

Modeling and measuring insurance risks for a hierarchical copula model  
considering IFRS 17 framework

Carlos Andrés Araiza Iturria

A Thesis  
in  
The Department  
of  
Mathematics and Statistics

Presented in Partial Fulfillment of the Requirements  
for the Degree of Master of Science (Mathematics) at  
Concordia University  
Montreal, Quebec, Canada

June 2019

© Carlos Andrés Araiza Iturria, 2019

# CONCORDIA UNIVERSITY

## School of Graduate Studies

This is to certify that the thesis prepared

By: Carlos Andrés Araiza Iturria

Entitled: Modeling and measuring insurance risks for a hierarchical copula model considering IFRS 17 framework

and submitted in partial fulfillment of the requirements for the degree of

### Master of Science (Mathematics)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final Examining Committee:

\_\_\_\_\_ Thesis Supervisor

Dr. M. Mailhot

\_\_\_\_\_ Thesis Supervisor

Dr. F. Godin

\_\_\_\_\_ Examiner

Dr. M. Pigeon

\_\_\_\_\_ Examiner

Dr. J. Garrido

Approved by \_\_\_\_\_

Chair of Department or Graduate Program Director

\_\_\_\_\_

Dean of Faculty

\_\_\_\_\_

Date

## **Abstract**

### **Modeling and measuring insurance risks for a hierarchical copula model considering IFRS 17 framework**

In this thesis, a stochastic approach to insurance risk modeling and measurement that is compliant with the new International Financial Reporting Standards (IFRS 17) is proposed. The compliance is achieved through the use of a semiparametric hierarchical copula which accounts for the dependence between the lines of business of the Canadian auto insurance industry. A model for the marginal unpaid claim liabilities of each line of business based on double generalized linear models is also developed. Development year and accident year effect factors along with an autoregressive feature for residuals enable modeling the dependence between the various entries of the loss triangles in a given line of business. Capital requirements calculations are then performed through simulation; numbers obtained with univariate and multivariate risk measures are compared. Moreover, a risk adjustment for non-financial risk required by IFRS 17 is also computed through a cost of capital approach.

## Acknowledgments

I would like to thank deeply Dr. Méлина Maillhot and Dr. Frédéric Godin for trusting and believing in me throughout this two year journey. By making a blind bet on me you have changed my life forever and for the better. Your constant support and motivation make you more than amazing mentors. Thank you very much.

This work was supported by the Institut des sciences mathématiques (ISM), the Faculty of Arts and Science Graduate Fellowship from Concordia University, Mitacs through the Mitacs Accelerate Program and Eckler Ltd.

I would also like to thank Dr. José Eliud Silva Urrutia, Professor in the Faculty of Actuarial Science at Anahuac University, Mrs. Cynthia M. Potts, FCIA, FCAS, Mr. Blair Manktelow, FCIA, FCAS, Dr. José Garrido, Professor in the Department of Mathematics and Statistics at Concordia University and Dr. Mathieu Pigeon, Professor in the Department of Mathematics at Université du Québec à Montréal for their valuable comments that helped improving my thesis.

## Dedication

In the famous painting *The Coronation of Napoleon* (1807) by Jacques-Louis David, the mother of Napoleon, Maria Letizia Ramolino was placed in the most important part of the painting even though she did not attend the coronation. Napoleon specifically asked the painter to add his mother in order to honor her for supporting him with all the burdens that came with office. In his words, *“The future destiny of a child is always the work of the mother”*.

There are not enough words to describe the gratitude I have towards my family. Every personal achievement has been a result of the never ending effort of my loving father, mother, brother and sister. For every time you have pushed and supported me to pursue my dreams, I thank you. I will always love you with all my heart.

*To the memory of my loving father, Andrés Humberto Araiza Baez.*

# Contents

List of Figures	xii
List of Tables	xiii
Introduction	1
<b>1 Background</b>	<b>2</b>
1.1 Reporting - IFRS 17 <i>Insurance Contracts</i>	3
1.1.1 Measurement	4
1.1.2 Risk adjustment for non-financial risk	7
1.1.2.1 Definition	8
1.1.2.2 Example	8
1.1.2.3 Objective	9
1.1.2.4 Characteristics	10
1.1.2.5 Disclosure	11
1.1.2.6 Contrasts with Standards of Practice	11
1.2 Capital allocation - OSFI	13
1.3 Justification of the copula model	15
<b>2 Review of Actuarial and Statistical Models</b>	<b>16</b>
2.1 Notation	16
2.2 Marginal distributions	18
2.2.1 Collective Risk Model	18

2.2.2	Compound Poisson-Gamma distribution . . . . .	18
2.2.3	Generalized Linear Models (GLM) . . . . .	20
2.2.4	Exponential Dispersion Family (EDF) . . . . .	21
2.2.4.1	Tweedie family . . . . .	24
2.2.5	Double Generalized Linear Models (DGLM) . . . . .	27
2.2.5.1	Estimation of DGLM . . . . .	28
2.2.5.2	DGLM estimation algorithm . . . . .	32
2.2.6	Goodness-of-fit . . . . .	32
2.2.6.1	Kolmogorov-Smirnov (K-S) . . . . .	32
2.2.6.2	Anderson-Darling (A-D) . . . . .	33
2.2.7	Generalized Estimating Equations (GEE) . . . . .	33
2.2.7.1	GEE estimation algorithm . . . . .	38
2.2.7.2	GEE Goodness-of-fit . . . . .	39
2.3	Copula models . . . . .	40
2.3.1	Dependence analysis . . . . .	42
2.3.2	Copula families . . . . .	44
2.3.2.1	Elliptical Copulas . . . . .	45
2.3.3	Hierarchical Copula Model (HCM) . . . . .	47
2.3.4	Estimation . . . . .	51
2.3.5	Copula selection and Goodness-of-fit . . . . .	52
2.4	Simulation . . . . .	54
2.4.1	Iman-Conover reordering algorithm . . . . .	54
<b>3</b>	<b>Model</b>	<b>59</b>
<b>4</b>	<b>Risk Assessment and Capital Requirements</b>	<b>66</b>
4.1	Reserving . . . . .	67
4.1.1	Risk measures . . . . .	67
4.1.1.1	Value-at-Risk (VaR) . . . . .	68
4.1.1.2	Tail Value-at-Risk (TVaR) or Expected shortfall (ES) . . . . .	68

4.1.1.3	Capital allocation . . . . .	69
4.1.2	Multivariate risk measures . . . . .	70
4.1.2.1	Capital allocation based on multivariate risk measures . . . . .	75
4.2	Cost of capital method . . . . .	80
<b>5</b>	<b>Numerical Application</b>	<b>82</b>
5.1	Descriptive statistics . . . . .	83
5.2	Marginal model . . . . .	85
5.3	Hierarchical copula model . . . . .	88
5.4	Risk assessment and capital requirements . . . . .	92
	<b>Conclusion</b>	<b>103</b>
	<b>Bibliography</b>	<b>103</b>
<b>A</b>	<b>Estimated marginal parameters</b>	<b>110</b>
<b>B</b>	<b>Other copula families</b>	<b>113</b>
B.1	Archimedean Copulas . . . . .	113
B.2	Plackett copula . . . . .	114



# List of Figures

1.1	Comparison of scenarios using a time value of money diagram. . . . .	9
2.1	Vectorial notation of loss ratios for GEE. . . . .	35
2.2	Kendall's $\tau$ in finite samples. . . . .	43
2.3	Dependence for accident year $i$ and lag $j$ between two lines of business. . .	45
2.4	Example of a hierarchical copula model for a portfolio with 6 lines of business.	48
2.5	Hierarchical copula model used in Burgi et al. (2008). . . . .	49
2.6	Example of nested Archimedean copula. . . . .	49
2.7	Example of a HCM with six lines of business, three levels and five bivariate copulas. . . . .	54
3.1	Loss triangle with GLM equations of mean model. . . . .	60
3.2	Distributional visualization of the DGLM considered for the run-off triangle.	62
3.3	First level copula structure in the HCM. . . . .	63
3.4	Second level copula structure in the HCM. . . . .	64
3.5	Complete dependence structure of the HCM. . . . .	64
4.1	Univariate $\text{VaR}_\alpha$ , bivariate lower and upper orthant VaR for Example 1. . .	72
4.2	Set of two curves composing the bivariate lower orthant $\overline{\text{TVaR}}_{\alpha, \mathbf{X}}(\mathbf{X})$ for Example 1. . . . .	74
4.3	Set of two curves composing the bivariate upper orthant $\overline{\text{TVaR}}_{\alpha, \mathbf{X}}(\mathbf{X})$ for Example 1. . . . .	75

4.4	Univariate and bivariate $\text{VaR}_\alpha$ and $\text{TVaR}_\alpha$ intersection with proportional allocation line for Example 1. . . . .	78
4.5	Zoom in of black box in Figure 4.4 to visualize intersection between univariate and bivariate risk measures with the proportional allocation line for Example 1. . . . .	79
5.1	Histograms of loss ratios for the 6 lines of business. . . . .	84
5.2	Behavior of the time series of the loss ratios by accident year for the automobile industry of Ontario. . . . .	84
5.3	Goodness-of-fit measured through the Anderson-Darling and Kolmogorov-Smirnov test for the personal auto line of Ontario before the GEE approach (first row) and after the GEE approach (second row). . . . .	87
5.4	Time series of residuals before (first row) and after (second row) using GEE for the commercial auto line of Alberta. . . . .	87
5.5	Personal auto vs Commercial auto for Ontario . . . . .	88
5.6	Personal auto vs Commercial auto for Alberta . . . . .	88
5.7	Comparison of empirical copula for Alberta vs rank-based simulations of different copula families. . . . .	90
5.8	HCM structure by province. . . . .	92
5.9	HCM structure by copula family. . . . .	92
5.10	Historical data vs historical data with one simulation of loss ratios for personal auto in Atlantic Canada. . . . .	93
5.11	Historical data vs historical data with one simulation of loss ratios for commercial auto in Atlantic Canada. . . . .	93
5.12	100,000 Simulations of the reserves for the province of Alberta under the HCM. Univariate $\text{VaR}$ (purple dotted lines) and $\text{TVaR}$ (green straight lines) along with bivariate lower and upper orthant $\text{TVaR}$ (green curved lines). Optimal couples shown with red dots. . . . .	96

# List of Tables

1.1	Liability measurement approaches under IFRS 17. . . . .	5
2.1	General notation for a run-off triangle with loss ratios for the $k$ -th line of business. . . . .	17
2.2	Some cases of the Tweedie family depending on the index parameter $p$ . . .	24
2.3	Iman-Conover reordering algorithm example for the first node of dependence structure (HCM) from Figure 2.7. Inspired by examples in <a href="#">Arbenz et al. (2012)</a> . . . . .	56
5.1	Goodness-of-fit and correlation test: $p$ -values for marginal models. . . . .	85
5.2	Independence tests by province. The Cramér-von Mises column shows the $p$ -values obtained. . . . .	89
5.3	Goodness-of-fit for copula models by province. . . . .	90
5.4	Goodness-of-fit for the second level of dependence: Alberta and Atlantic Canada. . . . .	91
5.5	Independence tests in the last node of the hierarchical copula model. . . .	91
5.6	Univariate VaR and TVaR at $\alpha = 99\%$ in billions (CAD) for the six lines of business. . . . .	94
5.7	Risk adjustments corresponding to Table 5.6 for the univariate VaR and TVaR at $\alpha = 99\%$ in billions (CAD) for the six lines of business. . . . .	95
5.8	Risk adjustments calculated through bivariate risk measures in billions (CAD) under different risk assumptions at a confidence level $\alpha = 99\%$ for the HCM. . . . .	98

5.9	Cost of capital method disclosed by accident year in billions (CAD) for personal auto in Ontario assuming a cost of capital rate $r_t = 8\%$ and discount rate $d_t = 2\%$ . . . . .	99
5.10	Risk adjustment for non-financial risks displayed in millions (CAD) for the six lines of business and equivalent confidence level for the VaR with two different assumptions for the cost of capital rate $r_t$ and discount rate $d_t$ . . .	100
A.1	Mean model - Accident year effects. . . . .	110
A.2	Mean model - Development Lag effects. . . . .	111
A.3	Dispersion submodel - Development lag effects. . . . .	112
A.4	Correlation parameter $\rho$ estimated using GEE. . . . .	112
A.5	Index parameter $p$ for the Tweedie distribution. . . . .	112
B.1	Common Archimedean copula families and some of their properties. . . . .	114

# Introduction

Actuaries are interested in reserving the appropriate amount to cover for future claims, ensuring solvency for the insurer while, most importantly, protecting the insureds. The capital allocated to a reserve is disclosed on the financial reports of the insurance entities. New international financial reporting regulations have been set with the new International Financial Reporting Standards (IFRS 17) to homogenize and facilitate their interpretation, making it simpler to compare insurance entities across jurisdictions.

Generalized linear models are commonly used in the industry to forecast future claims due to their accuracy and simple interpretation. In the recent actuarial literature, awareness to model the dispersion jointly with the mean has increased through double generalized linear models. Moreover, it is essential to verify that the assumptions of the statistical models are satisfied. Insurance portfolios are represented by multivariate distributions and due to the increasing computational power and development of the theory for copulas, modeling and measuring the risk associated with the multivariate distribution while accounting for dependence has become fundamental.

This thesis is structured as follows. In Chapter 1, we discuss the new IFRS 17 framework for financial reporting, the capital requirements set by the insurance regulator in Canada and a justification for the proposed model. Chapter 2 explains the statistical and actuarial concepts needed to understand the model in Chapter 3. Chapter 4 describes the concept of reserving through univariate and multivariate risk measures. Furthermore, the cost of capital method is presented to account for a risk adjustment. Then, in Chapter 5 the model is applied to a dataset from the Canadian automobile industry.

# Chapter 1

## Background

In this chapter, basic insurance definitions are presented, along with laws and the entities responsible of regulating the insurance industry in Canada to justify the model suggested in this thesis.

An insurance contract or policy is issued by an entity called the insurer for exchange of a monetary consideration also known as a premium. The policy protects the owner of the contract, also called the insured or policyholder, against a possible unfavorable event with economic consequences. Specifically in this thesis, we are interested in Property and Casualty (P&C), an insurance entity that protects the policyholder from costs arising by loss or damage to tangible or intangible property. Examples of P&C insurance contracts include but are not limited to fire, marine, legal expenses and automobile insurance. The numerical results presented in Chapter 5 are an application in automobile insurance.

In order to protect the policyholder, laws exist to regulate the financial management of the insurance companies through financial reporting and accounting standards. Laws regulating the insurance industry for Canada are found in the *Insurance Companies Act* [Government of Canada \(1991\)](#). The *Insurance Companies Act* states the Office of the Superintendent of Financial Institutions (OSFI) as the primary regulator of insurance companies with a federal charter in Canada. OSFI sets guidelines with regard to the capital and solvency of the insurer. The *Insurance Companies Act* also states that all

financial statements must be prepared in accordance with generally accepted accounting principles, the primary source of which is the Handbook of the Chartered Professional Accountants Canada. In May 2018, OSFI announced in [OSFI \(2018a\)](#) that IFRS 17 *Insurance Contracts*<sup>1</sup> was endorsed by the Canadian Accounting Standards Board and thus, it is now incorporated into the Handbook.

It is of vital importance to highlight that capital requirements imposed by OSFI or other insurance regulatory entities in other jurisdictions serve a different purpose than for IFRS reporting. While a regulatory entity establishes principles in order to protect the policyholder and ensure solvency by the insurer at all times, IFRS has the objective of creating comparable financial reporting across international boundaries for benchmarking purposes. Thus, in Section 1.1 specific information is presented regarding the new IFRS international framework for financial reporting and in Section 1.2 we deal with capital allocation requirements set by the OSFI in Canada. In Section 1.3, we justify the model used throughout this thesis related to the capital requirements and the financial reporting framework.

## 1.1 Reporting - IFRS 17 *Insurance Contracts*

The International Accounting Standards Board (IASB), an independent international non-profit group of experts in accounting and financial reporting, issued IFRS 17 *Insurance Contracts*, a new accounting standard for insurance contracts in May 2017, superseding the current regulatory framework IFRS 4. IFRS 17 *Insurance Contracts* establishes principles for the recognition, measurement, presentation and disclosure of insurance contracts.

IFRS 4 worked well reflecting national requirements because it allowed for different accounting practices. But having dissimilar standards across countries made it difficult for investors, analysts and decision makers to compare insurers' results. IFRS 17 is intended to be the key towards a common international insurance accounting standard [IASB](#)

---

<sup>1</sup>IFRS stands for International Financial Reporting Standards.

(2017a). The effective date of IFRS 17 has officially been set by the IASB to January 1<sup>st</sup>, 2021<sup>2</sup>, meaning March 31<sup>st</sup>, 2021 is the first quarter of reporting under IFRS 17.

Although in this thesis the numerical results presented in Chapter 5 are from the Canadian industry and the capital requirement standards are considered under the Canadian regulator (OSFI), IFRS 17 is an international framework working under several jurisdictions, meaning the model described in Chapter 3 can be applied in other countries while respecting the specific national capital requirements.

### 1.1.1 Measurement

Measurements under IFRS 17 seek to faithfully disclose the insurers' obligations arising from the portfolios of insurance contracts. IFRS 17 establishes three approaches to measure a liability depending on the duration (short and long term contracts), type of liability, and whether the contract depends on an underlying item. There are two types or classifications of liabilities under IFRS 17, liability for remaining coverage (LRC) and liability for incurred claims (LIC). LRC represents the unearned portion of risk from insurance contracts which are in force and LIC represent insurance events that already occurred but the claims have not been reported or have not been fully settled. Three measurement approaches and some examples of applicable contracts are presented in Table 1.1 which is adapted from CIA-CAS (2018). None of the measurement methods explained in what follows are mentioned in the superseded framework IFRS 4.

We are interested in P&C contracts, which under IFRS 17, can be measured with the General Model or with the PAA (as presented in Table 1.1). The General Model is the default approach for LIC because the horizon is usually more than one year. The PAA is an optional simplification which can apply for LRC if the duration of the contract is less or equal than one year (short-term contract). An example of a multi-year P&C contract is home insurance, where the insurer protects the policyholder against possible defects in the

---

<sup>2</sup>The IASB has proposed delaying the implementation of IFRS 17 by one year to 2022, subject to public consultation. As of the date of this thesis, the delay is not yet official.



construction of a home. Most automobile insurance contracts are short-term given that the duration of the policies is for a one year period. Consequently, on initial recognition<sup>3</sup>, the measurement of the liabilities for remaining coverage is performed under the PAA.

	General Model	Premium Allocation Approach (PAA)	Variable Fee Approach (VFA)
Applicability	Default approach	To simplify short-term contracts (optional).	Direct Participation Contracts: where policy cash flows are linked to underlying items.
Examples of applicable contracts	<ul style="list-style-type: none"> <li>• Life insurance</li> <li>• Life annuities</li> <li>• Universal life (UL)</li> <li>• Reinsurance contracts</li> <li>• Multi-year P&amp;C</li> </ul>	<ul style="list-style-type: none"> <li>• Most P&amp;C contracts</li> <li>• Short-term group contracts</li> </ul>	<ul style="list-style-type: none"> <li>• Segregated funds</li> <li>• Unit-linked contracts</li> <li>• Index-linked UL</li> <li>• Not applicable to P&amp;C</li> </ul>

Table 1.1: Liability measurement approaches under IFRS 17.

According to the PAA in paragraph 55 of [IASB \(2017b\)](#), the liability is measured on initial recognition as:

- The premiums,
- Minus acquisition cash flows (e.g. commissions paid to agents or taxation over the premiums),
- Plus or minus any amount arising from the derecognition (at the date of recognition of the insurance contracts) of prepayments, incurred expenses or any other acquisition cash flow the insurer pays or receives before the date of recognition.

P&C short-term contracts are accounted under the PAA which does not involve a risk ad-

<sup>3</sup>Recognition (derecognition) is the addition (removal) of an asset or liability from the balance sheet.

justment for non-financial risks component (described in Section 1.1.2). However, events not settled during the duration of the contract sometimes occur. Thus, the cost of the insurance event for the insurer could extend for years. The unknown amount that will be paid in upcoming years is also known as unpaid claims liabilities for expired coverage under short-term contracts. Although the original insurance contracts were accounted for upon initial recognition with the PAA, the unpaid claim liabilities fall under the general measurement due to the duration of these liabilities, as mentioned in IAA (2018). Contrary to the PAA, the General Model requires a risk adjustment for non-financial risks (described in Section 1.1.2). Therefore, even for short-term contracts, the general approach is most of the time required in order to take into consideration the unpaid claim liabilities.

We highlight the importance of the unpaid claim liabilities which are formed of two important reserves known as the Incurred But Not Reported (IBNR) reserve and the Reported But Not Settled (RBNS) reserve. IBNR and RBNS are reserve accounts representing the amount of money the insurer has to set aside to fulfill future liabilities as consequence of the insurance contracts. Notation and methods to calculate the unpaid claim liabilities are described in Chapter 2.

Since the unpaid claim liabilities or LIC fall under the General Model, we need to understand the measurement requirements. The General Model establishes in paragraph 32 of IASB (2017b) that upon initial recognition, a group of insurance contracts should be measured as the sum of:

- The fulfillment cash flow (FCF), which include:
  - Estimates of future cash flows,
  - An adjustment to reflect the time value of money and the financial risks related to the future cash flows,
  - A risk adjustment for non-financial risk.
- The contractual service margin (CSM).

*“The contractual service margin (CSM) represents the unearned profit the entity will recognize as it provides services in the future”* as stated in paragraph 38 of [IASB \(2017b\)](#). CSM applies for unexpired coverage (LRC) and it is not within the scope of this work.

### **1.1.2 Risk adjustment for non-financial risk**

To calculate the unpaid claim liabilities that fall under the General Model within the IFRS 17 framework, in this section we describe the definition, specific characteristics and other considerations for the risk adjustment for non-financial risks. There exist non-financial risks that can arise from insurance contracts that all insurance entities share but there are risks that are entity specific. The shared non-financial risks are, as mentioned in [IAA \(2018\)](#),

- Model risks: in practice, the true model of the unpaid claim amounts is unknown, thus, the difference between the true model and the model used to estimate the FCF is known as model risk. Additionally, we rely on limited variables to predict the unpaid claim amounts, augmenting the model risk.
- Parameter risk: given that only a sample of the phenomena is observed, the estimation of the parameters of the model could be biased or differ from the true parameters
- Process risk: assuming the model and parameters are correctly specified, the random nature of the phenomena can lead to a difference between the observed and estimated FCF.

Examples of entity specific non-financial risks for life insurance include mortality risk, lapse risk, etc. In this thesis, since we deal with automobile insurance (P&C), the entity specific non-financial risks considered are frequency and severity risks. Frequency and severity risks are the uncertainty associated to the number of claims and their cost, respectively.

### 1.1.2.1 Definition

As stated in paragraph 37 of [IASB \(2017b\)](#), for the definition of a risk adjustment for non-financial risks, “*An entity shall adjust the estimate of the present value of the future cash flows to reflect the compensation that the entity requires for bearing the uncertainty about the amount and timing of the cash flows that arises from non-financial risk.*” Thus, the actuary shall exclude incorporating into the calculation of the risk adjustment financial risks and risks that do not arise from insurance contracts. Examples of the aforementioned risks are, but are not limited to, investment risk, credit risk, operational risk, interest rate risk and underwriting risk which are included in other sections of the financial reports.

The risk adjustment can be understood as the price assigned by the insurer for bearing the non-financial risks associated with the portfolio of insurance contracts, more specifically, the risks that arise from unfavorable outcomes on a long-term horizon. This assigned price has to meet the objective and characteristics described in the following Sections [1.1.2.3](#) and [1.1.2.4](#).

### 1.1.2.2 Example

From paragraph B87 of [IASB \(2017b\)](#), “*The risk adjustment for non-financial risk for insurance contracts measures the compensation that the entity would require to make the entity indifferent between:*

- *Fulfilling a liability that has a range of possible outcomes arising from non-financial risk; and*
- *Fulfilling a liability that will generate fixed cash flows with the same expected present value as the insurance contracts.*”

To improve the understanding of the indifference principle, we present a simplifying example. We use a time value of money diagram in [Figure 1.1](#), where we consider the following simplifying assumptions: an annual interest rate of 5%, the range of outcomes arising from the non-financial risks in the portfolio of insurance contracts, which are equiproba-

ble over a 3 year period, the insurer has no profit nor any other expense and we assume equal fixed cash flows.

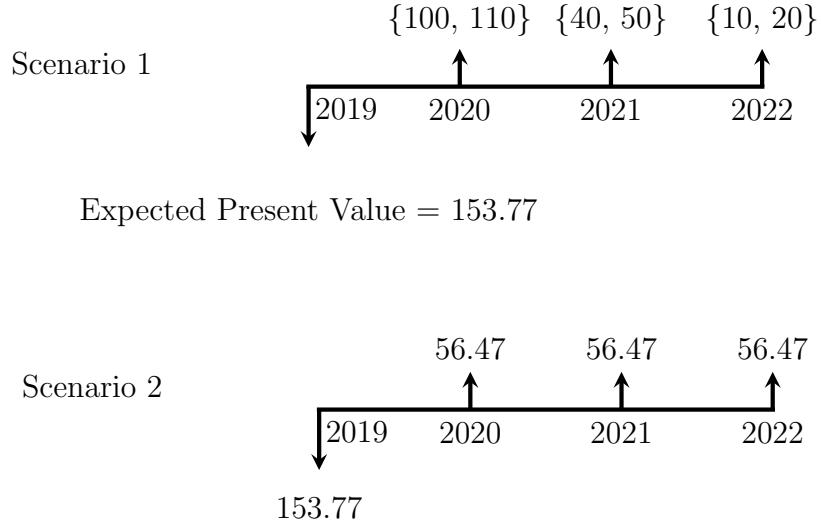


Figure 1.1: Comparison of scenarios using a time value of money diagram.

In Figure 1.1, we can observe in Scenario 1 a liability with a range of outcomes arising from non-financial risk and in Scenario 2, using the same expected present value from the insurance contracts in Scenario 1, a liability that will generate fixed cash flows. Therefore, the risk adjustment is the price in excess of the 153.77 currency units that accounts for the uncertainty in Scenario 1 and thus, causing the insurer to be indifferent from selecting either of the scenarios.

### 1.1.2.3 Objective

*“The objective of a risk adjustment is to provide a quantitative assessment of risk based on the entity’s risk preferences”* IAA (2018). To meet the objective, the following elements mentioned in IAA (2018) can be considered in developing the adjustment:

- Risk preferences,
- Key drivers of risk,

- The complexity of the probability or stochastic models,
- The ability to explain and quantify the risk adjustments in the context of financial statements.

#### 1.1.2.4 Characteristics

The risk adjustment is a point estimate, not a range, and is stated in the same monetary terms as the other monetary values in the entity's financial statement [IAA \(2018\)](#). Paragraph B88 of [IASB \(2017b\)](#) states that the risk adjustment for non-financial risk also reflects:

- *“The degree of diversification benefit the entity includes when determining the compensation it requires for bearing that risk,*
- *Both favorable and unfavorable outcomes, in a way that reflects the entity's degree of risk aversion.”*

In paragraph B91 of [IASB \(2017b\)](#), it is stated that the estimation technique(s) used to determine the risk adjustment for non-financial risks are not specified. However, the risk adjustment should comply with the following characteristics:

Risks with respectively,

- Low frequency and high severity,
- Longer duration,
- Higher variance,
- Higher parameter uncertainty,

will result in a higher risk adjustment for non-financial risks than risks with respectively high frequency and low severity, shorter duration, lower variance and lower parameter uncertainty. Additionally:

- New information reducing (increasing) uncertainty about the amount and/or timing

of cash flows, will decrease (increase) the value of the risk adjustments for non-financial risks.

#### 1.1.2.5 Disclosure

Once the risk adjustment for non-financial risks has been calculated with the appropriate technique(s) to reflect the entity's view of compensation for bearing the uncertainty of the insurance contracts considering the objective and characteristics presented in Sections 1.1.2.3 and 1.1.2.4, then, disclosure is required. More specifically, paragraph 119 of IASB (2017b) states that if the insurer uses a technique different from the Value-at-Risk<sup>4</sup>, commonly abbreviated as VaR (see Chapter 4 for definition) for determining the risk adjustment for non-financial risks, then, the insurer has to show the technique(s) used and disclose the corresponding level of confidence associated to the result obtained through the corresponding technique. For example, if the risk adjustment is set to  $X$  using the cost of capital method (described in Chapter 4) then the actuary has to disclose that  $X$  is equivalent to the VaR at a  $\alpha\%$  confidence level.

#### 1.1.2.6 Contrasts with Standards of Practice

The current accepted actuarial practice framework in Canada is established in a document called Standards of Practice ASB (2018) from the Actuarial Standards Board, established by the Canadian Institute of Actuaries (CIA). The Standards of Practice has a concept called Provision for Adverse Deviations (commonly known as PfAD) which measures the effect of uncertainty of the assumptions and data in determining the liability. The method to calculate PfAD can also be used to determine the risk adjustment but only if all the proper precautions are taking into account since both have different purposes. The following examples accompanied by a list that summarizes the contrasts between the PfAD and the risk adjustment will help understanding the important differences between the two concepts.

---

<sup>4</sup>In IFRS 17, the Value-at-Risk is referred to as the confidence level technique.

- Example A. Under IFRS 17, an insurer with a risk aversion policy has set the risk adjustment for non-financial risks to \$10 million using the cost of capital method (described in Chapter 4),
- Example B. Following the Standards of Practice, the appointed actuary of an insurer might expect the interest rate to be 5% but assumed for the calculations a 4% interest rate. This difference in the interest rate could value the liabilities in \$110 million and \$100 million, respectively. Thus, the provision for adverse deviations is the difference of \$10 million.

