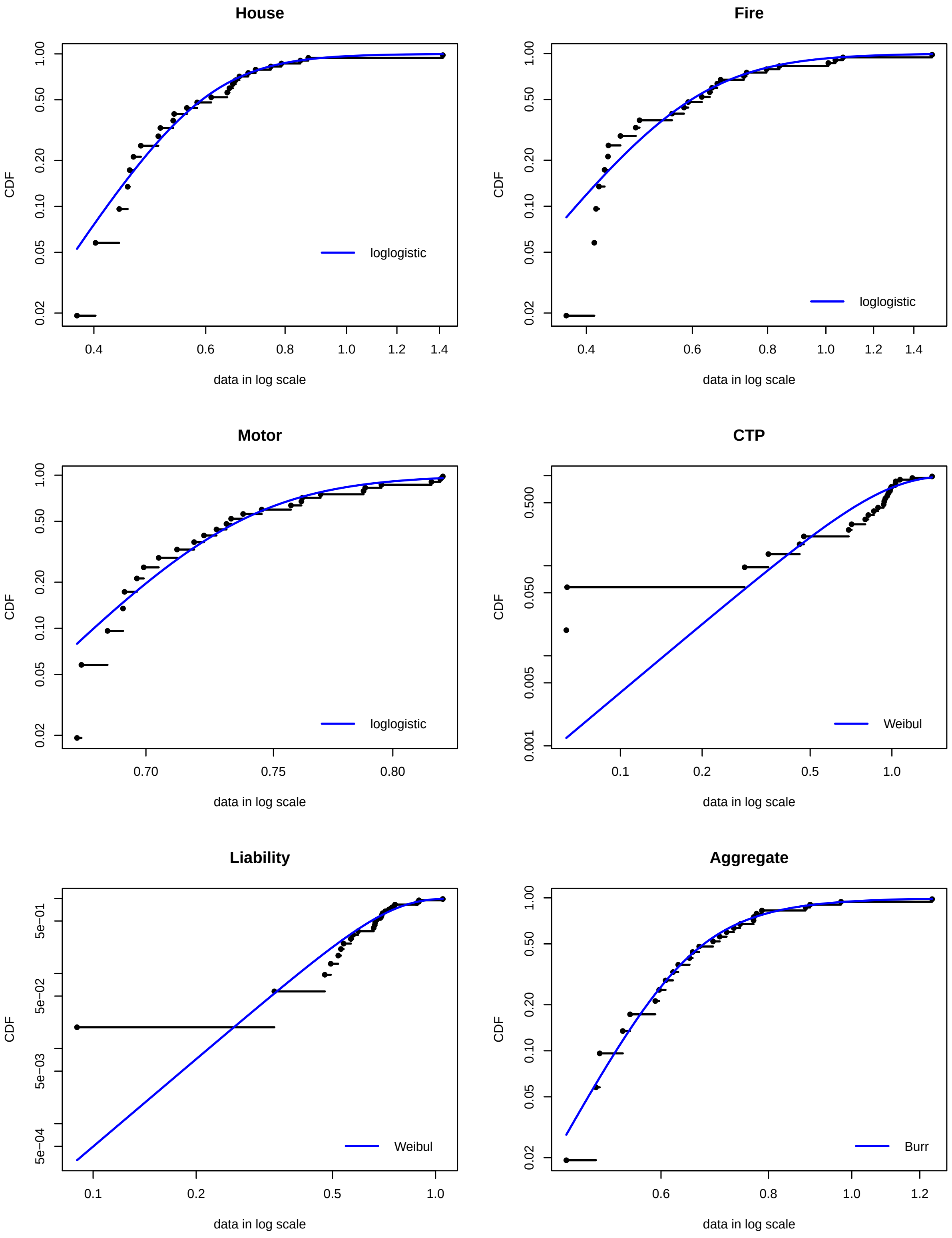


FIGURE 4.4: Subplots of empirical cumulative distribution with theoretical distributions corresponding to respective fitted distributions for net loss ratios.

4.4.3 Copula selection

In the construction of the hierarchical copula aggregation model for the loss ratios we have to select a copula family for (the two children of) each branching node of the tree structure determined in Section 4.4.1. As possible families of copulas we consider the most commonly used copulas, such as, Gaussian copula, the Student-t, the Frank, the Clayton, the Gumbel, the mixtures of Clayton and Gumbel copulas and corresponding survival copulas. To help informing the choice of copula family we calculate, and report in Table 4.5, non-parametric estimates of the upper and lower tail dependence coefficient (see Sibuya, 1960, and Schmid and Schmidt, 2007) for the pair of loss ratios associated with the children of each branching node. As the risk of extreme events is one of the main concerns when it comes to capital requirements, it is important to pay particular attention to the tails of the copula distributions in the modelling process. Table 4.5 summarizes the results of the copula selection process related to the four branching nodes for gross and net loss ratios. For the gross loss ratios the first node we consider is House and Fire, as shown in Figure 4.2. From Table 4.5 we can see that both lower (λ_L) and upper (λ_U) tail coefficient estimates are different from zero. The copula with the highest p-value (using the goodness of fit test statistic S_n from Genest et al., 2009) is a mixture of 40% Clayton and 60% survival Clayton copulas. As the Clayton copula allows for tail dependence the mixture model seems to be a reasonable choice. The p-value of the S_n goodness of fit test, the parameters and standard errors estimates are also listed Table 4.5. For the business lines CTP and Liability the best copula is a mixture of 25% Clayton and 75% survival Clayton copulas. The same copula mixture is again the best for Motor and Fire plus House but with only 10% weight on the Clayton component of the mixture. The estimates for the tail coefficients for the two root node children, Motor plus Fire plus House and CTP plus Liability, are zero. Indeed the best copula, according to the goodness of fit test, is the Gaussian copula which has no tail dependence.

For the net loss ratios the best copula for the House and Fire pair is 60% Gumbel plus 40% survival Gumbel. The resulting copula has both upper and lower tail dependence which is in line with the non-parametric estimates. CTP and Liability is best modelled by a Student-t which also allows for both upper and lower tail dependence. A mixture of 70% survival Gumbel with 30% survival Clayton has the highest p-value for Motor and Fire plus House. The estimate for the Kendall's tau for the pair Motor plus Fire plus House and CTP plus Liability is close to zero but negative. Hence we flip the variable Motor plus Fire plus House after transforming it into the $[0, 1]$ interval. By flipping we mean subtract the variable from one. The best copula for the resulting pair is then a Gumbel copula which allows for tail dependence between low values of Motor plus House plus Fire and high values of CTP and Liability.

4.4.4 Simulation of the aggregate loss ratios

In order to estimate VaR and TVaR from the hierarchical copula aggregation model we can now simulate observations of aggregate loss ratios using the model constructed in the previous sections. We implement the sample reordering algorithm from Section 4.2.2 for the gross and net loss ratios using $N = 1,000$. Using the estimator from equation (4.1) we estimate the TVaR for each business line, gross and net, loss ratios for the confidence levels of 90%, 95% and 99%. The VaR estimate for a given confidence level is the corresponding empirical quantile. We repeat the process $1,001$ times and obtain parametric bootstrap estimates for the VaR, TVaR and corresponding confidence intervals. The results are presented in Table 4.6 and its analysis follows in the next sections.

	λ_L	λ_U	Copula	p-value	$\hat{\theta}_1$ (s.e.)	$\hat{\theta}_2$ (s.e.)
			Gross loss ratios			
(X_h, X_f)	0.5218	0.5694	0.4 Clayton + 0.6 SurvClayton	0.4640	4.886	2.148
					(4.161)	(1.966)
(X_c, X_l)	0.1496	0.2742	0.25 Clayton + 0.75 SurvClayton	0.5410	1.022	1.482
					(3.194)	(1.596)
$(X_m, X_f + X_h)$	0.2772	0.4383	0.1 Clayton + 0.9 SurvClayton	0.8986	1.160	1.029
					(5.796)	(0.548)
$(X_m + X_f + X_h, X_c + X_l)$	0.0000	0.0000	Gaussian	0.9815	0.013036	
					(0.285)	
			Net loss ratios			
(X_h, X_f)	0.5390	0.5401	0.6 Gumbel + 0.4 SurvGumbel	0.7298	2.126	2.801
					(1.265)	(2.083)
(X_c, X_l)	0.2772	0.1070	Student-t	0.5549	0.7376	1.2910
					(0.115)	(0.593)
$(X_m, X_f + X_h)$	0.3977	0.4038	0.7 SurvGumbel + 0.3 SurvClayton	0.7607	1.750	1.047
					(0.954)	(2.884)
$(X_m + X_f + X_h, X_c + X_l)$	0.0143	0.1531	90° Rotated Gumbel	0.5569	1.0865	
					(0.186)	

TABLE 4.5: Upper (λ_U) and lower (λ_L) tail coefficient non-parametric estimates for the pairs of children of each branching node of the copula hierarchical model tree. The best fitting copula, corresponding goodness of fit test p-value, and parameter estimates (with standard errors in parenthesis) are also listed. For the mixture copulas θ_1 is the parameter estimate of the first component of the mixture and θ_2 corresponds to the second component of the mixture. For the last pair of net loss ratios, $(X_m + X_f + X_h, X_c + X_l)$, λ_L measures the tail coefficient in the second quadrant of the sample space and λ_U measures the tail coefficient in the fourth quadrant.

	House	Fire	Motor	CTP	Liability	Weighted Sum of risk measures	Risk measure of aggregate loss, LR_t
				Gross loss ratios			
90% VaR	0.8284	1.4422	0.8283	1.1991	0.8915	0.9603	0.8806
	(0.8281)	(1.437)	(0.8283)	(1.1991)	(0.8916)	(0.9596)	(0.88)
	[0.800,0.856]	[1.301,1.60]	[0.814,0.843]	[1.172,1.225]	[0.879,0.903]	[0.940,0.981]	[0.859,0.902]
95% VaR	0.9693	2.2516	0.8931	1.3088	0.9417	1.1377	1.0184
	(0.9669)	(2.2418)	(0.8928)	(1.3085)	(0.9415)	(1.1361)	(1.0157)
	[0.925,1.024]	[1.959,2.593]	[0.872,0.916]	[1.278,1.341]	[0.926,0.957]	[1.099,1.182]	[0.979,1.064]
99% VaR	1.365	6.2017	1.0602	1.5049	1.0346	1.8101	1.5937
	(1.3592)	(6.0118)	(1.058)	(1.5043)	(1.0346)	(1.7869)	(1.5706)
	[1.227,1.534]	[4.463,8.642]	[1.008,1.122]	[1.451,1.56]	[1.008,1.061]	[1.603,2.095]	[1.385,1.891]
90%TVaR	1.063	4.1271	0.9299	1.3413	0.9576	1.4060	1.2644
	(1.06)	(3.7259)	(0.9289)	(1.3414)	(0.9576)	(1.3571)	(1.2145)
	[1.007,1.128]	[2.873,6.202]	[0.906,0.957]	[1.313,1.37]	[0.944,0.972]	[1.256,1.652]	[1.118,1.518]
95%TVaR	1.2353	6.4776	1.0026	1.4322	1.0003	1.7755	1.5895
	(1.2299)	(5.6178)	(1.0015)	(1.4327)	(1.0001)	(1.6741)	(1.4845)
	[1.144,1.341]	[4.096,10.578]	[0.966,1.042]	[1.397,1.466]	[0.983,1.019]	[1.488,2.276]	[1.304,2.094]
99%TVaR	1.7244	18.4861	1.1898	1.6042	1.0836	3.4412	3.1437
	(1.6978)	(13.8818)	(1.1845)	(1.6032)	(1.0835)	(2.9017)	(2.609)
	[1.459,2.074]	[8.037,37.647]	[1.1,1.299]	[1.54,1.669]	[1.051,1.118]	[2.178,5.761]	[1.897,5.461]
				Net loss ratios			
90% VaR	0.835	0.9313	0.7961	1.2386	0.8843	0.8821	0.801
	(0.8349)	(0.9302)	(0.7961)	(1.2382)	(0.8839)	(0.8822)	(0.801)
	[0.813,0.857]	[0.9,0.965]	[0.791,0.801]	[1.207,1.273]	[0.869,0.899]	[0.874,0.89]	[0.792,0.81]
95% VaR	0.9383	1.0821	0.8177	1.3737	0.9462	0.9563	0.844
	(0.9377)	(1.0809)	(0.8175)	(1.3734)	(0.9459)	(0.9562)	(0.844)
	[0.904,0.973]	[1.033,1.134]	[0.811,0.825]	[1.334,1.414]	[0.927,0.965]	[0.945,0.967]	[0.832,0.857]
99% VaR	1.2087	1.4985	0.8662	1.6234	1.0549	1.1271	0.9443
	(1.2053)	(1.491)	(0.8655)	(1.622)	(1.0542)	(1.1263)	(0.9435)
	[1.124,1.311]	[1.366,1.668]	[0.851,0.883]	[1.56,1.693]	[1.026,1.083]	[1.101,1.156]	[0.916,0.976]
90%TVaR	0.9987	1.1803	0.8273	1.4158	0.9638	0.9916	0.8651
	(0.9979)	(1.1776)	(0.8271)	(1.416)	(0.9635)	(0.9913)	(0.8646)
	[0.959,1.04]	[1.121,1.246]	[0.821,0.835]	[1.379,1.453]	[0.948,0.981]	[0.979,1.004]	[0.853,0.878]
95%TVaR	1.1164	1.3622	0.8487	1.5303	1.0144	1.0674	0.9098
	(1.1152)	(1.3587)	(0.8484)	(1.53)	(1.014)	(1.0668)	(0.9093)
	[1.055,1.183]	[1.266,1.468]	[0.839,0.859]	[1.483,1.579]	[0.995,1.034]	[1.049,1.087]	[0.891,0.93]
99%TVaR	1.4311	1.8778	0.8983	1.7503	1.1077	1.2516	1.021
	(1.4181)	(1.8558)	(0.8984)	(1.7488)	(1.1066)	(1.2487)	(1.0175)
	[1.27,1.636]	[1.605,2.216]	[0.876,0.924]	[1.664,1.837]	[1.074,1.146]	[1.201,1.312]	[0.976,1.075]

TABLE 4.6: VaR and TVaR for the five business lines estimated by parametric bootstrap. The values in curved parenthesis are the median of the estimates obtained for the risk measures. The values in squared parenthesis are 95% bootstrap confidence intervals. The column labelled 'Weighted Sum of risk measures' corresponds to the weighted sum of the risk measures (VaR or TVaR) from each business line with weights as at June 2017. The column labelled 'Risk measure of aggregate loss' has the values obtained from the hierarchical aggregation copula model with weights for each business line as at June 2017.

4.5 Analysis of the results: effect of reinsurance

From Table 4.6 we can see that Fire has the largest VaR and TVaR among the five business lines for the gross loss ratios, followed by CTP except for the 99% TVaR, where House has the second largest. When we consider reinsurance, CTP has the largest risk measure values while Fire has the second largest except for the 99% TVaR where Fire still has the largest value. Nevertheless the 99% TVaR for Fire has a staggering reduction after reinsurance. Overall Motor has the lowest values for the risk measures in terms of both gross and net loss ratios, implying the least risky business line. The VaR and TVaR 95% confidence intervals for gross and net losses overlap in the cases of House, CTP and Liability. For Fire, Motor and (copula) aggregate losses the confidence intervals for gross and net losses do not overlap. We conclude that reinsurance is effectively reducing the level of risk only for Fire and Motor. And this reduction is strong enough to carry on to the (copula) aggregate loss. The effect of reinsurance in changing the risk level for House, CTP and Liability is much less pronounced. We come back to this point later in this thesis. It is worthwhile recalling here that the average loss is also not significantly different with and without reinsurance.

Comparing the two right columns of Table 4.6 we can see that the weighted sum of the risk measures, VaR and TVaR, is larger than the value obtained using the hierarchical aggregation copula model. This is true both for VaR and TVaR at all the probability levels considered, and for gross and net loss ratios. The risk measures obtained using the hierarchical copula model incorporate the dependence between the different business lines while the weighted sum of VaR and TVaR does not. Hence, the result obtained is clear evidence that there is a risk reduction effect in the tail when combining the five business lines. This reduction of risk by pooling different business lines (risks) corresponds to the notion of diversification well known in financial portfolio selection and allocation.

Our goal in the following is to explore the effect of reinsurance on the diversification effect by drawing some parallel between a portfolio of financial assets and the set of business lines. When addressing diversification in terms of portfolio selection we can think of two aspects. First, the risk reduction due to the effect of the interaction between the different components, or due to the dependence structure in terms of multivariate probability distributions. Second, for the same components risk reduction is also due to different weights hold by each component. We address these two aspects below separately.