Even though the monetary value of the PfAD and the risk adjustment in Example A and B are exactly the same, we compare and summarize their differences with the following list:

- Objective: While the PfAD is a provision to account for uncertainty in the assumptions of the liability estimation, the risk adjustment is meant to reflect the entity's view of compensation for risk. In Example A, the risk preferences of the insurer were taken into consideration. Another insurer with the same group of insurance contracts could have set a higher risk adjustment to reflect the entity's view for the compensation for bearing the uncertainty,
- Scope: The risk adjustment only accounts for uncertainty arising from the insurance contracts, contrary to the PfAD, which in Example B, clearly also accounts for uncertainty in financial risks,
- Responsible: Under the Standards of Practice, the entity responsible for the PfAD is the appointed actuary whereas IFRS 17 intends to involve management since the definition of the risk adjustment is associated with the entity's view of compensation for bearing the uncertainty,
- Method: as stated in Section 1.1.2.4, the technique(s) to calculate the risk adjustment for non-financial risks are not specified under IFRS 17. Under the Standards of Practice the PfAD is the difference obtained from using more conservative as-



sumptions to protect the insurer for adverse deviations (hence, the name PfAD),

- **Diversification Benefit:** Under IFRS 17, the level of aggregation depends solely on the insurer’s view of diversification within an insurance portfolio. In other words, when considering similar risks, the level of aggregation to calculate the diversification benefit can be done at a line of business level (low level) or across entities among the same group (highest level). Under the Standards of Practice no explicit consideration was given until 2017 where in section 2120, paragraph 07 was added. This paragraph states, “*The provision resulting from the application of all margins for adverse deviations<sup>5</sup>, in addition to increasing the net liability, should be appropriate in the aggregate.*” Thus, even if in practice it is considered (not often considered according to [CIA \(2018\)](#)), the diversification benefit would also involve financial risks, contrary to the IFRS 17 specifications.
- **Disclosure:** A entirely new requirement under IFRS 17 is to disclose the confidence level equivalent to the VaR technique (as described in [Section 1.1.2.5](#)).

## 1.2 Capital allocation - OSFI

The previous [Section 1.1](#) deals with financial reporting for insurance contracts, which is managed by IFRS and applies to all IFRS jurisdictions. In this section, the regulatory principles set by OSFI are described and thus, the information presented is specifically for Canadian federally regulated insurers. These principles have the goal of ensuring an insurance entity maintains adequate capital levels in order to always be able to pay their liabilities to policyholders and creditors. The document *Guideline A: Minimum Capital Test* (MCT) [OSFI \(2018b\)](#) provides the framework within which OSFI assesses whether a P&C insurer maintains an adequate level of capital. The capital allocated with IFRS 17 needs to be compared with the capital levels set by OSFI since the latter is the minimum amount permitted by an insurance company in Canada.

---

<sup>5</sup>The margin for adverse deviations (MfAD) is the PfAD in a relative percentage term.

OSFI sets a minimum amount of capital that needs to be available at all times to address and support insurance risks called minimum capital which works as an indicator for the regulator. If the insurer capital falls below this threshold, the continuity of the entity becomes questionable. To alert OSFI before the occurrence of this event, there exists a supervisory target capital providing a margin of 50% above the minimum capital. The target capital requirements has been set by OSFI in [OSFI \(2018b\)](#) to be the conditional tail expectation at a 99% level for insurance risk, over a period of one year or alternatively, if deemed not practical by an expert, the VaR at 99.5% confidence level (see Chapter 4 for a description of such risk measures). Thus, due to the 50% margin mentioned, the minimum capital becomes the target capital divided by 1.5.

The target capital requirement (TCR) is calculated as

$$\text{TCR} = \begin{array}{ccccc} \text{Insurance} & & \text{Market} & & \text{Credit} & & \text{Operational} & & \text{Diversification} \\ & + & & + & & + & & - & \\ & \text{risk} & & \text{risk} & & \text{risk} & & \text{risk} & & \text{credit} \end{array}$$

In this thesis, we focus on the capital required for insurance risk, which corresponds to the risk that arises purely from the insurance contracts. Insurance risk breaks down as the following four components:

- Capital required for unpaid claims liabilities (or LIC),
- Capital required for premium liabilities (or LRC),
- Margin required for reinsurance ceded to unregistered reinsurers,
- Catastrophe reserves.

Unpaid claims according to OSFI have to be calculated by line of business. An insurer should not rely only on these regulatory capital measures but should conduct its Own Risk and Solvency Assessment (ORSA) [OSFI \(2018c\)](#) and, based on the entity's risk composition, determine its particular capital requirements and establish Internal Capital Targets [OSFI \(2018d\)](#).

### 1.3 Justification of the copula model

In order to obtain the diversification benefit related to the aggregation and pooling of risks, it is necessary to apply appropriate statistical techniques. In the specific case of P&C, naturally offsetting product lines may not exist, contrary to products in life insurance that compensate with payout annuities where the insurer pays if the policyholder survives, and thus, mortality is offset by longevity as stated in [IAA \(2018\)](#). Therefore, in P&C, diversification benefits are available by aggregating types of commercial and personal products, through geographical dispersion of risk dependent on legislation [Burgi et al. \(2008\)](#) or by dependence strength measured with a distance-based method [Côté et al. \(2016\)](#).

IFRS 17 allows an entity to set the risk adjustment at a level of aggregation that represents the compensation the entity requires for bearing the uncertainty regarding the liabilities. *“For example, the risk adjustment might be set at the entity level, thus, incorporating all diversification benefits in the organization aggregated across its product lines”* as mentioned in [IAA \(2018\)](#). The use of this level of aggregation would produce a high level of diversification of risk, therefore, a small risk adjustment amount.

In order to comply with IFRS 17 and set the appropriate reserve amount to protect the policyholders under the requirements set by OSFI, we need a statistical model which can capture the characteristics established in this chapter and provide the risk adjustment for non-financial risks needed for the unpaid claim liabilities. An appropriate statistical method of risk aggregation is the hierarchical copula model reviewed in Chapter [2](#).

# Chapter 2

## Review of Actuarial and Statistical Models

In order to calculate the risk adjustment for non-financial risks from IFRS 17, described in Section 1.1 and to calculate the minimum capital required by OSFI to demonstrate financial strength towards the policyholder and creditors described in Section 1.2, this Chapter describes common actuarial and statistical models which are of use in the construction of the model presented in this thesis in Chapter 3. First, we introduce the notation. Then, in Section 2.2 the marginal distributions considered for each line of business are described. In Section 2.3, the multivariate model which accounts for the diversification benefit is presented, and finally, in Section 2.4, the simulation technique for the estimation of the unpaid claim liabilities is described.

### 2.1 Notation

Claims can be presented graphically in an upper triangle array as in Table 2.1, commonly known as a run-off triangle. This presentation can be done in two different ways, with cumulative claims  $C_{i,j}$  or incremental claims  $C_{i,j} - C_{i,j-1}$ , where  $C_{i,0} = 0$ ; the latter is preferred in this thesis. Incremental claims are arranged in Table 2.1, where  $i =$

$\{1, 2, \dots, I\}$  represents the accident year,  $j = \{1, 2, \dots, J\}$  the development lag, and  $I, J$  represent the last year of available information (data can also be measured by semester, quarterly or monthly). The development lag is the number of years between the occurrence of the accident and the date in which the final payment is made (closure of case).

To standardize claims and have comparable data, either premiums or volume (number of claims) are used for each line of business, depending on available information. In this thesis, the premiums are used to create what is known as a loss ratio. We denote the loss ratio for accident year  $i$ , development lag  $j$  and for line of business  $k$  as  $Y_{i,j}^{(k)}$  and the premiums of accident year  $i$  with  $p_i^{(k)}$ :

$$Y_{i,j}^{(k)} = \frac{C_{i,j}^{(k)} - C_{i,j-1}^{(k)}}{p_i^{(k)}}. \quad (2.1.1)$$

The superscript  $k = \{1, 2, \dots, K\}$  denotes the line of business, where  $K$  is the total number of business lines available.

$i \backslash j$	1	2	...	$j$	...	...	...	$J$
1								
2								
$\vdots$								
$i$				$Y_{i,j}^{(k)}$				
$\vdots$								
$\vdots$								
$\vdots$								
$\vdots$								
$I$								

Table 2.1: General notation for a run-off triangle with loss ratios for the  $k$ -th line of business.

Note that after each year that goes by, one diagonal is added to the loss triangle. Furthermore, in this thesis there is no explicit correction for inflation; it is assumed to be captured within the development lag factors.

## 2.2 Marginal distributions

### 2.2.1 Collective Risk Model

A common practice in the insurance industry is to use the collective risk model which provides a way to understand the aggregate loss as the sum of the individual claims. Following the notation from [Klugman et al. \(2008\)](#), let  $S$  be the random variable which represents the aggregate loss of a random number of claims  $N$ , of independent and identically distributed random variables  $X_i$ ,  $i = 1, 2, \dots, n$ . In the insurance industry,  $N$  is known as the frequency component and  $X_i$  as the severity (risks also described in [Section 1.1.2](#)). Then,  $S$  has the following additive representation,

$$S = \begin{cases} 0 & N = 0, \\ X_1 + X_2 + \dots + X_N & N > 0 \end{cases} \quad (2.2.1)$$

The collective risk model has the following independence assumption:

- Conditionally on  $N = n$ ,  $N$  and the i.i.d. sequence  $X_i$ ,  $i = 1, 2, \dots, n$  are independent.

$N$  has to be a discrete count random variable and in the insurance context  $X$  follows a positive continuous random variable. In this thesis, a Poisson distribution is assumed for the count random variable and a gamma distribution for the severity.

### 2.2.2 Compound Poisson-Gamma distribution

The following definitions are useful in the construction of the objective distribution  $S$ .

**Definition 2.2.1.** *Let  $X$  be a positive continuous random variable with probability density function,*

$$f_X(x) = \frac{\beta^\alpha x^{\alpha-1}}{\Gamma(\alpha)} e^{-x\beta}, \quad (2.2.2)$$

where  $\alpha, \beta > 0$ . Then,  $X$  is said to follow a gamma distribution with shape parameter  $\alpha$  and rate parameter  $\beta$  and it is denoted  $X \sim \text{Gamma}(\alpha, \beta)$ . A gamma distribution has moment generating function (m.g.f.)  $M_X(t) = \left(1 - \frac{t}{\beta}\right)^{-\alpha}$ , for  $t < \beta$ .

**Definition 2.2.2.** Let  $N$  be a discrete random variable with support on the set of non-negative integers  $\mathbb{N} \cup \{0\}$  and probability mass function,

$$P_N(n) = \frac{e^{-\lambda} \lambda^n}{n!}, \quad (2.2.3)$$

where  $\lambda > 0$ . Then,  $N$  is said to follow a Poisson distribution with parameter  $\lambda$  and it is denoted  $N \sim \text{Poisson}(\lambda)$ . The moment generating function of a Poisson distribution is  $M_N(t) = \exp(\lambda(e^t - 1))$ , for any  $t \in \mathbb{R}$ .

**Theorem 2.2.1.** (*Klugman et al., 2008*) Let  $S$  represent the aggregate loss from the collective risk model as presented in Section 2.2.1, with frequency  $N \sim \text{Poisson}(\lambda)$  and severity  $X \sim \text{Gamma}(\alpha, \beta)$ , then the moment generating function of  $S$  is,

$$M_S(t) = \exp \left( \lambda \left[ \left(1 - \frac{t}{\beta}\right)^{-\alpha} - 1 \right] \right), \quad t < \beta. \quad (2.2.4)$$

*Proof.*

$$\begin{aligned} M_S(t) &= \mathbb{E}[e^{St}] = \mathbb{E}[\mathbb{E}[e^{St} | N]] = \mathbb{E}[\mathbb{E}[e^{(X_1 + X_2 + \dots + X_N)t} | N]] \\ &\stackrel{\text{i.i.d.}}{=} \mathbb{E}[\mathbb{E}[e^{X_1 t} | N]^N] = \mathbb{E}[M_X^N] = \mathbb{E}[e^{N \log(M_X(t))}] \\ &= M_N(\log(M_X(t))) \\ &= \exp \left( \lambda \left[ \left(1 - \frac{t}{\beta}\right)^{-\alpha} - 1 \right] \right). \end{aligned}$$

□

**Definition 2.2.3.** A distribution  $S$  with m.g.f. of the form (2.2.4) is known as a Compound Poisson-Gamma distribution, denoted as  $S \sim \text{CPG}(\lambda, \alpha, \beta)$ .

### 2.2.3 Generalized Linear Models (GLM)

With the construction of our objective distribution  $S$  in Section 2.2.2 arises the need for a methodology to link the expected aggregate claim amounts with explanatory variables. Generalized linear modeling was introduced by McCullagh (1984), allowing to focus on the effects of explanatory variables, and generalizing the classical normal linear model by relaxing some of its restrictive assumptions. The GLM framework imposes for a random variable  $Y_i$ , fixed covariates  $\mathbf{X}_i \in \mathbb{R}^p$  and for observations  $i = \{1, 2, \dots, n\}$ , that

$$g(\mathbb{E}[Y_i|\mathbf{X}_i]) = \beta_0 + \sum_{j=1}^p X_{i,j}\beta_j = \mathbf{X}_i^T \boldsymbol{\beta}, \quad (2.2.5)$$

where  $g$  is a function called the link function, which defines the relationship between the linear predictors and the mean, and  $\boldsymbol{\beta}$  is the vector of parameters that we want to estimate. If  $g$  is the identity function, we have a classic linear model. To simplify the notation of equation (2.2.5), the relationship is usually denoted by,

$$g(\mu_i) = \eta_i, \quad (2.2.6)$$

where  $\mu_i$  is the conditional expected value of the response variable for observation  $i$  and  $\eta_i = \mathbf{X}_i^T \boldsymbol{\beta}$  is the additive relation of parameters and covariates.

GLMs rely on the following assumptions:

- $Y_1, Y_2, \dots, Y_n$  are conditionally independent given  $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ ,
- $g$  must be monotonic and differentiable,
- Given the covariates  $\mathbf{X}_i$ , the  $Y_i$  are all distributed from the same member of the exponential family.

**Definition 2.2.4.** *Let  $Y$  be a random variable that belongs to the exponential family, then  $Y$  has probability density function of the form,*

$$f_Y(y; \theta, \phi) = \exp \left( \frac{y\theta - b(\theta)}{a(\theta)} + c(y; \phi) \right), \quad (2.2.7)$$



with canonical parameter  $\theta$ , dispersion parameter  $\phi$  and some specific functions  $a(\cdot)$ ,  $b(\cdot)$  and  $c(\cdot)$ . Furthermore,  $b(\cdot)$  has to be twice differentiable and the random variable  $Y$  satisfies,

$$\mathbb{E}[Y] = b'(\theta), \quad (2.2.8)$$

$$\text{Var}[Y] = b''(\theta)a(\theta). \quad (2.2.9)$$

If the dispersion parameter  $\phi$  is known, then (2.2.7) is referred to as a one-parameter exponential family. If  $\phi$  is unknown, then, the distribution is part of the two-parameter exponential family or exponential dispersion family defined in the following section.

For the selection of the link function  $g$ , it is a common practice in actuarial applications to use the log link because every covariate enters in the mean equation through a multiplicative structure which allows to provide a simple interpretation of the parameters. Another convenient option is the canonical link function since it establishes a direct connection between the canonical parameter of the exponential family and the covariates, i.e.  $\theta_i = \mathbf{X}_i^T \boldsymbol{\beta}$ . But in some cases like the Gamma distribution, it allows the mean to vary on the real numbers and thus, to enforce positive means the log link is used in this thesis.

## 2.2.4 Exponential Dispersion Family (EDF)

GLMs defined in 2.2.3 were originally developed for the exponential family. Then, Jørgensen (1997) introduced the exponential dispersion family by analyzing the error distribution of the GLMs. To draw a comparison between the exponential family and the exponential dispersion family, we rewrite a random variable  $Y$  from the exponential family with density function as in equation (2.2.7), in its canonical parametrization:

$$f_Y(y; \theta) = c(y) \exp(y\theta - \kappa(\theta)), \quad (2.2.10)$$

where  $\theta$  remains the canonical parameter and  $\kappa(\cdot)$  is the cumulant generator. Equation (2.2.10) lacks the parameter  $\phi$  in contrast with equation (2.2.7) because for the usual

members of the exponential family (e.g. Normal, Poisson, Gamma and Binomial distribution), the dispersion parameter  $\phi$  is assumed to be known. The definition for the EDF is as follows.

**Definition 2.2.5.** *The distribution of a random variable  $Y$  is part of the exponential dispersion family if it has probability function of the form*

$$f_Y(y; \theta, \phi) = c(y, \phi) \exp \left( \phi^{-1}(y\theta - \kappa(\theta)) \right), \quad (2.2.11)$$

where  $c(\cdot, \cdot)$  and  $\kappa(\cdot)$  are given functions,  $\theta \in \mathbb{R}$  is the canonical parameter and  $\phi > 0$  is the dispersion parameter.  $\kappa(\cdot)$  is known as the cumulant function and is assumed to be twice differentiable.

Comparing the density function of the canonical parametrization of the exponential family (2.2.10) with the density function of the exponential dispersion family (2.2.11), we can deduce the latter is a generalization since they are equal up to an additional parameter, the dispersion parameter  $\phi$ . The intention of the dispersion parameter, as interpreted by Madsen and Thyregod (2010), is to separate the mean from dispersion features like the sample size or common over-dispersion effects not related to the mean. Thus, it is a perfect fit for insurance claims because an assumption of the collective risk model (described in 2.2.1) is that conditionally on the number of claims, the frequency and the severity are independent.

The moment generating function of a random variable of the form (2.2.11) is

$$M_Y(t) = \exp \left( \frac{\kappa(\theta + t\phi) - \kappa(\theta)}{\phi} \right). \quad (2.2.12)$$

Following the notation from the original author Jørgensen (1987), any random variable  $Y$  member of the EDF can be parametrized in terms of the location  $\mu$  and dispersion  $\phi$  denoted as  $Y \sim ED(\mu, \phi)$ . The random variable  $Y$  has first moment  $\mathbb{E}[Y] = \mu = \kappa'(\theta)$  and variance  $\text{Var}[Y] = \phi V(\mu)$ , where  $V(\mu) = \kappa''(\theta)$  as presented in Jørgensen (1987). The function  $V(\mu)$  is known as the variance function and captures the mean-variance relationship of the data as it is explicitly seen for the Tweedie family in Section 2.2.4.1.

Another generalized concept is the residual sum of squares from the analysis of variance, for a member of the exponential dispersion family, it is called analysis of deviance and is equivalent to sums of unit deviances [Jørgensen \(1997\)](#).

**Definition 2.2.6.** *A function  $d : \Omega \rightarrow \mathbb{R}$  is called a unit deviance if it satisfies*

$$d(y; y) = 0 \quad \forall y \in \Omega$$

*and*

$$d(y; \mu) > 0 \quad \forall y \neq \mu.$$

*where  $\Omega$  is the domain of the parameter  $\mu$ .*

The unit deviance  $d$  can be understood as a measure of the distance from  $y$  to the mean  $\mu$ . Therefore, to analyze the dispersion features, it is necessary to calculate the unit deviances from the estimated mean. Additionally, with the unit deviance a random variable  $Y$  from the exponential dispersion family can be reparametrized in the standard form,

$$f_Y(y; \theta, \phi) = c(y, \phi) \exp(\phi^{-1} d(y; \mu)). \quad (2.2.13)$$

Once the mean and dispersion have been estimated for different lines of business, in this thesis, the following theorem is essential to homogenize and study the dependence structure.

**Theorem 2.2.2.** *If  $Y \sim ED(\mu, \phi)$  for some  $\mu \in \mathbb{R}$ , we have*

$$\frac{Y - \mu}{\sqrt{\phi V(\mu)}} \xrightarrow{d} N(0, 1) \quad \text{as } \phi \rightarrow 0. \quad (2.2.14)$$

*Proof.* Refer to [Jørgensen \(1997\)](#).

Theorem [2.2.2](#) shows that any member of the EDF is asymptotically normal for a small dispersion parameter  $\phi$ .

### 2.2.4.1 Tweedie family

The Tweedie family, in honor of [Tweedie \(1984\)](#), is a member of the EDF where the variance function is of the form,

$$V(\mu) = \mu^p, \quad (2.2.15)$$

for some index  $p \in (-\infty, 0] \cup [1, \infty)$ <sup>1</sup>. A random variable of the Tweedie family is denoted  $Y \sim TW_p(\mu, \phi)$ .

Index	Distribution
$p = 0$	Normal
$p = 1$	Poisson
$p \in (1, 2)$	Compound Poisson-Gamma
$p = 2$	Gamma
$p = 3$	Inverse Gaussian

Table 2.2: Some cases of the Tweedie family depending on the index parameter  $p$ .

Table 2.2 shows how some very well known distributions are part of the Tweedie family. We pay special interest to the case where  $p \in (1, 2)$ . For  $p = 1$ , we have a Poisson distribution, which is discrete, and for  $p = 2$  we have a continuous Gamma distribution. When  $p \in (1, 2)$  the distribution is mixed, continuous for positive values with a point of mass at zero, and as  $p \rightarrow 2$ , the distribution starts losing the point of mass. This mixed domain makes it very relevant for actuarial analysis, given that for certain years, specially several years after the accident year, the claims converge to zero as the development lag increases, and in some cases, is exactly zero.

To obtain the probability density function and the moment generating function of a

---

<sup>1</sup>The case  $p \in (0, 1)$  has been shown by [Jørgensen \(1997\)](#) to have null variance for some cases of the canonical parameter making it a degenerate distribution and thus, concluding there are no members of the Tweedie family for these values of  $p$ .

Tweedie distribution, we use the property of the EDF where  $V(\mu) = \kappa''(\theta)$  and the procedure described by [Dunn and Smyth \(2004\)](#). Thus, the canonical parameter  $\theta$  is obtained by integrating the following differential equation and setting the constant equal to zero.

$$\begin{aligned}\mu^p &= \kappa''(\theta) = \frac{\partial \mu}{\partial \theta} \\ \int \frac{1}{\mu^p} d\mu &= \int \frac{\partial \theta}{\partial \mu} \\ \Rightarrow \theta &= \begin{cases} \frac{\mu^{1-p}}{1-p} & p \neq 1, \\ \log \mu & p = 1. \end{cases}\end{aligned}\tag{2.2.16}$$

And the cumulant function  $\kappa_p(\theta)$  (with subscript  $p$  to denote the functional dependence on the index) is obtained by integrating  $\kappa'_p(\theta) = \mu$  with respect to  $\theta$ :

$$\kappa_p(\theta) = \begin{cases} \frac{1}{2-p} ((1-p)\theta)^{\frac{2-p}{1-p}} & p \neq \{1, 2\}, \\ e^\theta & p = 1, \\ -\log(-\theta) & p = 2. \end{cases}\tag{2.2.17}$$

Thus, to obtain the connection between the Tweedie family (member of the EDF) with a CPG distribution, we substitute the canonical parameter (2.2.16) and the cumulant function (2.2.17) (when  $p \neq \{1, 2\}$ ) into the moment generating function of  $Y \sim ED(\mu, \phi)$  shown in equation (2.2.12),

$$M_Y(t) = \exp \left( \frac{1}{\phi} \left[ \frac{1}{2-p} \left( (\mu^{1-p} + (1-p)\phi t)^{\frac{2-p}{1-p}} - \mu^{2-p} \right) \right] \right).\tag{2.2.18}$$

Then, [Quijano Xacur \(2011\)](#) proved the existence of Tweedie families for  $p \in (1, 2)$  and shows that they are equivalent to a CPG( $\lambda, \alpha, \beta$ ) distribution with the following transformations,

$$\lambda = \frac{1}{\phi} \frac{\mu^{2-p}}{2-p}, \quad \alpha = -\frac{2-p}{1-p}, \quad \beta = -\frac{1}{\phi} \frac{\mu^{1-p}}{1-p}.\tag{2.2.19}$$

Then, the moment generating function (2.2.18) becomes:

$$\begin{aligned} M_Y(t) &= \exp \left( \frac{(\mu^{1-p} + (1-p)\phi t)^{-\alpha}}{\phi(2-p)} - \lambda \right) \\ &= \exp \left( \lambda \left[ \left( 1 - \frac{t}{\beta} \right)^{-\alpha} - 1 \right] \right). \end{aligned} \quad (2.2.20)$$

Finally, we have showed the connection between the CPG distribution and the Tweedie family with  $p \in (1, 2)$  through the moment generating function by showing that (2.2.12) is equal to (2.2.20), provided the appropriate transformations (2.2.19) are used.

Another important result with actuarial applications from Jørgensen (1997) is the scale invariance of the Tweedie family,

**Theorem 2.2.3.** *Let  $Y \sim TW_p(\mu, \phi)$ , then*

$$cY \sim TW_p(c\mu, c^{2-p}\phi). \quad (2.2.21)$$

*Proof.* Refer to Jørgensen (1997).

Theorem 2.2.3 is fundamental to homogenize the data and sample from the Tweedie family as done in Section 2.4.1 when the dispersion parameter  $\phi$  is assumed constant across observations. In this thesis, a GLM is used to estimate the dispersion parameter and therefore, another approach is taken.

**Definition 2.2.7.** *The probability density function of a Tweedie random variable  $Y \sim TW_p(\mu, \phi)$  is of the form,*

$$f_Y(y; \mu, \phi, p) = a(y; \phi, p) \exp \left[ \frac{1}{\phi} \left( y \frac{\mu^{1-p}}{1-p} - \frac{\mu^{2-p}}{2-p} \right) \right] \quad (2.2.22)$$

$$\text{where } a(y; \phi, p) = \sum_{r=1}^{\infty} \left[ \frac{\phi^{p-1} y^{\ell}}{(2-p)(p-1)^{\ell}} \right]^r \frac{1}{r! \Gamma(r\ell) y}, \quad \ell = -\frac{2-p}{1-p}.$$

Thus, the point of mass at zero has probability given by,

$$f_Y(0; \mu, \phi, p) = \exp \left[ -\frac{\mu^{2-p}}{\phi(2-p)} \right].$$

### 2.2.5 Double Generalized Linear Models (DGLM)

We have introduced GLMs in Section 2.2.3 to link the expected insurance claims with explanatory variables. Then, by analyzing the error distribution of the GLMs Jørgensen (1997) introduced the EDF, as described in Section 2.2.4. The EDF led us to the Tweedie family in Section 2.2.4.1 as a link with the CPG distribution under the appropriate transformations. In this section, to account for the dispersion features we study DGLMs. These double generalized linear models allow the mean and the dispersion to be modeled simultaneously using GLMs. DGLM handles the case where only the aggregate claims is available but the number of claims has not been recorded or is unreliable, see Smyth and Jørgensen (2002). Furthermore, DGLM is a more flexible model since we add a new set of parameters to be estimated for the dispersion and thus, should be used carefully to avoid overfitting. Research in actuarial science like Smyth and Jørgensen (2002); Boucher and Davidov (2011), and more recently Andersen and Bonat (2017); Smolárová (2017) have highlighted the importance of modeling the mean and dispersion structures for claims reserving.

Introduced by Smyth (1989), the simultaneous estimation for the mean and dispersion for a Tweedie distribution is possible due to the statistical orthogonality of the parameters  $\phi$  and  $p$  to  $\mu$ , as explained in what follows.

**Definition 2.2.8.** *The parameter orthogonality introduced by Cox and Reid (1987), states that if we partition the parameter vector of interest  $\boldsymbol{\theta} = (\mu, \phi, p)$  into  $\theta_1 = \mu$  and  $\boldsymbol{\theta}_2 = (\phi, p)$ , and the information matrix of the corresponding log-likelihood  $\mathcal{L}$  satisfies,*

$$\mathbb{E} \left( -\frac{\partial^2 \mathcal{L}}{\partial \theta_1 \partial \boldsymbol{\theta}_2}; \boldsymbol{\theta} \right) = 0, \quad (2.2.23)$$

*then,  $\theta_1$  and  $\boldsymbol{\theta}_2$  are said to be locally orthogonal.*

Equation (2.2.23) implies mainly that the maximum likelihood estimates of  $\theta_1$  and  $\boldsymbol{\theta}_2$  are asymptotically independent. Refer to Cox and Reid (1987) for further implications.

### 2.2.5.1 Estimation of DGLM

In this section, we explain the estimation procedure used to obtain the parameters of the mean model and the dispersion submodel for independent observations from a Tweedie distribution through a DGLM. The algorithm procedure described in what follows serves as a building block for algorithms in the presence of correlation between observations as illustrated in subsequent sections. The procedure consists in an alternating scheme presented in [Smyth \(1989\)](#) where the dispersion is assumed to be known when estimating the mean and then, the mean is assumed to be known when estimating the dispersion. The alternation is possible due to the parameter orthogonality (Definition [2.2.8](#)) of the parameter vector. The estimation of the index parameter  $p$ , the power of the variance function for a member of the Tweedie family, is a more difficult problem than estimating  $\phi$  or  $\mu$ . Most authors using Tweedie densities have taken  $p$  to be specified a priori as in [Dunn and Smyth \(2004\)](#). In other words,  $p$  is fixed in the interval  $(1, 2)$  to guarantee the existence of the connection with the CPG distribution, and then, the process to obtain the maximum likelihood estimates of  $\mu$  and  $\phi$  begins. The process is repeated with several values of the index  $p$  until the likelihood function is maximized within a given threshold. This procedure to estimate  $\mu$  and  $\phi$  is explained in what follows.

For observations  $\mathbf{y} = \{y_1, y_2, \dots, y_n\}$ , from the random variable  $Y_i | \mathbf{X}_i \sim \text{TW}_p(\mu_i, \phi_i)$  we assume a DGLM with the following equations:

$$g(\mu_i) = \mathbf{X}_i^T \boldsymbol{\beta}, \quad g(\phi_i) = \mathbf{Z}_i^T \boldsymbol{\gamma}, \quad (2.2.24)$$

where  $\mathbf{X}, \mathbf{Z}$  are the matrix of fixed covariates for the mean and dispersion submodel, respectively. The same log-link function  $g(x) = \log(x)$  is used for both GLMs (reasons explained in Section [2.2.3](#)) and  $\boldsymbol{\beta}, \boldsymbol{\gamma}$  are the parameter vectors we wish to estimate. Furthermore, to simplify notation we denote the vector of mean parameters as  $\boldsymbol{\mu} = \{\mu_1, \dots, \mu_n\}$  and the vector of dispersion parameters as  $\boldsymbol{\phi} = \{\phi_1, \dots, \phi_n\}$ . Initially, we assume the dispersion parameters  $\boldsymbol{\phi}$  are known and using maximum likelihood we estimate the vector



$\beta$ . We denote the log-likelihood function as,

$$\mathcal{L}(\beta|\mathbf{y}, \gamma, p) = \sum_{i=1}^n \log f_Y(y_i; \mu_i, \phi_i, p). \quad (2.2.25)$$

We proceed to establish a recursive algorithm to estimate  $\beta$  as presented in [Hardin and Hilbe \(2013\)](#). The estimates of the parameter vector  $\beta$  are the solution of the estimating equation given by,

$$\frac{\partial \mathcal{L}}{\partial \beta} = 0. \quad (2.2.26)$$

The solution denoted as  $\beta^*$  can be obtained by a Taylor series expansion,

$$\frac{\partial \mathcal{L}}{\partial \beta}(\beta^*) - (\beta - \beta^*) \frac{\partial^2 \mathcal{L}}{\partial \beta \partial \beta^T} + \dots = 0, \quad (2.2.27)$$

by solving for  $\beta^*$  and substituting the expected value of the Hessian matrix given by,

$$- \mathbb{E} \left[ \frac{\partial^2 \mathcal{L}}{\partial \beta \partial \beta^T} \right] = X^T W X. \quad (2.2.28)$$

Results in the weighted ordinary least squares equation for iteration  $k + 1$ ,

$$\beta^{(k+1)} = (X^T W X)^{-1} X^T W \mathbf{z}, \quad (2.2.29)$$

where  $W$  is the diagonal matrix of weights. To introduce  $W$ , we first denote  $\text{diag}(w_i)$  as the notation to represent a diagonal matrix whose entries are the elements  $\{w_1, \dots, w_n\}$ .