4.5.1 Reinsurance and (dependence) diversification

In order to analyse the effect of reinsurance in the risk reduction due to the dependence structure we calculate the percentage risk reduction from the weighted sum of the risk measures for each business line and the risk of the weighted sum of the business lines. The risk measures for the weighted sum of business lines loss ratios are the ones obtained by the hierarchical aggregation copula model. The weights are fixed and based on the premiums as at June 2017. The results are reported in Table 4.7.

	Weighted Sum of risk measures	Risk measure of aggregate loss, LR_t	Δ (%)
	Gross loss ratios		
90% VaR	0.9603	0.8806	8.30
95% VaR	1.1377	1.0184	10.49
99% VaR	1.8101	1.5937	11.96
90%TVaR	1.4060	1.2644	10.07
95%TVaR	1.7755	1.5895	10.48
99%TVaR	3.4412	3.1437	8.65
	Net loss ratios		
90% VaR	0.8821	0.8010	9.19
95% VaR	0.9563	0.8440	11.74
99% VaR	1.1271	0.9443	16.22
90%TVaR	0.9916	0.8651	12.76
95%TVaR	1.0674	0.9098	14.76
99%TVaR	1.2516	1.0210	18.42

TABLE 4.7: Weighted sum of risk measures for the five business lines compared with the risk measure of the weighted sum of business lines obtained by the hierarchical copula model. The last column of the table gives the percentage reduction, Δ, of the risk measures.

The difference of the VaR and TVaR between the weighted sum and the copula aggregated loss ratios is most striking for the net loss ratios. Note that, from Table 4.6, the 90% confidence intervals for the net loss ratio risk measures of the weighted sum and the copula aggregated sum do not even overlap. The difference in the risk reduction gross and net loss ratios is larger for higher probability risk measures. This indicates that, from a multivariate or portfolio point of view, reinsurance is reducing the upper tail dependence of some of the loss ratios across the different business lines (the estimates for λ_U in Table 4.5 do not contradict this assertion) and consequently is increasing the diversification effect. As far as we know this effect of reinsurance on the multivariate overall portfolio of business lines has not been previously reported in the literature.

4.5.2 Reinsurance and (weighted premiums) diversification

Here we evaluate the effect of reinsurance in the diversification due to unequal weights between the different business lines. We compare the risk measure obtained by the copula aggregation model for the equally weighted aggregated loss ratios and the loss ratios aggregated using the weights as at June 2017. The weights as at June 2017 are reported in Table 4.1 where we can see that reinsurance reduces the proportion of the business line Fire (mainly) and House, and increases the weight of Motor.

We quantify a measure of (weights) diversification as the percentage reduction in risk between an equally weighted portfolio, in this case of business lines, and the portfolio we want to measure the level of diversification. Using this notion we estimate the VaR and TVaR of an equally weighted sum of the five business lines. This corresponds to assume equal premiums across the five business lines. The results, reported in Table 4.8, are striking. Reinsurance vastly reduces the diversification originated from having different weights on the different business lines. As we can see in Table 4.1 reinsurance does not have such a strong effect on the change of weights. But on the other hand, reinsurance changes the multivariate dependence structure making the aggregated loss ratios much less sensitive, in terms of risk, to changes in the relative proportions of each business line.

4.5.3 Cession ratio and risk

The cession ratio is the outwards reinsurance expense divided by the gross earned premium. Our goal here is to explore the relation between the cession rate and the change in the risk implied by reinsurance and measured on the loss ratios. We would expect reinsurance to have the effect of reducing the VaR and TVaR for individual business lines. If we calculate the percentage change for the risk measures at each probability level we can see in Table 4.9 that this is indeed the case for House, Fire and Motor. But remarkably reinsurance increases VaR and TVaR at almost all probability levels for CTP and Liability. We plot in Figure 4.5 the cession rate versus the reduction in risk for each business line. There is a clear positive relation between the two. Remarkably for low cession rates the risk increases with reinsurance.

	Aggregate loss, LR_t Equal Weights	Aggregate loss, LR_t June 2017 Weights	(%) Δ (%)
Gross loss ratios			
90% VaR	0.9364	0.8806	5.96
95% VaR	1.1204	1.0184	9.10
99% VaR	1.9901	1.5937	19.92
90%TVaR	1.4782	1.2644	14.46
95%TVaR	1.9429	1.5895	18.19
99%TVaR	4.1556	3.1437	24.35
Net loss ratios			
90% VaR	0.8380	0.8010	4.42
95% VaR	0.8918	0.8440	5.36
99% VaR	1.0163	0.9443	7.08
90%TVaR	0.9187	0.8651	5.83
95%TVaR	0.9748	0.9098	6.67
99%TVaR	1.1178	1.0210	8.66

Table 4.8: Comparison between the risk measures for the loss ratio aggregated using the copula model. In the second column of the table the business lines have been aggregated using equal weighs, while in the third column the business lines have been aggregated using the weights as at June 2017. The right column of the table gives the percentage reduction of the risk measures.

	House	Fire	Motor	CTP	Liability
90% VaR	-0.80	35.43	3.89	-3.29	0.81
95% VaR	3.20	51.94	8.44	-4.96	-0.48
99% VaR	11.45	75.84	18.44	-7.87	-1.96
90%TVaR	6.05	71.40	11.03	-5.55	-0.65
95%TVaR	9.63	78.97	15.35	-6.85	-1.41
99%TVaR	17.01	89.84	24.50	-9.11	-2.22
Average Δ (%)	7.76	67.24	13.61	-6.27	-0.99
Average cession ratio (%)	30.65	43.50	16.95	11.00	20.34
	(2.83)	(3.90)	(1.35)	(7.35)	(6.83)
Average recovery ratio (%)	21.01	37.73	11.81	13.47	23.33
	(10.37)	(18.88)	(2.80)	(16.52)	(16.74)

Table 4.9: Percentage reduction (Δ) on the risk measures from gross to net loss ratios. The cession ratio is outwards reinsurance expense divided by the gross earned premium averaged across the sample period. The recovery ratio is the reinsurance recoveries revenue divided by the gross incurred claims. The values in parenthesis are the standard deviation of the corresponding rates.

4.6 Conclusion

It is important for every insurance company to determine and maintain the right amount of capital to keep as a solvency margin against the risk of not being able

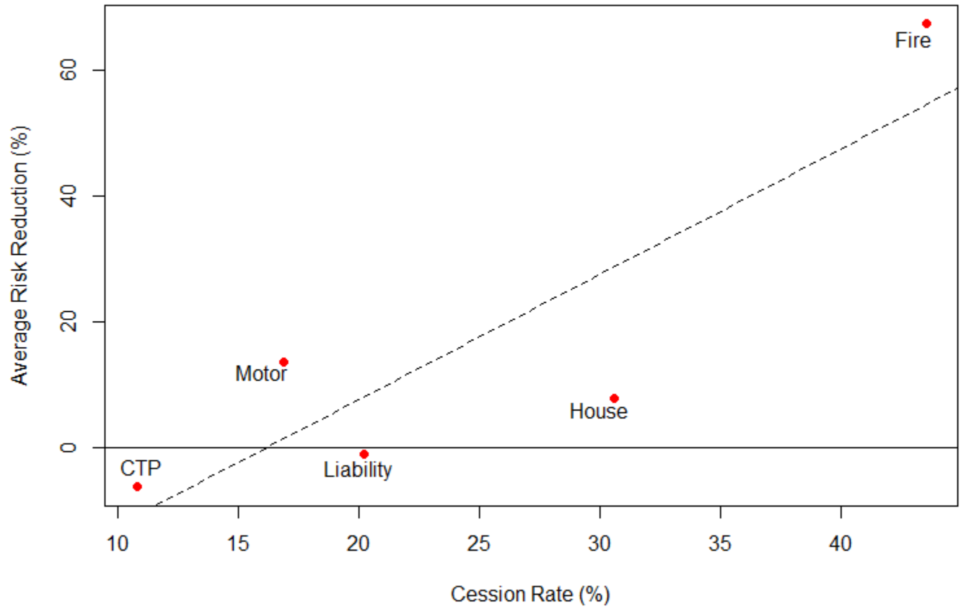


FIGURE 4.5: Cession rate versus the reduction on risk measures caused by reinsurance for each business line. The dashed line results from regressing risk reduction from reinsurance on the cession rate.

of covering the insurance company's liabilities. This calls for adequate methods of aggregating all risks and the use of appropriate risk measures to determine the capital requirement. In this chapter we use a hierarchical aggregation copula model to address the dependence structure of the different insurance business lines. We use several copula families to model the aggregated loss with particular emphasis on the tail dependence. We consider a range of copulas asymmetric, symmetric, with and without tail dependence as the Gaussian and Student-t, and Archimedian copulas Clayton, Gumbel, and Frank. Selecting the best copula families for the hierarchical aggregation model is crucial as it influences the estimated level of risk and consequently avoids overestimation or underestimation of the capital required.

A very important tool for risk management is reinsurance. Insurance companies diversify away part of its underwriting risk to reinsurance companies. In this chapter we investigate the effect and relevance of reinsurance on the risk of individual business lines and importantly on the aggregate risk. These effects are measured in this chapter by considering both gross and net loss ratios, where gross loss ratios are used to measure the insurance risk without considering the reinsurance business, while the net loss ratios are used to determine the insurance risk taking into account the reinsurance business.

Reinsurance can increase the risk even when measured by the standard deviation as we can see in Table 4.1. Unless there is a positive shift in the relation between the premium cession rate and the risk reduction from reinsurance, it is less risky to have a higher cession rate or not to have reinsurance at all than to have a low cession rate. For House, Fire and Motor the reinsurance premium cession rate is higher than the claims recovery rate. We would assume that this is to compensate for the expenses of the reinsurance company. But interestingly for CTP and Liability the recovery rate is higher is than the cession rate which counteracts the increase in risk brought by reinsurance for these two business lines.

Another aspect of reinsurance has to do with diversification. Reinsurance increases the diversification due to the dependence between the business lines and reduces the sensitivity of the aggregate risk to changes in the proportions of the different business lines. Hence, if the goal is to manage risk by changing the proportion of premiums between business lines, reinsurance might mitigate the reduction of risk of the aggregated loss ratio. Instead, a risk management strategy focussed on the dependence structure between the business lines should be more successful in reducing the risk of the aggregated loss ratio when reinsurance is present.

Chapter 5

Vine copula

5.1 Introduction

In the search for a flexible copula model for a high dimension, one might want to consider the Pair-Copula Construction (PCC) or also known as vine copula. Vine copula can be used for high dimensional data sets and can incorporate complex dependence structures. These include negative dependence and different types of copula dependence such as 90%, 180% and 270% rotated copulas. The building block of vine copula is bivariate copulas which is similar to hierarchical model in Chapter 4.

In this chapter, we develop a new model to determine the capital requirement for general insurance companies in the United Kingdom (UK) using a vine copula. We develop the model by incorporating the most commonly used copulas from the Archimedean copula family such as Clayton copula, Gumbel copula, and Frank copula. The results from these analysis are presented in Section 5.5. The rest of this chapter is organised as follows. In Section 5.2, we discuss the recent literature on vine copula and highlight the key contribution of vine copula to the financial and insurance industries. In Section 5.3, we discuss and provide a comprehensive explanation on the concepts of vine copula which are useful in developing the vine copula model for capital requirement. The modelling procedures are explained in Section 5.4 and the empirical analysis to determine the capital requirement for UK general insurance industry is presented in Section 5.5. We end this chapter with the Conclusion in Section 5.6.

5.2 The growing literature on vine copula

The Pair-Copula Construction (PCC), also known as vine copula, has been evolving in the recent years. Vine copula was first introduced by Joe (1996) and later developed by Bedford and Cooke (2001), Bedford and Cooke (2002) and more recently by Aas et al. (2009), Hobæk Haff et al. (2010), Czado et al. (2012), Brechmann and Schepsmeier (2013), Dißmann et al. (2013), Killiches et al. (2017). Motivated from the first research on vine copula by Joe (1996), Bedford and Cooke (2001) analyse the construction of the copula based on its density function and present vine copula in a graphical form with two of its subclasses known as Drawable-vine (D-vine) and Canonical-vine (C-vine). These vine copulas are different from one another by their respective tree structures. In a D-vine, each node in a tree has a maximum of two edges while in a C-vine, each tree has a unique node that connects with other remaining nodes by an edge. Bedford and Cooke (2001) show that, the structure of a vine copula is developed by generalizing the concept of Markov trees. However, the conditional independence in vine copula structure is weakened in order to allow for different kind of conditional dependence.

According to Acar et al. (2012), vine copulas are graphical models and through systematic procedures a high dimensional copula can be decomposed into a lower dimensional copula (bivariate copula). Unlike the hierarchical model in Chapter 4, the decomposition of vine copula includes conditional copulas. As mentioned in the previous section, vine copula has the advantage over other multivariate copula models due to its high flexibility. Specifically, it is adequate to model a range of complex dependencies such as asymmetric dependence or strong joint tail dependence (Joe et al., 2010). In particular, Aas et al. (2009) provide a comprehensive studies on the superiority of vine copula against Nested Archidemedean Copula (NAC). NAC is originally proposed by Joe (1997) and also studied by Embrechts et al. (2003), Whelan (2004), Nelsen (2006), Mcneil (2008), and Hofert (2008). The results of Aas et al. (2009) suggest that vine copula is preferable to model high dimensional multivariate data as it is more efficient than NAC. Moreover, vine copula provides more flexibility by allowing free specification of $d(d-1)/2$ copulas for the d-dimensional case and provide a wide selection of copula families. These results match those observed in the earlier study by Fischer et al. (2009).

The key concept of vine copula is on the simplifying assumption. In vine copula model, it is assumed that every conditional bivariate copula is independent of the conditioning variable unless through its own marginal distributions (Hobæk Haff et al., 2010). This assumption is important so that it is possible to get a good inference. In fact, a growing literature on vine copula recently have been focusing on this assumption (Hobæk Haff et al., 2010; Acar et al., 2012; Kraus and Czado, 2017 and Killiches et al., 2017). Hobæk Haff et al. (2010) show that the simplifying assumption on vine copula model is a good approximation and propose the conditions for a multivariate distribution to be in this simplified form. On the contrary, Acar et al. (2012) graphically prove that a pair of variables are mistakenly considered as conditionally independent although otherwise. However, it is important to note that Acar et al. (2012) mainly rely on a graphical tool for the study. This shows that more work is needed to confirm this claim. Interestingly, a recent study by Killiches et al. (2017) provide evidence that vine copula with simplifying assumption is preferable to model high dimensional data with the view to avoid over fitting without failing to accurately measure the dependence structure. Further details on this assumption are discussed in Section 5.3.1.

The superiority of vine copula also relies on its statistical inference which has received great attention in the recent literature (Barthel et al., 2018; Hobæk Haff and Segers, 2015; Gruber and Czado, 2015; Erhardt et al., 2015; Vaz De Melo Mendes and Accioly, 2014; Schmidl et al., 2013; Brechmann and Schepsmeier, 2013; Hobæk Haff, 2013; Czado et al., 2012; Min and Czado, 2010, and many others). However, the main inference methods discussed in the literature are centrally focused on maximum likelihood and sequential estimation. In this chapter, we use the improvised sequential estimation method proposed by Czado et al. (2012). This method allows for two-step parameter estimation in each vine copula tree structure. In the first step, the parameters for the marginal distribution of each random variable in the first tree are estimated. Then, the copula parameters linking two random variables are estimated accordingly. This method proceeds sequentially to the next step by estimating the parameters for marginals and copulas for the next tree using the estimates from the previous tree. This sequential estimation method, which was originally introduced by Aas et al. (2009), provides a good starting value for the maximum likelihood estimation in vine copulas. From the computational perspective, this method is the best fitting estimation as confirmed by the algorithm developed by Brechmann and Schepsmeier (2013).

Vine copula has great flexibility in modelling high dimensional problem making it suitable for solving statistical problems not just in Finance and Insurance but also in sociology (Cooke et al., 2015), biology (Barthel et al., 2018; Schellhase and Spanhel, 2018) and hydrology (Erhardt and Czado, 2018; Hobæk Haff and Segers, 2015; Killiches and Czado, 2015). In Finance, Czado et al. (2012) study the inference of mixed C-vines using maximum likelihood (ML) estimation and apply the model to solve the problem involving the USD exchange rate. Dißmann et al. (2013) use R-vines and develop a new algorithm to investigate the joint-density of financial asset classes returns from equity, fixed income and commodity indices.

The past literature in insurance applications of vine copula mainly focus on individual business line, e.g., Shi and Yang (2018) investigate the property insurance claims using mixed D-vine copula to model the temporal dependence among recurring observations. The result is used to incorporate policyholders' past experience into future premiums. In addition, Erhardt and Czado (2012) analyse the dependence between yearly claim totals and different coverages. These include age, sex, and other factors. It is noticed that there have been limited applications in modelling capital requirement and dependence structure for insurance, especially the general insurance segment in the UK. This chapter provides a new contribution to the literature by empirically analyse and subsequently model the capital requirement for UK general insurance industry. We begin our research by providing a comprehensive overview of the vine copula key concepts in the following sections.

5.3 Vine copula

Let F be the joint probability distribution function with marginals $F_1, F_2, ..., F_n$ of the vector $\boldsymbol{X} = (X_1, X_2, ..., X_n)$ of n random variables. According to Theorem 3.2 of Sklar (1959) in Section 3.1, the joint probability function is given by

$$F(x_1, x_2, ..., x_n) = C(F_1(x_1), F_2(x_2)..., F_n(x_n)), \tag{5.1}$$

where C is the n-dimensional copula with the following expression

$$C(u_1, ..., u_n) = F\big(F_1^{-1}(u_1), ..., F_n^{-1}(u_n)\big). \tag{5.2}$$

The F^{-1} denotes the inverse distribution function of the marginals. Consequently, the joint density function of vector $\boldsymbol{X}$ is

$$f(x_1, x_2, ..., x_n) = c_{12...n}(F_1(x_1), F_2(x_2), ..., F_n(x_n)) \cdot f_1(x_1) \cdots f_n(x_n), \tag{5.3}$$

assuming that the copula density function $c_{12...n}$ exists. We decompose the density function in Equation (5.3) as

$$f(x_1, x_2, ..., x_n) = f_n(x_n) \cdot f(x_{n-1}|x_n) \cdot f(x_{n-2}|x_{n-1}, x_n) \cdots f(x_1|x_2, ..., x_n). \tag{5.4}$$

For 2-dimensional ($n = 2$), we write

$$f(x_1, x_2) = f_2(x_2) \cdot f(x_1|x_2). \tag{5.5}$$

Similarly, we can re-write Equation (5.5) as

$$f(x_1|x_2) = \frac{f(x_1, x_2)}{f_2(x_2)}, \tag{5.6}$$

and by Sklar Theorem we can write

$$\begin{aligned} f(x_1|x_2) &= \frac{c_{12}[F_1(x_1), F_2(x_2)] \cdot f_1(x_1) \cdot f_2(x_2)}{f_2(x_2)} \\ &= c_{12}[F_1(x_1), F_2(x_2)] \cdot f_1(x_1). \end{aligned} \tag{5.7}$$

For further example consider 3-dimensional ($n = 3$) and using Equation (5.4), we write

$$f(x_1, x_2, x_3) = f_3(x_3) \cdot f(x_2|x_3) \cdot f(x_1|x_2, x_3), \tag{5.8}$$

and we decompose the conditional density function as in the following equation

$$\begin{aligned} f(x_1|x_2, x_3) &= \frac{f(x_1, x_3|x_2)}{f(x_3|x_2)} \\ &= \frac{c_{13|2}[F(x_1|x_2), F(x_3|x_2); x_2] \cdot f(x_1|x_2) \cdot f(x_3|x_2)}{f(x_3|x_2)} \\ &= c_{13|2}[F(x_1|x_2), F(x_3|x_2); x_2] \cdot f(x_1|x_2) \\ &= c_{13|2}[F(x_1|x_2), F(x_3|x_2); x_2] \cdot c_{12}[F(x_1), F(x_2)] \cdot f_1(x_1). \end{aligned} \tag{5.9}$$

Finally, by replacing the conditional density function for 2-dimensional in Equation (5.7) and 3-dimensional in Equation (5.9) into Equation (5.8), we get

$$\begin{aligned} f(x_1, x_2, x_3) = f_1(x_1) \cdot f_2(x_2) \cdot f_3(x_3) \\ \cdot c_{12}[F(x_1), F(x_2)] \\ \cdot c_{23}[F(x_2), F(x_3)] \\ \cdot c_{13|2}[F(x_1|x_2), F(x_3|x_2); x_2]. \end{aligned} \tag{5.10}$$

Note that the joint density function is decomposed into marginal terms, unconditional pairs and also conditional pair. The general formula

$$f(x|\boldsymbol{v}) = c_{xv_j|\boldsymbol{v}_{-j}}[F(x|\boldsymbol{v}_{-j}), F(v_j|\boldsymbol{v}_{-j}); \boldsymbol{v}_{-j}] \cdot f(x|\boldsymbol{v}_{-j}), \tag{5.11}$$

where v_j is a component, arbitrary chosen from a d-dimensional vector $\boldsymbol{v}$ and $\boldsymbol{v}_{-j}$ denotes the $(d-1)$-dimensional vector $\boldsymbol{v}$ excluding v_j.

5.3.1 The simplifying assumptions

To model a vine copula, it is important to assume that the bivariate copula $C(\theta_1, \theta_2)$ does not depend on the conditional variables but only through its distribution function F (Hobæk Haff et al., 2010).

For illustration, consider the conditional bivariate copula of 3-dimensional case from Equation (5.9). The conditional copula density is given by

$$c_{13|2}\left[F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2); x_2\right] = c_{13|2}\left[F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2)\right]. \tag{5.12}$$

Generally, the conditional bivariate copula density from Equation (5.11) becomes as

$$c_{xv_j|\boldsymbol{v}_{-j}}[F(x|\boldsymbol{v}_{-j}), F(v_j|\boldsymbol{v}_{-j}); \boldsymbol{v}_{-j}] = c_{xv_j|\boldsymbol{v}_{-j}}[F(x|\boldsymbol{v}_{-j}), F(v_j|\boldsymbol{v}_{-j})]. \tag{5.13}$$

Clearly, we assume that the families of each bivariate copulas are constant or independent over the values of its corresponding conditioning variables. These assumptions are important for fast and robust vine copula inference. We discuss vine copula inference in Section 5.4.3.

5.3.2 Conditional distribution function for modelling vine copulas

From the previous section, it is clear that the conditional distribution function (cdf) in the form $F(x|v)$ involves in modelling the PCC. To resolve the cdf, Joe (1996) proves that for every v_j from the vector $\boldsymbol{v}$, $F(x|\boldsymbol{v})$ is given by

$$F(x|\boldsymbol{v}) = \frac{\partial C_{xv_j|\boldsymbol{v}_{-j}}(F(x|\boldsymbol{v}_{-j}), F(v_j|\boldsymbol{v}_{-j}))}{\partial F(v_j|\boldsymbol{v}_{-j})}, \tag{5.14}$$

where $C_{xv_j|v_{-j}}$ is the bivariate copula distribution function. In the case of univariate distribution where $\boldsymbol{v}$ is univariate and for simplification consider $x = x_1$ and $v = v_j = x_2$, it follows that

$$F(x_1|x_2) = \frac{\partial C_{x_1x_2}(F(x_1), F(x_2))}{\partial F(x_2)}. \tag{5.15}$$

5.3.3 *h*-function

For uniform margin where $x, v \sim U(0,1)$, we can get $f(x_1) = f(x_2) = 1, F(x_1) = x_1$ and $F(x_2) = x_2$. Subsequently, following h-function introduced by Aas et al. (2009), Equation (5.15) is written as

$$h(x_1, x_2) = \frac{\partial C_{x_1x_2}(x_1, x_2)}{\partial x_2}. \tag{5.16}$$

To explain the importance of solving the conditional distribution functions, we consider a multivariate distribution with d=4. In this case $x_1, x_2, x_3, x_4 \sim U(0,1)$ and using Equation (5.14) and the h-function described in (5.16), we have

$$\begin{aligned} F(x_1|x_2, x_3, x_4) &= \frac{\partial C_{x_1x_2|x_3x_4}(F(x_1|x_3, x_4), F(x_2|x_3, x_4))}{\partial F(x_2|x_3, x_4)} \\ &= h_{x_1x_2|x_3x_4}\big(F(x_1|x_3, x_4), F(x_2|x_3, x_4)\big). \end{aligned} \tag{5.17}$$

Equation (5.17) is one possible outcome for a 4-dimensional case. In total, $d!/2$ can be constructed from a d-dimensional random vector (Aas et al., 2009). Based on Equation (5.17) to derive $F(x_1|x_2, x_3, x_4)$ we need to evaluate the conditional distribution function of $F(x_1|x_3, x_4)$ and $F(x_2|x_3, x_4)$. These are

$$\begin{aligned} F(x_1|x_3, x_4) &= \frac{\partial C_{x_1|x_3x_4}\big(F(x_1|x_3), F(x_1|x_4)\big)}{\partial F(x_1|x_4)}, \\ &= h_{x_1|x_3x_4}\big(F(x_1|x_3), F(x_1|x_4)\big) \end{aligned} \tag{5.18}$$

$$\begin{aligned} F(x_2|x_3, x_4) &= \frac{\partial C_{x_2|x_3x_4}\big(F(x_2|x_3), F(x_2|x_4)\big)}{\partial F(x_2|x_4)}. \\ &= h_{x_2|x_3x_4}\big(F(x_2|x_3), F(x_2|x_4)\big) \end{aligned} \tag{5.19}$$

Next, following the same procedure, we need to evaluate the univariate distribution functions: $F(x_1|x_3), F(x_1|x_4), F(x_2|x_3), F(x_2|x_4)$. We get the following

$$\begin{aligned} F(x_1|x_3) &= \frac{\partial C_{x_1x_3}\big(F(x_1), F(x_3)\big)}{\partial F(x_3)}, \\ &= h_{x_1x_3}\big(F(x_1), F(x_3)\big) \end{aligned} \tag{5.20}$$

$$\begin{aligned} F(x_1|x_4) &= \frac{\partial C_{x_1x_4}\big(F(x_1), F(x_4)\big)}{\partial F(x_4)}, \\ &= h_{x_1x_4}\big(F(x_1), F(x_4)\big) \end{aligned} \tag{5.21}$$

and

$$\begin{aligned} F(x_2|x_3) &= \frac{\partial C_{x_2x_3}\big(F(x_2), F(x_3)\big)}{\partial F(x_3)}, \\ &= h_{x_2x_3}\big(F(x_2), F(x_3)\big) \end{aligned} \tag{5.22}$$

$$F(x_2|x_4) = \frac{\partial C_{x_2x_4}\big(F(x_2), F(x_4)\big)}{\partial F(x_4)}. \\ = h_{x_2x_4}\big(F(x_2), F(x_4)\big) \tag{5.23}$$

Then, we summarize the above equations as a nested h-functions

$$\begin{aligned} F(x_1|x_2, x_3, x_4) &= \frac{\partial C_{x_1x_2|x_3x_4}\big(F(x_1|x_3, x_4), F(x_2|x_3, x_4)\big)}{\partial F(x_2|x_3, x_4)} \\ &= h_{x_1x_2|x_3x_4}\big[F(x_1|x_3, x_4), F(x_2|x_3, x_4)\big] \\ &= h_{x_1x_2|x_3x_4}\big[h_{x_1|x_3x_4}\big(F(x_1|x_3), F(x_1|x_4)\big), h_{x_2|x_3x_4}\big(F(x_2|x_3), F(x_2|x_4)\big)\big] \\ &= h_{x_1x_2|x_3x_4}\Big[h_{x_1|x_3x_4}\Big(h_{x_1x_3}\big(F(x_1), F(x_3)\big), h_{x_1x_4}\big(F(x_1), F(x_4)\big)\Big), \\ &\quad h_{x_2|x_3x_4}\Big(h_{x_2x_3}\big(F(x_2), F(x_3)\big), h_{x_2x_4}\big(F(x_2), F(x_4)\big)\Big)\Big]. \end{aligned} \tag{5.24}$$

For $d = 5$, using the same procedure we have the following

$$\begin{aligned} F(x_1|x_2, x_3, x_4, x_5) &= \frac{\partial C_{x_1x_5|x_2x_3x_4}\big[F(x_1|x_2, x_3, x_4), F(x_5|x_2, x_3, x_4)\big]}{\partial F(x_5|x_2, x_3, x_4)} \\ &= h_{x_1x_5|x_2x_3x_4}\big[F(x_1|x_2, x_3, x_4), F(x_5|x_2, x_3, x_4)\big]. \end{aligned} \tag{5.25}$$

Then using Equation (5.24) we have

$$\begin{aligned} F(x_1|x_2, x_3, x_4) &= \frac{\partial C_{x_1x_2|x_3x_4}\big(F(x_1|x_3, x_4), F(x_2|x_3, x_4)\big)}{\partial F(x_2|x_3, x_4)} \\ &= h_{x_1x_2|x_3x_4}\big[F(x_1|x_3, x_4), F(x_2|x_3, x_4)\big] \\ &= h_{x_1x_2|x_3x_4}\big[h_{x_1|x_3x_4}\big(F(x_1|x_3), F(x_1|x_4)\big), h_{x_2|x_3x_4}\big(F(x_2|x_3), F(x_2|x_4)\big)\big] \\ &= h_{x_1x_2|x_3x_4}\Big[h_{x_1|x_3x_4}\Big(h_{x_1x_3}\big(F(x_1), F(x_3)\big), h_{x_1x_4}\big(F(x_1), F(x_4)\big)\Big), \\ &\quad h_{x_2|x_3x_4}\Big(h_{x_2x_3}\big(F(x_2), F(x_3)\big), h_{x_2x_4}\big(F(x_2), F(x_4)\big)\Big)\Big]. \end{aligned} \tag{5.26}$$

Similarly,

$$
\begin{aligned}
F(x_5|x_2,x_3,x_4) &= \frac{\partial C_{x_5x_2|x_3x_4}\big(F(x_5|x_3,x_4),F(x_2|x_3,x_4)\big)}{F(x_2|x_3,x_4)} \\
&= h_{x_5x_2|x_3x_4}\big[F(x_5|x_3,x_4),F(x_2|x_3,x_4)\big] \\
&= h_{x_5x_2|x_3x_4}\big[h_{x_5|x_3x_4}\big(F(x_5|x_3),F(x_5|x_4)\big),h_{x_2|x_3x_4}\big(F(x_2|x_3),F(x_2|x_4)\big)\big] \\
&= h_{x_5x_2|x_3x_4}\Big[h_{x_5|x_3x_4}\Big(h_{x_5x_3}\big(F(x_5),F(x_3)\big),h_{x_2x_4}\big(F(x_2),F(x_4)\big)\Big), \\
&\qquad h_{x_2|x_3x_4}\Big(h_{x_2x_3}\big(F(x_2),F(x_3)\big),h_{x_2x_4}\big(F(x_2),F(x_4)\big)\Big)\Big].
\end{aligned}
\tag{5.27}
$$

Finally, we get

$$
\begin{aligned}
F(x_1|x_2,x_3,x_4,x_5) &= \frac{\partial C_{x_1x_5|x_2x_3x_4}\big[F(x_1|x_2,x_3,x_4),F(x_5|x_2,x_3,x_4)\big]}{\partial F(x_5|x_2,x_3,x_4} \\
&= h_{x_1x_5|x_2x_3x_4}\big[F(x_1|x_2,x_3,x_4),F(x_5|x_2,x_3,x_4)\big] \\
&= h_{x_1x_5|x_2x_3x_4}\Bigg[h_{x_1x_2|x_3x_4}\Big[h_{x_1|x_3x_4}\Big(h_{x_1x_3}\big(F(x_1),F(x_3)\big),h_{x_1x_4}\big(F(x_1),F(x_4)\big)\Big), \\
&\qquad h_{x_2|x_3x_4}\Big(h_{x_2x_3}\big(F(x_2),F(x_3)\big),h_{x_2x_4}\big(F(x_2),F(x_4)\big)\Big)\Big], \\
&\qquad h_{x_5x_2|x_3x_4}\Big[h_{x_5|x_3x_4}\Big(h_{x_5x_3}\big(F(x_5),F(x_3)\big),h_{x_2x_4}\big(F(x_2),F(x_4)\big)\Big), \\
&\qquad h_{x_2|x_3x_4}\Big(h_{x_2x_3}\big(F(x_2),F(x_3)\big),h_{x_2x_4}\big(F(x_2),F(x_4)\big)\Big)\Big]\Bigg].
\end{aligned}
\tag{5.28}
$$

5.3.4 h-function of selected copula families

Here we present the h-function of Clayton, Gumbel and Frank copula.