Consequently, the weights to estimate the parameters of the mean model are,

$$W = \text{diag} \left[ \left( \frac{\partial g(\mu_i)}{\partial \mu} \right)^{-2} \frac{1}{\text{Var}(y_i)} \right] = \text{diag} \left( \frac{\mu_i^{2-p}}{\phi_i} \right), \quad (2.2.30)$$

where  $\text{Var}(y_i) = \phi_i V(\mu_i)$ ,  $\mu_i = g^{-1}(X_i^T \beta^{(k)})$  and the right side equation is obtained through the log-link function and the Tweedie distributional assumption with variance function of the form (2.2.15). The vector  $\mathbf{z}$  from equation (2.2.29) has components,

$$z_i = \frac{\partial g(\mu_i)}{\partial \mu} (y_i - \mu_i) + g(\mu_i) = \frac{y_i - \mu_i}{\mu_i} + \log \mu_i. \quad (2.2.31)$$

After estimating the parameters of the mean model  $\beta$ , we proceed to estimate the parameters of the dispersion submodel  $\gamma$ . As stated in Section 2.2.4, from the theory of dispersion models by Jørgensen (1997), the unit deviance is the equivalent of the sum of squares in analysis of variance. The unit deviance 2.2.6 for the Tweedie family is given by,

$$d_i(y_i; \mu_i) = \begin{cases} 2 \left( y_i \frac{y_i^{1-p} - \mu_i^{1-p}}{1-p} - \frac{y_i^{2-p} - \mu_i^{2-p}}{2-p} \right) & y_i \neq 0, \\ 2 \frac{\mu_i^{2-p}}{2-p} & y_i = 0. \end{cases} \quad (2.2.32)$$

By assuming the mean parameters  $\beta$  are fixed, Smyth (1989) showed that the log-likelihood given in equation (2.2.25) can be reparametrized in terms of the unit deviance, as done with equation (2.2.13) to obtain the standard form, causing the dispersion submodel to have the form of a GLM with observations  $d_i$ , canonical parameter  $\phi_i$  and dispersion parameter 2. Thus, analogous to the mean model with observations  $y_i$  and canonical parameter  $\mu_i$ , we use the unit deviance as the response vector of the dispersion submodel.

Furthermore, Smyth and Verbyla (1999) showed that  $d_i \sim \phi_i \chi_1^2$  approximately as  $\phi_i \rightarrow 0$  with the saddlepoint approximation (refer to Jørgensen (1997) for more on the saddlepoint approximation). Assuming a gamma distribution instead of a chi-square simplifies the procedure for the weight matrix of the dispersion submodel as shown in what follows. This assumption is possible given that the  $\chi_1^2$  distribution is a special case of the gamma distribution (2.2.1), therefore, a gamma GLM is fitted for the dispersion submodel. Maximum likelihood estimators for variance parameters in regression models are generally biased, thus, a restricted maximum likelihood (REML) approach is used as in Smyth and Verbyla (1999) to adjust for degrees of freedom and produce estimators that are approximately unbiased.

Consequently, the unit deviances (2.2.32) are calculated and the parameters for the dispersion submodel  $\gamma$  are obtained by using the same technique as for the mean model: Taylor series expansion for  $\mathcal{L}(\gamma|\mathbf{y}, \beta, p)$ , substitute expected value of Hessian matrix and

thus, for iteration  $k + 1$  the weighted ordinary least squares equation for the dispersion submodel is,

$$\boldsymbol{\gamma}^{(k+1)} = (Z^T W_d Z)^{-1} Z^T W_d \mathbf{z}_d, \quad (2.2.33)$$

where  $W_d$  is the matrix of weights for the dispersion submodel using the same notation as [Smyth and Jørgensen \(2002\)](#). The matrix  $W_d$  is given by,

$$W_d = \text{diag} \left[ \left( \frac{\partial g(\phi_i)}{\partial \mu} \right)^{-2} \frac{1}{2V(\phi_i)} \right]. \quad (2.2.34)$$

Since we are fitting a gamma GLM, member of the Tweedie family with  $p = 2$  (refer to Table 2.2), the variance function becomes  $V(\phi_i) = \phi_i^2$ . Thus, the weight matrix (2.2.34) simplifies to a diagonal matrix with entries  $\frac{1}{2}$ . Furthermore in equation (2.2.33), the vector  $\mathbf{z}_d$  has the same structure as equation (2.2.31) but with response  $d_i$ . A REML approach is used as [Smyth and Jørgensen \(2002\)](#) to correct for the bias of the maximum likelihood estimators for the variance parameter as mentioned in Section 2.2.5. The dispersion submodel is modified with,

$$\boldsymbol{\gamma}^{(k+1)} = (Z^T W_d^* Z)^{-1} Z^T W_d^* \mathbf{z}_d^*, \quad (2.2.35)$$

where

$$W_d^* = \text{diag} \left( \frac{1 - h_i}{2} \right), \quad \mathbf{z}_d^* = \frac{d_i^* - \phi_i}{\phi_i} + \log \phi_i. \quad (2.2.36)$$

The modified responses are  $d_i^* = \frac{d_i}{1 - h_i}$  and  $h_i$  are the diagonal elements of the projection matrix  $H$ , known as leverages from the mean model. The projection matrix  $H$  is given by,

$$H = \text{diag}(h_i) = W^{1/2} X (X^T W X)^{-1} X^T W^{1/2}. \quad (2.2.37)$$

This REML approach is possible because of the extension made by [Cox and Reid \(1987\)](#) and the simplifications on the convergence algorithm provided by [Lee and Nelder \(1998\)](#).

### 2.2.5.2 DGLM estimation algorithm

To summarize, the estimation algorithm for the DGLM is as follows:

1. Set initial values  $\mu_i^{(0)} = y_i, \quad \phi_i^{(0)} = 1,$
2. For iteration  $k$ , obtain  $\beta^{(k)}$  with the equation (2.2.29),
3. Calculate the unit deviances  $d_i$  with equation (2.2.32),
4. Obtain  $\gamma^{(k)}$  with equation adjusted for the dispersion submodel (2.2.35) using the REML approach,
5. Calculate the maximum likelihood  $\mathcal{L}(\mathbf{y}, p, \boldsymbol{\mu}^{(k)}, \boldsymbol{\phi}^{(k)})$  with the estimated parameters,
6. Set  $k = k + 1$  and repeat steps 2-5 until convergence.

## 2.2.6 Goodness-of-fit

The goal is to determine a model that is good enough for the marginal distributions to continue with the modeling procedure. We have to remember that “*All models are wrong, but some models are useful.*”<sup>2</sup>

A model deemed *good enough* in this thesis is defined as a marginal model which will not reject the null hypothesis  $H_0$ : the data came from a population with the stated model. This is done through the Kolmogorov-Smirnov and Anderson-Darling tests to assess the fit on the center and tails of the distribution, respectively. The tests are defined as follows,

### 2.2.6.1 Kolmogorov-Smirnov (K-S)

For a random variable  $X$  and independent observations  $\{x_1, \dots, x_n\}$  from  $X$ , let  $F_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}(x_i \leq x)$  be the empirical cumulative distribution function, where  $\mathbb{1}(A)$  denotes the indicator function which is set to 1 if  $A$  occurs and zero otherwise. Additionally, let

---

<sup>2</sup>Usually attributed to George Box.

$F(x)$  be the model cumulative distribution function. Then, the statistic of the K-S test is calculated as,

$$D = \sup_x |F_n(x) - F(x)|. \quad (2.2.38)$$

The model distribution function  $F(x)$  is assumed to be continuous over the relevant range [Klugman et al. \(2008\)](#).

#### 2.2.6.2 Anderson-Darling (A-D)

The test is similar to the Kolmogorov-Smirnov test but it uses a different measure of the distance between the two cumulative distribution functions. The test statistic is

$$A^2 = n \int_0^1 \frac{(F_n(x) - F(x))^2}{F(x)(1 - F(x))} f(x) dx. \quad (2.2.39)$$

The Anderson-Darling test assesses whether the estimated model has a good fit close to the tails of the distribution. This can be seen in the denominator of the statistic; as  $F(x)$  gets closer to 0 and 1, the quotient reaches its maximum value.

The combination of the Kolmogorov-Smirnov test in Section [2.2.6.1](#) and the Anderson-Darling test in Section [2.2.6.2](#) ensures that the center and the tails of the fitted distribution are being assessed statistically.

### 2.2.7 Generalized Estimating Equations (GEE)

An important inspection for a fitted GLM is a residual plot in order to verify the independence assumption of the response variable given the fixed covariates as stated in the GLM Section [2.2.3](#). For the loss ratio  $Y_{i,j}$  and the vector of fixed covariates  $\mathbf{X}_{i,j}$ , it is reasonable to suspect dependence between  $Y_{i,j}|\mathbf{X}_{i,j}$  and the loss ratio of the next development lag  $Y_{i,j+1}|\mathbf{X}_{i,j+1}$ . To address the issue when there is correlation between observations, GEE suggest a way to estimate efficient parameters with the same concepts of a GLM model. Thus, a two-stage estimation method is proposed in this section by first estimating the

DGLM from Section 2.2.5 and then, to account for correlation, apply the estimation of the GEE presented in what follows. An example of claims reserving with GEE is mentioned in Hudecová and Pešta (2013) and this problem applied to non-life insurance data is mentioned in Smolárová (2017), but there are differences in this thesis regarding to the correlation estimator presented in what follows and the estimator of  $\phi$  given that we consider a DGLM, compared to the moment estimator applied in Smolárová (2017).

GEE is a an *ad hoc* method which can make it unattractive for some researchers. There are several research articles dealing with GLM models for non-life insurance, e.g. Shi and Frees (2011) and Côté et al. (2016), where correlation is not an issue raised by the authors. The usual application of GEE is in biostatistics, where the correlation is assumed for each subject  $i$  between the different  $j$  measurements. The concepts are borrowed from this field and adapted to our context to capture the correlation for every accident year  $i$  between the different  $j$  development lags. The theory is also developed to account for unbalanced data which is exactly the case for a loss triangle given that for every accident year  $i$ , there is exactly one less development lag  $j$  due to the triangular nature of the data as it can be seen in what follows. First, we introduce the linear correlation considered for the generalized estimating equations.

**Definition 2.2.9.** *For two square integrable random variables  $X$  and  $Y$ , Pearson's correlation is defined as,*

$$\rho(X, Y) = \frac{\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}. \quad (2.2.40)$$

The correlation parameter  $\rho$  is an additional parameter for the marginal distributions as it is described in what follows. The loss ratios for accident year  $i = \{1, 2, \dots, I\}$  are grouped in a vector to ease the notation as in Figure 2.1,

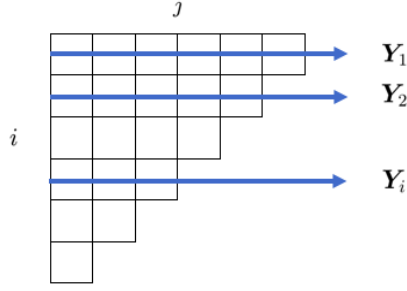


Figure 2.1: Vectorial notation of loss ratios for GEE.

Thus, from Figure 2.1, the random vector  $\mathbf{Y}_i = \{Y_{i,1}, \dots, Y_{i,n_i}\}$  corresponds to the vector of  $n_i$  loss ratios with each component distributed  $Y_{i,j}|\mathbf{X}_{i,j} \sim \text{TW}_p(\mu_{i,j}, \phi_{i,j})$  for accident year  $i$ , development lag  $j$  and correlation parameter  $\rho$  which is explained in what follows. The mean vector of  $\mathbf{Y}_i$  is  $\boldsymbol{\mu}_i = \{\mu_{i,1}, \dots, \mu_{i,n_i}\}$  and has dispersion vector  $\boldsymbol{\phi}_i = \{\phi_{i,1}, \dots, \phi_{i,n_i}\}$ . The following results and notation are taken and adapted from Liang and Zeger (1986). Let  $X, Z$  be the same matrices of fixed covariates for the mean model and dispersion vector submodel, respectively, as in Section 2.2.5.1. The conditional variance matrix of the random vector is denoted  $\mathbf{V}_i = \text{Var}[\mathbf{Y}_i|\mathbf{X}_i]$ . The following generalized estimating equation obtained from Liang and Zeger (1986) corresponds to the weighted least squares estimator, which ensures the estimator of  $\boldsymbol{\beta}$  to remain consistent,

$$\sum_{i=1}^{n_i} D_i^T \mathbf{V}_i^{-1} (\mathbf{Y}_i - \boldsymbol{\mu}_i) = 0, \quad (2.2.41)$$

where  $D_i = \frac{\partial \boldsymbol{\mu}_i}{\partial \boldsymbol{\beta}}$  is a matrix of derivatives with dimension  $(n_i \times q)$ ,  $\boldsymbol{\beta}$  the parameter vector of dimension  $q$  and variance matrix  $\mathbf{V}_i = A_i^{1/2} R_i(\rho) A_i^{1/2}$ , where the matrix  $A_i$  is presented in what follows. The variance matrix  $\mathbf{V}_i$  of dimension  $(n_i \times n_i)$  is key to capture the correlation between observations. The diagonal matrix  $A_i$  is given by,

$$A_i = \begin{bmatrix} \phi_{i,1} V(\mu_{i,1}) & 0 & \dots & 0 \\ 0 & \phi_{i,2} V(\mu_{i,2}) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \phi_{i,n_i} V(\mu_{i,n_i}) \end{bmatrix}_{(n_i \times n_i)}, \quad (2.2.42)$$

where the elements  $V$  represent the variance function of the Tweedie family with equation (2.2.15). Furthermore,  $R_i(\rho)$  is the correlation matrix of the random vector  $\mathbf{Y}_i|\mathbf{X}_i$  and in this thesis, we assumed the correlation matrix  $R_i(\rho)$  to be functionally related to the scalar  $\rho$  which can be easily extended to be a vector depending on the independence assumptions. If the correlation matrix  $R_i(\rho) = I_{n_i}$ , where  $I_{n_i}$  is the identity matrix of dimension  $(n_i \times n_i)$ , then independence is assumed and the model is equivalent to the normal GLM estimation. Common correlation matrices can be found in [Hardin and Hilbe \(2013\)](#) with their respective estimators. In this thesis, an autoregressive model of order 1 or AR(1), is used to minimize the number of estimated parameters while providing a reasonable correlation matrix. An AR(1) assumes  $\text{cor}(Y_{i,j}, Y_{i,j'}|\mathbf{X}_{i,j}, \mathbf{X}_{i,j'}) = \rho^{|j-j'|}$ . Thus, the further apart the observations are in time, a smaller correlation is assumed. The correlation matrix for claims in accident year  $i$ , has the form,

$$R_i(\rho) = \begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^{n_i-1} \\ \rho & 1 & \rho & \dots & \rho^{n_i-2} \\ \rho^2 & \rho & 1 & \dots & \rho^{n_i-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{n_i-1} & \rho^{n_i-2} & \rho^{n_i-3} & \dots & 1 \end{bmatrix}_{(n_i \times n_i)} \quad (2.2.43)$$

The correlation matrix in (2.2.43) assumes independence across accident years but for a given accident year, loss ratios are dependent between development lags. A parallel can be drawn between the general estimating equation (2.2.41) introduced in this section and the equation (2.2.29) used for the DGLM estimation. Instead of having a diagonal matrix of weights  $W$ , we use the variance matrix  $\mathbf{V}_i$  with non-zero off-diagonal entries in the presence of correlation. Thus, to obtain the estimator of  $\beta$  for the GEE equation, we follow the same procedure as in the DGLM Section 2.2.5 by using a Taylor expansion and replacing the expected value of the Hessian matrix with the variance matrix presented in this section. After fitting the DGLM with the algorithm in Section 2.2.5.1, we need to



check for autocorrelation. Therefore, we calculate the residuals,

$$\hat{r}_{i,j} = \frac{y_{i,j} - \hat{\mu}_{i,j}}{\sqrt{\hat{\phi}_{i,j} \hat{\mu}_{i,j}^p}}. \quad (2.2.44)$$

Relying on the asymptotic convergence stated in Theorem 2.2.2, the residuals  $\{\hat{r}_{i,j}\} \xrightarrow{d} N(0, 1)$  converge asymptotically in distribution to a standard normal random variable as  $\phi \rightarrow 0$ . Since we are dealing with loss ratios (i.e. usually, numbers between 0 and 1) it is reasonable to assume a dispersion parameter close to 0 as seen more explicitly with the assumptions made for the model presented in Chapter 3. We calculate the residuals in (2.2.44) to test if the independence assumption is valid. To test this we order the residuals first by accident year and then by development lag, leading to the following notation, where we denote the ordered set,

$$\hat{\mathbf{r}} = \{\hat{r}_{1,1} \dots, \hat{r}_{1,n_1}, \hat{r}_{2,1}, \dots, \hat{r}_{2,n_2}, \dots, \hat{r}_{I,n_I}\} = \{\hat{r}_1, \hat{r}_2, \dots, \hat{r}_N\}, \quad (2.2.45)$$

where  $N = \frac{(n_I+1) \cdot n_I}{2}$  is the total number of residuals. Then, proceed with a Ljung & Box (L-B) test (for more technical details refer to Ljung and Box (1978)), with the following test statistic,

$$Q(\hat{\mathbf{r}}) = N(N+2) \sum_{i=1}^H \frac{\hat{\rho}^2(i)}{N-i}, \quad (2.2.46)$$

where  $H$  is the number of autocorrelation lags to be tested, which in this case, an AR(1) is assumed, and  $\hat{\rho}(i)$  is the sample autocorrelation function. If the null hypothesis,  $H_0$  : The series is a square integrable strong white noise<sup>3</sup> of the series  $\hat{\mathbf{r}}$ , is rejected, then, estimate the correlation parameter  $\rho$  for the correlation matrix  $R_i(\rho)$ . The  $p$ -value of the statistical test is obtained with an approximative chi-square distribution as the residuals are a proxy of the white noise. In this thesis, to estimate the correlation parameter  $\rho$ , the

---

<sup>3</sup>A strong white noise  $X$  is a sequence of integrable i.i.d random variables such that  $\mathbb{E}[X_t] = 0 \forall t$ .

ordinary least squares estimator is used,

$$\hat{\rho} = \frac{\sum_{i=1}^{I-1} \sum_{j=2}^{n_i} \hat{r}_j \hat{r}_{j-1}}{\sum_{i=1}^{I-1} \sum_{j=2}^{n_i} \hat{r}_{j-1}^2}. \quad (2.2.47)$$

### 2.2.7.1 GEE estimation algorithm

In this section, we provide an algorithm which handles the estimation of the DGLM parameters in the presence of correlation between observations. Once the correlation has been confirmed through the Ljung-Box test (2.2.46), the algorithm is as follows.

1. Set the estimates from the DGLM algorithm in Section 2.2.5.2 as initial values  $\hat{\boldsymbol{\mu}}_i^{(0)}$  and  $\hat{\boldsymbol{\phi}}_i^{(0)}$ ,
2. For iteration  $k$ , compute the residuals  $r_{i,j}^{(k)}$  with equation (2.2.44) under the asymptotic normality assumption,
3. Compute the correlation parameter  $\hat{\rho}$  with equation (2.2.47) and set the block diagonal correlation matrix:

$$R = \begin{bmatrix} R_1(\hat{\rho}) & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & R_2(\hat{\rho}) & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & R_{n_i}(\hat{\rho}) \end{bmatrix}_{(N \times N)},$$

4. Compute the variance matrix  $\mathbf{V} = A^{1/2} R A^{1/2}$ , where  $A$  has block diagonal entries  $A_i$  as in equation (2.2.42),
5. Compute the matrix of derivatives  $D = \frac{\partial \boldsymbol{\mu}_i}{\partial \boldsymbol{\beta}} = \hat{\boldsymbol{\mu}}^{(k)T} X$  (due to the log-link function),
6. Re-estimate the parameters of the mean model,

$$\hat{\boldsymbol{\beta}}^{(k+1)} = \hat{\boldsymbol{\beta}}^{(k)} + (D^T \mathbf{V}^{-1} D)^{-1} (D^T \mathbf{V}^{-1} (\mathbf{y} - \hat{\boldsymbol{\mu}}^{(k)})) \quad (2.2.48)$$

7. Compute  $\hat{\boldsymbol{\mu}}^{(k+1)}$  with the new estimated parameters,
8. Do steps 3-4 from the DGLM algorithm in Section 2.2.5.2 to re-adjust the parameters of the dispersion submodel  $\boldsymbol{\gamma}^{(k+1)}$ ,
9. Set  $k = k + 1$  and repeat steps 2-8 until convergence of the correlation parameter  $\hat{\rho}$ .

Before we proceed to Section 2.3 with the dependence analysis, compute the uncorrelated residuals  $\mathbf{r}^*$ ,

$$\mathbf{r}^* = L^{-1} \cdot \mathbf{r}, \quad (2.2.49)$$

where  $L$ , called the Cholesky factor, is the lower triangular matrix of the Cholesky decomposition of the full correlation matrix  $R$  and  $\mathbf{r}$  is the ordered vector of residuals defined in (2.2.45). Thus,  $L^{-1}$  is the block diagonal matrix of dimension  $(N \times N)$ , containing  $L_i^{-1}(\rho)$  on entry  $i$ , with the following form,

$$L_i^{-1}(\rho) = \frac{1}{\sqrt{1 - \rho^2}} \begin{bmatrix} \sqrt{1 - \rho^2} & 0 & 0 & \dots & 0 & 0 \\ -\rho & 1 & 0 & \dots & 0 & 0 \\ 0 & -\rho & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & -\rho & 1 \end{bmatrix}_{(n_i \times n_i)}. \quad (2.2.50)$$

Matrix  $L_i^{-1}(\rho)$  can be shown to have the form as in (2.2.50) by using the Cholesky-Banachiewicz algorithm if and only if  $R$  is positive definite.

### 2.2.7.2 GEE Goodness-of-fit

To measure the fit using GEE, in this thesis, we perform the goodness-of-fit tests presented in Section 2.2.6, i.e. the Kolmogorov-Smirnov and the Anderson-Darling tests as if the parameters obtained from using GEE were from a GLM which is done in Swan (2006) with quantile residuals. As mentioned in Swan (2006), using diagnostics of GLMs when using GEE is not entirely rigorous but it is the best approach available to test the

distribution assumed for the response variable. There are limitations to this approach due to the fact that observations are not independent but we still do it as no other consensual approach currently exists, specially when using GEE considers a varying dispersion parameter (DGLM).

The Correlation Information Criterion (CIC) introduced in [Hin and Wang \(2009\)](#) and discussed in [Hudecová and Pešta \(2013\)](#) is considered in the analysis but not further pursued as it works better to compare different correlation structures which is not the goal in this thesis. Furthermore, in [Hudecová and Pešta \(2013\)](#) it is mentioned that even if the CIC suggests the independence structure, it does not imply the claim amounts are independent, which under misspecification, results in misleading reserve estimates.

## 2.3 Copula models

In Section 2.2, we removed the marginal effects to have homogeneous data and to be able to account for the diversification benefit. In this section, we describe the multivariate model that has as marginal inputs the uncorrelated residuals  $\mathbf{r}^*$  obtained with the GEE algorithm in Section 2.2.7.1. In our context, assuming  $K$  lines of business with distribution  $F = C(F^{(1)}, F^{(2)}, \dots, F^{(K)})$ , where  $C$  is the copula associated with  $F$ , a distribution function  $C : [0, 1]^K \rightarrow [0, 1]$  that satisfies

$$F(\mathbf{x}) = C(F_1(x_1), \dots, F_L(x_K)), \mathbf{x} \in \mathbb{R}^K. \quad (2.3.1)$$

To understand the definition of a copula we refer to the most important theorem in the field, Sklar's theorem.

**Theorem 2.3.1.** *Sklar's theorem ([Sklar, 1959](#)). Let  $F$  be a joint distribution with margins  $F_{X_1}, \dots, F_{X_d}$ . Then, there exists a copula  $C : [0, 1]^d \rightarrow [0, 1]$  such that, for all  $X_1, \dots, X_d$ ,*

$$F(x_1, \dots, x_d) = C(F_{X_1}(x_1), \dots, F_{X_d}(x_d)). \quad (2.3.2)$$

If the margins are continuous, then  $C$  is unique. A clever interpretation of Sklar's theorem is made in [Derendinger \(2015\)](#), and it can be interpreted as follows,

- We can decompose any multivariate distribution function into its margins and a copula. This allows us to study multivariate distributions independently of the margins,
- With a backward analysis, a copula together with marginal distribution functions can be used to construct a new multivariate distribution.

The following two theorems were obtained from [Nelsen \(2006\)](#) and are presented in their bivariate form but can be generalized for  $d$  dimensions. These theorems describe some properties of copulas which are fundamental in following sections for dependence analysis and copula selection.

**Theorem 2.3.2.** *Fréchet-Hoeffding bounds: For every copula  $C$  and every  $(u, v)$  in the unit square  $[0, 1]^2$ ,*

$$W(u, v) \leq C(u, v) \leq M(u, v),$$

where  $W(u, v) = \max(0, u + v - 1)$  is the counter-monotonic copula and  $M(u, v) = \min(u, v)$  is the comonotonic copula.

*Proof.* Refer to [Nelsen \(2006\)](#). Theorem 2.3.2 is used in the definition of the Plackett copula introduced in Section [B.2](#).

**Theorem 2.3.3.** *Let  $X$  and  $Y$  be continuous random variables, then  $X$  and  $Y$  are independent if and only if,*

$$C(u, v) = uv.$$

*In this case,  $C$  is known as the product copula and denoted by  $\Pi(u, v)$ .*

*Proof.* Let  $u = F_X(x)$  and  $v = F_Y(y)$  be the marginal functions of the independent random variables  $X$  and  $Y$  with joint distribution  $H$ . By the independence assumption,  $H(x, y) = F_X(x)F_Y(y)$ . Then, by Sklar's theorem [2.3.1](#),

$$\begin{aligned} C(u, v) &= H(x, y) \\ &= F_X(x)F_Y(y) \\ &= uv. \end{aligned}$$

The proof in the other direction follows the same way using Sklar's theorem [2.3.1](#). □

Theorem 2.3.3 is summoned when the dependence between lines of business is not statistically significant and thus, a product copula is used. A fundamental tool used in Chapter 4 is the survival copula, denoted as  $\hat{C}(u, v)$  and equal to,

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

The survival copula  $\hat{C}(\bar{F}_X(x), \bar{F}_Y(y))$  is the joint probability  $\bar{H}(x, y) = P[X > x, Y > y]$  which is key to analyze the tail of the distribution, i.e., in the insurance context, the survival copula is the probability of two lines of business with losses that simultaneously exceed a given threshold  $(x, y)$ .

### 2.3.1 Dependence analysis

In Section 2.3, we describe copula models with the purpose of using multivariate distributions for multiple lines of business and ultimately, be able to model the expected losses and calculate the capital requirements described in Chapter 1. In this section, we describe common dependence measurement tools in order to explain the relationship between random variables and assign a proper member of a copula family. Copula families used in this thesis are presented in the following section.

The first dependence measurement is presented in Section 2.2.7, with Pearson's correlation  $\rho$ . However, the major disadvantage of Pearson's correlation is that it only measures linear dependence, when  $X$  and  $Y$  could still be strongly dependent with another type of association. Referring back to Sklar's theorem (2.3.1), the inputs of the selected copula are the marginals and the Pearson's correlation  $\rho$  between the random variables  $X$  and  $Y$  is not the same as for the cumulative distribution functions  $F_X$  and  $F_Y$ . To circumvent this problem, we study dependence with rank correlations. Rank correlations have an advantage over the usual Pearson's correlation because they are invariant under monotonic transformations Joe (1997). Furthermore, Spearman's  $\rho_S$  and Kendall's  $\tau$  described in what follows, are equal to 1 for the upper copula bound and  $-1$  for the lower counter-monotonic copula; a desirable property not held by Pearson's correlation  $\rho$ .

**Definition 2.3.1.** Spearman's  $\rho_S$  is the linear correlation for the cumulative distribution functions of  $X$  and  $Y$ . Let  $F_X(x) = u$  be the cdf of  $X$  and  $F_Y(y) = v$  the cdf of the r.v.  $Y$ , then,

$$\begin{aligned}\rho_S(X, Y) &= \rho(F_X(X), F_Y(Y)) \\ &= -3 + 12 \int_0^1 \int_0^1 uv \, dC(u, v).\end{aligned}$$

**Definition 2.3.2.** Kendall's  $\tau$  (refer to [Joe \(1997\)](#) for more details) is another dependence measure between two random variables. It can be understood as a measure of concordance. For any two points  $(X_1, Y_1)$  and  $(X_2, Y_2)$  from the random variables  $X$  and  $Y$ , Kendall's  $\tau$  is defined as

$$\begin{aligned}\tau(X, Y) &= \mathbb{P}((X_1 - X_2)(Y_1 - Y_2) > 0) - \mathbb{P}((X_1 - X_2)(Y_1 - Y_2) < 0) \\ &= \mathbb{P}(\text{concordance}) - \mathbb{P}(\text{discordance}) \\ &= -1 + 4 \int_0^1 \int_0^1 C(u, v) dC(u, v).\end{aligned}$$

In Figure 2.2, we have a set of three points  $\{P_1, P_2, P_3\}$  from the random vector  $(X, Y)$ . Kendall's  $\tau$  is defined as the probability of concordance minus the probability of discordance. Thus, from the figure we can observe the pair of points  $\{P_1, P_2\}$  and  $\{P_1, P_3\}$  are concordant pairs while  $\{P_2, P_3\}$  are discordant pairs. Thus,  $\tau = \frac{2}{3} - \frac{1}{3} = \frac{1}{3}$ .

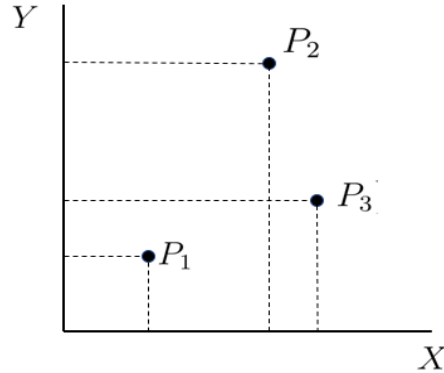


Figure 2.2: Kendall's  $\tau$  in finite samples.

In loss reserving, analyzing the probability of two random variables being jointly large

is fundamentally important. Thus, we introduce what is known as measuring the tail dependence of a random vector.

**Definition 2.3.3.** *Tail dependence coefficient: Let  $X$  and  $Y$  be random variables with a copula  $C(u, v) = C(F_X(x), F_Y(y))$ . The coefficient of upper tail dependence of  $X$  and  $Y$  is given by,*

$$\begin{aligned}\lambda_u &= \lim_{q \rightarrow 1^-} \mathbb{P}[Y > F_Y^{\leftarrow}(q) | X > F_X^{\leftarrow}(q)] \\ &= \lim_{q \rightarrow 1^-} \frac{1 - 2q + C(q, q)}{1 - q}\end{aligned}$$

*If the limit exists. Where  $F_X^{\leftarrow}(q) = \inf\{x \in \mathbb{R} : F_X(x) \geq q\}$ . The coefficient  $\lambda_u \in [0, 1]$ ,  $\lambda_u = 0$  for asymptotically independent random variables in the upper tail and  $\lambda_u = 1$  for  $X$  and  $Y$  comonotonic.*

### 2.3.2 Copula families

The multivariate model to account for the diversification benefit has been chosen to be a copula model and the strength of the dependence is measured with rank correlations due to the advantages presented in Section 2.3.1. In Figure 2.3, we show the dependence considered in this thesis using the same approach as in Shi and Frees (2011) for the observations of accident year  $i$  and development lag  $j$  for two lines of business  $k$  and  $k'$ . The dependence between the random variables is measured using the empirical distribution function (further details in Section 2.3.4 and 2.3.5) of the residuals that have been calculated to remove the marginal effects.



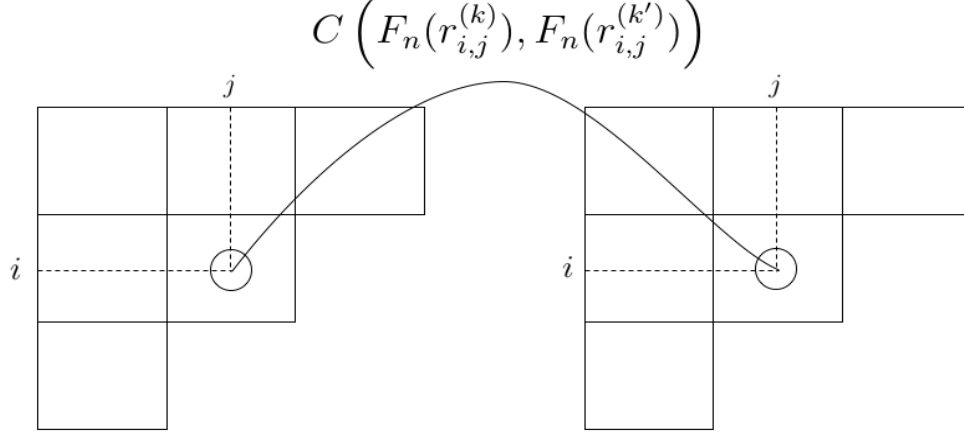


Figure 2.3: Dependence for accident year  $i$  and lag  $j$  between two lines of business.

In order to select the parametric copula  $C$  from Figure 2.3, in this section we describe some copula families considered in the modeling procedure. For other copula families considered in the modeling procedure refer to Appendix B.

### 2.3.2.1 Elliptical Copulas

For an elliptical distribution  $\mathbf{X}$ , an elliptical copula is the copula associated with  $\mathbf{X}$  Embrechts et al. (2001). The definition of an elliptical distribution follows,

**Definition 2.3.4.** *Elliptical distributions.* The vector  $\mathbf{X} = \{X_1, \dots, X_d\}$  of dimension  $d$  is an elliptical distribution if and only if there exist  $R$ ,  $\mathbf{L}$  and  $\mathbf{U}$  such that,

$$\mathbf{X} \stackrel{d}{=} \boldsymbol{\mu} + R \cdot \mathbf{L} \cdot \mathbf{U}, \quad (2.3.3)$$

where  $\boldsymbol{\mu}$  is the mean vector,  $R$  is a positive random variable,  $\mathbf{L}$  is the Cholesky factor of the covariance matrix  $\boldsymbol{\Sigma}$  and  $\mathbf{U}$  is uniformly distributed on the unit sphere of dimension  $d$ .

The density of the elliptical vector  $\mathbf{X}$  is given by,

$$H(X_1, \dots, X_d) = \frac{1}{|\boldsymbol{\Sigma}|^{\frac{1}{2}}} \cdot g((\mathbf{X} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{X} - \boldsymbol{\mu})), \quad (2.3.4)$$

where  $g$  is the density generator of the multivariate spherical distribution associated with  $\mathbf{X}$  (refer to [McNeil et al. \(2005\)](#) for more technical details).

The normal or Gaussian copula is based on the multivariate normal distribution. The density generator is given by  $g(x) = (2\pi)^{-\frac{d}{2}} e^{-\frac{x}{2}}$ . Although the normal or Gaussian copula is a benchmark model, it is not used nor discussed in this thesis because the upper tail coefficient is  $\lambda_u = 0$  independently of the dependence parameter [McNeil et al. \(2005\)](#). Therefore, it underestimates the probability of two random variables being simultaneously large, which makes it unattractive for loss reserving.

**Definition 2.3.5.** *The  $t$ -copula is the copula associated with the elliptical vector in equation (2.3.3), with  $R \sim \sqrt{d \cdot \mathcal{F}_{d,\nu}}$ , where  $\nu$  represent the degrees of freedom,  $\mathcal{F}_{d,\nu}$  is a random number from an  $F$ -distribution and density generator given by,*

$$g(x) = \frac{\Gamma(\frac{d+\nu}{2})}{\Gamma(\frac{\nu}{2}) \cdot (\pi\nu)^{\frac{d}{2}}} \cdot \left(1 + \frac{x}{\nu}\right)^{-\frac{d+\nu}{2}}.$$

The bivariate  $t$ -copula with  $\nu$  degrees of freedom and shape parameter  $\rho$  is given by,

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

where  $\mathbf{t}_{\nu}$  and  $t_{\nu}$  are the multivariate and univariate distribution functions of a student- $t$ , respectively.

The shape parameter  $\rho$  is the standard linear correlation parameter if  $\nu > 2$ . The gamma function  $\Gamma(x)$  appears in the density generator because a  $t$ -distribution is a mixture of normal distributions with weights from an inverse gamma distribution as mentioned in [McNeil et al. \(2005\)](#). Since the inverse gamma distribution is a heavy-tailed distribution, the  $t$ -copula is a better fit for loss models. For example, modeling a catastrophic event for personal and commercial auto insurance  $X$  and  $Y$ , the losses are highly likely placed in the upper corner of the unit square (when analyzing the joint cumulative distribution) due to the number of policyholders affected and the amount of their losses. The underlying event

is the same and therefore, the observations should be highly dependent. A Gaussian copula would underestimate the probability of the catastrophic event due to the asymptotic independence compared to the heavy tailed t-copula which allocates a higher probability to the unfavorable outcome.

The upper tail coefficient for a t-copula of a random vector  $(X, Y)$  (which is the same as the lower tail coefficient because elliptical copulas are radially symmetric [McNeil et al. \(2005\)](#)) is given by,

$$\lambda_u = 2 \cdot t_{\nu+1} \left( -\sqrt{\frac{(\nu+1)(1-\rho)}{1+\rho}} \right),$$

where  $\rho$  is the off-diagonal element of the bivariate correlation matrix associated with  $(X, Y)$  and  $t_{\nu+1}$  is the density of a univariate t-distribution with  $\nu+1$  degrees of freedom. For fixed  $\rho$ , the strength of the tail dependence increases as the degrees of freedom  $\nu$  decreases. And for fixed  $\nu$ , the tail dependence increases as  $\rho$  increases.

### 2.3.3 Hierarchical Copula Model (HCM)

This section is inspired by the research of [Burgi et al. \(2008\)](#), [Arbenz et al. \(2012\)](#) and [Côté \(2014\)](#); [Côté et al. \(2016\)](#) in hierarchical copula modeling where the aggregation of risks plays an important role. The first idea that comes to mind when learning copula models presented in Section 2.3, is to capture the dependence of a whole portfolio with one copula. As mentioned in [Derendinger \(2015\)](#), *“the main advantage of hierarchical risk aggregation is that we do not need to specify the copula of all risks. It is extremely unlikely to find a copula model which adequately describes the dependence structure between a large number of risks. Joint observations between all risks are too rare, and the attainable dependence structures of common parametric copula models are too limited”*. The rigorous mathematical foundation of a HCM can be found in [Arbenz et al. \(2012\)](#) but in what follows, we explain through a graphical representation of a dependence structure in Figure

2.4, a more intuitive explanation of a HCM.

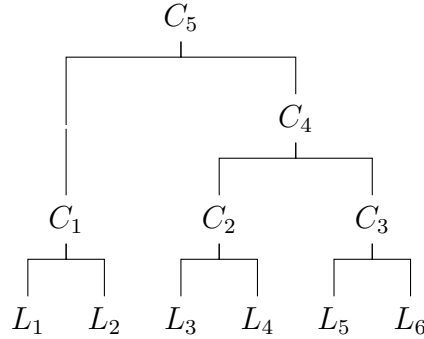


Figure 2.4: Example of a hierarchical copula model for a portfolio with 6 lines of business.

The idea behind a hierarchical copula model (an example can be seen in Figure 2.4) is to use a divide and conquer approach, where the portfolio (in an insurance setting) has to be subdivided either by geographical standards with dependence on legislation as in Burgi et al. (2008), by pooling similar risks as done in Shi and Frees (2011) or by using a dependence-distance criteria as done in Côté et al. (2016). Thus, if the HCM is pursued, risks have to be shown that they can be aggregated with respect to any of the aforementioned criteria. Once the first level of copulas have been fitted ( $C_1$ ,  $C_2$  and  $C_3$  in Figure 2.4), risks are aggregated to then fit the second level copulas ( $C_4$  in Figure 2.4), afterwards, risks are aggregated until last dependence level to attain the model accounting for the total risk in the portfolio. Thus, higher level copulas ( $C_4$  and  $C_5$  in example of Figure 2.4) are meant to represent the dependence structure in the sum of risks when this aggregations is a reasonable choice. The objective of aggregating different risks is to reduce the overall risk, providing a diversification benefit as established by IFRS 17, see characteristics in Section 1.1.2. The idea for the aggregation of risks for a copula model first came from Burgi et al. (2008) where they set an example with a P&C reinsurer that sells policies against fire and windstorm events.

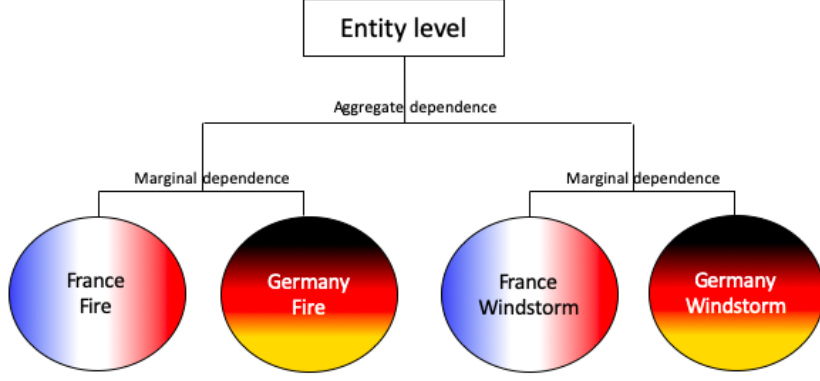


Figure 2.5: Hierarchical copula model used in [Burgi et al. \(2008\)](#).

Figure 2.5 (recreated from Figure 5 in [Burgi et al. \(2008\)](#)), shows the hierarchical dependence structure considered by the authors. It may seem strange to model the dependence between a fire in France and a fire in Germany with a copula due to the fact that these unfavorable events occur kilometers from each other and are physically independent of each other but the interpretation is not as straightforward. As said by the original author, significant dependence between products do not arise through the underlying peril itself but are caused by changes in legislation and insurance practices. In this example, since both are countries of the European Union, it may imply dependence between risks. A specific case of a hierarchical copula model is called a nested Archimedean copula (NAC) model, where the copulas at each node of the hierarchy are part of the Archimedean family presented in Appendix B.1.

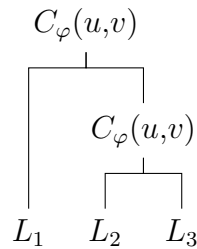


Figure 2.6: Example of nested Archimedean copula.

Figure 2.6 shows three lines of business where the dependence is accounted for with a NAC model with bivariate Archimedean copulas  $C_\varphi$ , where  $\varphi$  is the copula generator.

As mentioned in [Abdallah et al. \(2015\)](#), one of the main advantages of a NAC model compared to any other hierarchical copula model is that they can be explicitly defined in terms of the generator  $\varphi$  mentioned in equation (B.1.1). Furthermore, sampling from a NAC model is a more simple task than sampling from a hierarchical copula model with copulas from other families because they rely on the Iman-Conover reordering algorithm [Côté et al. \(2016\)](#); presented in Section 2.4.1. Described in [Hofert and Pham \(2013\)](#) and [Abdallah et al. \(2015\)](#), nested Archimedean copulas have to satisfy the following:

- The degree of dependence must decrease at each ascending level,
- The same copula has to be used for each level due to the convexity on nested Archimedean copulas.

Thus, NAC models are not considered in this thesis because they are more restrictive than considering a hierarchical copula model with other copula families at each node. In the examples presented in Figures 2.4, 2.5 and 2.6, the hierarchical copula models considered have bivariate copulas at each dependence level. This is not a necessary condition for a hierarchical copula model given that at each level,  $d$ -dimensional copulas can be fitted. In this thesis, we focus on bivariate copulas due to the simplicity in the interpretability of the results and moreover, we share the idea presented in [Derendinger \(2015\)](#) that one copula is highly unlikely to reflect the complete dependence structure of more than two risks. In a nutshell, the proposed hierarchical copula model has the flexibility of fitting any bivariate copula family at each node, allowing for more complex dependence structures than what can be achieved with a NAC model. Additionally, all formal goodness-of-fit tests are rank-based (refer to 2.3.5) and the bivariate approach allows for graphical and more interpretable results. An important limitation mentioned in [Côté et al. \(2016\)](#), due to the aggregation step at each node, the copulas need to have the same number of components, implying the procedure done in [Abdallah et al. \(2015\)](#) to incorporate calendar year dependence with a copula cannot be implemented with this hierarchical copula model. To circumvent this issue, a calendar year covariate could be easily added into the marginal distribution with the disadvantage of increasing the number of estimated

parameters more than the copula approach. The calendar year effect is not included in this thesis because it is deemed not significant for the numerical application presented in Chapter 5.

### 2.3.4 Estimation

The estimation procedure for the HCM is done with a frequentist approach as for the marginal distributions in Section 2.2. To obtain the set of dependence parameters of the HCM we use maximum pseudo-likelihood (MPL). MPL is first used in Oakes (1994) and further analyzed in Genest et al. (1995) where they show the estimator is consistent and asymptotically normal. This estimation approach consists in obtaining the dependence parameter  $\theta$  of a parametric copula density denoted by  $c_\theta$  for a random vector  $(X, Y)$  by maximizing the following log pseudo-likelihood,

$$\mathcal{L}(\theta) = \sum_{i=1}^n \log (c_\theta(\hat{u}, \hat{v})), \quad (2.3.5)$$

where  $\hat{u}$  and  $\hat{v}$  are the empirical cumulative distribution functions applied to the observations of  $X$  and  $Y$ , respectively. This method differs from the traditional maximum likelihood estimate by not including the marginal distributions in the function to be maximized and instead use a nonparametric estimate of the marginal cumulative distribution functions which does not depend on the parameter  $\theta$ . The maximum likelihood estimate would result in an estimate with smaller variance than with MPL only if the margins were known exactly. The HCM has a conditional independence assumption established in the rigorous mathematical foundation of a HCM in Arbenz et al. (2012), meaning that given a sum at each node, then, the elements that made the sum are independent of the rest of the hierarchy. Thus, equation (2.3.5) is used for each node of the HCM to obtain the set of all dependence parameters. In other words, due to the conditional independence assumption, the estimation procedure is independent for each copula composing the HCM.

### 2.3.5 Copula selection and Goodness-of-fit

The most common model selection criteria is Akaike's Information Criteria (AIC). It is meant to reflect a parsimony and accuracy trade-off,

$$AIC = 2(q - \log \mathcal{L}), \quad (2.3.6)$$

where  $q$  is the number of parameters in the model and  $\mathcal{L}$  is the value of the maximum likelihood function with the estimated parameters. The best model is then selected with respect to the smallest AIC, since we are trying to use the least number of parameters, thus, minimizing the term  $q$  and with the highest accuracy, i.e., the largest value of  $\log \mathcal{L}$ . In [Grønneberg and Hjort \(2014\)](#), it is pointed out that using the AIC when estimating the copula through MPL is not appropriate since we are using a pseudo-likelihood function instead of the likelihood function. Thus, the copula selection for each node of the semiparametric HCM is more heuristic than analytical due to the current lack of formal criterions to evaluate the best copula when considering MPL. From the copula families presented in [Section 2.3.2](#), taking advantage of the bivariate approach, graphical comparisons are drawn between the empirical copula  $C_n$  and the parametric copula with the MPL estimate denoted  $C_{\theta_n}$ . The empirical copula defined in [Deheuvels \(1979\)](#) for a random vector  $(X, Y)$  with marginal cumulative distribution functions  $u = F_X(x)$  and  $v = F_Y(y)$  is given by

$$C_n(u, v) = \frac{1}{n} \sum_{i=1}^n \mathbb{1} \left( \frac{R_i}{n+1} \leq u, \frac{S_i}{n+1} \leq v \right), \quad (2.3.7)$$

where  $R_i$  and  $S_i$  are the ranks of  $x_i$  and  $y_i$ , respectively. For the goodness-of-fit process for the HCM we rely on more formal mathematical tools. The goal is to statistically test the null hypothesis  $H_0 : C \in C_\theta$ , i.e., the copula  $C$  is indeed part of the parametric family  $C_\theta$ . An underlying copula  $C$  of a random vector is invariant by continuous, strictly increasing transformations of its components, thus, [Genest et al. \(2009\)](#) proposed to base the inference on the maximally invariant statistics<sup>4</sup> with respect to this set of transforma-

---

<sup>4</sup>As established in [Young and Smith \(2005\)](#), a statistic is maximally invariant if every other invariant statistic is a function of it.



tions, i.e., the ranks. Motivated by selecting the appropriate copula family, [Genest and Rémillard \(2008\)](#) introduced the Cramér-Von Mises statistic for copula models which can be calculated as follows,

$$S_n = n \int_{[0,1]^d} (C_n(\mathbf{u}) - C_{\theta_n}(\mathbf{u}))^2 dC_n(\mathbf{u}), \quad (2.3.8)$$

where  $C_n$  is the empirical copula and  $C_{\theta_n}$  is the parametric copula with a rank-based estimate of  $\theta$ . The Cramér-von Mises statistic  $S_n$  is used to assess if the sample can be assigned to a member of a certain parametric copula family (some families presented in [Section 2.3.2](#)). The validity of the  $p$ -value relies on performing the Cramér-Von Mises test by parametric bootstrapping, see [Genest and Rémillard \(2008\)](#). Furthermore, our approach is rank-based due to the recommendation of [Genest et al. \(2009\)](#) which states: *“Based on the present state of knowledge: Overall, the statistic  $S_n$  [...] yield the best blanket<sup>5</sup> goodness-of-fit test procedure for copula models.”*

When the goodness-of-fit is meant to be assessed between two copulas without the need of assuming a parametric copula, the following test is appropriate. The null hypothesis is  $H_0 : C = D$  for two copulas  $C$  and  $D$ , which in other words is a test to try to determine if two copulas are identical in distribution. The test statistic [\(2.3.9\)](#) for the equality between two copulas relies on the empirical process [Rémillard and Scaillet \(2009\)](#), given by,

$$\mathbb{E}_{n_1, n_2} = \frac{C_{n_1} - D_{n_2}}{\sqrt{n_1^{-1} + n_2^{-1}}}, \quad (2.3.9)$$

where  $n_1$  and  $n_2$  are the sample size of the respective copulas  $C$  and  $D$ . Similarly as in [equation \(2.3.8\)](#), the approach considered in [Rémillard and Scaillet \(2009\)](#) depends on the Cramér-von Mises principle, providing the test statistic:

$$S_{n_1, n_2} = \int_{[0,1]^d} \mathbb{E}_{n_1, n_2}^2(\mathbf{u}) d\mathbf{u}. \quad (2.3.10)$$

---

<sup>5</sup>By “blanket” the author means the statistical test for  $S_n$  does not involve any parameter tuning or strategic choices.

## 2.4 Simulation

The purpose of the HCM as stated in Chapter 1 is to estimate the unpaid claim liabilities while accounting for the diversification benefit for financial reporting and reserving objectives. Throughout Chapter 2, the mathematical justification, estimation procedure and goodness-of-fit techniques for the marginal and multivariate model are described. This section describes the algorithm used to simulate from the bottom-right part of the run-off triangle corresponding to the unpaid claim liabilities. Due to the aggregation involved at each node of the hierarchical copula modeling, simulating from a HCM is not a trivial task. The Iman-Conover reordering algorithm proposed by [Iman and Conover \(1982\)](#) and adapted by [Arbenz et al. \(2012\)](#) is the solution to the problem. An additional step is considered in this thesis to account for the marginal distributions where the aggregate claims are of the form  $Y_{i,j}|\mathbf{X}_{i,j} \sim TW_p(\mu_{i,j}, \phi_{i,j})$  presented in Section 2.2.5 and the distributional assumption of the residuals from Theorem 2.2.2.

### 2.4.1 Iman-Conover reordering algorithm

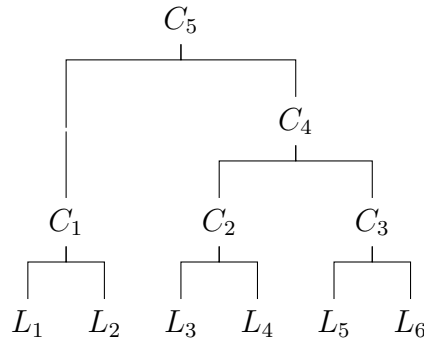


Figure 2.7: Example of a HCM with six lines of business, three levels and five bivariate copulas.

In Figure 2.7, we have the visual representation of a given dependence structure in a hierarchical copula model. We assume the aggregate claims of the six lines of business

each follow Tweedie marginal distributions of the form  $Y_{i,j}^{(k)} | \mathbf{X}_{i,j}^{(k)} \sim TW_{p^{(k)}}(\mu_{i,j}^{(k)}, \phi_{i,j}^{(k)})$  for accident year  $i$ , development lag  $j$  and lines  $k = \{1, 2, 3, 4, 5, 6\}$ . Furthermore, we assume for each line of business the variance-covariance structure  $\mathbf{V}^{(k)} = A^{(k)1/2} R(\rho^{(k)}) A^{(k)1/2}$  as in Section 2.2.7, where  $\mathbf{V}^{(k)}$  is a block diagonal matrix that assumes correlation between development lags but independence across accident years. The vector of uncorrelated residuals for all accident years and development lags without the dependence structure are denoted  $\mathbf{r}_{\text{ind}}^{*(k)}$  and conversely,  $\mathbf{r}^{*(k)}$  when the dependence structure is obtained for each line of business  $k$ . Correspondingly, the correlated residuals are denoted  $\mathbf{r}^{(k)}$  to preserve the notation from Section 2.2.7.1. In this example, we use the Iman-Conover reordering algorithm to generate i.i.d. samples of size  $n = \frac{J(J-1)}{2}$  from the lower part of the loss triangle with indices  $i + j > I + 1$ , assuming the HCM observed in Figure 2.7. The residuals  $\mathbf{r}_{\text{ind}}^{*(k)}, \mathbf{r}^{*(k)}$  and  $\mathbf{r}^{(k)}$  are to be associated with the corresponding line of business  $L_k$  from Figure 2.7 but should not be confused mathematically speaking. In other words, Figure 2.7 is a graphical representation of the HCM assumed for the example, which is meant to ease the understanding of the dependence structure, but the  $L_k$  have no place in the algorithm. The algorithm goes as follows,

1. Simulate  $k$  independent standard normal random samples of size  $m \gg n$ <sup>6</sup>.

$$\mathbf{r}_{\text{ind}}^{*(k)} \sim N(0, 1), \quad k = \{1, 2, 3, 4, 5, 6\}.$$

2. Simulate independent copula samples of size  $m$  from each bivariate copula  $C_1, \dots, C_5$ .
3. Reorder the samples of each bivariate vector by merging the observed marginal ranks with the joint ranks in the copula sample. A brief example follows for the first node of the HCM.

---

<sup>6</sup>In Côté (2014) it is pointed out that the empirical distribution functions of the marginals and the copula converge asymptotically to the true distributions. Thus, a larger sample size  $m$  provides a better estimate of the HCM sample.

$\mathbf{r}_{\text{ind}}^{*(1)}$	Rank	$\mathbf{r}_{\text{ind}}^{*(2)}$	Rank	$C_1$	Ranks	Reordered Sample
1.27	2	3.71	3	(0.7, 0.4)	(3, 2)	(2.80, 0.40)
-0.10	1	-2.19	1	(0.2, 0.9)	(1, 3)	(-0.10, 3.71)
2.80	3	0.40	2	(0.5, 0.3)	(2, 1)	(1.27, -2.19)

Table 2.3: Iman-Conover reordering algorithm example for the first node of dependence structure (HCM) from Figure 2.7. Inspired by examples in [Arbenz et al. \(2012\)](#).

Then, the reordered data is a sample from the copula  $(\mathbf{r}^{*(1)}, \mathbf{r}^{*(2)}) \sim C_1$ .

4. Repeat step 3 for the first level copulas  $C_2$  and  $C_3$ .
5. Aggregate the reordered data following the dependence structure to obtain samples from  $\mathbf{r}^{*(1)} + \mathbf{r}^{*(2)}$  and respectively for  $\mathbf{r}^{*(3)} + \mathbf{r}^{*(4)}$  and  $\mathbf{r}^{*(5)} + \mathbf{r}^{*(6)}$ .
6. Repeat step 3 to obtain sample from  $(\mathbf{r}^{*(3)} + \mathbf{r}^{*(4)}, \mathbf{r}^{*(5)} + \mathbf{r}^{*(6)}) \sim C_4$ .
7. Aggregate the reordered sample from  $C_4$  to obtain a sample from  $\sum_{k=3}^6 \mathbf{r}^{*(k)}$ , and repeat step 3 for  $\left(\mathbf{r}^{*(1)} + \mathbf{r}^{*(2)}, \sum_{k=3}^6 \mathbf{r}^{*(k)}\right) \sim C_5$ .
8. To obtain a joint sample of  $(\mathbf{r}^{*(1)}, \mathbf{r}^{*(2)}, \mathbf{r}^{*(3)}, \mathbf{r}^{*(4)}, \mathbf{r}^{*(5)}, \mathbf{r}^{*(6)})$ , perform the permutations applied to  $\mathbf{r}^{*(1)} + \mathbf{r}^{*(2)}$  back to  $\mathbf{r}^{*(1)}$  and  $\mathbf{r}^{*(2)}$ , the permutations applied to  $\mathbf{r}^{*(3)} + \mathbf{r}^{*(4)}$  back to  $\mathbf{r}^{*(3)}$  and  $\mathbf{r}^{*(4)}$ , and finally, the permutations applied to  $\mathbf{r}^{*(5)} + \mathbf{r}^{*(6)}$  back to  $\mathbf{r}^{*(5)}$  and  $\mathbf{r}^{*(6)}$ .
9. Get a subsample of size  $n$  from the reordered sample of size  $m$ .

Formally, the Iman-Conover reordering algorithm is finished after step 9, but given the model presented in Chapter 3 we want to induce the covariance structure  $\mathbf{V}_i^{(k)}$  for each

line of business  $k$ . Thus, apply the Cholesky block matrix factor  $L^{(k)}$  given by

$$L^{(k)} = \begin{bmatrix} L_1(\hat{\rho}^{(k)}) & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & L_2(\hat{\rho}^{(k)}) & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & L_{n_i}(\hat{\rho}^{(k)}) \end{bmatrix}_{(n \times n)},$$

to the lower part of the loss triangle (inverse of matrix in equation (2.2.49)), as follows,

$$L^{(k)} \mathbf{r}^{*(k)} = \mathbf{r}^{(k)}. \quad (2.4.1)$$

Thus, we have obtained the vector  $\mathbf{r}^{(k)}$  with variance  $\mathbf{V}^{(k)}$  and the dependence structure from the HCM represented in Figure 2.7 for  $k = \{1, 2, 3, 4, 5, 6\}$ . The additional step described in what follows is useful when calculating reserves (for more details refer to Chapter 4). The objective by invoking Theorem 2.2.2 is having all observations from the same distribution, without mean and dispersion differences between each cell  $(i, j)$  of the loss triangle in Table 2.1. This approach is useful for obtaining the rank estimate of the dependence parameter  $\theta_n$  discussed in the copula Section 2.3. Unfortunately, applying the inverse transformation, i.e, multiplying by the standard deviation and adding the mean to obtain Tweedie samples is not a good idea since in this case, the objective distribution is a non-negative random variable while a standard normal variable  $Z$  has support in  $\mathbb{R}$ . Thus, the inverse transformation has non-zero probability of being a negative number, or,

$$\mathbb{P} \left[ Z \sqrt{\phi_{i,j} \mu_{i,j}^p} + \mu_{i,j} < 0 \right] \neq 0.$$

To avoid approximation errors using the aforementioned technique, instead, we calculate the empirical distribution for each marginal distribution and apply the appropriate inverse quantiles. The empirical distribution is calculated as,

$$\hat{F}_n^{(k)}(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{1} \left( r_i^{(k)} \leq x \right), \quad k = \{1, 2, 3, 4, 5, 6\},$$

where  $r_i^{(k)}$  is the  $i$ -th entry of the vector  $\mathbf{r}^{(k)}$  obtained with equation (2.4.1) with sample size  $n$  obtained for the marginal distribution  $k$  after applying the Iman-Conover reordering

algorithm [2.4.1](#) and correlating the residuals. Then, apply the inverse quantile transformation to obtain samples from the desired Tweedie distribution,

$$F^{-1}(\hat{F}_n^{(k)}; \mu_{i,j}^{(k)}, \phi_{i,j}^{(k)}, p^{(k)}) \sim TW_{p^{(k)}}(\mu_{i,j}^{(k)}, \phi_{i,j}^{(k)}).$$

# Chapter 3

## Model

A HCM is considered with semi-parametric estimation of the dependence parameter and a two-step estimation for the marginals. The two-step procedure consists in (1) a DGLM with Tweedie distributions is assumed for the loss ratio of the aggregate claims  $Y_{i,j}$  and in the presence of correlation, (2) rely on the GEE with an autoregressive factor of lag 1 or AR(1) to obtain uncorrelated residuals. The aforementioned model has the objective of generating unpaid claim liabilities for financial reporting and the reserving purposes mentioned in Chapter 1.