Clayton copula

Recall that

$$C(u,v) = (u^{-\theta}+v^{-\theta}-1)^{-\frac{1}{\theta}}. \tag{5.29}$$

Then, we have

$$
\begin{aligned}
h(u_1, u_2; \theta_{12}) &= \frac{\partial C(u_1, u_2; \theta_{12})}{\partial u_2} \\
&= \frac{\partial}{\partial u_2}(u_1^{-\theta_{12}} + u_2^{-\theta_{12}} - 1)^{-\frac{1}{\theta_{12}}} \\
&= -\frac{1}{\theta_{12}}(u_1^{-\theta_{12}} + u_2^{-\theta_{12}} - 1)^{-\frac{1}{\theta_{12}}-1}(-\theta_{12}u_1^{-\theta_{12}-1}) \\
&= -(\theta_{12}^{-1})(-\theta_{12}u_1^{-\theta_{12}-1})(u_1^{-\theta_{12}} + u_2^{-\theta_{12}} - 1)^{-\frac{1}{\theta_{12}}-1} \\
&= u_1^{-\theta_{12}-1}(u_1^{-\theta_{12}} + u_2^{-\theta_{12}} - 1)^{-\frac{1}{\theta_{12}}-1}.
\end{aligned}
\tag{5.30}
$$

Gumbel copula

Recall that

$$
C(u_1, u_2; \theta_{12}) = \exp\Big[- \big((-\log u_1)^{\theta_{12}} + (-\log u_2)^{\theta_{12}}\big)^{\frac{1}{\theta_{12}}}\Big], \tag{5.31}
$$

then h-function is given by

$$
\begin{aligned}
h(u_1, u_2; \theta_{12}) &= \frac{\partial C(u_1, u_2; \theta_{12})}{\partial u_2} \\
&= \frac{\partial}{\partial u_2}\left(\exp\Big[- \big((-\log u_1)^{\theta_{12}} + (-\log u_2)^{\theta_{12}}\big)^{\frac{1}{\theta_{12}}}\Big]\right) \\
&= \left\{ \left(\exp\Big[- \big((-\log u_1)^{\theta_{12}} + (-\log u_2)^{\theta}_{12}\big)^{\frac{1}{\theta_{12}}}\Big]\right)\right. \\
&\quad \cdot \left(\frac{1}{u_2}(-\log u_2)^{\theta_{12}-1} \right) \\
&\quad \cdot \left.\left((-\log u_1)^{\theta}_{12} + (-\log u_2)^{\theta}_{12} \right)^{\frac{1}{\theta_{12}}-1} \right\}.
\end{aligned}
\tag{5.32}
$$

Frank copula

Recall that

$$
C(u_1, u_2) = -\frac{1}{\theta}\log\Big[1 + \frac{\big(\exp(-\theta u_1) - 1\big)\big(\exp(-\theta u_2) - 1\big)}{\exp(-\theta) - 1}\Big], \tag{5.33}
$$

then the h-function is given by

$$\begin{aligned} h(u_1, u_2; \theta_{12}) &= \frac{\partial C(u_1, u_2; \theta_{12})}{\partial u_2} \\ &= -\frac{1}{\theta}\left[\frac{\exp(-\theta)-1}{\big(\exp(-\theta)-1\big)+\big(\exp(-\theta u_1)-1\big)\big(\exp(-\theta u_2)-1\big)}\right] \\ &\quad \cdot \frac{\big(\exp(-\theta)-1\big)\exp(-\theta u_2)(-\theta)}{\big(\exp(-\theta)-1\big)} \\ &= \frac{\big(\exp(-\theta u_1)-1\big)\exp(-\theta u_2)}{\big(\exp(-\theta)-1\big)+\big(\exp(-\theta u_1)-1\big)(\exp(-\theta u_2)-1)\big)}. \end{aligned} \tag{5.34}$$

5.4 Vine copula models

Vine copula has been very useful in constructing multivariate distribution involving high dimensional random variables. In this regard, Bedford and Cooke (2001) introduce three different vine copulas and the constructions of these vine copulas are in the graphical form. The first type of vine copula with broad range of possible pair-copula decompositions is known as Regular-Vine (R-Vine). The structure of a R-vine is quite complex. However, to simplify the multivariate modelling using vine copula, R-vine can be further divided into two simpler types: Canonical-vine (C-vine) and Drawable-vine (D-vine). As discussed in Section 5.2, C-vine and D-vine are different by their respective structure and construction method. D-vine limits number of edges to every node to a maximum of two edges while in a C-vine, all nodes are connected with a root node to form a pairwise of random variables. In this chapter, we concentrate on C-vine copula to model the dependence structure of losses from multiple insurance business lines. We choose C-vine due to the fact that it is relatively new in insurance application especially in modelling insurance risks.

5.4.1 C-vine

To build a C-vine tree structure, we need to determine the root node of each tree through sequential estimation tree-by-tree. In a d-dimensional C-vine, each root node in every tree, $T_n, n = 1, ..., d-1$ is connected to $d-n$ edges. The density function for d-dimensional C-vine is given by

$$f(x_1,...,x_d)=\prod_{k=1}^{d}f(x_k)\prod_{n=1}^{d-1}\prod_{m=1}^{d-n}c_{n,n+m|1,...,n-1}\Big[F(x_n|x_1,...,x_{n-1}),F(x_{n+m}|x_1,...,x_{n-1})\Big]. \tag{5.35}$$

For the 5-dimensional case we have the following density function

$$\begin{aligned} f(x_1,x_2,x_3,x_4,x_5) = {} & f_1(x_1)\cdot f_2(x_2)\cdot f_3(x_3)\cdot f_4(x_4)\cdot f_5(x_5) \\ & \cdot c_{12}[F(x_1),F(x_2)]\cdot c_{13}[F(x_1),F(x_3)] \\ & \cdot c_{14}[F(x_1),F(x_4)]\cdot c_{15}[F(x_1),F(x_5)] \\ & \cdot c_{23|1}[F(x_2|x_1),F(x_3|x_1)]\cdot c_{24|1}[F(x_2|x_1),F(x_4|x_1)] \\ & \cdot c_{25|1}[F(x_2|x_1),F(x_5|x_1)]\cdot c_{34|12}[F(x_3|x_1,x_2),F(x_4|x_1,x_2)] \\ & \cdot c_{35|12}[F(x_3|x_1,x_2),F(x_5|x_1,x_2)] \\ & \cdot c_{45|123}[F(x_4|x_1,x_2,x_3),F(x_5|x_1,x_2,x_3)]. \end{aligned} \tag{5.36}$$

Figure 5.1 shows a complete structure of a C-vine based on 5-dimensional random variable $(X_1,X_2,X_3,X_4,X_5)'$. In the first tree, $C_{12},C_{13},C_{14},C_{15}$ are the unconditional bivariate copulas for pairwise $(X_1,X_2),(X_1,X_3),(X_1,X_4)$ and (X_1,X_5) respectively. The conditional bivariate copulas are presented in the second, third and forth tree only with notation $C_{23|1},C_{24|1},C_{25|1}$ for the second tree, $C_{34|12},C_{34|12}$ for third tree and $C_{45|123}$ for the forth (final) tree. We explain step-by-step in Section 5.4.2 the procedure to construct the structures.

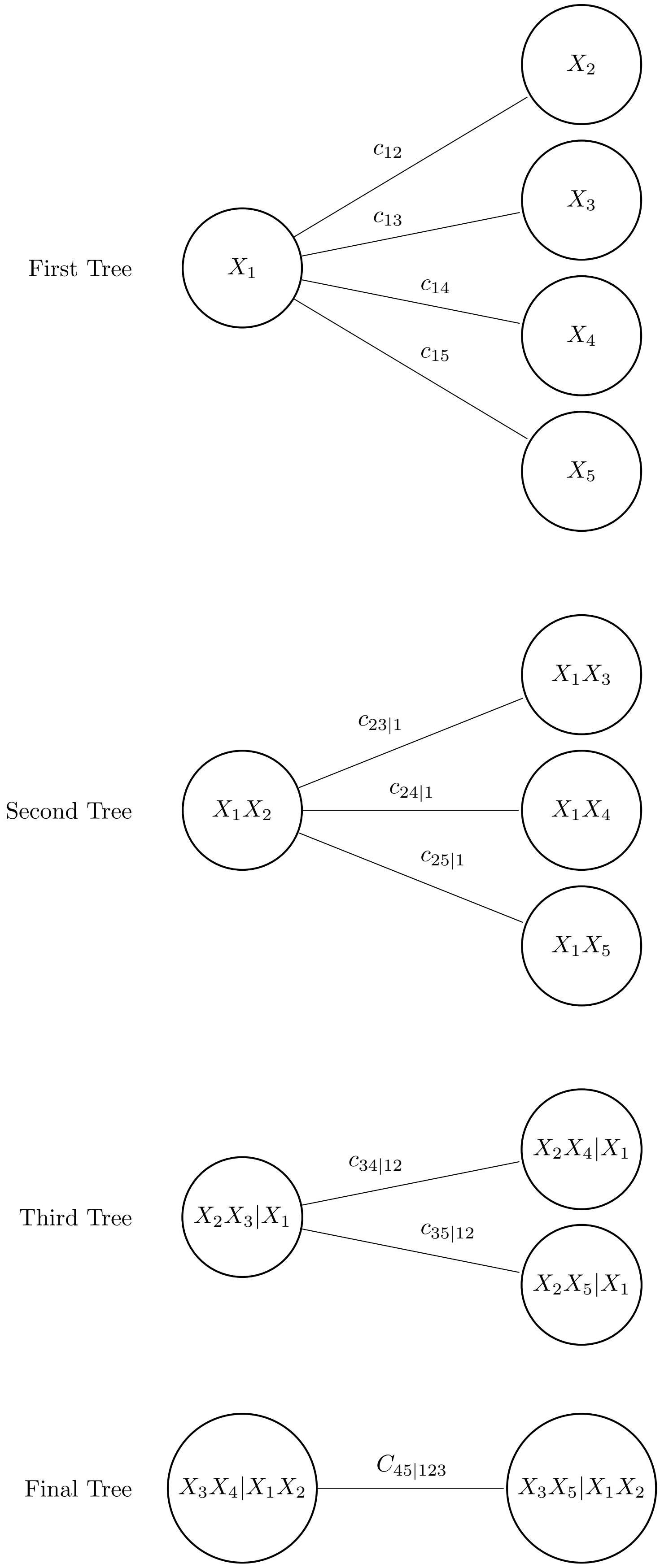


FIGURE 5.1: Illustration of a complete structure of a 5-dimensional C-vine

.

5.4.2 Building a C-vine tree

Here we explain the main procedure to build a tree structure of a C-vine copula. A C-vine copula model is different from a hierarchical model, discussed in Chapter 4. In a vine copula model, each stage is considered as a tree and the number of trees is determined by the total number of dimension, d, minus 1 or specifically, $d-1$. Each variable in every tree is represented by a node. Each tree consists of a root node which is paired to other remaining nodes by an edge and a copula. A root node is the key node which is linked to other remaining nodes. The basic concept of building a C-vine goes back to the Graph Theory discussed in Section 3.5. As explained in Section 5.3, two types of copulas are considered in vine copula. The first type is unconditional copula and only present in the first tree. The second type is conditional copula for the remaining trees. At this point, it is important to note that, the conditional copulas are assumed to be independent from the conditioning variables except through its marginal distribution.

The first step to build a C-vine model is to choose a variable which represents the root node of the first tree. The selection of the root node is performed by selecting the variable with the highest sum of absolute value of Kendall's tau against other variables. In order words, we estimate the pairwise Kendall's tau value for all possible pairs (X_m, X_n), and select the variable X_m that maximizes the following

$$\hat{T}_m := \sum_{n=1}^{d} |\hat{\tau}_{m,n}|, \tag{5.37}$$

where $\hat{\tau}_{m,m} = 1$ for $m = 1, 2, ..., d$. Once X_m is identified, we rearrange the order of all variables. In this way, the variable X_m as the root node will be the first variable and can be linked to other variables. Then we need to select the appropriate unconditional bivariate copulas $c_{k,n}$ with $n = 1, ..., d-1$ for each pair before estimating the next root node for the second tree. At this stage, we also estimate the bivariate copula parameter $\theta_{1,0}^{SE}$. We denote $\theta_{n,0}^{SE}$ as the bivariate copula parameter at $T_n, n = 1, ..., d-1$ using sequential estimation (SE). The selection of the bivariate copula is based on Akaike Information Criterion(AIC) of Akaike (1974). We discuss the bivariate copula selection later in Section 5.4.4.

We proceed using the same procedure to determine the root node in the second tree. However, the root node is determined using bivariate copula parameter $\hat{\theta}_{1,0}^{SE}$ chosen in the first tree and the transformed variables using h-function as explained earlier in Section 5.3.3 and 5.3.4. The $(d-1)th$ transformed variables and sample t for the second tree are determined by

$$\hat{v}_{n+2,t} := h(u_{n+2,t}|u_{k,t};\hat{\theta}^{SE}_{n+1,0}), \tag{5.38}$$

where $n = 0, ..., d-2$ and $t = 1, ..., T$. Now, we re-arrange the order so that the variable maximizes Equation (5.37) assuming l becomes the first variable followed by the remaining variables. Then we select the bivariate conditional copula $c_{l,n+2|k}$ for $n = 1, ..., d-2$ with single conditioning variable.

Following the same procedure, we continue sequentially to determine the next tree until we have all root nodes for every tree together with its corresponding bivariate copulas and associated sequential estimates $\hat{\theta}^{SE}$.

Example. Consider C-vine trees depicted in Figure 5.1 with five random variables: X_1, X_2, X_3, X_4 and X_5. Now assume X_1 is the variable that maximizes Equation (5.37) and therefore chosen as the root node for the first tree. We re-arrange the order for all remaining variables with X_1 as the first variable . We get the pairwise and the first tree structure as follows

$$\text{First Tree:} \qquad \overbrace{(X_1, X_2)}^{C_{1,2}}, \overbrace{(X_1, X_3)}^{C_{1,3}}, \overbrace{(X_1, X_4)}^{C_{1,4}}, \overbrace{(X_1, X_5)}^{C_{1,5}}$$

The copulas, $C_{1,2}, C_{1,3}$, $C_{1,4}$ and $C_{1,5}$ are determined based on Akaike Information Criterion (AIC). As mentioned earlier, we discuss further the copula selection in Section 5.4.4. The structure for the first tree is illustrated in Figure 5.2.

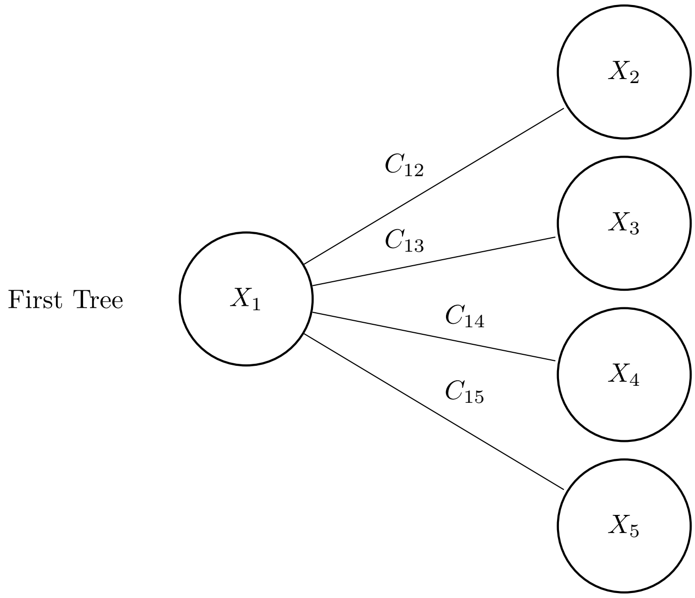


FIGURE 5.2: The first tree with X_1 is chosen as the root node linked to remaining nodes, X_2, X_3, X_4, X_5. The dependence structures between each pairwise $(X_1, X_2), (X_1, X_3), (X_1, X_4)$ and (X_1, X_5) are modelled by bivariate copula C_{12}, C_{13}, C_{14} and C_{15} respectively

.

To build the second tree, we transform all variables using h-function as in Equation (5.24) and (5.38). Then, using the copulas selected in the first tree together with its corresponding parameters, we determine the root node for the second tree that maximizes Equation (5.37). For simplicity, we assume the chosen variable for the root node in the second tree is X_2. The copulas used in this tree are conditional copulas with distribution functions conditioning on X_1 (the root node of the first tree). Now we re-arrange the order of the remaining variables as follows

$$\text{Second Tree:} \qquad \overbrace{(X_2, X_3|X_1)}^{C_{2,3|1}}, \overbrace{(X_2, X_4|X_1)}^{C_{2,4|1}}, \overbrace{(X_2, X_5|X_1)}^{C_{2,5|1}}$$

Once we have determined the variables order and the corresponding bivariate copula, we get the tree structure for the second tree as presented in Figure 5.3.