For all lines of business and loss ratio denoted  $Y_{i,j}^{(k)} | \mathbf{X}_{i,j} \sim TW_{p^{(k)}}(\mu_{i,j}^{(k)}, \phi_j^{(k)})$ , a DGLM is considered with the following mean model from equation (2.2.6),

$$g(\mu_{i,j}^{(k)}) = \iota^{(k)} + \alpha_i^{(k)} + \delta_j^{(k)}, \quad (3.0.1)$$

where for line of business  $k = \{1, 2, \dots, K\}$ , the parameter  $\iota^{(k)}$  is the intercept,  $\alpha_i^{(k)}$  accounts for the effects of the accident year  $i = \{1, 2, \dots, I\}$ ,  $\delta_j^{(k)}$  considers the effects of the development lag  $j = \{1, 2, \dots, J\}$  and to avoid non-identifiability issues we set  $\alpha_1^{(k)} = \delta_1^{(k)} = 0$ . Furthermore, we assume  $I = J$  and a log-link function  $g$  as mentioned in Section 2.2.3 to have an interpretable and multiplicative structure. In matrix notation, the mean model has the following equation,

$$g(\mu_{i,j}^{(k)}) = \mathbf{X}_{i,j}^T \boldsymbol{\beta}^{(k)}, \quad (3.0.2)$$

where  $\mathbf{X}_{i,j}$  is the vector of covariates for accident year  $i$  and development lag  $j$ . The covariates in this case are dummy variables, which for example in equation (3.0.2), the parameter of accident year  $i$  has covariates set to 1 for the loss ratios in accident year  $i$  and 0 otherwise. Correspondingly, for the parameter of development lag  $j$ , the covariates are set to 1 for loss ratios in column  $j$  of the loss triangle and 0 otherwise. Thus, the matrix of covariates  $X$  and the parameter vector  $\boldsymbol{\beta}^{(k)}$  have the following general form,

$$X\boldsymbol{\beta}^{(k)} = \begin{bmatrix} & & & I & & & & J \\ 1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & \dots & 0 & 0 & 0 & \dots & 1 \\ 1 & 1 & \dots & 0 & 0 & 0 & \dots & 0 \\ 1 & 1 & \dots & 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & \dots & 1 & 0 & 0 & \dots & 0 \end{bmatrix}_{(n \times q)} \begin{bmatrix} \iota^{(k)} \\ \alpha_2^{(k)} \\ \vdots \\ \alpha_I^{(k)} \\ \delta_2^{(k)} \\ \vdots \\ \delta_J^{(k)} \end{bmatrix}_{(q \times 1)} = \begin{bmatrix} \iota^{(k)} \\ \iota^{(k)} + \delta_2^{(k)} \\ \vdots \\ \iota^{(k)} + \delta_J^{(k)} \\ \iota^{(k)} + \alpha_2^{(k)} \\ \iota^{(k)} + \alpha_2^{(k)} + \delta_2^{(k)} \\ \vdots \\ \iota^{(k)} + \alpha_I^{(k)} \end{bmatrix}_{(n \times 1)},$$

where  $n = \frac{I(I+1)}{2}$  and  $q = 1 + (I - 1) + (J - 1) = 2I - 1$ . In Figure 3.1 we place the equations into a loss triangle to have a better understanding of the GLM structure.

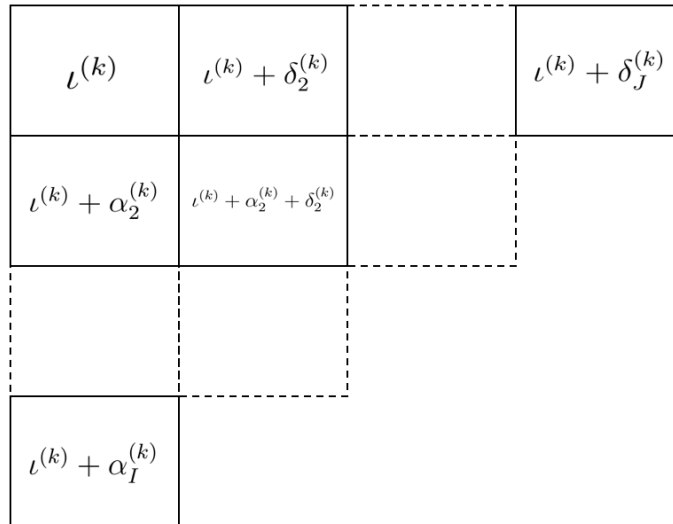


Figure 3.1: Loss triangle with GLM equations of mean model.



Simultaneously, for the dispersion parameter  $\phi$ , we consider another log-link relation to obtain the following dispersion submodel (presented in Section 2.2.5),

$$g(\phi_j^{(k)}) = \mathbf{Z}_j^T \boldsymbol{\gamma}^{(k)}. \quad (3.0.3)$$

The matrix of covariates  $Z$  and the link function  $g$  for convenience and simplicity, are usually assumed to be the same as for the mean model. But practically, it can be whatever the user considers statistically relevant. In this thesis, to avoid overfitting and due to statistical evidence, the dispersion submodel only varies with the development lags and not with the accident years as for the mean model, i.e,  $\phi_{i,j} = \phi_j \ \forall i$ . From Figure 3.2 we can infer that although every observation is assumed from a Tweedie distribution with its own mean, we expect the dispersion (closely related to the number of claims)  $j$  years after the accident to be equal across accident years. Thus, for a development lag we assume the dispersion has common effects across accident years. In matrix notation, we have the following equations,

$$Z\boldsymbol{\gamma}^{(k)} = \begin{bmatrix} 1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & \dots & 1 & 0 & 0 & \dots & 0 \end{bmatrix}_{(n \times J)} \begin{bmatrix} \iota_d^{(k)} \\ \gamma_2^{(k)} \\ \vdots \\ \gamma_J^{(k)} \end{bmatrix}_{(J \times 1)} = \begin{bmatrix} \iota_d^{(k)} \\ \iota_d^{(k)} + \gamma_2^{(k)} \\ \vdots \\ \iota_d^{(k)} + \gamma_J^{(k)} \end{bmatrix}_{(n \times 1)},$$

where  $\iota_d^{(k)}$  is the intercept for the dispersion submodel and we also set  $\gamma_1^{(k)} = 0$  to avoid non-identifiability of the parameter vector  $\boldsymbol{\gamma}$ . The DGLM can be visualized in Figure 3.2 for accident year  $i$  and  $i'$ , and development lag  $j$  and  $j'$ .

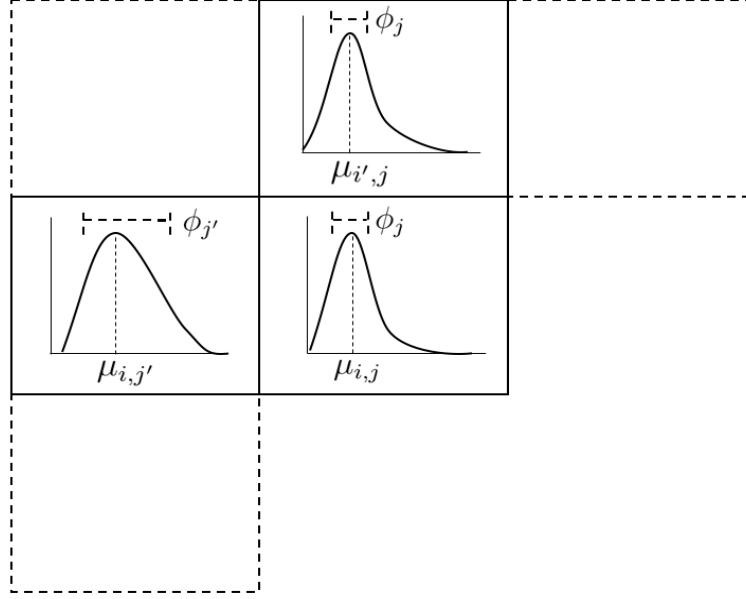


Figure 3.2: Distributional visualization of the DGLM considered for the run-off triangle.

Then, the set of parameters  $\{\beta^{(k)}, \gamma^{(k)}\}$  are obtained with the DGLM estimation algorithm presented in Section 2.2.5.2 for each  $k$ . When the goodness-of-fit tests presented in Section 2.2.6 are satisfactory, we proceed to the second step of the marginal fitting. To check for correlated residuals we use the Ljung-Box test (2.2.46) in the residuals  $\hat{r}_{i,j}^{(k)}$  computed with equation (2.2.44). In the presence of dependent residuals we estimate the correlation parameter  $\rho^{(k)}$  for each line of business with the GEE algorithm presented in Section 2.2.7.1. Upon convergence and validation with the techniques presented in Section 2.2.7.2, calculate the uncorrelated residuals with equation (2.2.49) to proceed to the fitting of the multivariate distribution.

For the  $K$  lines of business denoted  $L_k, k = \{1, 2, \dots, K\}$ , a HCM presented in Section 2.3.3 is considered. This thesis deals specifically with automobile insurance and thus, we assume the insurer does business in  $K/2$  (we assume  $K$  even) geographical regions where for each region the insurer has a personal auto line and a commercial auto line. Thus, the first level copulas are split by region to consider the dependence between the personal and commercial line as done in Shi and Frees (2011). The first level copulas have the following structure,

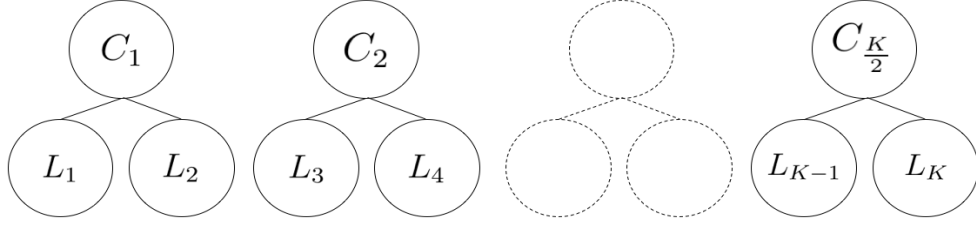


Figure 3.3: First level copula structure in the HCM.

The set of parametric copulas  $\{C_1, C_2, \dots, C_{\frac{K}{2}}\}$  from Figure 3.3 are estimated with the MPL approach from Section 2.3.4, from the selected parametric families  $C_\theta$  presented in Section 2.3.2. Furthermore, to test the null hypothesis  $H_0 : C_k \in C_\theta$ , we perform the Cramér-von Mises test described in Section 2.3.5. Once the first level copulas have been statistically validated, we move up the dependence structure by computing the  $K/2$  aggregates of residuals for each region denoted  $\{R_i\}_{i=1}^{K/2}$  to analyze the second level of copulas. The fitting procedure now considers the regions with stronger dependence as done in Côté et al. (2016). The dependence is measured with Kendall's  $\tau$  from Section 2.3.1 to all possible region pairings, the result is a matrix of the following form,

$$\begin{bmatrix} 1 & \tau(R_1, R_2) & \tau(R_1, R_3) & \dots & \tau(R_1, R_{\frac{K}{2}}) \\ \tau(R_1, R_2) & 1 & \tau(R_2, R_3) & \dots & \tau(R_2, R_{\frac{K}{2}}) \\ \tau(R_1, R_3) & \tau(R_3, R_2) & 1 & \dots & \tau(R_3, R_{\frac{K}{2}}) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \tau(R_{\frac{K}{2}}, R_1) & \tau(R_{\frac{K}{2}}, R_3) & \tau(R_{\frac{K}{2}}, R_3) & \dots & 1 \end{bmatrix}_{\left(\frac{K}{2} \times \frac{K}{2}\right)} \quad (3.0.4)$$

The pair of regions with the highest off-diagonal absolute value for Kendall's  $\tau$  in (3.0.4) are selected for the second level of the HCM which has exactly  $\lfloor K/4 \rfloor^1$  copulas to maintain the bivariate approach.

---

<sup>1</sup>The notation  $\lfloor x \rfloor$  represents the floor function of  $x$ .

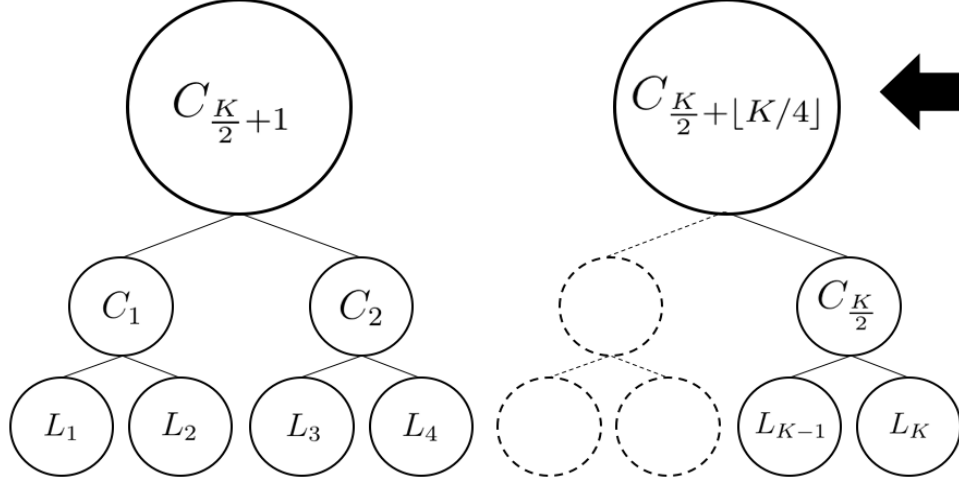


Figure 3.4: Second level copula structure in the HCM.

The dependence parameter rank-based estimation, copula selection and goodness-of-fit statistical tests for the set of second level copulas  $\{C_{\frac{K}{2}+1}, \dots, C_{\frac{K}{2}+\lfloor K/4 \rfloor}\}$  is the same as for the first level copulas. This process is repeated until fitting the last bivariate copula  $C_{K-1}$  at the top of the dependence structure as observed in Figure 3.5.

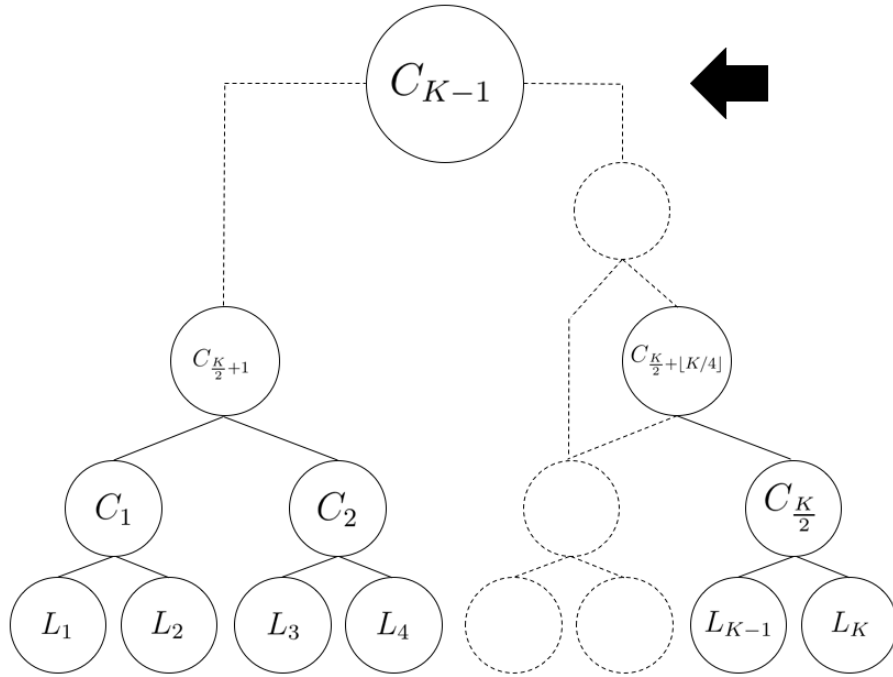


Figure 3.5: Complete dependence structure of the HCM.

Once the bivariate copula  $C_{K-1}$  has been statistically validated, the HCM fitting procedure is finished and we are able to proceed to the simulation of the unpaid claim liabilities. To generate simulations we rely on the Iman-Conover reordering algorithm presented in Section [2.4.1](#).

## Chapter 4

# Risk Assessment and Capital Requirements

In this chapter, we explain the need for reserving through the model presented in Chapter 3. Furthermore, to calculate the risk adjustment for non-financial risk defined in Chapter 1, we present techniques involving risk measures and the cost of capital method. There is special emphasis in the risk measures section given that not only they are useful for the risk adjustment required by IFRS 17, but insurance companies' regulators commonly define them as the entity's level of solvency.

Insurance companies differ from other non-financial enterprises given that capital from shareholders is not directly linked with produced goods or a provided service, but with the ability to maximize returns by maintaining financial solvency while paying contingent claims produced during the lifetime of insurance contracts. This amount of capital needed to maintain financial solvency is called a reserve.

## 4.1 Reserving

The goal for actuaries is to forecast future claims  $\hat{Y}_{i,j}$ , where  $i+j > J+1$ . In other words, to use the upper triangle of Table 2.1 to predict the lower triangle corresponding to the unpaid claim liabilities. A procedure also known as squaring the triangle. The sum of the lower triangle  $\sum_{i=2}^I \sum_{j=J+2-i}^J \hat{Y}_{i,j}$  is known liabilities for incurred claims reserve.

### 4.1.1 Risk measures

In Chapter 3, the objective of the stochastic model is to generate simulations of the unpaid claim liabilities while considering the dependence in the insurance portfolio. The simulation of the unpaid claim liabilities form themselves a loss distribution for the insurer. In order to evaluate the risk inherent of the insurance portfolio, in this section, we describe risk measures which are a measurement with respect to a certain criteria of the loss distribution. The risk adjustment for non-financial risk as established in Chapter 1 can be determined using a risk measure.

**Definition 4.1.1.** *Artzner et al. (1999) coined as coherent, the risk measures that satisfy the following properties.*

We consider a measure  $\varrho : \mathcal{G} \rightarrow \mathbb{R}$ , where  $\mathcal{G} \in \Omega$  is the set of all risks.

1. *Translation invariance.* For all  $X \in \mathcal{G}$  and all real numbers  $\alpha$ ,

$$\varrho(X + \alpha) = \varrho(X) + \alpha.$$

2. *Subadditivity.* For all  $X$  and  $Y \in \mathcal{G}$ ,

$$\varrho(X + Y) \leq \varrho(X) + \varrho(Y).$$

3. *Positive homogeneity.* For all  $\lambda > 0$  and all  $X \in \mathcal{G}$ ,

$$\varrho(\lambda X) = \lambda \varrho(X).$$

4. *Monotonicity.* For all  $X$  and  $Y \in \mathcal{G}$  with  $X \leq Y$ ,

$$\varrho(X) \leq \varrho(Y).$$

One important characteristic of a coherent measure of risk is that the aggregate risk measure is less than or equal to the sum of the measures for the individual risks being aggregated.

#### 4.1.1.1 Value-at-Risk (VaR)

The VaR or confidence level technique, represents the minimum loss incurred within a given time horizon if one of the  $(1 - \alpha)\%$  worst-case scenarios occurs. Consider the loss random variable  $X$ , then,

$$\text{VaR}_\alpha(X) = \inf\{x \in \mathbb{R} \mid \mathbb{P}[X \leq x] \geq \alpha\}. \quad (4.1.1)$$

This risk measure has a blind spot and is not sufficient to understand the worst possible scenarios. The worst scenarios are located on the right tail of the loss distribution, and it is what the insurer should be most concerned about. The VaR fails to satisfy the sub-additivity property and thus, is not a coherent risk measure.

#### 4.1.1.2 Tail Value-at-Risk (TVaR) or Expected shortfall (ES)

The Tail Value-at-Risk (TVaR)<sup>1</sup> is defined as the expected loss of the worst  $(1 - \alpha)\%$  cases and thus, its equation is related to the VaR (4.1.1), as it can be seen in the following,

$$\text{TVaR}_\alpha(X) = \frac{1}{1 - \alpha} \int_\alpha^1 \text{VaR}_u(X) du. \quad (4.1.2)$$

The TVaR is meant to correct for the blind spot in the VaR by providing more insight of the behavior in the right tail of the loss distribution, thus, it reflects better the possibility

---

<sup>1</sup>For a continuous random variable the TVaR and the Conditional Tail Expectation (CTE) (mentioned in Chapter 1 by OSFI capital requirements) are equal.



of extreme losses which have small probabilities of occurring because it takes into account outcomes beyond any chosen threshold. Furthermore, the TVaR complies with the four properties established by Artzner et al. (1999), making it a coherent risk measure.

#### 4.1.1.3 Capital allocation

The capital allocation problem as explained by McNeil et al. (2005) consists on having several lines of business where we calculate the overall risk capital and then, we wish to assign a portion of the overall risk capital into each individual line of business. In mathematical notation, we have an insurance portfolio with  $d$  lines of business represented by a random vector  $\{X_1, X_2, \dots, X_d\}$ . The overall risk capital has the form  $\varrho(X)$ , where  $X = \sum_{i=1}^d X_i$  and  $\varrho$  is a risk measure like the VaR described in (4.1.1) or the TVaR described in (4.1.2), and then, we wish to allocate properly a portion or weight of  $\varrho(X)$  into each line of business. Thus, for simplicity we denote the weights  $\mathbf{u} = \{u_1, \dots, u_d\}$  and we define the mapping  $f : \Lambda \subset \mathbb{R}^d \rightarrow \mathbb{R}$ , with the following form:

$$f_{\varrho, X}(\mathbf{u}) = \varrho \left( \sum_{i=1}^d u_i X_i \right). \quad (4.1.3)$$

**Definition 4.1.2.** (Tasche, 2008) *Let  $\varrho$  be a risk measure that satisfies the positive homogeneity property and  $f_{\varrho}$  be continuously differentiable on  $\Lambda$ , then, the Euler capital allocation principle is the mapping,*

$$\varrho_{Euler}(X_i|X) = \frac{\partial f_{\varrho}}{\partial u_i}(1, \dots, 1). \quad (4.1.4)$$

Theorems 4.1.1 and 4.1.2 stated in what follows are consequences of Definition 4.1.2 and are the most common capital allocation approaches for the VaR and TVaR because these methods have two desirable economic properties. These properties are the full allocation property, i.e.,  $\sum_{i=1}^d \varrho(X_i|X) = \varrho(X)$  and the RORAC (Return on Risk Adjusted Capital) compatible property (refer to Tasche (2008) for more details).

**Theorem 4.1.1.** *VaR contribution.* Let  $\rho$  be the value-at-risk with equation (4.1.1) satisfying the Euler capital allocation principle, then, the capital allocation for  $X_i$  is given by,

$$\mathbb{E}[X_i|X = \text{VaR}_\alpha(X)]. \quad (4.1.5)$$

*Proof.* Refer to [Tasche \(2000\)](#).

**Theorem 4.1.2.** *TVaR contribution.* Let  $\rho$  be the tail value-at-risk with equation (4.1.2) satisfying the Euler capital allocation principle, then, the capital allocation for  $X_i$  is given by,

$$\mathbb{E}[X_i|X \geq \text{VaR}_\alpha(X)]. \quad (4.1.6)$$

*Proof.* Refer to [McNeil et al. \(2005\)](#).

For more capital allocation techniques refer to [McNeil et al. \(2005\)](#).

## 4.1.2 Multivariate risk measures

The HCM presented in Section 2.3.3 enables to generate simulations of the unpaid claim liabilities or as we renamed it in this section, the loss distribution, which by construction, accounts for the dependence of the lines of business that make up the insurance portfolio. The univariate risk measures presented in Section 4.1.1 allow to individually evaluate the risk associated with each line of business and by analyzing the sum (as in Section 4.1.1.3) we can allocate capital and obtain a diversification benefit if the risk measure  $\rho$  satisfies the subadditivity property. In this section, we present some multivariate risk measures. Specifically, due to the bivariate approach in the HCM, we describe bivariate risk measures that consider the dependence of the lines of business to provide additional financial security such as the lower and upper orthant VaR presented in [Cossette et al. \(2016\)](#). The additional security provided by the multivariate approach for a risk adverse entity would be a favorable thing in terms of reserving and/or calculating the risk adjustment for non-financial risks. In a multivariate setting, risk measures can be defined in several ways due to partial ordering.

In the following, we describe the bivariate lower and upper orthant VaR introduced in [Embrechts and Puccetti \(1996\)](#), with a capital allocation motivation as described in [Cossette et al. \(2013\)](#). Consider a bivariate random vector  $\mathbf{X} = (X_1, X_2)$  with marginal cumulative distribution functions  $F_{X_1}, F_{X_2}$ , marginal survival distribution functions  $\bar{F}_{X_1}, \bar{F}_{X_2}$ , joint cumulative distribution function  $F_{\mathbf{X}}$  and joint survival distribution function  $\bar{F}_{\mathbf{X}}$ . With the same notation as [Cossette et al. \(2013\)](#), we denote for a fixed  $x_1$ ,  $x_2 \rightarrow F_{x_1}(x_2) = F_{\mathbf{X}}(x_1, x_2)$  and  $x_2 \rightarrow \bar{F}_{x_1}(x_2) = \bar{F}_{\mathbf{X}}(x_1, x_2)$ . Then, the lower orthant VaR is given by,

$$\underline{\text{VaR}}_{\alpha}(\mathbf{X}) = \{(x_1, F_{x_1}^{-1}(\alpha)), x_1 \geq \text{VaR}_{\alpha}(X_1)\}, \quad (4.1.7)$$

or equivalently

$$\underline{\text{VaR}}_{\alpha}(\mathbf{X}) = \{(F_{x_2}^{-1}(\alpha), x_2), x_2 \geq \text{VaR}_{\alpha}(X_2)\},$$

where  $F_{x_i}^{-1}(\alpha) = \inf\{t \in \mathbb{R} \mid F_{x_i}(t) \geq \alpha\}$ , and thus,  $F_{\mathbf{X}}(x_1, F_{x_1}^{-1}(\alpha)) = F_{\mathbf{X}}(F_{x_2}^{-1}(\alpha), x_2) = \alpha$ . The  $\underline{\text{VaR}}_{\alpha}$  curve can be interpreted as the threshold that protects  $\mathbf{X}$  at level  $\alpha$ , meaning that events above this curve are considered tail events [Mailhot and Mesfioui \(2016\)](#). The bivariate upper orthant VaR at confidence level  $\alpha$  is denoted by the set,

$$\overline{\text{VaR}}_{\alpha}(\mathbf{X}) = \{(x_1, \bar{F}_{x_1}^{-1}(1 - \alpha)), x_1 \leq \text{VaR}_{\alpha}(X_1)\}, \quad (4.1.8)$$

or equivalently

$$\overline{\text{VaR}}_{\alpha}(\mathbf{X}) = \{(\bar{F}_{x_2}^{-1}(1 - \alpha), x_2), x_2 \leq \text{VaR}_{\alpha}(X_2)\},$$

where  $\bar{F}_{x_i}^{-1}(\alpha) = \inf\{t \in \mathbb{R} \mid \bar{F}_{x_i}(t) \leq \alpha\}$ . The  $\overline{\text{VaR}}_{\alpha}$  can be interpreted as the curve that protects the joint survival probability of  $\mathbf{X}$  at level  $\alpha$ . The following example is carried throughout the current section and the following [Section 4.1.2.1](#) to visualize graphically the bivariate risk measures.

*Example 1* Consider  $X_1 \sim \text{Exponential}(0.2)$ ,  $X_2 \sim \text{Exponential}(0.4)$ , confidence level  $\alpha = 95\%$  and a Clayton copula (refer to [Appendix B.1](#) for more information on the Clayton copula) with dependence parameter  $\theta = 2$ . Then, the univariate, bivariate lower orthant and bivariate upper orthant VaR can be visualized as follows,

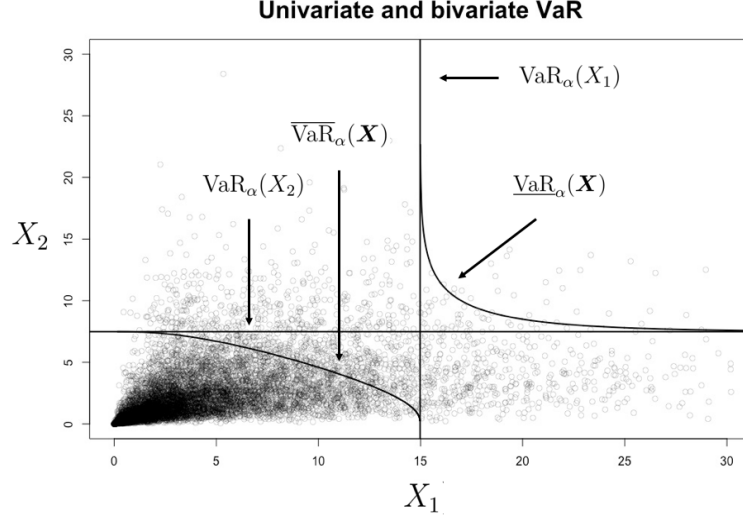


Figure 4.1: Univariate  $\text{VaR}_\alpha$ , bivariate lower and upper orthant VaR for Example 1.

Figure 4.1 has 10,000 observations from the Clayton copula in the background just for didactic purposes. The risk measures displayed are the theoretical values by using equations (4.1.1), (4.1.7) and (4.1.8).