Following the same procedures as in building the second tree, we proceed sequentially to determine the root node for the third and forth tree. The number of tree for a C-vine is defined based on $d - 1$ and since in this example $d = 5$, therefore the maximum number of tree is only 4. The pairwise and tree structure for the third tree are as follows:

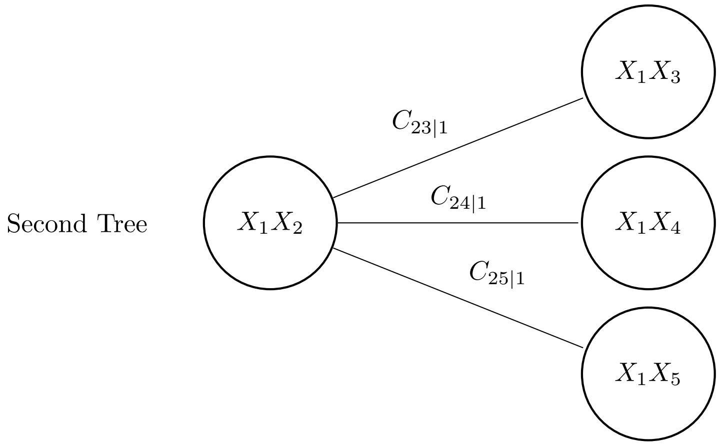


FIGURE 5.3: The second tree with X_2 is chosen as the root node linked to remaining pairwise nodes. Given X_1 as the conditional variable, the dependence structures between each pairwise (X_1, X_2) and (X_1, X_3) is modelled by bivariate conditional copula $C_{23|1}, (X_1, X_2)$ and (X_1, X_4) by $C_{24|1}$, (X_1, X_2) and (X_1, X_5) by $C_{25|1}$.

Third Tree: $\overbrace{(X_3, X_4|X_1, X_2)}^{C_{3,4|1,2}}, \overbrace{(X_3, X_5|X_1, X_2)}^{C_{3,5|1,2}}$

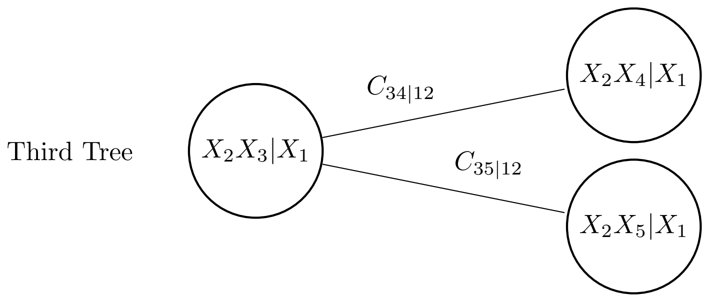


FIGURE 5.4: The third tree with X_3 is chosen as the root node linked to remaining pairwise nodes. Given X_2 as the conditional variable, the dependence structures between each pairwise $(X_2, X_3|X_1)$ and $(X_2, X_4|X_1)$ is modelled by bivariate conditional copula $C_{34|12}$ The remaining pairwise $(X_2, X_3|X_1)$ and $(X_2, X_5|X_1)$ is modelled by $C_{35|12}$.

And finally the forth (last) tree we only left with two nodes and root node is not required. The pairwise for the forth tree is as follows

$$\text{Third Tree:} \qquad \overbrace{(X_4, X_5|X_1, X_2, X_3)}^{C_{4,5|1,2,3}}$$

Figure 5.5 show the final tree structure for the 5-dimensional vine copula.

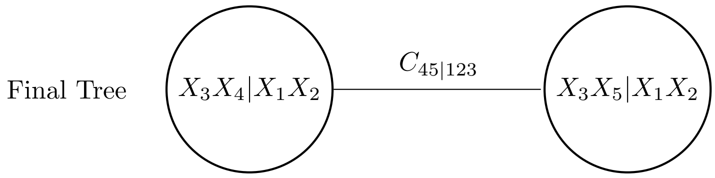


FIGURE 5.5: Illustration of the final tree of a C-Vine with conditional bivariate copula $C_{45|123}$ is chosen to model the dependence structure of pairwise $(X_3, X_4|X_1, X_2)$ and $(X_3, X_5|X_1, X_2)$

.

5.4.3 Vine copula inference

In this section, we discuss the methods for statistical inference of a C-vine copula. We focus on two widely used methods in the literatures (Czado et al., 2012). We start by introducing the method that we will use later in our empirical analysis called the sequential estimation (SE). For comparison we also discuss the second method known as maximum likelihood estimation (MLE).

C-vine copula consists of of bivariate copulas in each tree. The bivariate copulas in the first tree are unconditional bivariate copulas followed by conditional bivariate copula with single conditioning variable in the second tree, conditional bivariate copula with two conditioning variables in the third tree and so on. The number of conditioning variable increases according to number of dimension. The copula parameters are estimated sequentially, with the parameters of unconditional copulas in the first tree to be used for estimation of conditional copulas parameters in second tree and these procedures continue until all copula parameters are estimated.

Consider independent and identically distributed (i.i.d) vector $u_t := (u_{1,t}, ...u_{d,t})'$ for sample size, $t = 1, ..., T$. In the first tree, the parameter $\theta_{1,n}$ of unconditional bivariate copula $c_{1,n}$ is estimated using pairwise $(u_{1,t}, u_{n+1,t})$ with $t = 1, ..., T$ for $n = 1, ..., d-1$. Having $\hat{\theta}_{1,n}^{SE}$ as the estimated parameter for the first tree, then for the second tree we estimate the parameter $\theta_{2,n}$ of conditional bivariate copula with single conditioning variable $c_{2,n+2|1}$ for $n = 1, ..., d-2$. Using h-function defined in Section 5.3.3, we write

$$\begin{aligned} \hat{v}_{2|1,t} &:= F(u_{2,t}|u_{1,t}; \hat{\theta}_{1,1}^{SE}), \\ &= h(u_{2,t}|u_{1,t}; \hat{\theta}_{1,1}^{SE}) \end{aligned} \tag{5.39}$$

and

$$\begin{aligned} \hat{v}_{n+2|1,t} &:= F(u_{n+2,t}|u_{1,t}; \hat{\theta}_{1,n+1}^{SE}), \\ &= h(u_{n+2,t}|u_{1,t}; \hat{\theta}_{1,n+1}^{SE}) \end{aligned} \tag{5.40}$$

where $n = 1, ..., d-2$. We use pairwise $(\hat{v}_{2|1,t}, \hat{v}_{n+2,|1,t})$ with $t = 1, ..., T$ to estimate the parameter of bivariate conditional copula in tree 2 denoted as $\hat{\theta}_{2,n}^{SE}$ for $n = 1, ..., d-2$.

For the third tree, we follow the same procedure as in the second tree to estimate the parameter $\theta_{3,n}$ of the conditional bivariate copula with 2 conditioning variables $c_{3,n+3|1,2}$ where $n = 1, ..., d-3$. We also use the following the transformed pairwise variables $\hat{v}_{3|1,2,t}$ and $\hat{v}_{n+3|1,2,t}$, $t = 1, ..., T$ to estimate $\theta_{3,n}$. The transformed pairwise are given by

$$\begin{aligned} \hat{v}_{3|1,2,t} &:= F(\hat{v}_{3|1,t}|\hat{v}_{2|1,t}; \hat{\theta}^{SE}_{2,1}), \\ &= h(\hat{v}_{3|1,t}|\hat{v}_{2|1,t}; \hat{\theta}^{SE}_{2,1}) \end{aligned} \tag{5.41}$$

$$\begin{aligned} \hat{v}_{n+3|1,2,t} &:= F(\hat{v}_{n+3|1,t}|v_{2|1,t}; \hat{\theta}^{SE}_{2,n+1}). \\ &= h(\hat{v}_{n+3|1,t}|\hat{v}_{2|1,t}; \hat{\theta}^{SE}_{2,n+1}) \end{aligned} \tag{5.42}$$

The estimation procedures explained earlier proceed sequentially until all parameters for unconditional and conditional bivariate copulas in a C-Vine are estimated.

In summary, the general case to estimate the parameters of a C-vine $\theta_{m,n}$ where $n = 1, ..., d-2$ can be derived by the estimation of the following pairwise:

$$\hat{v}_{n+1|1,...,n,t} := F(\hat{v}_{n+1|1,...,n-1,t}|\hat{v}_{n|1,....n-1,t}; \hat{\theta}^{SE}_{1,n}), \tag{5.43}$$

$$\hat{v}_{n+m+1|1,...,n,t} := h(\hat{v}_{m+n+1|1,...,n-1,t}|\hat{v}_{m+1|1,...,m-1,t}; \hat{\theta}^{SE}_{n+1,m-1}). \tag{5.44}$$

We discussed earlier the sequential estimation procedures for bivariate unconditional and conditional copulas of a C-vine model. This estimation provides starting values of parameters for numerical maximisation of the log-likelihood for a C-vine. To complete discussion on inference and estimation of a C-vine model, we now explain the MLE procedure.

Consider a random vector $x_m = (x_{m,1}, ..., x_{m,T})'$ where $m = 1, ..., d$ with sample size T to be independent over time and uniformly distributed on $[0, 1]$.

The log-likelihood of a C-vine introduced by Aas et al. (2009) can be written as

$$\sum_{n=1}^{d-1}\sum_{m=1}^{d-n}\sum_{t=1}^{T}\log\left[c_{n,n+m|1,...,n-1}\Big\{F(x_{n,t}|x_{1,t},...,x_{n-1,t}),F(x_{n+m,t}|x_{1,t},...,x_{n-1,t})\Big\}\right]. \tag{5.45}$$

The conditional distributions, $F(x_{n,t}|x_{1,t},...,x_{n-1,t})$ and $F(x_{n+m,t}|x_{1,t},...,x_{n-1,t})$ in Equation (5.45) are determined using Equation (5.14) and h-function in (5.16).

5.4.4 Copula selection

We consider the same copula families as in Chapter 4. However, the copula for each pairwise loss ratio in every C-vine tree is selected based on Akaike Information Criterion (AIC) proposed by Joe (1997). The AIC is defined as

$$AIC = -2l(x_{n,1},x_{n,2};\hat{\boldsymbol{\theta}}) + 2p, \tag{5.46}$$

where

$$\begin{aligned} l(x_{n,1},x_{n,2};\hat{\boldsymbol{\theta}}) &= \sum_{n=1}^{d}\left[\log c\{(F_{X_{n,1}}(x_{n,1}),F_{X_{n,2}}(x_{n,2});\boldsymbol{\theta})\}\right] \\ &= \sum_{n=1}^{d}\left[\log c(u_{n,1},u_{n,2};\boldsymbol{\theta})\right] \end{aligned}$$

is the log-likelihood, p denotes the number of parameters θ. For a bivariate copula with $p=1$ indicates one parameter bivariate copula ($\boldsymbol{\theta}=\theta_1$) and $p=2$ indicates a bivariate copula with 2 parameters ($\boldsymbol{\theta}=(\theta_1,\theta_2)'$), which avoid over-fitting by penalizing the log-likelihood. Based on this test, the copula with the lowest AIC value is chosen.

The AIC for a C-Vine is given by

$$AIC = -2\sum_{n=1}^{d-1}\sum_{m=1}^{d-n}\sum_{t=1}^{T}\log\left[c_{n,n+m|1,...,n-1}\Big\{F(x_{n,t}|x_{1,t},...,x_{n-1,t}),F(x_{n+m,t}|x_{1,t},...,x_{n-1,t})\Big\}\right]+2p, \tag{5.47}$$

where p denotes the number of bivariate copula parameters.

An alternative to AIC is Bayesian Information Criterion (BIC) proposed by Schwarz (1978). BIC has a stronger penalty term than the AIC and can be written as

$$BIC = -2l(x_{n,1}, x_{n,2}; \hat{\boldsymbol{\theta}}) + \log(d)p, \tag{5.48}$$

We can see the similarity between AIC and BIC from the first term of Equation (5.47) and (5.48) that both AIC and BIC use maximum likelihood. The second term of these equations is the penalty term. It consists of the number of parameters for both AIC and BIC and number of observations for only BIC. Therefore, BIC depends on sample size and penalty for additional parameters is stronger. It is also proved by Burnham and Anderson (2004) that AIC is practically superior to BIC. Hans (2007) tests on copula fitting by considering both the upper and lower tail dependence and concludes that AIC has the highest accuracy in most of the tests. Hence, AIC is used in this study to determine the appropriate bivariate copula for every pairwise loss ratios.

5.4.5 Simulation of a C-vine

We folllow the algorithm introduced by Aas et al. (2009) that provide procedure for sampling a C-Vine copula in a simplified way, based on earlier algorithm discussed in Bedford and Cooke (2001), Bedford and Cooke (2002) and Kurowicka and Cooke (2007).

Algorithm 2:

1. Define the number of simulations $N \in \mathbb{N}$.
2. Sample $s_1, ..., s_N$ independent uniform observations on [0,1].
3. Set:
   $x_1 = v_{1,1} = s_1$
   $x_2 = F^{-1}(s_2|x_1)$
   $x_3 = F^{-1}(s_3|x_1, x_2)$
   $... = ...$
   $x_N = F^{-1}(s_N|x_1, ..., x_{N-1})$
4. Compute the conditional distribution functions, $F(x_j|x_1, ..., x_{j-1})$ for sampling $(N+1)$th variable using h-function defined in Equation 5.16 and 5.14

recursively. We obtain the $F(x_j|x_1,...,x_{j-1})$ for every j recursively using the following

$$F(x_i|x_1,...,x_{j-1}) = \frac{\partial C_{j,j-1|1,...,j-2}(F(x_j|x_1,...,x_{j-2}), F(x_{j-1}|x_1,...,x_{j-2})}{\partial F(x_1|x_2,...,x_{j-1})}. \tag{5.49}$$

Example:

To illustrate example for Algorithm 2 above, consider the following steps:

1. For a 3-dimensional dataset, x_1, x_2, x_3, we first sample 3 independent uniform on [0,1], s_1, s_2, s_3.
2. Then we set $x_1 = s_1$.
3. Next we have the first conditional distribution function with x_1 as the first conditional variable $F(x_2|x_1) = h(x_2, x_1, \theta_{11})$. This give us $x_2 = h^{-1}(s_2, x_1, \theta_{11})$.
4. $F(x_3|x_1, x_2) = h\{h(x_3, x_1, \theta_{12}), h(x_2, x_1, \theta_{21}), \theta_{21}\}$ and give us $x_3 = h^{-1}[h^{-1}\{s_3, h(x_2, x_1, \theta_{11}), \theta_{21}\}, x_1, \theta_{21}]$

5.5 Empirical analysis on the UK general insurance industry data

5.5.1 Analysis of the data

To investigate the real impact of capital requirement for insurance companies, we require real data from insurance industry. In particular, we use the UK general insurance incurred claims and written premiums. We select 5 major business lines in the UK general insurance industry such as Motor, Property, Accident & Health, Liability and Miscellenous insurance. Miscellenous insurance includes other insurance business lines such as assistance, creditor, extended warranty, legal expenses, mortgage indemnity, pet, other personal financial loss, fidelity and contract guarantee, all "bond" business, credit, suretyship, commercial contingency, trade indemnity, special indemnity, licence business, foot and mouth and finally rainfall (pluvius).

The type of these data are similar to the data used in Chapter 4. We are using data on incurred claims. However, for premiums data, we use written premiums instead of earned premiums due to a different reporting format used in the UK versus its counterpart in Australia. In this chapter, we focus on practical insurance business with reinsurance exposure. In other words, we use net data for both incurred claims and written premiums to derive the loss ratios.

We also consider availability of the insurance data in the UK as one of our research limitations. Historical data for net incurred claims is only available from 2001 to 2015 and limited to annual frequency. Therefore, the data used in this chapter comprise of 15 observations from 2001 to 2015. These data are not publicly available and were purchased from the Association of British Insurers (ABI)[1] with the rights to use for our research.

Since our main objective in this chapter is to develop a new capital requirement model for insurance companies in the UK, we derive the loss ratios using the incurred claims and written premiums data as explained in Section 2.3.

Table 5.1 provides descriptive statistics of the loss ratios for each business line.

[1] ABI is an association established in 1985 made up of insurance companies in the UK. This association collects extensive data from insurance companies covering all insurance business lines in the UK.

	Motor	Property	Accident & Health	Liability	Miscellaneous	Aggregate
Mean	0.7999	0.5656	0.7064	0.7393	0.4199	0.6693
Standard deviation	0.0627	0.0779	0.1396	0.1037	0.0786	0.0372
Skewness	1.8617	1.3771	2.6628	0.1470	0.6708	0.6233
Excess kurtosis	1.9770	1.5825	6.7306	0.6450	-0.8842	-0.5654

TABLE 5.1: Summary statistics of Motor, Property, Accident & Health, Miscellaneous and Aggregate Loss ratio.

In this chapter, we start modelling the capital requirement by testing if the loss ratios are stationary. As mentioned earlier, the data sourced from ABI are limited to annual frequency. In other words, this type of data do not present seasonal factor therefore does not affect all 5 business lines loss ratios. However, trend can be seen in Accident & Health, Liability and Miscellaneous. We conclude that these three loss ratios are non-stationary and require further time-series adjustments. We eliminate the trend by a fitting linear regression model. (see Montgomery et al. (2007) for details on time-series adjustments).