To improve the understanding of the bivariate lower and upper orthant VaR and show a very interesting application, in what follows we relate them to a ruin probability as done in Cossette et al. (2013). Consider two lines of business  $(X_1, X_2)$  with initial allocated reserves  $u_1$  and  $u_2$ , respectively. We denote the ruin probabilities of line  $i$  for the next period as,

$$\Psi_{or}(u_1, u_2) = \mathbb{P} \left[ \bigcup_{i=1}^2 \text{Ruin}_i \right], \quad \Psi_{and}(u_1, u_2) = \mathbb{P} \left[ \bigcap_{i=1}^2 \text{Ruin}_i \right], \quad (4.1.9)$$

where  $\text{Ruin}_i$  is defined as the next period aggregate claims  $S_i$  minus the corresponding premium income  $p_i$  being larger than the initial reserves, or in mathematical notation  $\text{Ruin}_i = \{S_i - p_i > u_i\}$ . Then, we fix the confidence level  $\alpha$  and assume fixed premium incomes such that  $p_i < \text{VaR}_\alpha(S_i)$ . Consequently, we express the set of initial reserves  $(u_1^{(or)}, u_2^{(or)})$  such that,

$$\Psi_{or}(u_1, u_2) = 1 - \alpha \quad \text{coincides with} \quad \underline{\text{VaR}}_\alpha(S_1 - p_1, S_2 - p_2). \quad (4.1.10)$$

The interpretation of equations (4.1.9) and (4.1.10) is that either  $S_1$  or  $S_2$  do not surpass their respective  $(1 - \alpha)\%$  largest value. On the other hand, we express the set of initial reserves  $\left(u_1^{(and)}, u_2^{(and)}\right)$  such that,

$$\Psi_{and}(u_1, u_2) = \alpha \quad \text{coincides with} \quad \overline{\text{VaR}}_\alpha(S_1 - p_1, S_2 - p_2). \quad (4.1.11)$$

Equations (4.1.9) and (4.1.11) mean that both  $S_1$  and  $S_2$  do not surpass the level  $\alpha$ . For examples and more details refer to [Cossette et al. \(2013\)](#).

Analogous to the univariate case, the bivariate upper and lower orthant VaR do not provide information about the tail of the joint distribution  $\mathbf{X}$ , thus, in what follows we describe the bivariate upper and lower orthant TVaR as the conditional expectation of a set as done in [Cossette et al. \(2016\)](#). The bivariate lower orthant TVaR is a set composed of two curves of the form,

$$\underline{\text{TVaR}}_{\alpha, \mathbf{X}}(\mathbf{X}) = \left( (x_1, \underline{\text{TVaR}}_{\alpha, x_1}(\mathbf{X})), (\underline{\text{TVaR}}_{\alpha, x_2}(\mathbf{X}), x_2) \right), \quad (4.1.12)$$

where

$$\underline{\text{TVaR}}_{\alpha, x_i}(\mathbf{X}) = \mathbb{E} \left[ X_j | X_j > F_{x_i}^{-1}(\alpha), X_i \leq x_i \right], x_i \geq \text{VaR}_\alpha(X_i), \quad i, j = 1, 2 \quad (i \neq j).$$

The  $\underline{\text{TVaR}}_{\alpha, x_i}(\mathbf{X})$  is the expectation of the random variable under study  $X_j$ , if it exceeds its lower orthant VaR curve at level  $\alpha$ , but remains below  $x_i \in [\text{VaR}_\alpha(X_i), \infty)$ . Thus, if the line of business  $X_j$  surpasses the lower orthant VaR,  $\underline{\text{TVaR}}_{\alpha, x_i}(\mathbf{X})$  ensures that both  $X_i$  and  $X_j$  are at least protected up to the level  $\alpha$  as stated in [Cossette et al. \(2016\)](#). The two curves that compose the bivariate lower orthant TVaR are observed for Example 1 in the following figure.

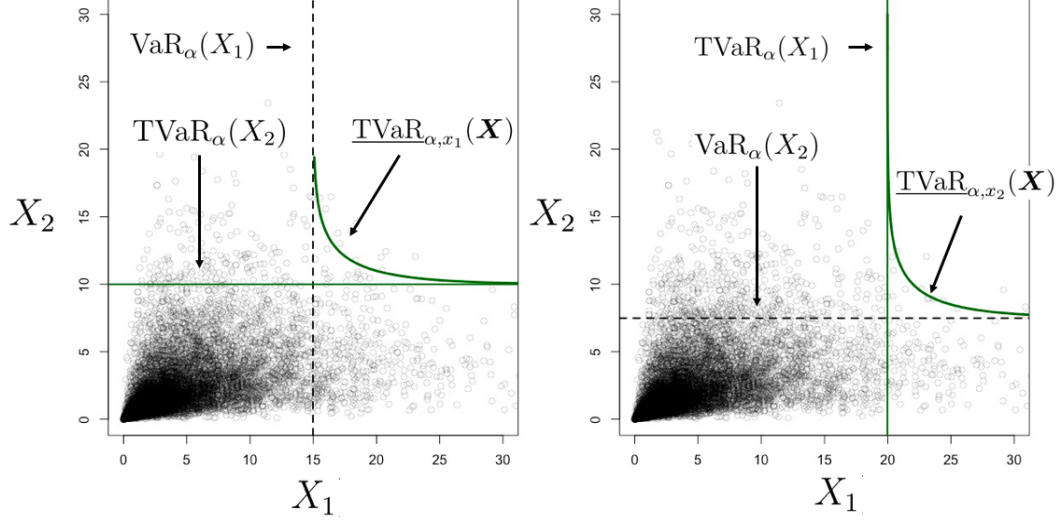


Figure 4.2: Set of two curves composing the bivariate lower orthant  $\overline{\text{TVaR}}_{\alpha, \mathbf{X}}(\mathbf{X})$  for Example 1.

Analogous to the lower orthant, the bivariate upper orthant TVaR is a set composed of two curves of the form,

$$\overline{\text{TVaR}}_{\alpha, \mathbf{X}}(\mathbf{X}) = ((x_1, \overline{\text{TVaR}}_{\alpha, x_1}(\mathbf{X})), (\overline{\text{TVaR}}_{\alpha, x_2}(\mathbf{X}), x_2)), \quad (4.1.13)$$

where

$$\overline{\text{TVaR}}_{\alpha, x_i}(\mathbf{X}) = \mathbb{E} [X_j | X_j > \bar{F}_{x_i}^{-1}(1 - \alpha), X_i \geq x_i], x_i \leq \text{VaR}_\alpha(X_i), \quad i, j = 1, 2 \quad (i \neq j).$$

The  $\overline{\text{TVaR}}_{\alpha, x_i}(\mathbf{X})$  is the expectation of  $X_j$ , if it exceeds its upper orthant VaR curve at level  $1 - \alpha$ , but remains below  $x_i \in [0, \text{VaR}_\alpha(X_i)]$ . Thus, if  $X_j$  exceeds the upper orthant VaR, the  $\overline{\text{TVaR}}_{\alpha, x_i}(\mathbf{X})$  ensures with probability  $1 - \alpha$  that both  $X_i$  and  $X_j$  are above the threshold. The two curves that compose the bivariate upper orthant TVaR are observed for Example 1 in the following figure.

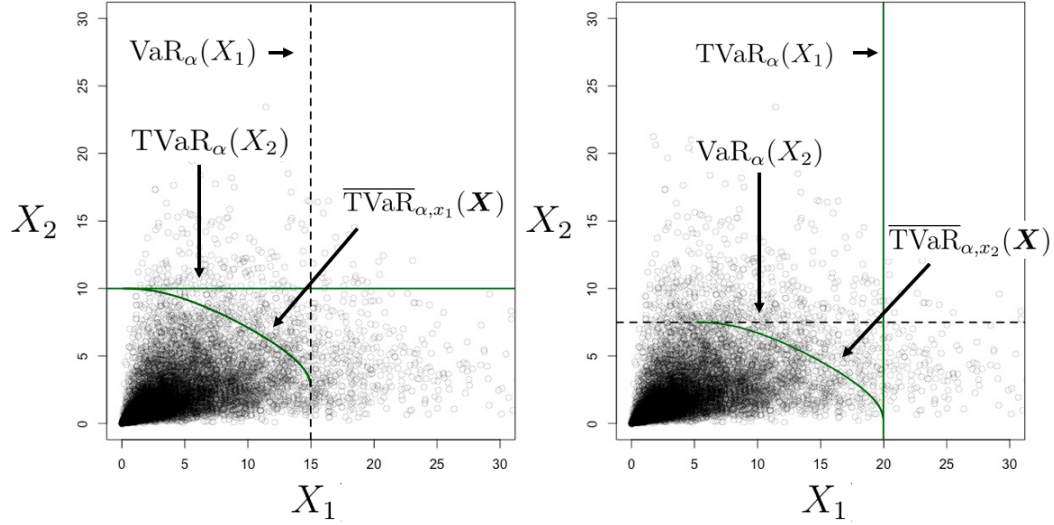


Figure 4.3: Set of two curves composing the bivariate upper orthant  $\overline{TVaR}_{\alpha, \mathbf{X}}(\mathbf{X})$  for Example 1.

#### 4.1.2.1 Capital allocation based on multivariate risk measures

In Section 4.1.1.3, we describe the capital allocation problem in the univariate setting. For bivariate risk measures, capital allocation becomes fundamental since we have a set of curves from which we have to select a single couple according to a certain criteria to satisfy the capital requirement or the risk adjustment depending on the objective. We focus on two criterion described in Cossette et al. (2013) for the lower and upper orthant VaR and in Mailhot and Mesfioui (2016) for the lower and upper orthant TVaR. The two criterion are the proportional allocation method and the orthogonal projection method. Furthermore, we seek to extend the proportional allocation method to the TVaR with a practical application.

From Figure 4.1, we can observe that the univariate VaR for  $X_1$  and  $X_2$  serve as a boundary for the bivariate upper and lower orthant VaR (refer to Cossette et al. (2013) for a proof). Thus, the orthogonal projection VaR, mathematically, seeks to minimize the euclidean distance from the curve to the two boundaries, but in our context, the interpretation is to find the closest couple to the stand-alone basis while accounting for

dependence. Thus, we have the following optimization problem,

$$\min_{x_1 > \text{VaR}_\alpha(X_1)} (x_1 - \text{VaR}_\alpha(X_1))^2 + (F_{x_1}^{-1}(\alpha) - \text{VaR}_\alpha(X_2))^2. \quad (4.1.14)$$

We denote as  $x_1^*$  the argument that minimizes expression (4.1.14). Thus, the couple  $(x_1^*, F_{x_1^*}^{-1}(\alpha))$  is the optimal couple using the orthogonal projection for the bivariate lower orthant VaR. Analogously, to obtain the optimal couple for the upper orthant VaR we replace  $F_{x_1}^{-1}(\alpha)$  with  $\bar{F}_{x_1}^{-1}(1 - \alpha)$  in (4.1.14) subject to values of  $x_1 < \text{VaR}_\alpha(X_1)$ . The orthogonal projection method relies on the assumption that the  $\underline{\text{VaR}}_\alpha(\mathbf{X})$  curve ( $\overline{\text{VaR}}_\alpha(\mathbf{X})$ , respectively) is convex (concave).

The idea for the orthogonal projection of the VaR can be used for the TVaR as done in Mailhot and Mesfioui (2016). In Figure 4.2 and 4.3 it can be seen that the  $\text{VaR}_\alpha(X_1)$  and  $\text{TVaR}_\alpha(X_2)$  serve as boundaries for the lower and upper orthant TVaR of  $x_1$  (refer to Mailhot and Mesfioui (2016) for a proof). Thus, the orthogonal projection for the lower orthant TVaR is modified from the VaR in the following way,

$$\min_{x_i > \text{VaR}_\alpha(X_i)} (x_i - \text{VaR}_\alpha(X_i))^2 + (\underline{\text{TVaR}}_{\alpha, x_i}(\mathbf{X}) - \text{TVaR}_\alpha(X_j))^2, \quad i, j = 1, 2 \quad (i \neq j). \quad (4.1.15)$$

After finding the optimal values  $(x_1^*, x_2^*)$  that solve the two optimization problems (4.1.15), the optimal couple is given by,

$$\underline{\text{TVaR}}_{\alpha, \mathbf{x}^*}(\mathbf{X}) = \left( \underline{\text{TVaR}}_{\alpha, x_2^*}(\mathbf{X}), \underline{\text{TVaR}}_{\alpha, x_1^*}(\mathbf{X}) \right). \quad (4.1.16)$$

The method is analogous for the bivariate upper TVaR.

The second method is the proportional allocation method described in Cossette et al. (2013) for the VaR. Its advantage is that it considers the possible difference of scale between the lines of business by selecting the couples that intersect with the linear equation  $x_1 = \frac{\text{VaR}_\alpha(X_1)}{\text{VaR}_\alpha(X_2)} x_2$  where  $\frac{\text{VaR}_\alpha(X_1)}{\text{VaR}_\alpha(X_2)}$  is the ratio measuring the difference in scales between  $X_1$  and  $X_2$ . Thus, the optimization problem for the bivariate lower VaR is given by,

$$\min_{x_1 > \text{VaR}_\alpha(X_1)} \left( x_1 - \frac{\text{VaR}_\alpha(X_1)}{\text{VaR}_\alpha(X_2)} F_{x_1}^{-1}(\alpha) \right)^2. \quad (4.1.17)$$



Similar to the orthogonal projection method, as stated in [Cossette et al. \(2013\)](#) the solution relies on the convexity of the bivariate lower orthant VaR (concavity for the respective upper orthant VaR) and the argument which minimizes (4.1.17) provides the optimal couple  $(x_1^*, F_{x_1^*}^{-1}(\alpha))$ . Respectively, for the bivariate upper orthant VaR we replace  $F_{x_1}^{-1}(\alpha)$  with  $\bar{F}_{x_1}^{-1}(1 - \alpha)$  subject to values of  $x_1 < \text{VaR}_\alpha(X_1)$ .

In this thesis, we seek to extend the proportional allocation method to the TVaR since we consider of the utmost importance to consider the tail events of the joint distribution of the random vector  $\mathbf{X}$ . Similarly to the optimization problem with the VaR, we now seek to find the intersection with the linear equation  $x_1 = \frac{\text{TVaR}_\alpha(X_1)}{\text{TVaR}_\alpha(X_2)}x_2$ , where the slope of the line homogenizes the scale in the data through the quotient of the univariate tail value-at-risks. Thus, the minimization problem for the bivariate lower orthant TVaR is given by,

$$\min_{x_1 > \text{VaR}_\alpha(X_1)} \left( x_1 - \frac{\text{TVaR}_\alpha(X_1)}{\text{TVaR}_\alpha(X_2)} \underline{\text{TVaR}}_{\alpha, x_1}(\mathbf{X}) \right)^2. \quad (4.1.18)$$

To obtain the optimal couple we rely on the convexity of the bivariate lower orthant TVaR and respectively, on the concavity of the bivariate upper orthant TVaR. Consequently, the argument  $x_1^*$  minimizing (4.1.18) is found by solving the following equation,

$$\frac{d}{dx_1} \underline{\text{TVaR}}_{\alpha, x_1}(\mathbf{X}) = \frac{\text{TVaR}_\alpha(X_1)}{\text{TVaR}_\alpha(X_2)}. \quad (4.1.19)$$

Yielding the optimal couple,

$$\left( x_1^*, \underline{\text{TVaR}}_{\alpha, x_1^*}(\mathbf{X}) \right) = \left( x_1^*, \frac{\text{TVaR}_\alpha(X_2)}{\text{TVaR}_\alpha(X_1)} x_1^* \right). \quad (4.1.20)$$

The minimization problem in (4.1.18) is then repeated for the curve  $\underline{\text{TVaR}}_{\alpha, x_2}(\mathbf{X})$  to find the optimal couple given by,

$$\underline{\text{TVaR}}_{\alpha, \mathbf{x}^*}(\mathbf{X}) = \left( \underline{\text{TVaR}}_{\alpha, x_2^*}(\mathbf{X}), \underline{\text{TVaR}}_{\alpha, x_1^*}(\mathbf{X}) \right). \quad (4.1.21)$$

The process is analogous for the bivariate upper TVaR subject to the restriction of  $x_1 < \text{VaR}_\alpha(X_1)$ .

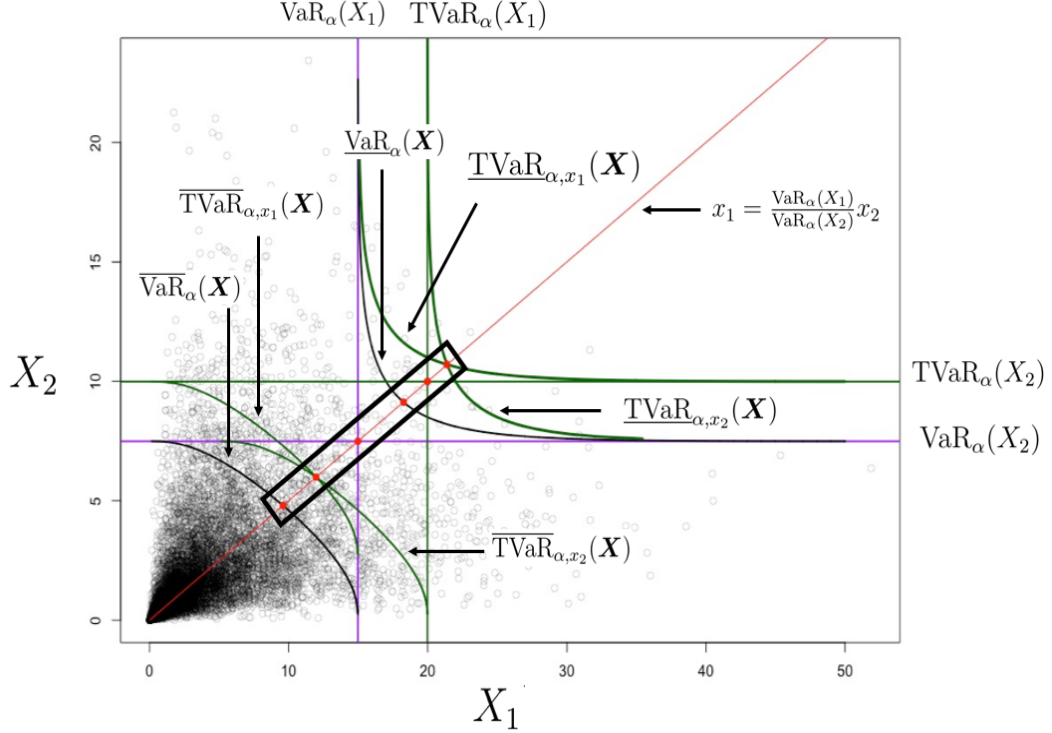


Figure 4.4: Univariate and bivariate  $\text{VaR}_\alpha$  and  $\text{TVaR}_\alpha$  intersection with proportional allocation line for Example 1.

In Figure 4.4 we continue with Example 1 where we can observe with straight lines the univariate  $\text{VaR}_\alpha$  (in purple) for  $X_1$  and  $X_2$ , the  $\text{TVaR}$  (in darkgreen) for  $X_1$  and  $X_2$  and the proportional allocation line (in red)  $x_1 = \frac{\text{VaR}_\alpha(X_1)}{\text{VaR}_\alpha(X_2)}x_2$  which for this example happens to be the same as  $x_1 = \frac{\text{TVaR}_\alpha(X_1)}{\text{TVaR}_\alpha(X_2)}x_2$ , although this is not always the case. Then, we have the bivariate lower and upper  $\text{VaR}$  (in black) and the bivariate lower and upper  $\text{TVaR}$  (in darkgreen). Furthermore, in the black box at the center of the graph we highlight with red points the intersections of all the given curves which are further explained in the following Figure 4.5.

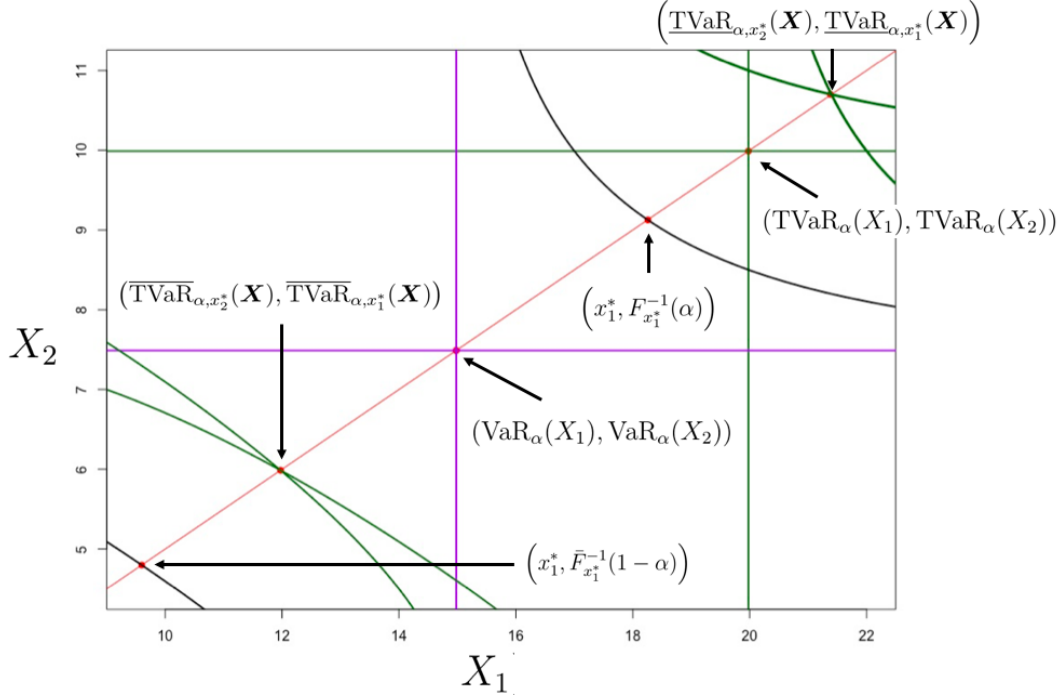


Figure 4.5: Zoom in of black box in Figure 4.4 to visualize intersection between univariate and bivariate risk measures with the proportional allocation line for Example 1.

Figure 4.5 is of the utmost importance because it shows empirically the ordering of the optimal couples when considering the proportional allocation method. Using the traditional stand-alone method the optimal couples are represented by the intersection of the univariate value-at-risk  $(\text{VaR}_{\alpha}(X_1), \text{VaR}_{\alpha}(X_2))$  or if the insurer decides to consider the tail events, the intersection of the tail value-at-risk  $(\text{TVaR}_{\alpha}(X_1), \text{TVaR}_{\alpha}(X_2))$ . If the insurer considers dependence between the lines of business without considering tail events and wants to protect the joint survival probability up to level  $\alpha$ , then the optimal couple with the proportional allocation method is  $(x_1^*, \bar{F}_{x_1^*}^{-1}(1 - \alpha))$  and when considering the tail events, the optimal reserving amounts are  $(\overline{\text{TVaR}}_{\alpha, x_2^*}(\mathbf{X}), \overline{\text{TVaR}}_{\alpha, x_1^*}(\mathbf{X}))$ . Analogously, if the insurer wants to protect the portfolio  $\mathbf{X}$  with a confidence level  $\alpha$  then, the optimal threshold with the proportional allocation method is  $(x_1^*, F_{x_1^*}^{-1}(\alpha))$  and more conservatively, when considering the tail events the optimal couple is  $(\underline{\text{TVaR}}_{\alpha, x_2^*}(\mathbf{X}), \underline{\text{TVaR}}_{\alpha, x_1^*}(\mathbf{X}))$ .

Constructing the proportional allocation line (observed in red in Figure 4.5) allows to

have a set of 6 possible reserving couples accounting for dependence, tail events, joint probability of survival or the threshold that protects the whole portfolio. Thus, its a tool providing the entity and the actuary with a risk management problem which can be solved according to the entity's view of risk as established by IFRS 17 in Chapter 1.

## 4.2 Cost of capital method

Another way of determining the risk adjustment for non-financial risk is using the cost of capital method described in IAA (2018), where the risk adjustment is the present value of the future cost of capital associated with the unpaid claim liabilities. The amount of capital is determined considering the probability distribution of future cash flows, however, it is not necessary since such amounts are not defined based on regulatory capital. The risk adjustment is calculated as,

$$\text{Risk adjustment} = \sum_{t=1}^n \frac{r_t \cdot C_t}{(1 + d_t)^t} , \quad (4.2.1)$$

where,

- $C_t$  is the assigned capital amount for the period ending at time  $t$ ,
- $r_t$  is the selected cost of capital rate for period ending at time  $t$ , and
- $d_t$  is the selected discount rate at time  $t$ , reflecting a yield curve, if appropriate.

The amount of capital  $C_t$ , considered in this thesis, is the difference between the amount from the probability distribution associated with the selected confidence level  $\alpha$  and the expected value. Thus, for a non-life insurance liability  $X_t$  at time  $t$  and risk measure  $\varrho$  at level  $\alpha$ , the amount of capital is given by,

$$C_t = \varrho_\alpha(X_t) - \mathbb{E}[X_t].$$

The main advantage of the cost of capital method is the simplicity with which an actu-

ary can explain the method to management and contrary to the quantile techniques (in Section 4.1.1 and 4.1.2), it considers the cost of bearing the uncertainty in the liabilities through the cost of capital rate. But, as mentioned in Section 1.1.2.5, in the end, the risk adjustment has to be converted to a confidence level. Moreover, this technique requires additional assumptions over the cost of capital rate which is not an easy task due to the long-term nature of the unpaid claim liabilities.

# Chapter 5

## Numerical Application

This chapter contains details on the data analysis which led to the model presented in Chapter 3 under the General Model measurement of IFRS 17 and the capital requirements set by OSFI, described in Chapter 1. The dataset used is provided by Eckler Ltd and corresponds to data from the Canadian automobile industry. All calculations are done with **R**, some specific packages are used and are mentioned in their specific section and furthermore, each statistical test is evaluated at a 95% confidence level. The dataset can be purchased online through the General Insurance Statistical Agency (GISA) and consists of two different P&C insurance lines (Reports AUTO7001 and AUT07002) for three regions in Canada, making a total of six lines of business to model. The lines of business are personal auto (PA) and commercial auto (CA) for the province of Ontario (ON), Alberta (AB) and Atlantic Canada (ATL)<sup>1</sup>. The incremental incurred claim amounts and earned premiums are not disclosed for confidentiality reasons.

Loss and Expense (L&E) semestrial claim amounts are available for each dataset from the first semester of 1997 (denoted from now on as 1997-1 and correspondingly, 1997-2 for the second semester) to the second semester of 2017. General notes on the historical claims are available starting from 2003-1 and thus, only fifteen years of information are taken into

---

<sup>1</sup>Atlantic Canada is made up of four provinces: Prince Edward Island, New Brunswick, Nova Scotia and Newfoundland and Labrador.

account. It is important to note that the Memorandum for the Appointed Actuary of a P&C company [OSFI \(2017\)](#) establishes that for the Appointed Actuary's Return specified in subsection 667(2) of the *Insurance Companies Act* [Government of Canada \(1991\)](#), the Appointed Actuary should assess the change in the unpaid claim liabilities for all lines of business for 10 years of data or move towards the 10 year standard. Furthermore, in the actuarial literature it is a common practice to analyze data with run-off triangles using the 10 year standard. Although that is the standard practice, in the following analysis, the industry aggregates are analyzed (the sum for each insurer of a given province) and thus, it is assumed that a large insurance company is being assessed which consequently has actual experience that surpasses the 10 year standard. Therefore, the semestrial data for 15 years of experience is modeled for demonstration purposes. For a specific company replicating the presented model, the data could be stored and modeled every four months and be equivalent to the size of the loss triangle in this analysis.

## 5.1 Descriptive statistics

After constructing the run-off triangle with incremental incurred claims as in [Table 2.1](#) and calculating the loss ratios for each accident year and development lag, we deal with a loss triangle of  $30 \times 30$  ( $I = J = 30$ ) making up for a total of actual experience with 465 cells where our goal is to estimate the corresponding 435 unpaid claim liabilities.

In [Figure 5.1](#) we can observe the empirical distribution of the loss ratios through the histograms. The personal auto line has higher loss ratios than for the commercial auto line. Furthermore, the concentration of zeros is higher for the province of AB and ATL compared to ON, highlighting the size of the auto insurance industry being larger for the latter.

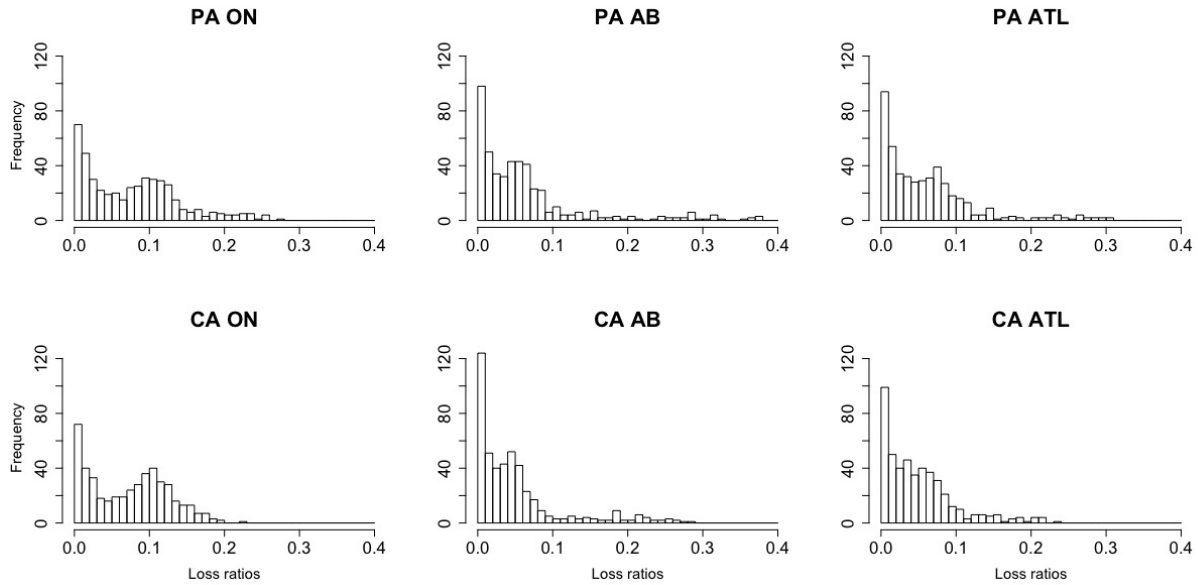


Figure 5.1: Histograms of loss ratios for the 6 lines of business.