We further examine if the loss ratios present serial dependence. The data consist of five different insurance business lines and we treat each business line as univariate time series data. The plots for each loss ratio are given in Appendix A. We test for serial dependence of each business line following procedures outlined in Section 2.3.3. The test on Motor and Property business lines are based on original observations while Accident & Health, Liability and Miscellaneous are based on residuals from the respective linear regression model. We refer to Ljung-Box test in Table 5.2 to confirm if the loss ratio exhibits serial dependence. The test hypothesis is rejected if the p-value is less than 0.05. All business lines loss ratios recorded p-values of greater than 0.05.

	Motor	Property	Accident & Health	Liability	Miscellaneous
Observations	0.2332	0.7120	-	-	-
Residuals	-	-	0.9976	0.6430	0.8925

TABLE 5.2: Ljung-Box test with p-values for all business lines. Observations on the left column represent the original loss ratio data and residuals obtained from de-trended time-series from the respective linear regression model.

From Table 5.3, we can see that Accident & Health, Liability and Miscellaneous business line is modelled by linear regression with respective parameters for the three business lines loss ratios. At 0.05 significant level, the linear regression models are statistically significant which provide low p-values of 0.0026, 0.0476 and 0.0045 for residuals of Accident & Health, Liability and Miscellaneous respectively. These tests show fitting the linear regression models to the residuals of loss ratios for the three business lines provide a satisfactory fit.

Business line	Linear regression model	p-value
Accident & Health	LR_{ah_t}=-5.1002+0.0026t+$\varepsilon_{ah,t}$	0.0026
Liability	$LR_{l_t} = 2.2616 + 0.0011t + \varepsilon_{l,t}$	0.0476
Miscellaneous	$LR_{mi_t} = -2.9873 + 0.0015t + \varepsilon_{mi,t}$	0.0045

Table 5.3: Linear regression models for Accident & Health, Liability and Miscellaneous loss ratios. LR_{ah_t}, LR_{l_t}, LR_{mi_t} are the loss ratio for Accident & Health, Liability and Miscellaneous respectively. The p-values are for the slope parameter of the linear regression.

Then, following (A-D) test, Normal distribution is best fitted to the residuals of the three business lines loss ratio. However, the remaining business lines loss ratios, Motor and Property do not present any serial dependence and trend. The selection of univariate distribution for these business lines is similar to family distribution and method used in Section 4.4.2. We fit a Burr distribution to these loss ratios following similar (A-D) test.

Business Line	Distribution	Parameters (s.e)
Motor	Burr	Shape 1=0.3870 (0.1921)
		Shape 2= 47.8949 (17.4649)
		Scale* = 0.2890 (0.0542)
Property	Burr	Shape 1=0.3133 (0.2011)
		Shape 2 =28.9751 (12.6449)
		Scale* = 0.1477 (0.2637)
Accident & Health (residual)	Normal	$\varepsilon_{ah,t} \sim N(\mu = 0, \sigma = 1.1731e^{-2})$
Liability (residual)	Normal	$\varepsilon_{l,t} \sim N(\mu = 0, \sigma = 7.7834e^{-3})$
Miscellaneous (residual)	Normal	$\varepsilon_{mi,t} \sim N(\mu = 0, \sigma = 6.3582e^{-3})$

Table 5.4: Family of distributions selected for each business line loss ratios. The parameter and corresponding standard errors estimates are listed for each business line. *In the case of the Burr distribution the value listed in the table as being the scale is in fact the estimate for the rate which is 1/scale.

5.5.2 The mixed C-vine tree structure

In building the C-vine tree structure to model the capital requirement for general insurance companies in the UK, we follow the procedure explained in Section 5.4.2.

At this point, we develop a mixed C-vine tree structure allowing for a mixture of copula families. The copulas are determined based on AIC as discussed in Section 5.4.4. Recall that a C-vine tree structure consists of trees with a root node linked to other nodes in each tree. Table 5.5 summarizes the sum of absolute value of Kendall's tau for each loss ratio and loss ratio pairs.

The first step to build mixed C-vine is to determine the root node for the first tree based on the highest sum of absolute value of Kendall's tau.

	First tree						
		LR_m	LR_p	LR_{ah}	LR_l	LR_{mi}	Sum
1	LR_m	1.0000	-0.0667	-0.2952	-0.2381	-0.2381	1.8381
2	LR_p	-0.0667	1.0000	0.5048	-0.1238	-0.0095	1.7048
3	LR_{ah}	-0.2952	0.5048	1.0000	-0.0476	0.2571	**2.1048**
4	LR_l	-0.2381	-0.1238	-0.0476	1.0000	0.3524	1.7619
5	LR_{mi}	-0.2381	-0.0095	0.2571	0.3524	1.0000	1.8571

	Second tree					
		LR_{ah}, LR_m	LR_{ah}, LR_p	LR_{ah}, LR_l	LR_{ah}, LR_{mi}	Sum
1	LR_{ah}, LR_m	1.0000	-0.1238	0.3333	-0.1238	1.5810
2	LR_{ah}, LR_p	-0.1238	1.0000	-0.0286	0.1238	1.2762
3	LR_{ah}, LR_l	0.3333	-0.0286	1.0000	-0.2571	**1.6190**
4	LR_{ah}, LR_{mi}	-0.1238	0.1238	-0.2571	1.0000	1.5048

	Third tree				
		LR_{ah},LR_l,LR_m	LR_{ah},LR_l,LR_p	LR_{ah},LR_l,LR_{mi}	Sum
1	LR_{ah},LR_l,LR_m	1.0000	-0.3333	-0.0667	1.4000
2	LR_{ah},LR_l,LR_p	-0.3333	1.0000	0.2000	**1.5333**
3	LR_{ah},LR_l,LR_{mi}	-0.0667	0.2000	1.0000	1.2667

TABLE 5.5: Dependence structure estimated using empirical Kendall's tau and the corresponding sum of absolute value of Kendall's taus on the right table denote as Sum. The highest sum of each tree is highlighted in bold and chosen as the root note for the first, second and third C-vine tree. LR_m, LR_p, LR_{ah}, LR_l, LR_{mi} represent loss ratio for Motor, Property, Accident & Health, Liability and Miscellaneous respectively.

As reported in Table 5.5, Accident & Health has the highest sum of absolute value of Kendall's tau of 2.1048 and therefore is selected as the root node for first tree. Then, Accident & Health is linked to the remaining nodes as can be visualised from top left panel of Figure 5.6. We use copula or specifically bivariate copula (represent by edges in Figure 5.6.) to model the dependency between the root node and other nodes. The selection of bivariate copula is based on AIC test explained in Section 5.4.4. Pairwise (Accident & Health, Motor) presents no upper and lower tail dependence and suitable for Frank copula. This selection is also confirmed by the selection criteria, AIC which provides the lowest test value. With strong non-parametric lower tail dependence of 0.73, Clayton copula is selected for pairwise (Accident & Health, Property). The best copula for pairwise (Accident & Health, Liability) is Gumbel while Survival Clayton for pairwise (Accident & Health, Miscellaneous). The copulas selected and its estimated parameters for each tree is presented in Table 5.6.

Pairwise	Copula	$\hat{\theta}$ (s.e.)	τ	λ_u	λ_l
First tree (No conditional variable)					
(LR_{ah}, LR_m)	Frank	-3.44 (1.79)	-0.34	-	-
(LR_{ah}, LR_p)	Clayton	2.18 (1.51)	0.52	-	0.73
(LR_{ah}, LR_l)	Gumbel	1.13 (0.59)	0.12	0.16	-
(LR_{ah}, LR_{mi})	Survival Clayton	1.33 (0.87)	0.40	0.59	-
Second tree (One conditional variable)					
$(LR_l, LR_m \mid LR_{ah})$	Frank	-2.66 (1.52)	0.28	-	-
$(LR_l, LR_p \mid LR_{ah})$	Frank	-1.28 (1.41)	-0.14	-	-
$(LR_l, LR_{mi} \mid LR_{ah})$	Gumbel	1.58 (0.88)	0.37	0.45	-
Third tree (Two conditional variables)					
$(LR_p, LR_m \mid LR_{ah}, LR_l)$	Clayton	0.22 (0.49)	0.10	-	0.04
$(LR_p, LR_{mi} \mid LR_{ah}, LR_l)$	Frank	-2.12 (1.55)	-0.23	-	-
Forth tree (Three conditional variables)					
$(LR_m, LR_{mi} \mid LR_{ah}, LR_l, LR_p)$	Gumbel	1.16 (0.29)	0.13	0.18	-

TABLE 5.6: C-vine copula and the estimated parameters $\hat{\theta}$ with the corresponding standard error in parentheses. τ is the value of estimated Kendall's tau and λ_u and λ_l are the upper and lower copula tail dependence respectively. LR_m, LR_p, LR_{ah}, LR_l, LR_{mi} represent loss ratio for Motor, Property, Accident & Health, Liability and Miscellaneous respectively.

The root node for the second tree is more complicated and involves one conditional variable. To facilitate the process, root node is determined using estimated bivariate copula parameter chosen in the first tree and transformed variables using h-function explained in Section 5.3.3. We also discuss earlier in Equation (5.39), (5.40), (5.41) and (5.42) of Section 5.4.3 procedures to transform the loss ratio. Referring to the second tree of Table 5.5 , pairwise (Accident & Health, Liability) has the highest sum of absolute value of Kendall's tau and selected as the root node. Given Accident & Health as the conditional variable, Frank copula is selected to model the dependence between pairwise (Liability, Motor) and (Liability, Property). For simplification, these pairwise notation can also be presented as (Liability, Motor | Accident & Health) and (Liability, Property | Accident & Health) respectively. Also, given Accident & Health as the conditional variable, pairwise (Liability, Miscellaneous) is best modelled by Gumbel copula which also exhibits upper tail dependence.

We proceed sequentially following the same procedures to determine the root node for third tree. In the third tree, given Accident & Health and Liability as the two conditional variables, the pairwise (Property, Motor) is best modelled by Clayton copula while pairwise (Property, Miscellaneous) is best modelled by Frank copula. The forth tree is the final tree leaving only two pairwise with one edge. Therefore the procedures stop at the third tree. Pairwise in the forth tree consist of three conditional variables and Gumbel copula is selected as the best copula to model the dependence between the pairwise.

The C-vine model can be presented in graphical form. Figure 5.6 illustrates four different structures of a C-vine model. Each structure is represented by a tree of different level following bottom-up approach. Tree 1 is the bottom level while tree 4 is the top level. The boxes contain numbers to represent the variables and nodes. Each node is linked to other node (pairwise) by an edge and represent the bivariate copula for every pairwise.

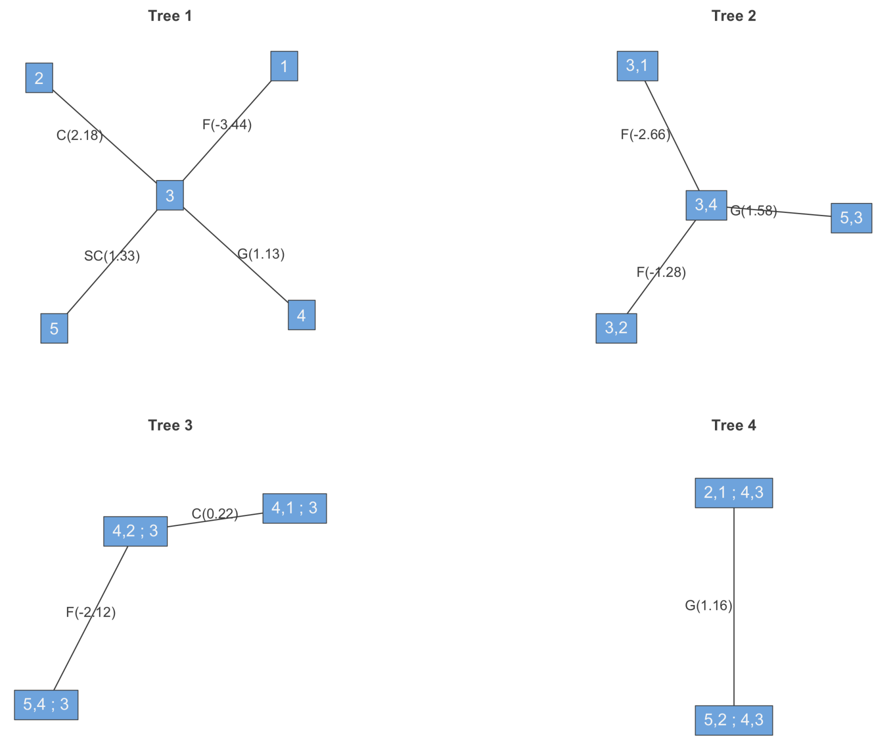


FIGURE 5.6: First to forth tree of C-vine model. Node 1,2,3,4 and 5 represent the loss ratio Motor, Property, Accident & Health, Liability and Miscellenous respectively. The nodes are connected by edge with corresponding copula. F, C, SC, and G is the bivariate copula for Frank, Clayton, Survival Clayton, and Gumbel copula respectively.

5.5.3 Models comparison

We test the robustness of the mixture of C-vine copula model constructed in the previous sections. To put things into perspective, we include additional models consisting of C-vine Gaussian copula and C-vine t-copula. Unlike the mixture of C-vine copula model, the C-vine Gaussian copula model and C-vine t-copula model are constructed with the same copula in all C-vine trees. Table 5.7 summarizes the models used to investigate the performance of all models.

Model	Copula	Selection criteria	Tail dependence
Mixed C-vine	Mixed	Sequential estimation	λ_L and λ_U
C-vine Gaussian	Gaussian	Sequential estimation	No tail dependence
C-vine t-copula	t	Sequential estimation	$\lambda_L = \lambda_U$

Table 5.7: Summary of three possible models to determine the capital requirement for insurance companies. The possible underlying copula in Mixed copula includes Gaussian copula, t-copula, Clayton copula, Gumbel copula and Frank copula. The selection of the copula is based on AIC. λ_L and λ_U represent the upper tail and lower tail dependence respectively.

5.5.4 Estimation of risks

At this point, we simulate new random observations from the aggregate loss ratios using c-vine model developed in the previous sections and estimate the VaR and TVaR to determine the level of risk in an insurance company. We follow Algorithm 2 for simulation procedures described in Section 5.4.5. We include VaR estimates as a comparison to TVaR which is determined by the empirical quantile of the new simulate random observations. We use parametric bootstrap to estimate both VaR and TVaR at confidence levels of 90%, 95%, and 99%.

The VaR and TVaR estimates from the simulation of the aggregate risk are reported in Table 5.8. The results provide evidence of the importance of dependence structure and proper practical model to determine the capital requirement. Specifically, in Table 5.8, we observe that C-vine Gaussian copula provides the lowest value both in VaR and TVaR. However, Embrechts et al. (2003) argues the relevancy of Gaussian copula as a symmetrical copula to model dependence structure of heavy tailed distribution especially insurance loss ratios. Further, the remaining copula models, Mixed C-vine and C-vine t-copula provide non-significant difference.

	Mixed C-vine	C-vine Gaussian copula	C-vine t-copula
90% VaR	0.5059	0.5021	0.5033
95% VaR	0.5248	0.5169	0.5217
99% VaR	0.5631	0.5565	0.5688
90% TVaR	0.5321	0.5258	0.5318
95% TVaR	0.5498	0.5430	0.5525
99% TVaR	0.5913	0.5868	0.6048

Table 5.8: VaR and TVaR estimates each at 90%, 95% and 99% confidence levels.

To further investigate C-vine models discussed earlier, we perform comparison based on AIC criterion. The results are presented in Table 5.9. We observe that the smallest AIC is recorded by mixed C-vine copula although the highest log-likelihood is recorded by C-vine t-copula. Clearly, based on these results, mixed C-vine copula is preferred over other models. It is also important to note that, C-vine Gaussian copula proves to be a better option than C-vine t-copula to model the dependence structure for insurance risks. This result is counter-intuitive with the fact that Gaussian copula is a symmetrical copula and does not present any tail dependence while t-copula has the ability to model both lower and upper tail dependence. However, this is justifiable by the following: as discussed in Section 5.5.1, the original observations of three business lines (Accident & health, Liability, and Miscellaneous) present trend component of time-series analysis. After filtering these data by first differencing, the residuals are best fitted by Normal distribution and consequently changed the best fitted bivariate copula for the pairwise that present Accident & Health, Liability and Miscellaneous. In other words, filtering process mentioned earlier caused the changes in the dependence structures of these three business lines. Consequently, we observed lower AIC criterion of -4.91 for C-vine Gaussian copula versus 5.51 recorded by C-vine t-copula.

	Mixed C-vine	C-vine Gaussian copula	C-vine t-copula
AIC	-10.71	-4.91	5.51
Loglikelihood	15.35	12.45	17.25

Table 5.9: AIC and Loglikelihood values for Mixed C-vine, C-vine Gaussian copula and C-vine t-copula.