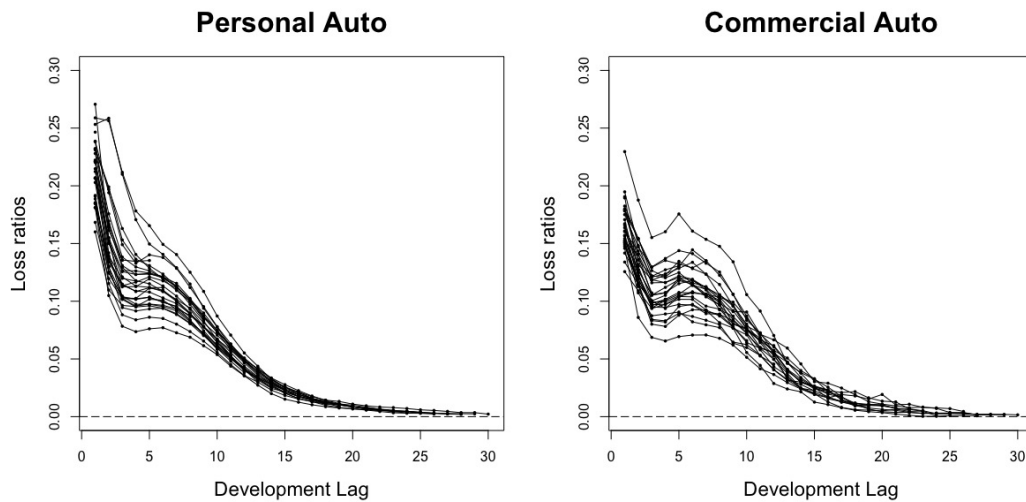


Figure 5.2: Behavior of the time series of the loss ratios by accident year for the automobile industry of Ontario.

Figure 5.2 shows the slow convergence of the loss ratios to zero for the whole Ontario automobile industry. Even after 10 years (lag 20), the loss ratios are considerably different than zero, confirming the decision to model with 15 years of data. Another conclusion we can draw from Figure 5.2 is the higher volatility observed for the commercial line as



compared to personal auto.

## 5.2 Marginal model

We assume that every loss ratio for line of business  $k = \{1, 2, 3, 4, 5, 6\}$  corresponding respectively to {PA ON, CA ON, PA AB, CA AB, PA ATL, CA ATL}, is of the form,

$$Y_{i,j}^{(k)} | \mathbf{X}_{i,j} \sim TW_{p^{(k)}}(\mu_{i,j}^{(k)}, \phi_j^{(k)}), \quad (5.2.1)$$

where the covariates  $\mathbf{X}_{i,j}$  are the latent variables accident semester  $i$  and development lag  $j$ . Furthermore, for each line of business and accident year we have the following covariance matrix  $\mathbf{V}_i^{(k)} = A_i^{(k)1/2} R_i(\rho^{(k)}) A_i^{(k)1/2}$ . The mean model and dispersion sub-model parameters including correlation  $\{\beta^{(k)}, \gamma^{(k)}, \rho^{(k)}\}$  are estimated with the two-step estimation method introduced in Section 2.2.7.1. The estimated parameters for the six lines of business are shown in Appendix A.

The goodness-of-fit tests to assess the distributional assumptions on the response variable with the respective log-link function along with the correlation test are shown in the following Table 5.1.

	ON	ON	AB	AB	ATL	ATL
	PA	CA	PA	CA	PA	CA
K-S	0.21	0.69	0.08	0.29	0.47	0.04
A-D	0.25	0.37	0.10	0.11	0.31	0.07
L-B	0.00	0.54	0.00	0.18	0.86	0.11

Table 5.1: Goodness-of-fit and correlation test:  $p$ -values for marginal models.

The goodness-of-fit tests of Table 5.1 are performed on the estimated cumulative distribution functions of the loss ratios which are calculated through the following formula,

$$\hat{U}_{i,j}^{(k)} = F(Y_{i,j}; \hat{\mu}_{i,j}^{(k)}, \hat{\phi}_j^{(k)}, \hat{p}^{(k)}), \quad (5.2.2)$$

where  $F$  is the cumulative distribution function of a Tweedie random variable with mean  $\hat{\mu}_{i,j}^{(k)}$ , dispersion  $\hat{\phi}_j^{(k)}$  and index parameter  $\hat{p}^{(k)}$  for all values of  $i, j$  and  $k$ . To evaluate the cumulative distribution function of a Tweedie distribution, the package **tweedie** constructed from the research of Tweedie (1984); Jørgensen (1987, 1997); Smyth and Verbyla (1999) is used. Then, the  $p$ -values from Table 5.1 are calculated under the assumption that the  $\hat{U}_{i,j}^{(k)}$  are independent observations with distribution  $U(0, 1)$  in spite of the covariance structure as discussed in Section 2.2.7.2. Both K-S and A-D tests assessing the fit in the center and tails of the distributions respectively, are satisfactory for the six lines of business. The goodness-of-fit with equation (5.2.2) is shown in Figure 5.3 for the personal auto line of Ontario given that it is the line with the highest financial impact of the insurance portfolio as seen in Section 5.4. The Ljung-Box test performed on the uncorrelated residuals calculated with equation (2.2.49) is satisfactory for four out of the six lines of business. For personal auto in Ontario and Alberta, the AR(1) assumption is insufficient and could be extended to account for higher lags but in order to maintain a parsimonious model and the assumptions equal for all lines of business, the residuals remain correlated. A graphical contrast between the correlated residuals computed through equation (2.2.44) before the GEE approach and the uncorrelated residuals obtained with equation (2.2.49) after using GEE is shown for the commercial auto line of Alberta in Figure 5.4.

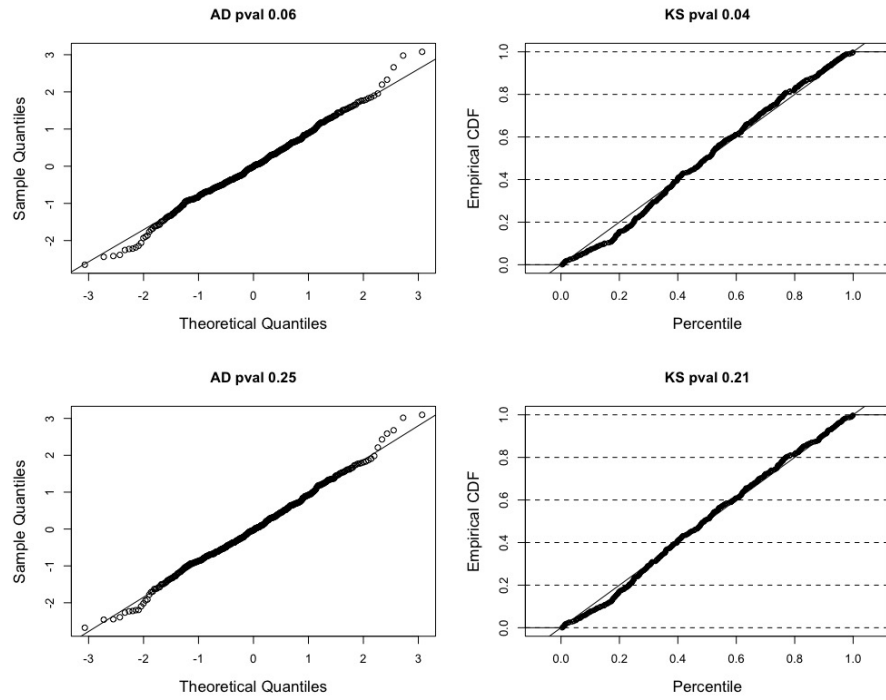


Figure 5.3: Goodness-of-fit measured through the Anderson-Darling and Kolmogorov-Smirnov test for the personal auto line of Ontario before the GEE approach (first row) and after the GEE approach (second row).

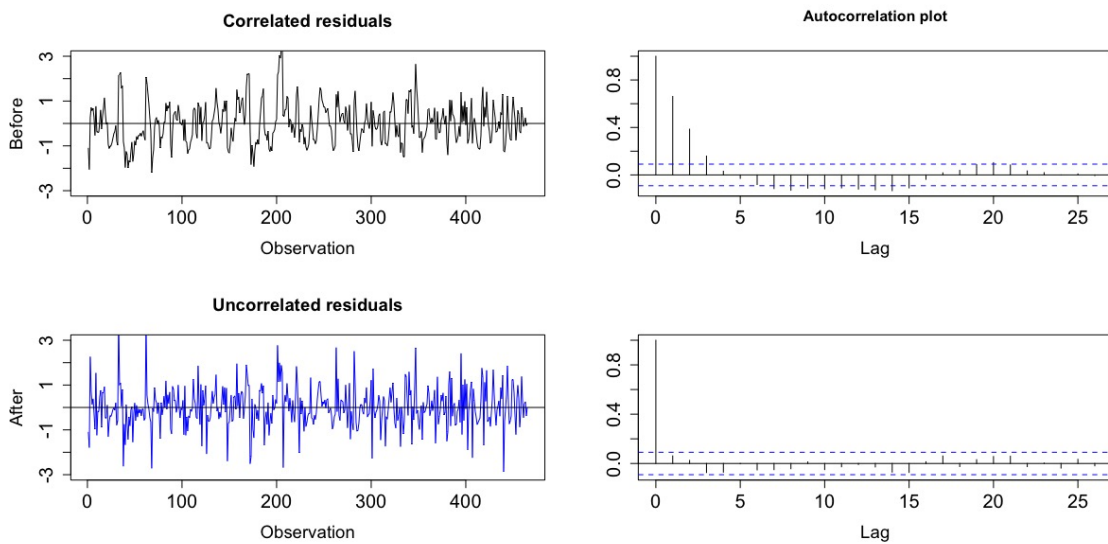


Figure 5.4: Time series of residuals before (first row) and after (second row) using GEE for the commercial auto line of Alberta.

### 5.3 Hierarchical copula model

The fitting of copulas and statistical tests performed in this section are done with help from the **copula** package. It is important to mention that Ontario, Alberta and Atlantic Canada are provinces with privately owned insurers, meaning they have to operate with a lower loss ratio relative to public regimes [Devlin \(2017\)](#). Moreover, the three provinces are regulated under the Canadian legislation, and thus, the risk aggregation technique by [Burgi et al. \(2008\)](#) described in Section 2.3.3 is a reasonable approach because of their shared characteristics. The first decision to account for the diversification benefit is to consider the dependence between personal and commercial auto for the same province as stated in Chapter 3. In the first column of Figure 5.5 and Figure 5.6 we have a marginal dependence contrast, in the residuals plot we observe the dependence once the marginal effects (temporal effects and correlation) have been removed and in the third column, we observe the empirical copula for the province of Ontario and Alberta, respectively.

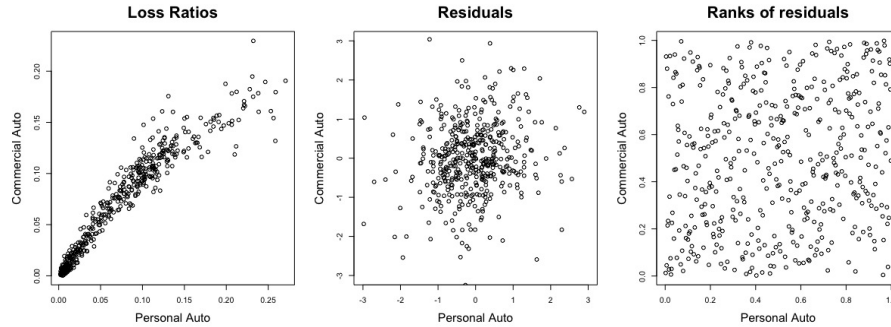


Figure 5.5: Personal auto vs Commercial auto for Ontario

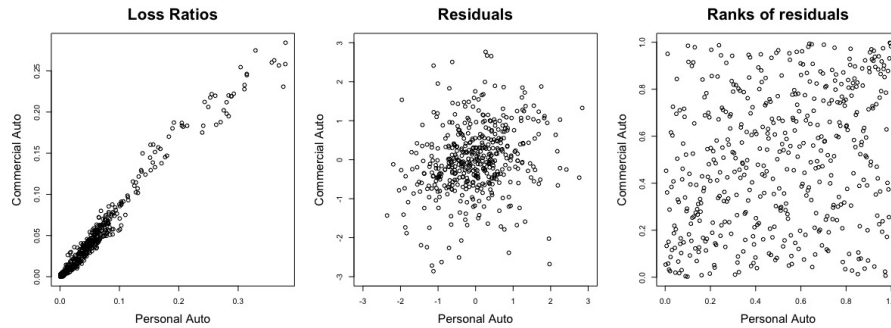


Figure 5.6: Personal auto vs Commercial auto for Alberta

For the province of Alberta, a stronger dependence can be observed as compared to Ontario. In the residuals plot of Figure 5.6, the cloud of points has an observable stronger positive dependence. Before assigning a copula to each province, we test for independence through the value of Kendall's  $\tau$  and testing if its statistically different than zero. Additionally, a Cramér-von Mises test with 1000 simulations of the statistic  $S_n$  under independence relying on the empirical copula process as in [Genest and Rémillard \(2004\)](#).

Province	Kendall's $\tau$	$p$ -value	Cramér-von Mises
ON	0.102	0.000	0.007
AB	0.188	0.000	0.000
ATL	0.058	0.058	0.073

Table 5.2: Independence tests by province. The Cramér-von Mises column shows the  $p$ -values obtained.

In Table 5.2 we can observe the three provinces have a Kendall's  $\tau$  that is statistically significant different than zero but the independence test using the Cramér-von Mises is not rejected for Atlantic Canada. To assess whether the independence copula is a good assumption for Atlantic Canada, an additional test is performed with 100 simulations of a product copula with help from the package **TwoCop** as in [Genest and Rémillard \(2008\)](#) resulting in a  $p$ -value of 0.11 confirming the hypothesis  $H_0 : C_n = \Pi$ , where  $\Pi$  is the independence copula.

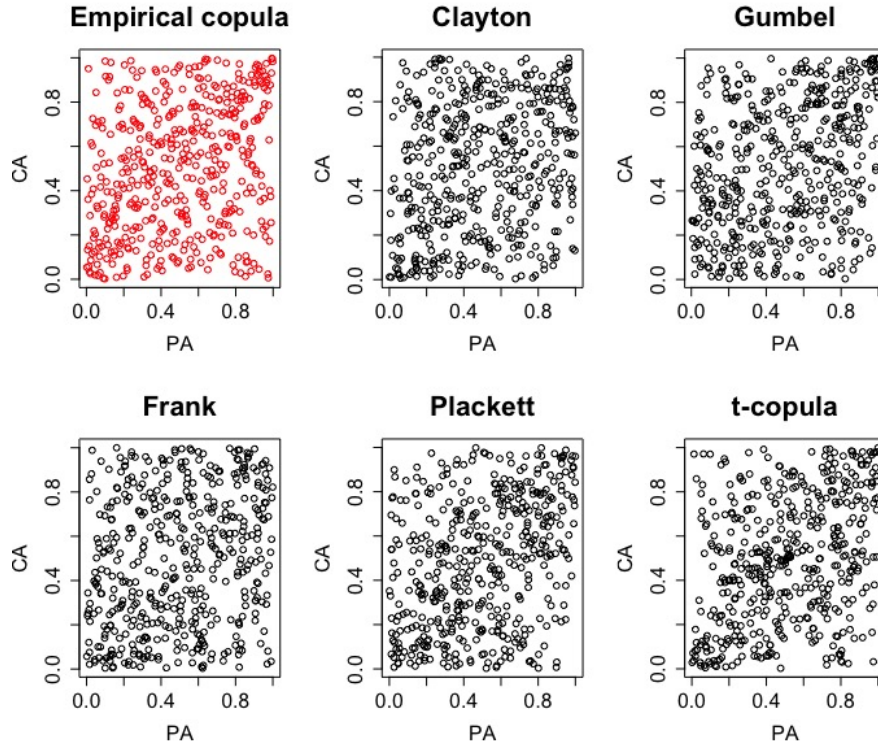


Figure 5.7: Comparison of empirical copula for Alberta vs rank-based simulations of different copula families.

Through visual comparisons as the one presented for Alberta in Figure 5.7 and satisfactory goodness-of-fit tests based on 500 parametric bootstraps of the Cramér-von Mises test, the following copulas were selected.

Province	Copula family	Dependence parameters	Standard deviation parameter	$p$ -value
ON	t	$\nu = 8, \rho = 0.166$	0.050	0.59
AB	t	$\nu = 5, \rho = 0.290$	0.049	0.77
ATL	Independence	-	-	-

Table 5.3: Goodness-of-fit for copula models by province.

The residuals are summed in order to continue with the bivariate semi-parametric estima-

tion approach for the HCM. Three permutations of pairs are available for the second level of the hierarchy: Alberta-Ontario, Alberta-Atlantic Canada or Ontario-Atlantic Canada. Alberta and Atlantic Canada have the stronger dependence as measured by Kendall's  $\tau$  with a value of 0.147 and moreover, they share characteristics as using common factors for pricing and equal possible minimum liabilities limits as mentioned in [Devlin \(2017\)](#). Additionally, for a same subject, prices are more alike between Alberta and Atlantic Canada than compared to Ontario, possibly related to population size and density.

Risk aggregation	Copula family	Dependence parameters	Standard deviation parameter	$p$ -value
AB+ATL	t	$\nu = 4, \rho = 0.228$	0.050	0.39

Table 5.4: Goodness-of-fit for the second level of dependence: Alberta and Atlantic Canada.

In Table 5.4 the dependence reflected by the risk aggregation of Alberta and Atlantic Canada is stronger than the dependence within the risks in the province of Ontario. Residuals are summed again for Alberta and Atlantic Canada and the same dependence analysis, copula selection and estimation techniques are applied in the last dependence level to merge the province of Ontario, but the independence test shows,

Risk aggregation	Kendall's $\tau$	$p$ -value	Cramér-von Mises
ON+AB+ATL	0.016	0.59	0.204

Table 5.5: Independence tests in the last node of the hierarchical copula model.

From Table 5.5, we can conclude there is independence between Ontario and the aggregation of Alberta and Atlantic Canada, therefore, a product copula is assumed for the highest level of our HCM. The estimation stage of the hierarchical copula model is over, and it can be represented with the following graphical structure,

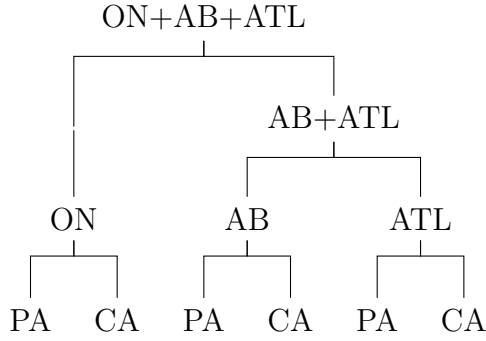


Figure 5.8: HCM structure by province.

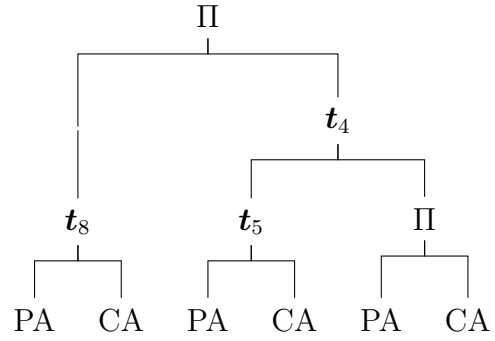


Figure 5.9: HCM structure by copula family.

## 5.4 Risk assessment and capital requirements

With the marginals from Section 5.2 and the HCM structure in Figure 5.9 we can proceed to generate copula samples from the lower triangle using the Iman-Conover reordering algorithm presented in Section 2.4.1 in order to compute the capital requirements and the risk adjustment for non-financial risks. Inverting a Tweedie distribution is not a trivial task, thus, to accelerate the simulations, the code of the **tweedie** package is used along with help from the **Rcpp** and **parallel** packages.

The simulation goes as follows,

1. Generate 100,000 copula samples from the HCM in Figure 5.9.
2. Apply appropriate inverse transformations from the fitted Tweedie marginals.
3. Multiply each loss ratio by the premium of the respective accident year to obtain the unpaid claim liabilities.



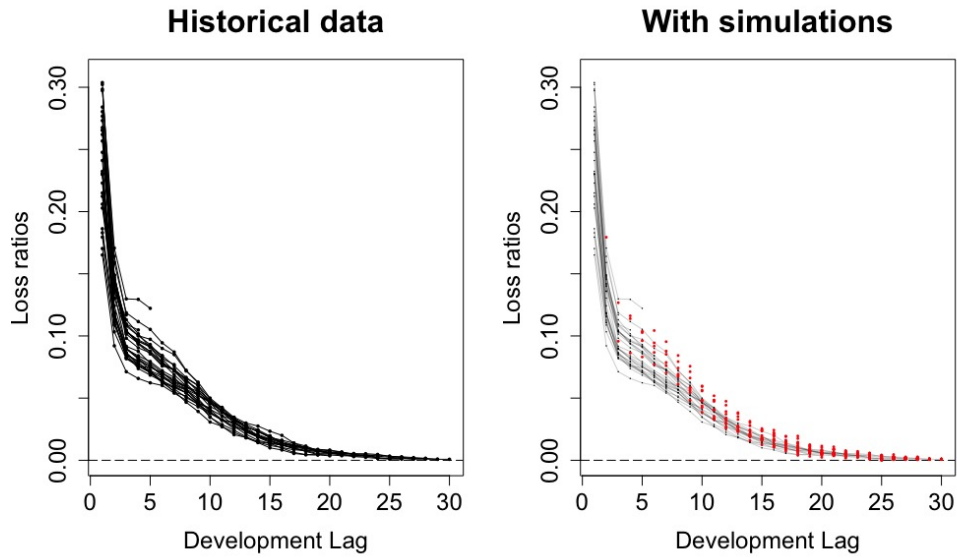


Figure 5.10: Historical data vs historical data with one simulation of loss ratios for personal auto in Atlantic Canada.

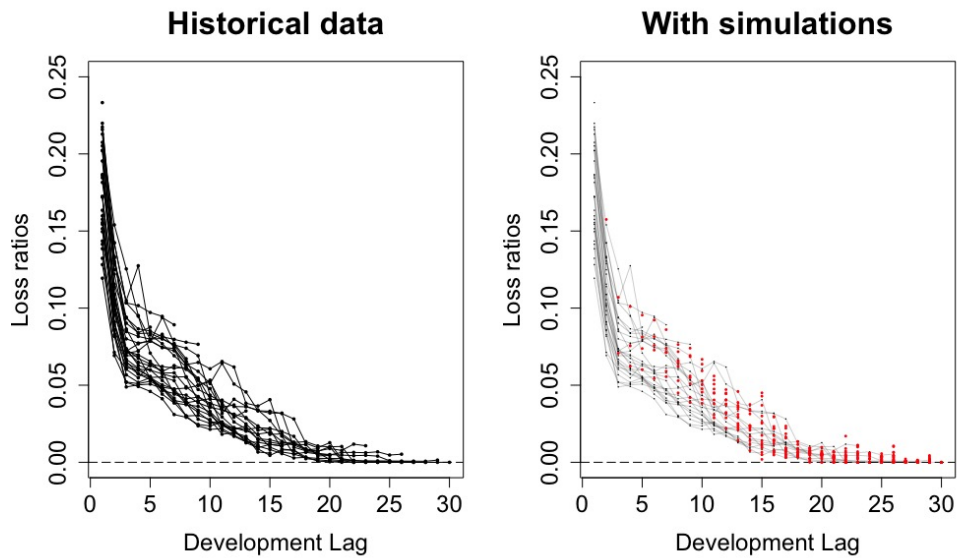


Figure 5.11: Historical data vs historical data with one simulation of loss ratios for commercial auto in Atlantic Canada.

In Figure 5.10 and 5.11, the historical data is contrasted with a plot showing one simulation (in red) drawn from the HCM for personal auto and commercial auto in Atlantic

Canada, respectively. Modeling with a DGLM the marginal distributions captured the different volatilities observed in the data and forecasted the loss ratios accordingly. After obtaining the 100,000 simulations of the unpaid claim liabilities for  $i + j > J + 1$  we can proceed to calculate risk measures as shown in Table 5.6.

		Capital allocation	ON PA	ON CA	AB PA	AB CA	ATL PA	ATL CA	Sum
VaR	SILO	-	95.24	6.99	18.03	1.96	8.73	0.72	131.67
	HCM	-	95.19	6.98	18.01	1.96	8.73	0.72	131.59
TVaR	SILO	-	95.63	7.03	18.12	1.98	8.77	0.73	132.26
	HCM	-	95.61	7.02	18.10	1.98	8.77	0.73	132.21
	SILO	Euler	95.54	6.69	17.43	1.81	8.42	0.65	130.55
	HCM	Euler	95.52	6.73	17.41	1.82	8.42	0.65	130.54

Table 5.6: Univariate VaR and TVaR at  $\alpha = 99\%$  in billions (CAD) for the six lines of business.

Table 5.6 shows the VaR and TVaR at a confidence level of  $\alpha = 99\%$  with 100,000 simulations of the reserves. The SILO method stands for single line of business as commonly known in the actuarial literature and is obtained by simulating each line of business independently, i.e. using the marginal distributions without accounting for any form of dependence within the insurance portfolio. For the TVaR, the traditional Euler capital allocation principle presented in Section 4.1.1.3 is obtained by summing across all business lines and then allocating capital to each line. In this example, the Euler capital allocation principle is shown for demonstration purposes but should not be selected as the reserve amount since we are dealing with insurance lines from different provinces. Table 5.6 shows the diversification benefit obtained from using the HCM presented in Chapter 3 compared to the SILO method by resulting in lower risk measures, as expected, since the simulation of the reserve using the SILO method do not consider any diversification

within the insurance portfolio. The diversification benefit is equal to \$82.8 million (CAD) with the VaR, \$53.6 million with the TVaR (CAD) without capital allocation techniques and \$7.4 million (CAD) using the Euler capital allocation principle. To calculate the risk adjustment we need to subtract the expected value of the reserves from the numbers in Table 5.6 resulting in the following:

		ON	ON	AB	AB	ATL	ATL	Sum
		PA	CA	PA	CA	PA	CA	
VaR	SILO	2.79	0.28	0.67	0.13	0.28	0.06	4.20
	HCM	2.74	0.27	0.65	0.13	0.28	0.06	4.12
TVaR	SILO	3.18	0.31	0.76	0.15	0.32	0.07	4.80
	HCM	3.16	0.31	0.75	0.15	0.32	0.07	4.75

Table 5.7: Risk adjustments corresponding to Table 5.6 for the univariate VaR and TVaR at  $\alpha = 99\%$  in billions (CAD) for the six lines of business.

Furthermore, if the capital allocation is pursued under a bivariate approach, because allocating capital through the aggregation of risks on different jurisdictions is deemed inappropriate, we can exploit measuring the insurance risks with bivariate risk measures taking advantage of the bivariate copulas at each node specified under the HCM. In Figure 5.12 the simulations of the t-copula for the province of Alberta under the HCM are shown along with the univariate risk measures and the bivariate upper and lower TVaR. Furthermore,  $P_1$  and  $P_2$  are the optimal couples under the proportional allocation method for the TVaR. The values for all optimal couples can be observed in Table 5.8.

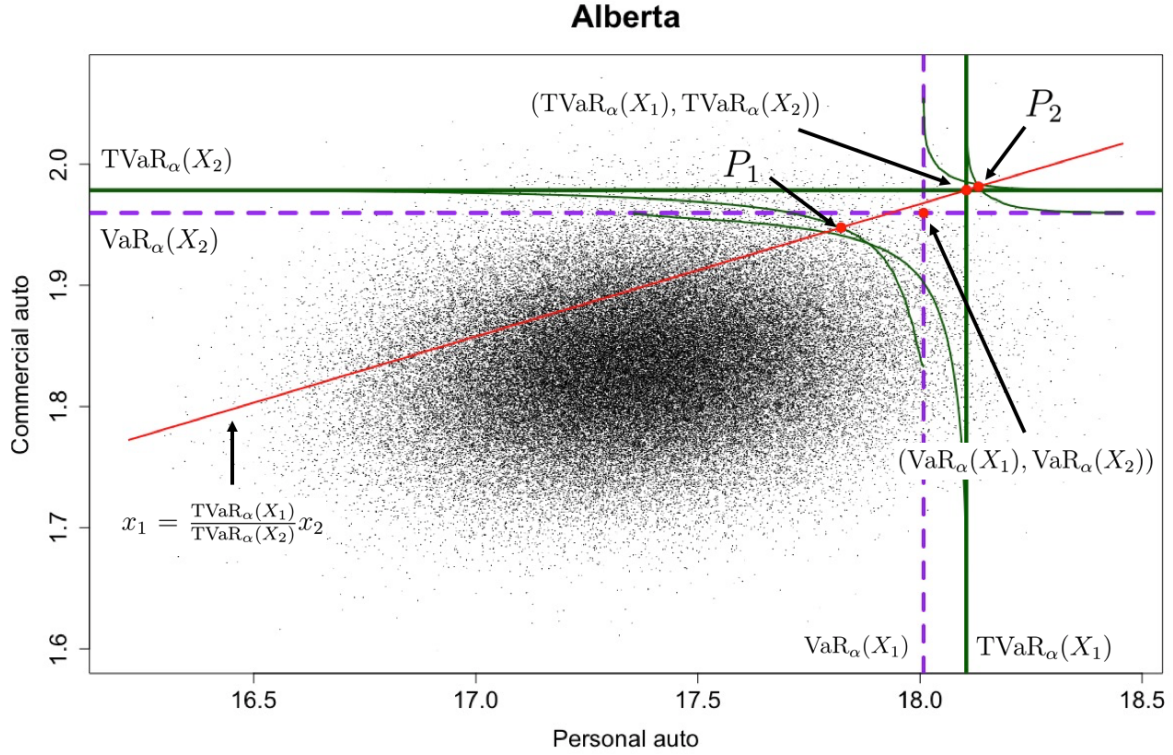


Figure 5.12: 100,000 Simulations of the reserves for the province of Alberta under the HCM. Univariate VaR (purple dotted lines) and TVaR (green straight lines) along with bivariate lower and upper orthant TVaR (green curved lines). Optimal couples shown with red dots.

From Table 5.8, any combination of optimal couples can be chosen as risk measures. For example, the traditional modeling scheme is to select either the univariate VaR or TVaR for each province which would ignore the dependence when measuring the risk associated with the joint loss distribution. In what follows, an example is given for each province with different couples to explain the possible justifications for selecting bivariate risk measures. Nonetheless, we recommend to the appointed actuary to select a homogeneous approach across the insurance portfolio to keep the capital selection as objective as possible. For Ontario, we could select the orthogonal couple of the bivariate upper TVaR because the insurer wants to protect the joint survival probability of both lines simultaneously exceeding the threshold with probability  $1 - \alpha$  while accounting for the size of the personal auto line which is the largest for the whole portfolio through the

orthogonal approach. For Alberta, we could select the proportional optimal couple for the bivariate lower orthant TVaR ( $P_2$  in Figure 5.12) due to the strong dependence observed for this province and we want to protect proportionally the joint probability of both lines not exceeding the threshold while accounting for tail events. For Atlantic Canada, we could select the univariate VaR given that the dependence and tail events (due to independence copula) are not as significant as to be considered for the risk measurements. The combination of these three couples makes a total of \$4.3 billion (CAD) corresponding to a smaller adjustment of risk compared to the univariate TVaR as seen in Table 5.7. The risk adjustment calculated with the optimal couple using the orthogonal method for the upper VaR in Table 5.8 is negative due to the fact that personal auto in Alberta is more than six times the volume of commercial auto. This difference in proportions is causing the orthogonal method to put more weight to the personal auto line in the optimization problem seen in equation (4.1.14). Although covering both lines of business is useful from a probability of ruin perspective, capital can not be negative, thus, we discard this specific couple from the possible selections of capital.