5.6 Concluding remarks

In this chapter, we introduce the theoretical foundation for vine copula. This includes the assumptions, statistical inference and method to build a C-vine copula model. C-vine copula is a type of vine copula and with bivariate copulas as the main building block a similar concept presented in Chapter 4. Bivariate copulas are selected based on AIC criterion from the list of copulas discussed in the previous sections to model the dependence structures of all pairwise. It is also important to note that C-vine copula is significantly different from Hierarchical aggregation model in terms of the type of copula used. In a C-vine copula, both conditional and unconditional copulas are used to build the tree structures.

To build a C-vine copula, we implement sequential estimation procedures where the first root node and copula parameter in the first tree are used to determine the second tree structure. The process proceed sequentially until all copula parameters and trees structures are estimated.

The choice of copulas used to model the dependence structures of insurance business play the central role to determine the capital requirement of an insurance company. The risk estimates determine by VaR and TVaR are highly influenced by types of copula used. For example, in this study, C-vine Gaussian copula provides the lowest risk estimates among other C-vine models. However, this model is not the best copula as Gaussian copula is a symmetrical copula and does not suitable to model dependence structure of insurance risks. We perform comparison study based on AIC criterion proposed by Joe (1997) to analyse the robustness of C-vine models with different underlying copula. A model constructed by mixed C-vine copula satisfies the test and selected as the best candidate to determine the capital requirement for insurance companies.

Chapter 6

Conclusions and future work

6.1 Conclusion

In this PhD thesis, we show that modelling multivariate distribution of insurance risks using copulas is practically useful for insurance companies to estimate the total risk exposure. In this context, insurance risks are determined by claims per unit of premium paid by policyholders and represented by loss ratios. More importantly, copula can be used to aggregate different marginal of insurance business line losses with any possible dependence structure.

From risk management perspective, it is important to estimate the insurance business losses. In particular, Chapter 2 introduces risk measures and the properties for a coherent risk measure that is important element to get a good estimate of insurance risk. We also discuss two possible methods for computing risk measures. In addition, we define loss ratios and discuss its application to multivariate modelling.

In Chapter 3, we present the fundamentals of copulas and its statistical properties. We further define dependence measures and provide a simple motivating example to understand the importance of copulas in measuring dependence structure. We conclude this chapter by giving an overview of Graph Theory which is an important concept that we used to build our hierarchical risk aggregation and vine copula models in Chapter 4 and 5 respectively.

The hierarchical risk aggregation model is introduced in Chapter 4 to address the complex multivariate insurance losses distributions and dependence structure of different insurance business lines. Based on this model, a multivariate distribution is decomposed into bivariate distribution on each aggregation level to avoid complex numerical computation. In building risk aggregation model, we consider loss of each insurance business line as a node to form a tree structure based on Graph Theory. Also, guided by respective tail dependence, insurance business lines losses are aggregated using the best copulas determined by the highest p-value of goodness-of-fit tests from Genest et al. (2009). As already mentioned in Chapter 1, insurance companies need to determine the level of capital requirement. We propose in Chapter 4 to use hierarchical risk aggregation model for aggregating insurance risk and determine the capital requirement using appropriate risk measures.

On the other hand, in Chapter 4, we also provide a comprehensive analysis on the effect of reinsurance to the insurance company's total risk. Theoretically, insurance companies share premiums collected from its policyholders to reinsurance companies. In returns, reinsurance companies absorb part of insurance companies underwriting risks. More importantly, reduce insurance companies total risks. We investigate the risk reduction effect in the tail dependence of aggregated gross and net loss ratios. In this regard, gross loss ratios are used to determine insurance risk without considering the reinsurance business, while the net loss ratios are used to determine insurance risk taking into account the reinsurance business. Risk reduction effect in the tail or diversification effect can be observed once all loss ratios are aggregated. Reinsurance can prevent the reduction of risk if risk is managed by changing the proportion of premiums between business lines. However, risk reduction is more prominent if risk is managed by focusing on the dependence structure between the business lines using the hierarchical risk aggregation model.

We extend the literature on risk aggregation in Chapter 5 by introducing pair-copula construction or also known as vine copula with special focus on developing a new capital requirement model for insurance companies. Even though the main building block is still using bivariate distribution as used in hierarchical aggregation model in Chapter 4, however, we incorporate both unconditional and conditional copulas to construct vine copula. Chapter 5 starts with reviews on past literatures on vine copula, followed by comprehensive overview of vine copula concepts. This includes assumptions on its conditional copulas, statistical inferences as well as copula selection and simulation of a C-vine. In this chapter, we use real data sourced from the Association of British Insurers (ABI) to investigate the dependence structure of multiple insurance business lines risks. In particular, we model the dependence structure of insurance business line losses using bivariate copula and aggregate the risks using C-vine copula. Further, we select the best C-vine copula models among mixed C-vine copula, C-vine Gaussian copula and C-vine t-copula using AIC criterion to represent the capital requirement model for insurance companies in the UK.

In general, our results show that mixed copulas models, used in hierarchical aggregation and vine copulas method are the key to adequately model the dependence structure of different insurance business lines. In particular, insurance business risks represented by loss ratios are suitable to model by mixture of copulas from the Archimedean copula families. This includes Clayton, Gumbel and Frank copula. The mixed models can incorporate variety of tail dependence of the insurance losses and consequently avoid underestimation or overestimation of capital requirement. Hence, these models are suitable for insurance companies to legally remain solvent. In addition, we show that the risk of aggregated losses can be successfully reduced by modelling the dependence structure of multiple insurance business lines risks. In this regard, we assume that reinsurance is present as part of insurance companies strategy to reduce their business risks. However, it is important to highlight that, reinsurance might not benefit an insurance company if the risk management strategy is by changing premium allocation between business lines. In other words, based on this strategy, reinsurance might mitigate the risk reduction of risk of the aggregate insurance losses.

6.2 Limitations and future research

We consider data availability as the main limitation for this study. Diers et al. (2012) highlight that the availability of insurance data are very limited, unlike data in finance which is publicly available especially the stocks market data. Data for the UK insurance industry are not publicly available and require subscription. If a research is funded by a research grant, this data could be purchased and extended for longer time series. However, data providers are not many and historical data are very limited which are limiting capabilities for real data analysis. Real data analysis is very important to investigate the practical impact from the research. It is also important to model insurance risk with appropriate data to reflect and observe real implication to insurance company. To this end, the analysis in this study use loss ratio as the proxy to insurance risk. However, we may also consider other factors to represent insurance risks such as ruin probability.

In this thesis, we show that the dependence structure of multiple insurance business lines in Australia and the UK are modelled by hierarchical risk aggregation and vine copulas respectively. However, we do not intend to make comparison between these models. This will be addressed in our future work.

While we do not estimate the level of capital requirement for insurance companies, our models can be applied with respective insurance company loss data for insurance companies to estimate the level of capital requirement. The capital requirement of an insurance company is different from one company to another and highly influenced by the total premium size of the insurance company. Generally, a big size insurance company requires higher capital requirement than a smaller size insurance company.

Appendix A

Figure A.1, A.2, A.3, A.4 and A.5 are the plots of the five business lines loss ratios in Chapter 5.

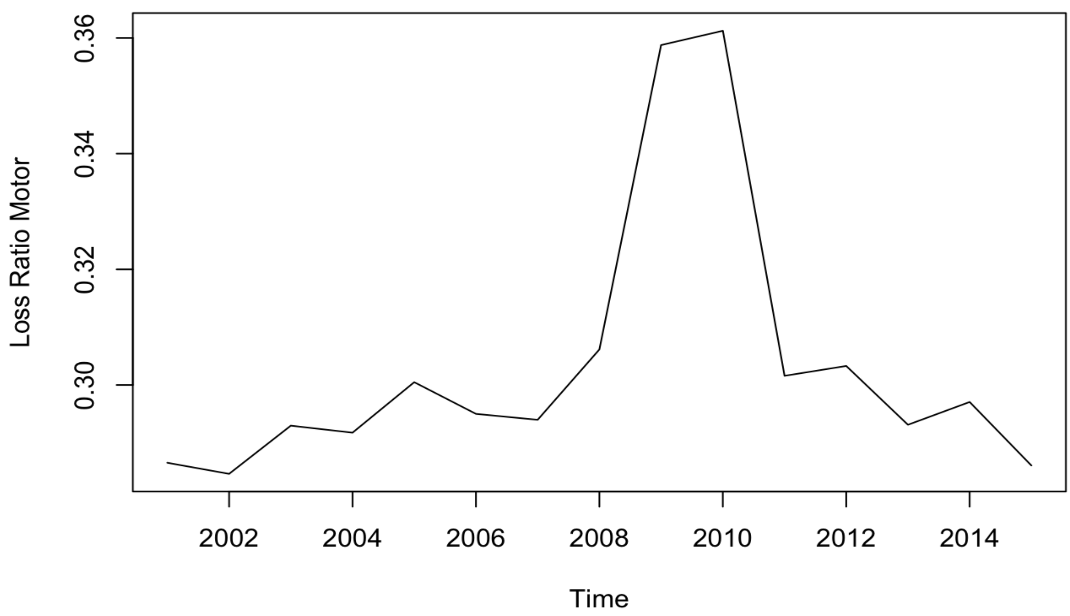


FIGURE A.1: Plot of Motor loss ratio.

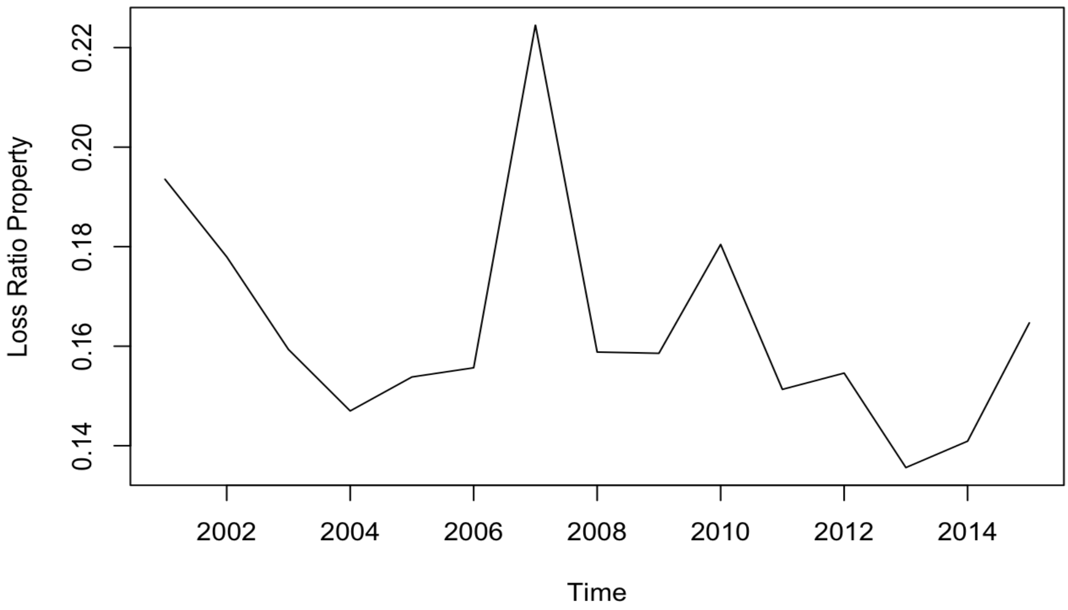


Figure A.2: Plot of Property loss ratio.

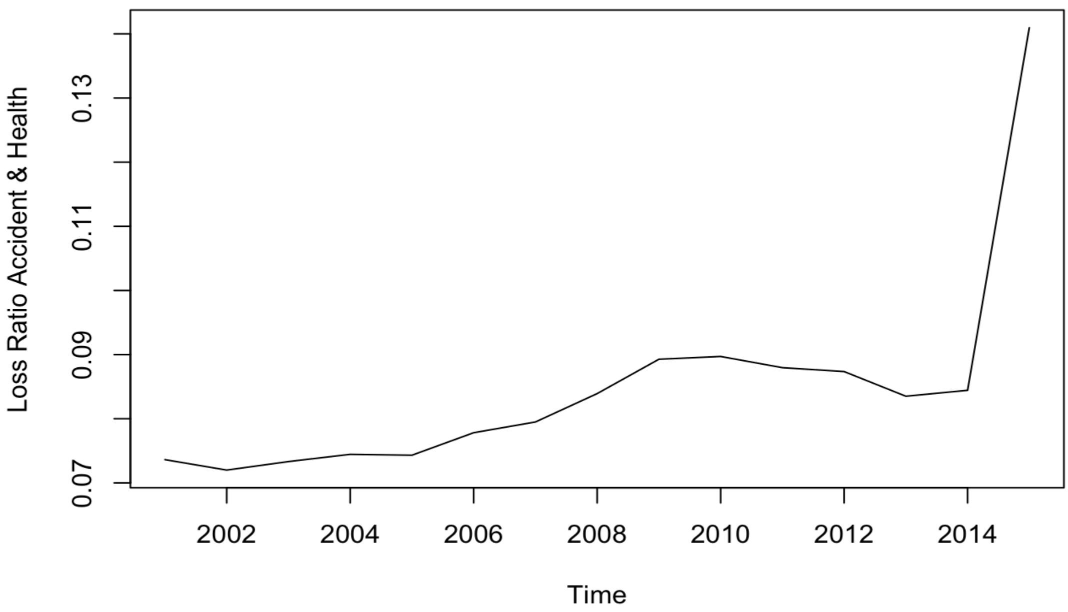


Figure A.3: Plot of Accident & Health loss ratio.

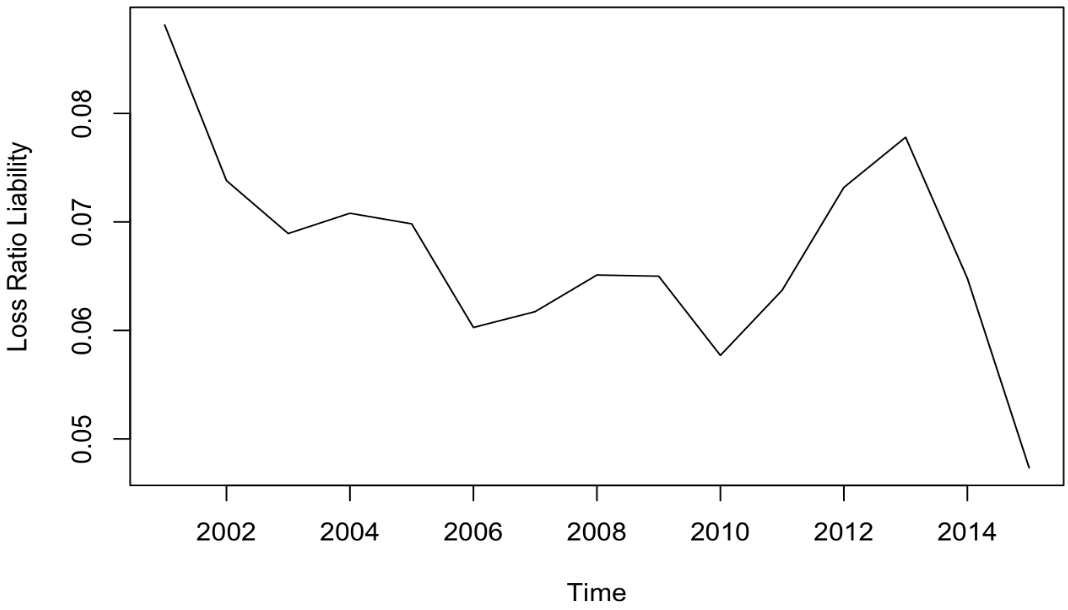


Figure A.4: Plot of Liability loss ratio.

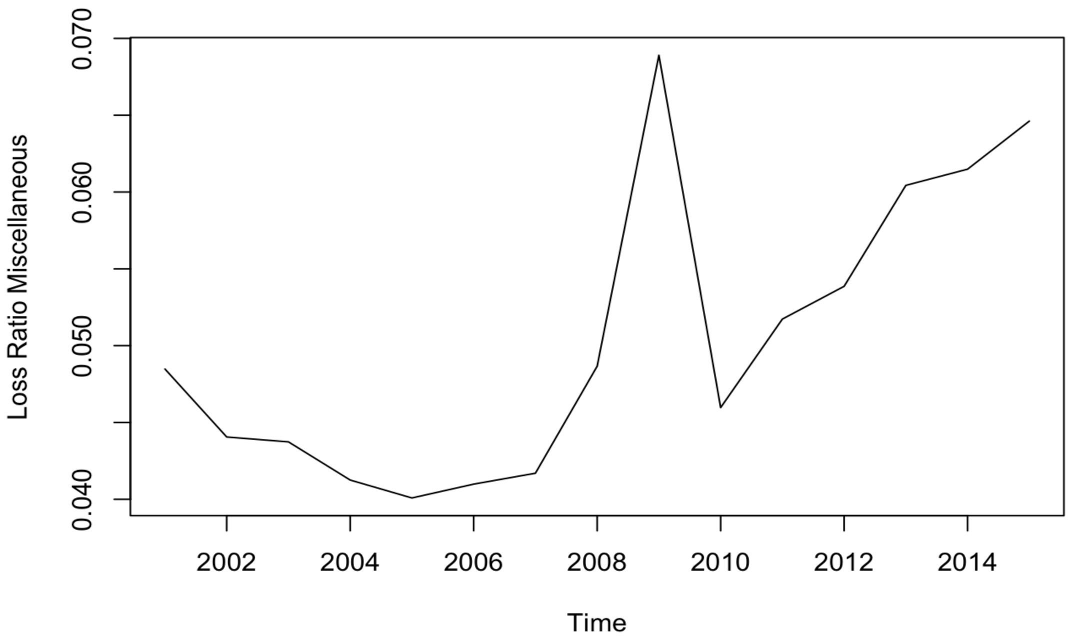


Figure A.5: Plot of Miscellaneous loss ratio.

Bibliography

Aas, K., Czado, C., Frigessi, A., and Bakken, H. (2009). Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics*, 44(2):182 – 198.

Acar, E. F., Genest, C., and Nešlehová, J. (2012). Beyond simplified pair-copula constructions. *Journal of Multivariate Analysis*, 110:74–90.

Acerbi, C. and Tasche, D. (2002). On the coherence of expected shortfall. *Journal of Banking & Finance*, 26(7):1487–1503.

Adam, A., Houkari, M., and Laurent, J.-P. (2008). Spectral risk measures and portfolio selection. *Journal of Banking & Finance*, 32(9):1870–1882.

Akaike, H. (1974). A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control*, 19(6):716–723.

Alexander, G. J. and Baptista, A. M. (2004). A comparison of var and cvar constraints on portfolio selection with the mean-variance model. *Management Science*, 50(9):1261–1273.

Anderson, T. W. and Darling, D. A. (1954). A test of goodness of fit. *Journal of the American Statistical Association*, 49(268):765–769.

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

Artzner, P., Delbaen, F., Eber, J.-M., and Heath, D. (1999). Coherent Measures of Risk. *Mathematical Finance*, 9(3):203–228.

Barthel, N., Geerdens, C., Killiches, M., Janssen, P., and Czado, C. (2018). Vine copula based likelihood estimation of dependence patterns in multivariate event time data. *Computational Statistics and Data Analysis*, 117:109–127.

Baur, P., Breutel-O'Donoghue, A., and Hess, T. (2004). *Understanding reinsurance: How reinsurers create value and manage risk*. Swiss Re, Swiss Reinsurance Company.

Bedford, T. and Cooke, R. M. (2001). Probability density decomposition for conditionally dependent random variables modeled by vines. *Annals of Mathematics and Artificial Intelligence*, 32(1-4):245–268.

Bedford, T. and Cooke, R. M. (2002). Vines : A new graphical model for dependent random variables. *the Annals of Mathematical Statistics*, 30(4):1031–1068.

Bickerstaff, D. R. (1972). Automobile collision deductibles and repair cost groups: the log-normal model. *Proceedings of the Casualty Actuarial Society. Available at:* `https://www.casact.org/pubs/proceed/proceed72/72068.pdf`, LIX:68.

Brechmann, E. C. (2014). Hierarchical Kendall copulas: Properties and inference. *The Canadian Journal of Statistics*, 42(1):78–108.

Brechmann, E. C. and Schepsmeier, U. (2013). Modeling dependence with C- and D-vine copulas: The R package CDVine. *Journal of Statistical Software*, 52(3):1–27.

Breymann, W., Dias, A., and Embrechts, P. (2003). Dependence structures for multivariate high frequency data in finance. *Quantitative Finance*, 3(1):1–14.

Bürgi, R., Dacorogna, M. M., and Iles, R. (2008). Risk aggregation, dependence structure and diversification benefit. *Stress Testing for Financial Institutions*. Available at SSRN: http://dx.doi.org/10.2139/ssrn.1468526.

Burnham, K. P. and Anderson, D. R. (2004). Multimodel inference: understanding AIC and BIC in model selection. *Sociological Methods & Research*, 33(2):261–304.

Cook, R. D. and Johnson, M. E. (1981). A family of distributions for modelling non-elliptically symmetric multivariate data. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 210–218.

Cooke, R. M., Joe, H., and Chang, B. (2015). Vine regression. *RFF working paper*, pages 1–34. Available at http://www.rff.org/research/publications/vine-regression.

Cossette, H., Gadoury, S.-P., Marceau, É., and Mtalai, I. (2017). Hierarchical Archimedean copulas through multivariate compound distributions. *Insurance: Mathematics and Economics*, 76(Supplement C):1–13.

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

Cummins, J. D., Dionne, G., Gagné, R., and Nouira, A. (2008). The costs and benefits of reinsurance. Available at `https://ssrn.com/abstract=1142954`.

Czado, C., Schepsmeier, U., and Min, A. (2012). Maximum likelihood estimation of mixed C-vines with application to exchange rates. *Statistical Modelling: An International Journal*, 12(3):229–255.

Darsow, W. F., Nguyen, B., and Olsen, E. T. (1992). Copulas and Markov processes. *Illinois Journal of Mathematics*, 36(4):600–642.

Denuit, M., Dhaene, J., Goovaerts, M., and Kaas, R. (2006). *Actuarial theory for dependent risks: measures, orders and models*. John Wiley & Sons.

Dias, A. and Embrechts, P. (2009). Testing for structural changes in exchange rates' dependence beyond linear correlation. *European Journal of Finance*, 15(7-8):619–637.

Dias, A., Ismail, I., and Zhang, A. (2018). Risk aggregation and insurance. *Working Paper*. University of Leicester.

Diers, D., Eling, M., and Marek, S. D. (2012). Dependence modeling in non-life insurance using the Bernstein copula. *Insurance: Mathematics and Economics*, 50(3):430–436.

Diestel, R. (2017). *Graph Theory*. Springer, Heidelberg, 5th edition.

Dißmann, J., Brechmann, E. C., Czado, C., and Kurowicka, D. (2013). Selecting and estimating regular vine copulae and application to financial returns. *Computational Statistics and Data Analysis*, 59(1):52–69.

Dowd, K. and Blake, D. (2006). After VaR: The theory, estimation, and insurance applications of quantile-based risk measures. *The Journal of Risk and Insurance*, 73(2):193–229.

Dropkin, L. B. (1964). Size of loss distribution in workmens compensation insurance. *Proceedings of the Casualty Actuarial Society. Available at:* `https://www.casact.org/pubs/proceed/proceed64/1964.pdf`, LI:198.

Eling, M. and Toplek, D. (2009). Risk and return of reinsurance contracts under copula models. *European Journal of Finance*, 15(7-8):751–775.

Embrechts, P., Lindskog, F., and McNeil, A. (2003). Modelling dependence with copulas and application to risk management. In Rachev, S., editor, *Handbook of Heavy Tailed Distributions in Finance*, chapter 8, pages 329–384. Elsevier.

Emmer, S., Klüppelberg, C., and Korn, R. (2001). Optimal portfolios with bounded capital at risk. *Mathematical Finance*, 11(4):365–384.

Engmann, S. and Cousineau, D. (2011). Comparing distributions: the two-sample Anderson-Darling test as an alternative to the Kolmogorov-Smirnoff test. *Journal of Applied Quantitative Methods*, 6(3):1–17.

Erhardt, T. M. and Czado, C. (2018). Standardized drought indices: a novel univariate and multivariate approach. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 67(3):643–664.

Erhardt, T. M., Czado, C., and Schepsmeier, U. (2015). Spatial composite likelihood inference using local c-vines. *Journal of Multivariate Analysis*, 138(C):74–88.

Erhardt, V. and Czado, C. (2012). Modeling dependent yearly claim totals including zero claims in private health insurance. *Scandinavian Actuarial Journal*, 2012(2):106–129.

Fischer, M., Köck, C., Schlüter, S., and Weigert, F. (2009). An empirical analysis of multivariate copula models. *Quantitative Finance*, 9(7):839–854.

Fortin, I. and Kuzmics, C. (2002). Tail-dependence in stock-return pairs. *International Journal of Intelligent Systems in Accounting, Finance & Management*, 11(2):89–107.

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

Genest, C. and Nešlehová, J. (2012). Copulas and copula models. In: *Encyclopedia of Environmetrics*, 2nd ed., vol. 2, A.H. El-Shaarawi and W.W. Piegorsch, editors. Wiley, Chichester, pages 541–553.

Genest, C., Rémillard, B., and Beaudoin, D. (2009). Goodness-of-fit tests for copulas: A review and a power study. *Insurance: Mathematics and economics*, 44(2):199–214.

Gruber, L. and Czado, C. (2015). Sequential bayesian model selection of regular vine copulas. *Bayesian Analysis*, 10(4):937–963.

Hans, M. (2007). Estimation and model selection of copulas with an application to exchange rates. Research Memorandum 056, Maastricht University, Maastricht Research School of Economics of Technology and Organization (METEOR). Available at https://EconPapers.repec.org/RePEc:unm:umamet:2007056.

Hobæk Haff, I. (2013). Parameter estimation for pair-copula constructions. *Bernoulli*, 19(2):462–491.

Hobæk Haff, I., Aas, K., and Frigessi, A. (2010). On the simplified pair-copula construction - Simply useful or too simplistic? *Journal of Multivariate Analysis*, 101(5):1296–1310.

Hobæk Haff, I. and Segers, J. (2015). Nonparametric estimation of pair-copula constructions with the empirical pair-copula. *Computational Statistics and Data Analysis*, 84(C):1–13.

Hofert, M. (2008). Sampling Archimedean copulas. *Computational Statistics and Data Analysis*, 52(12):5163–5174.

Hofert., M., Kojadinovic, I., Mächler, M., and Yan, J. (2017). *Elements of Copula Modeling with R*. CRoNoS Winter Course on Copula-based modeling with R, London.

Hofert, M. and Scherer, M. (2011). CDO pricing with nested Archimedean copulas. *Quantitative Finance*, 11(5):775–787.

Hutchinson, T. and Lai, C. (1990). *Continuous bivariate distributions, emphasising applications*. Rumsby Scientific Publishing, Adelaide.

Iman, R. and Conover, W. (1982). A distribution-free approach to including rank correlation among input variables. *Communications in Statistics - Simulation and Computation*, 11(3):311–334.

Joe, H. (1996). Families of m-variate distributions with given margins and m(m-1)/2 bivariate dependence parameters. *Lecture Notes-Monograph Series*, 28:120–141.

Joe, H. (1997). *Multivariate models and dependence concepts*, volume 73 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, London.

Joe, H., Li, H., and Nikoloulopoulos, A. K. (2010). Tail dependence functions and vine copulas. *Journal of Multivariate Analysis*, 101(1):252–270.

Jondeau, E. and Rockinger, M. (2006). The Copula-GARCH model of conditional dependencies: An international stock market application. *Journal of International Money and Finance*, 25(5):827–853.

Killiches, M. and Czado, C. (2015). Block-maxima of vines. *Extreme Value Modelling and Risk Analysis: Methods and Applications*, pages 109–130.

Killiches, M., Kraus, D., and Czado, C. (2017). Examination and visualisation of the simplifying assumption for vine copulas in three dimensions. *Australian and New Zealand Journal of Statistics*, 59(1):95–117.

Kleiber, C. (2003). *Statistical size distributions in economics and actuarial sciences*. Wiley series in probability and statistics. Wiley-Interscience, Hoboken, N.J.

Klugman, S. A., Panjer, H. H., and Willmot, G. E. (2012). *Loss Models : From Data to Decisions*. John Wiley amp; Sons, Incorporated, Somerset.

Klugman, S. A. and Parsa, R. (1999). Fitting bivariate loss distributions with copulas. *Insurance: Mathematics and Economics*, 24(1-2):139–148.

Kolmogorov, A. (1933). Sulla determinazione empirica di una lgge di distribuzione. *Inst. Ital. Attuari, Giorn.*, 4:83–91.

Kraus, D. and Czado, C. (2017). Growing simplified vine copula trees: improving dißmann's algorithm. Available at https://arxiv.org/abs/1703.05203.

Kurowicka, D. and Cooke, R. M. (2007). Sampling algorithms for generating joint uniform distributions using the vine-copula method. *Computational Statistics and Data Analysis*, 51(6):2889–2906.

Kurowicka, D. and Joe, H. (2010). *Dependence Modeling. Vine Copula Handbook*. World Scientific, Singapore.

Ljung, G. M. and Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65(2):297–303.

Mai, J.-F. and Scherer, M. (2012). H-extendible copulas. *Journal of Multivariate Analysis*, 110(Supplement C):151–160. Special Issue on Copula Modeling and Dependence.

Mcneil, A. J. (2008). Sampling nested archimedean copulas. *Journal of Statistical Computation and Simulation*, 78(6):567–581.

McNeil, A. J., Frey, R., and Embrechts, P. (2015). *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton University Press, Princeton, NJ, second edition.

Mikosch, T. (2009). *Non-Life Insurance Mathematics: An Introduction with the Poisson Process*. Universitext. Springer Berlin Heidelberg, Berlin, Heidelberg.

Min, A. and Czado, C. (2010). Bayesian inference for multivariate copulas using pair-copula constructions. *Journal of Financial Econometrics*, 8(4):511–546.

Montgomery, D. C., Jennings, C. L., and Kulahci, M. (2007). *Introduction to time series analysis and forecasting*. Wiley series in probability and statistics. Wiley, Hoboken, N.J.

Nelsen, R. B. (2006). *An Introduction to Copulas*. Springer-Verlag, New York, second edition.

Nguyen, T. and Molinari, R. D. (2011). Risk aggregation by using copulas in internal models. *Journal of Mathematical Finance*, 1(03):50–57.

OECD (2017). International comparisons. *OECD Insurance Statistics 2016, OECD Publishing, Paris*, pages 45–79.

Pekasiewicz, D. A. (2016). Bootstrap estimation methods of Value at Risk. *The Central European Review of Economics and Management*, 16(3):123–132.

Rosenberg, J. V. and Schuermann, T. (2006). A general approach to integrated risk management with skewed, fat-tailed risks. *Journal of Financial Economics*, 79(3):569–614.

Schellhase, C. and Spanhel, F. (2018). Estimating non-simplified vine copulas using penalized splines. *Statistics and Computing*, 28(2):387–409.

Schmid, F. and Schmidt, R. (2007). Multivariate conditional versions of spearman's rho and related measures of tail dependence. *Journal of Multivariate Analysis*, 98:1123–1140.

Schmidl, D., Czado, C., Hug, S., and Theis, F. J. (2013). A vine-copula based adaptive MCMC sampler for efficient inference of dynamical systems. *Bayesian Analysis*, 8(1):1–22.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6(2):461–464.

Shi, P. and Yang, L. (2018). Pair copula constructions for insurance experience rating. *Journal of the American Statistical Association*, 113(521):122–133.

Sibuya, M. (1960). Bivariate extreme statistics. *Annals of the Institute of Statistical Mathematics*, 11(3):195–210.

Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de l'Université de Paris*, 8:229–31.

Smirnoff, N. (1939). Sur les écarts de la courbe de distribution empirique. *Matematicheskii Sbornik*, 48(1):3–26.

Smirnov, N. (1948). Table for estimating the goodness of fit of empirical distributions. *Annals of Mathematical Statistics*, 19:279–281.

Tang, A. and Valdez, E. A. (2006). Economic capital and the aggregation of risks using copulas. *Proceedings of the 28th International Congress of Actuaries, Paris, France. Available at:* `http://www.ica2006.com/282.html`.

Taylor, J. M. (1997). *Claim Reserving Manual*. Faculty and Institute of Actuaries, Great Britain, 1997 rev. edition.

Tsai, J. T., Wang, J. L., and Tzeng, L. Y. (2010). On the optimal product mix in life insurance companies using conditional Value at Risk. *Insurance: Mathematics and Economics*, 46(1):235–241.

Vaz De Melo Mendes, B. and Accioly, V. B. (2014). Robust pair-copula based forecasts of realized volatility. *Applied Stochastic Models in Business and Industry*, 30(2):183–199.

Wang, S. (1998). Aggregation of correlated risk portfolios: models and algorithms. In *Proceedings of the Casualty Actuarial Society*, volume 85, pages 848–939.

Whelan, N. (2004). Sampling from Archimedean copulas. *Quantitative Finance*, 4(3):339–352.

Ye, J. and Li, T. (2012). The optimal mean–variance investment strategy under Value-at-Risk constraints. *Insurance: Mathematics and Economics*, 51(2):344–351.