Risk measure	Optimal couple	ON	ON	AB	AB	ATL	ATL
		PA	CA	PA	CA	PA	CA
VaR	Univariate	(2.74, 0.27)		(0.65, 0.13)		(0.28, 0.06)	
<u>VaR</u>	Orthogonal	(2.79, 0.36)		(0.67, 0.16)		(0.29, 0.07)	
<u>VaR</u>	Proportional	(3.06, 0.29)		(0.74, 0.14)		(0.33, 0.06)	
$\overline{\text{VaR}}$	Orthogonal	(2.57, -0.02)		(0.58, 0.05)		(0.25, 0.00)	
$\overline{\text{VaR}}$	Proportional	(1.64, 0.19)		(0.41, 0.10)		(0.09, 0.04)	
TVaR	Univariate	(3.16, 0.31)		(0.75, 0.15)		(0.32, 0.07)	
<u>TVaR</u>	Orthogonal	(3.20, 0.39)		(0.77, 0.17)		(0.33, 0.08)	
<u>TVaR</u>	Proportional	(3.29, 0.32)		(0.77, 0.15)		(0.34, 0.07)	
$\overline{\text{TVaR}}$	Orthogonal	(2.98, 0.06)		(0.68, 0.07)		(0.29, 0.02)	
$\overline{\text{TVaR}}$	Proportional	(1.85, 0.21)		(0.46, 0.11)		(0.11, 0.05)	

Table 5.8: Risk adjustments calculated through bivariate risk measures in billions (CAD) under different risk assumptions at a confidence level  $\alpha = 99\%$  for the HCM.

As an alternative from the quantile methods, in the following we use the cost of capital method presented in Section 4.2 to calculate the risk adjustment for non-financial risks. To use the cost of capital method, we need assumptions for the cost of capital rate and the discount rate. In practice, the rates should vary with time but in this thesis, the analysis is over the whole automobile industry for three provinces in Canada instead of a single entity. Therefore, the cost of capital rate is entity specific and for demonstration purposes we assume a constant cost of capital rate of  $r_t = 8\%$ . Moreover, modeling the discount rate is out of the scope of this research and thus, we assume a constant discount rate of  $d_t = 2\%$ . We used the cost of capital method with the risk measure  $\varrho$  being the univariate TVaR at a confidence level  $\alpha = 99\%$  to account for tail events.

Period	$\mathbb{E}[X]$	$\text{TVaR}_\alpha$	$C_t$	Cost	$\frac{\text{Cost}}{(1+d_t)^t}$	Period	$\mathbb{E}[X]$	$\text{TVaR}_\alpha$	$C_t$	Cost	$\frac{\text{Cost}}{(1+d_t)^t}$
2018-1	13.43	14.40	0.97	0.08	0.08	2025-2	0.99	1.11	0.11	0.01	0.01
2018-2	11.80	12.54	0.74	0.06	0.06	2026-1	0.83	0.93	0.10	0.01	0.01
2019-1	10.52	11.03	0.51	0.04	0.04	2026-2	0.70	0.79	0.10	0.01	0.01
2019-2	9.33	9.71	0.38	0.03	0.03	2027-1	0.59	0.67	0.09	0.01	0.00
2020-1	8.12	8.43	0.31	0.02	0.02	2027-2	0.49	0.57	0.08	0.01	0.00
2020-2	6.94	7.20	0.26	0.02	0.02	2028-1	0.41	0.48	0.08	0.01	0.00
2021-1	5.83	6.05	0.22	0.02	0.02	2028-2	0.34	0.41	0.07	0.01	0.00
2021-2	4.83	5.03	0.20	0.02	0.01	2029-1	0.28	0.34	0.06	0.01	0.00
2022-1	3.96	4.15	0.20	0.02	0.01	2029-2	0.22	0.27	0.05	0.00	0.00
2022-2	3.22	3.42	0.19	0.02	0.01	2030-1	0.17	0.22	0.05	0.00	0.00
2023-1	2.62	2.81	0.19	0.02	0.01	2030-2	0.13	0.17	0.04	0.00	0.00
2023-2	2.14	2.32	0.18	0.01	0.01	2031-1	0.09	0.12	0.03	0.00	0.00
2024-1	1.75	1.91	0.16	0.01	0.01	2031-2	0.06	0.09	0.03	0.00	0.00
2024-2	1.44	1.59	0.15	0.01	0.01	2032-1	0.03	0.06	0.03	0.00	0.00
2025-1	1.19	1.33	0.14	0.01	0.01						

Table 5.9: Cost of capital method disclosed by accident year in billions (CAD) for personal auto in Ontario assuming a cost of capital rate  $r_t = 8\%$  and discount rate  $d_t = 2\%$ .

In Table 5.9 the procedure to calculate the risk adjustment is shown for the personal auto line of Ontario. The risk adjustment is respectively the sum of the discounted costs which totals to \$395 million (CAD). The risk adjustment for non-financial risks using this method is shown for all lines of business in Table 5.10 along with a comparison when the cost of capital rate is  $r_t = 10\%$  and with undiscounted cost of capital, i.e.  $d_t = 0\%$ , to make the data comparable with the risk measures from Table 5.8.

	Province	Line	Risk adjustment	Equivalent $\alpha$	Risk adjustment	Equivalent $\alpha$
			$r_t = 8\%$		$r_t = 10\%$	
$d_t = 2\%$	ON	PA	395	64.74%	493	68.18%
	ON	CA	50	69.28%	62	73.61%
	AB	PA	106	66.30%	132	70.06%
	AB	CA	24	69.62%	30	73.88%
	ATL	PA	45	66.82%	57	70.67%
	ATL	CA	11	68.89%	13	73.06%
$d_t = 0\%$	ON	PA	457	65.49%	571	69.05%
	ON	CA	60	69.96%	75	74.33%
	AB	PA	122	67.01%	153	70.82%
	AB	CA	29	70.61%	36	74.97%
	ATL	PA	54	67.41%	68	71.26%
	ATL	CA	13	69.67%	16	73.82%

Table 5.10: Risk adjustment for non-financial risks displayed in millions (CAD) for the six lines of business and equivalent confidence level for the VaR with two different assumptions for the cost of capital rate  $r_t$  and discount rate  $d_t$ .

In Table 5.10 we disclose the confidence level  $\alpha$  equivalent to calculating the risk adjustment through the  $\text{VaR}_\alpha$  as specified by IFRS 17 described in Section 1.1.2.5. In this specific example, the TVaR at a 99% confidence level is used with two different assumptions over the cost of capital rate and the discount rate which resulted in a smaller risk adjustment than using risk measures (see Table 5.7 and Table 5.8). In Canada, if the risk adjustment is computed through the cost of capital method, the insurance entity has to make sure the total amount of the liability does not fall below the supervisory target capital requirement set by OSFI, discussed in Section 1.2. Modifying the assumptions accordingly to the entity's view of risk might result in a larger risk adjustment than with the risk measures but the important conclusion is drawn from understanding the difference between these methods. The cost of capital method takes into account the specific cost of



capital for the insurance entity which can vary across companies and as shown in Table 5.10, a 2% increase in the cost of capital rate can increase the risk adjustment significantly. Furthermore, using risk measures as done in Table 5.8 considers the possible tail events in each province directly through the measurement of risk and depending on the couple selected, additional financial security can be obtained for the insurance portfolio.

# Conclusion

In this thesis, we developed a model for the assessment of non-financial risks of a P&C insurer complaint with the new International Financial Reporting Standards (IFRS 17). A numerical application is presented with automobile data for three provinces from the Canadian insurance industry which simultaneously has to comply with the regulator's (OSFI) capital requirements.

The Tweedie family provides a clever method to fit aggregate insurance claims even when the number of claims is not recorded or is unreliable and moreover, through a transformation of the parameters it is linked to the classical Compound Poisson-Gamma distribution. Generalized linear modeling for the marginal effects is a common practice in the industry, hence, the same approach is used in this thesis. Moreover, in scenarios where the volatility of the data is significant or due to long development of claims, a double generalized linear model is a powerful tool to solve the problem. Long development of claims could carry along correlated data which left unchecked, would change the ranks of the residuals and consequently, the dependence analysis. Thus, the autoregressive lag through generalized estimating equations just adds one parameter to be estimated and solves the problem in most scenarios.

IFRS 17 seeks to globally homogenize financial reports for insurance companies making of the utmost importance to spread the possible techniques currently available to measure the risk adjustment for non-financial risks required to measure an insurance portfolio while complying with the risk adjustment characteristics and objective. In this thesis, the diversification benefit is obtained with a semi-parametric HCM and through multivariate

risk measures when aggregation is deemed not appropriate. Although there exist several possible ways to account for dependence in an insurance portfolio, the bivariate approach simplifies understanding through visual representations of the joint distribution and its risk measures.

# Bibliography

- Abdallah, A., Boucher, J.-P., and Cossette, H. (2015). Modeling dependence between loss triangles with hierarchical archimedean copulas. *Astin Bulletin*, 45(3):1–23.
- Andersen, D. A. and Bonat, W. H. (2017). Double generalized linear compound poisson models to insurance claims data. *Electronic Journal of Applied Statistical Analysis*, 10 (2):384–407.
- Arbenz, P., Hummel, C., and Mainik, G. (2012). Copula based hierarchical risk aggregation through sample reordering. *Insurance: Mathematics and Economics*, 51:122–133.
- Artzner, P., Delbaen, F., Eber, J.-M., and Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, 9 (3):203–228.
- ASB (2018). Standards of practice. <http://www.cia-ica.ca/docs/default-source/standards/sc010118e.pdf>.
- Boucher, J.-P. and Davidov, D. (2011). On the importance of dispersion modeling for claims reserving: An application with the Tweedie distribution. *Casualty Actuarial Society*, 5 (2):158–172.
- Burgi, R., Dacorogna, M., and Iles, R. (2008). Risk aggregation, dependence structure and diversification benefit. Stress testing for financial institutions. <http://ssrn.com/abstract=1468526>.
- CIA (2018). Comparison of ifrs 17 to current cia standards of practice (draft educational note). <http://www.cia-ica.ca/docs/default-source/2018/218117e.pdf>.

- CIA-CAS (2018). Introduction to IFRS 17 for P&C Actuaries [Webinar]. <https://bit.ly/2TFVpDP>.
- Cossette, H., Mailhot, M., Étienne Marceau, and Mesfioui, M. (2013). Bivariate lower and upper orthant value-at-risk. *European Actuarial Journal*, 3(2):321–357.
- Cossette, H., Mailhot, M., Étienne Marceau, and Mesfioui, M. (2016). Vector-valued tail value-at-risk and capital allocation. *Methodology and Computing in Applied Probability*, 18(3):653–674.
- Côté, M. P. (2014). Copula-based risk aggregation modelling. Master’s thesis, McGill University, Quebec, Canada.
- Côté, M. P., Genest, C., and Abdallah, A. (2016). Rank-based methods for modeling dependence between loss triangles. *European Actuarial Journal*, 6(2):377–408.
- Cox, D. and Reid, N. (1987). Parameter orthogonality and approximate conditional inference. *Royal Statistical Society*, 49 (1):1–39.
- Deheuvels, P. (1979). La fonction de dépendance empirique et ses propriétés. un test non paramétrique. *Bulletin Royal Belge de l’Académie des Sciences*, 65:274–292.
- Derendinger, F. (2015). Copula-based hierarchical risk aggregation. Master’s thesis, Swiss Federal Institute of Technology Zurich, Zurich, Switzerland.
- Devlin, R. A. (2017). A comparison of automobile insurance regimes in canada. <https://bit.ly/2RAwuo5>.
- Dunn, P. K. and Smyth, G. K. (2004). Series evaluation of tweedie exponential dispersion model densities. *Statistics and Computing*, 15(4):267–280.
- Embrechts, P., Lindskog, F., and J. Mcneil, E. (2001). Modelling dependence with copulas and applications to risk management. *Handbook of Heavy Tailed Distributions in Finance*, 8.

- Embrechts, P. and Puccetti, G. (1996). Bounds for functions of multivariate risks. *Journal of Multivariate Analysis*, 97(2):526–547.
- Genest, C., Ghoudi, K., and Rivest, L.-P. (1995). A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika*, 82(3):543–552.
- Genest, C. and Rémillard, B. (2004). Tests of independence and randomness based on the empirical copula process. *Sociedad de Estadística e Investigación Operativa*, 13(2):335–369.
- Genest, C. and Rémillard, B. (2008). Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models. *Annales de l’Institut Henri Poincaré*, 44(6):1096–1127.
- Genest, C., Rémillard, B., and Beaudoin, D. (2009). Goodness-of-fit tests for copulas: A review and a power study. *Insurance: Mathematics and Economics*, 44:199–213.
- Government of Canada (c. 47, S.C. 1991). *Insurance Companies Act*. Minister of Justice. <https://laws-lois.justice.gc.ca/PDF/I-11.8.pdf>.
- Grønneberg, S. and Hjort, N. L. (2014). The copula information criteria. *Scandinavian Journal of Statistics*, 41(2).
- Hardin, J. W. and Hilbe, J. M. (2013). *Generalized Estimating Equations*. Chapman and Hall.
- Hin, L.-Y. and Wang, Y.-G. (2009). Working-correlation-structure identification in generalized estimating equations. *Statistics in Medicine*, 28:642–658.
- Hofert, M. and Pham, D. (2013). Densities of nested archimedean copulas. *Journal of Multivariate Analysis*, pages 37–52.
- Hudecová, S. and Pešta, M. (2013). Modeling dependencies in claims reserving with gee. *Insurance: Mathematics and Economics*, 53:786–794.

- IAA (2018). *Risk Adjustments for Insurance Contracts under IFRS 17*. Canada.
- IASB (2017a). *Basis for Conclusions on IFRS 17 Insurance Contracts*. IFRS Foundation. <https://bit.ly/2XEKfnW>.
- IASB (2017b). *IFRS 17 Insurance Contracts*. IFRS Foundation. <https://bit.ly/2X1XFG0>.
- Iman, R. L. and Conover, W. J. (1982). A distribution-free approach to inducing rank correlation among input variables. *Communications in Statistics - Simulation and Computation*, 11 (3):311–334.
- Joe, H. (1997). *Multivariate Models and Dependence Concepts*. Chapman Hall, London.
- Jørgensen, B. (1987). Exponential dispersion models. *Royal Statistical Society*, 49 (2):127–162.
- Jørgensen, B. (1997). *The Theory of Dispersion Models*. CRC Press.
- Klugman, S. A., Panjer, H. H., and Willmot, G. E. (2008). *Loss Model: From Data to Decisions Third Edition*. Wiley, USA.
- Lee, Y. and Nelder, J. (1998). Generalized linear models for the analysis of quality-improvement experiments. *The Canadian Journal of Statistics*, 26 (1):95–105.
- Liang, K.-Y. and Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, 73 (1):13–22.
- 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.
- Madsen, H. and Thyregod, P. (2010). *Introduction to General and Generalized Linear Models*. CRC Press.
- Mailhot, M. and Mesfioui, M. (2016). Multivariate tvar-based risk decomposition for vector-valued portfolios. *Risks*, 4(4):1–16.

- Mardia, K. V. (1967). Contributions to contingency-type bivariate distributions. *Biometrika*, 54(1):235–249.
- McCullagh, P. (1984). Generalized linear models. *European Journal of Operational Research*, 16(3):285–292.
- McNeil, A. J., Frey, R., and Embrechts, P. (2005). *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton University Press, New Jersey.
- Nelsen, R. B. (2006). *An Introduction to Copulas*. Springer, second edition.
- Oakes, D. (1994). Multivariate survival distributions. *Journal of Nonparametric Statistics*, 3:343–354.
- OSFI (2017). *Property and Casualty Memorandum to the Appointed Actuary*. Canada. [http://www.osfi-bsif.gc.ca/Eng/Docs/PC\\_AA\\_Memo\\_2017.pdf](http://www.osfi-bsif.gc.ca/Eng/Docs/PC_AA_Memo_2017.pdf).
- OSFI (2018a). *IFRS 17 Transition and Progress Report Requirements for Federally Regulated Insurers*. Canada. <http://www.osfi-bsif.gc.ca/Eng/Docs/ifrs17.pdf>.
- OSFI (2018b). *Minimum Capital Test*. Canada. <http://www.osfi-bsif.gc.ca/Eng/Docs/mct2018.pdf>.
- OSFI (2018c). *Own Risk and Solvency Assessment (OSRA)*. Canada. <http://www.osfi-bsif.gc.ca/Eng/Docs/e1918.pdf>.
- OSFI (2018d). *Regulatory Capital and Internal Capital Targets*. Canada. [http://www.osfi-bsif.gc.ca/Eng/Docs/a4\\_gd18.pdf](http://www.osfi-bsif.gc.ca/Eng/Docs/a4_gd18.pdf).
- Plackett, R. L. (1965). A class of bivariate distributions. *Journal of the American Statistical Association*, 60(310):516–522.
- Quijano Xacur, O. A. (2011). Property and casualty premiums based on tweedie families of generalized linear models. Master’s thesis, Concordia University, Quebec, Canada.
- Rémillard, B. and Scaillet, O. (2009). Testing for equality between two copulas. *Journal of Multivariate Analysis*, 100:377–386.



- Shi, P. and Frees, E. W. (2011). Dependent loss reserving using copulas. *Astin Bulletin*, 41(2):449–486.
- Sklar, A. (1959). Fonctions de répartition à  $n$  dimensions et leurs marges. *Publications de l'Institut de Statistique de l'Université de Paris*, 8:229–231.
- Smolárová, T. (2017). Tweedie models for pricing and reserving. Master's thesis, Charles University, Prague, Czech Republic.
- Smyth, G. K. (1989). Generalized linear models with varying dispersion. *Journal of the royal Statistical Society*, 51(1):47–60.
- Smyth, G. K. and Jørgensen, B. (2002). Fitting Tweedie's compound poisson model to insurance claims data: Dispersion modelling. *Astin Bulletin*, 32 (1):143–157.
- Smyth, G. K. and Verbyla, A. P. (1999). Adjusted likelihood methods for modelling dispersion in generalized linear models. *Environmetrics*, 10:695–709.
- Swan, T. (2006). Generalized estimating equations when the response variable has a tweedie distribution: An application for multi-site rainfall modelling. Master's thesis, The University of Southern Queensland, Queensland, Australia.
- Tasche, D. (2000). Conditional expectation as quantile derivative. *Preprint. TU München*. <https://arxiv.org/pdf/math/0104190.pdf>.
- Tasche, D. (2008). Capital allocation to business units and sub-portfolios: the euler principle. *Quantitative Finance Papers*, page 423453.
- Tweedie, M. (1984). An index which distinguishes between some important exponential families. *Statistics: Applications and New Directions. Proceedings of the Indian Statistical Institute Golden Jubilee International Conference (Eds. J. K. Ghosh and J. Roy)*, pages 579–604.
- Young, G. and Smith, R. (2005). *Essentials of Statistical Inference*. Cambridge University Press.

# Appendix A

## Estimated marginal parameters

The standard deviation of the parameters is shown in parenthesis.

No.	Parameter	PA ON	PA ON	CA ON	CA ON	PA AB	PA AB	CA AB	CA AB	PA ATL	PA ATL	CA ATL	CA ATL
1	Intercept	-1.55	(0.00)	-1.80	(0.00)	-1.05	(0.01)	-1.12	(0.01)	-1.40	(0.01)	-1.48	(0.01)
2	AY = 2003-2	-0.13	(0.00)	0.03	(0.00)	0.01	(0.00)	-0.14	(0.00)	-0.10	(0.00)	-0.07	(0.00)
3	AY = 2004-1	-0.29	(0.00)	-0.24	(0.00)	-0.21	(0.00)	-0.36	(0.00)	-0.27	(0.00)	-0.29	(0.00)
4	AY = 2004-2	-0.10	(0.00)	-0.27	(0.00)	-0.12	(0.00)	-0.10	(0.00)	-0.09	(0.00)	-0.07	(0.00)
5	AY = 2005-1	-0.23	(0.00)	-0.40	(0.00)	-0.21	(0.00)	-0.42	(0.00)	-0.10	(0.00)	-0.48	(0.00)
6	AY = 2005-2	-0.06	(0.00)	0.04	(0.00)	-0.12	(0.00)	-0.24	(0.00)	0.05	(0.00)	-0.07	(0.00)
7	AY = 2006-1	-0.11	(0.00)	-0.23	(0.00)	-0.26	(0.00)	-0.37	(0.00)	-0.17	(0.00)	-0.34	(0.00)
8	AY = 2006-2	0.09	(0.00)	0.00	(0.00)	-0.01	(0.00)	-0.07	(0.00)	0.01	(0.00)	-0.38	(0.00)
9	AY = 2007-1	0.02	(0.00)	-0.16	(0.00)	-0.31	(0.00)	-0.42	(0.00)	-0.19	(0.00)	-0.54	(0.00)
10	AY = 2007-2	0.08	(0.00)	0.10	(0.00)	-0.11	(0.00)	-0.19	(0.00)	-0.02	(0.00)	-0.33	(0.00)
11	AY = 2008-1	-0.02	(0.00)	-0.07	(0.00)	-0.25	(0.00)	-0.45	(0.00)	-0.23	(0.00)	-0.47	(0.00)
12	AY = 2008-2	0.09	(0.00)	0.36	(0.00)	-0.13	(0.00)	-0.29	(0.00)	-0.25	(0.00)	-0.39	(0.00)
13	AY = 2009-1	0.06	(0.00)	0.18	(0.00)	-0.38	(0.00)	-0.84	(0.00)	-0.21	(0.00)	-0.59	(0.00)
14	AY = 2009-2	0.27	(0.00)	0.16	(0.00)	-0.20	(0.00)	-0.51	(0.00)	0.01	(0.00)	-0.54	(0.00)
15	AY = 2010-1	0.16	(0.00)	0.03	(0.00)	-0.53	(0.00)	-0.61	(0.00)	-0.10	(0.00)	-0.57	(0.00)
16	AY = 2010-2	0.15	(0.00)	0.09	(0.00)	-0.23	(0.00)	-0.66	(0.00)	0.01	(0.00)	-0.12	(0.00)
17	AY = 2011-1	-0.08	(0.00)	-0.12	(0.00)	-0.44	(0.00)	-0.60	(0.00)	-0.18	(0.00)	-0.64	(0.00)
18	AY = 2011-2	0.00	(0.00)	0.00	(0.00)	-0.18	(0.00)	-0.34	(0.00)	0.01	(0.00)	-0.06	(0.00)
19	AY = 2012-1	-0.13	(0.00)	-0.06	(0.00)	-0.29	(0.00)	-0.63	(0.00)	-0.16	(0.00)	-0.73	(0.00)
20	AY = 2012-2	-0.03	(0.00)	0.03	(0.00)	-0.09	(0.00)	-0.21	(0.00)	0.13	(0.00)	-0.45	(0.00)
21	AY = 2013-1	-0.13	(0.00)	-0.14	(0.00)	-0.26	(0.00)	-0.24	(0.00)	-0.15	(0.00)	-0.36	(0.00)
22	AY = 2013-2	0.06	(0.00)	0.02	(0.00)	-0.02	(0.00)	-0.30	(0.00)	0.17	(0.00)	0.00	(0.00)
23	AY = 2014-1	-0.10	(0.00)	-0.06	(0.00)	-0.25	(0.00)	-0.63	(0.00)	-0.09	(0.00)	-0.07	(0.00)
24	AY = 2014-2	0.07	(0.01)	0.18	(0.00)	0.06	(0.00)	-0.25	(0.00)	0.07	(0.00)	-0.03	(0.00)
25	AY = 2015-1	-0.03	(0.01)	-0.09	(0.00)	-0.13	(0.01)	-0.48	(0.00)	0.05	(0.00)	-0.17	(0.00)
26	AY = 2015-2	0.13	(0.01)	0.12	(0.01)	0.05	(0.01)	-0.43	(0.01)	0.31	(0.00)	-0.05	(0.01)
27	AY = 2016-1	-0.02	(0.01)	-0.09	(0.01)	-0.18	(0.01)	-0.66	(0.01)	0.09	(0.00)	-0.31	(0.01)
28	AY = 2016-2	0.09	(0.02)	0.02	(0.01)	0.02	(0.01)	-0.33	(0.01)	0.16	(0.01)	-0.11	(0.01)
29	AY = 2017-1	-0.12	(0.02)	-0.15	(0.01)	-0.27	(0.02)	-0.47	(0.02)	0.01	(0.01)	-0.12	(0.02)
30	AY = 2017-2	0.15	(0.02)	0.13	(0.02)	-0.06	(0.05)	-0.17	(0.09)	0.20	(0.05)	0.03	(0.04)

Table A.1: Mean model - Accident year effects.

No.	Parameter	PA ON	PA ON	CA ON	CA ON	PA AB	PA AB	CA AB	CA AB	PA ATL	PA ATL	CA ATL	CA ATL
31	DL = 2	-0.33	(0.00)	-0.24	(0.00)	-0.63	(0.01)	-0.46	(0.01)	-0.60	(0.01)	-0.47	(0.00)
32	DL = 3	-0.55	(0.00)	-0.44	(0.00)	-1.21	(0.01)	-1.14	(0.01)	-0.90	(0.01)	-0.81	(0.01)
33	DL = 4	-0.61	(0.00)	-0.43	(0.00)	-1.36	(0.01)	-1.30	(0.01)	-0.98	(0.01)	-0.87	(0.01)
34	DL = 5	-0.61	(0.00)	-0.37	(0.00)	-1.43	(0.01)	-1.37	(0.01)	-1.05	(0.01)	-0.94	(0.01)
35	DL = 6	-0.65	(0.00)	-0.38	(0.00)	-1.51	(0.01)	-1.49	(0.01)	-1.15	(0.01)	-0.99	(0.01)
36	DL = 7	-0.71	(0.00)	-0.43	(0.00)	-1.56	(0.01)	-1.55	(0.01)	-1.24	(0.01)	-1.07	(0.01)
37	DL = 8	-0.82	(0.00)	-0.50	(0.00)	-1.66	(0.01)	-1.66	(0.01)	-1.35	(0.01)	-1.20	(0.01)
38	DL = 9	-0.97	(0.00)	-0.63	(0.00)	-1.77	(0.01)	-1.78	(0.01)	-1.49	(0.01)	-1.30	(0.01)
39	DL = 10	-1.14	(0.00)	-0.78	(0.00)	-1.90	(0.01)	-1.94	(0.01)	-1.67	(0.01)	-1.46	(0.01)
40	DL = 11	-1.34	(0.00)	-0.96	(0.00)	-2.08	(0.01)	-2.14	(0.01)	-1.84	(0.01)	-1.62	(0.01)
41	DL = 12	-1.56	(0.00)	-1.15	(0.00)	-2.25	(0.01)	-2.37	(0.01)	-2.02	(0.01)	-1.75	(0.01)
42	DL = 13	-1.78	(0.00)	-1.42	(0.00)	-2.46	(0.01)	-2.61	(0.01)	-2.22	(0.01)	-1.91	(0.01)
43	DL = 14	-2.02	(0.00)	-1.68	(0.00)	-2.68	(0.01)	-2.79	(0.01)	-2.42	(0.01)	-2.14	(0.01)
44	DL = 15	-2.25	(0.00)	-1.91	(0.00)	-2.89	(0.01)	-3.01	(0.01)	-2.62	(0.01)	-2.21	(0.01)
45	DL = 16	-2.45	(0.00)	-2.11	(0.00)	-3.15	(0.01)	-3.24	(0.01)	-2.82	(0.01)	-2.40	(0.01)
46	DL = 17	-2.64	(0.00)	-2.34	(0.00)	-3.39	(0.01)	-3.39	(0.01)	-3.05	(0.01)	-2.70	(0.01)
47	DL = 18	-2.84	(0.00)	-2.60	(0.00)	-3.58	(0.01)	-3.83	(0.01)	-3.25	(0.01)	-3.14	(0.01)
48	DL = 19	-2.99	(0.00)	-2.77	(0.00)	-3.76	(0.01)	-4.03	(0.01)	-3.49	(0.01)	-3.39	(0.01)
49	DL = 20	-3.15	(0.00)	-2.79	(0.00)	-3.98	(0.01)	-4.23	(0.01)	-3.63	(0.01)	-3.45	(0.01)
50	DL = 21	-3.33	(0.00)	-3.03	(0.00)	-4.22	(0.01)	-4.42	(0.01)	-3.74	(0.01)	-3.91	(0.01)
51	DL = 22	-3.48	(0.00)	-3.25	(0.00)	-4.51	(0.01)	-4.74	(0.01)	-3.95	(0.01)	-4.11	(0.01)
52	DL = 23	-3.63	(0.00)	-3.46	(0.00)	-4.70	(0.01)	-4.89	(0.01)	-4.08	(0.01)	-4.15	(0.01)
53	DL = 24	-3.74	(0.00)	-3.72	(0.00)	-5.10	(0.01)	-5.05	(0.01)	-4.34	(0.01)	-4.67	(0.01)
54	DL = 25	-3.88	(0.00)	-3.78	(0.00)	-5.37	(0.01)	-4.96	(0.01)	-4.69	(0.01)	-5.07	(0.01)
55	DL = 26	-4.03	(0.00)	-3.88	(0.00)	-5.82	(0.01)	-5.22	(0.01)	-4.74	(0.01)	-4.95	(0.01)
56	DL = 27	-4.15	(0.00)	-4.58	(0.00)	-5.83	(0.01)	-5.28	(0.01)	-5.08	(0.01)	-6.56	(0.01)
57	DL = 28	-4.25	(0.00)	-4.46	(0.00)	-5.84	(0.01)	-9.87	(0.01)	-5.09	(0.01)	-6.80	(0.01)
58	DL = 29	-4.23	(0.00)	-4.55	(0.00)	-6.09	(0.01)	-13.40	(0.01)	-5.62	(0.01)	-5.73	(0.01)
59	DL = 30	-4.57	(0.00)	-4.67	(0.00)	-6.16	(0.01)	-13.48	(0.01)	-5.72	(0.01)	-12.12	(0.01)

Table A.2: Mean model - Development Lag effects.

No.	Parameter	PA ON	PA ON	CA ON	CA ON	PA AB	PA AB	CA AB	CA AB	PA ATL	PA ATL	CA ATL	CA ATL
60	Intercept	-4.80	(0.28)	-5.78	(0.27)	-4.29	(0.27)	-2.94	(0.27)	-4.53	(0.27)	-4.80	(0.28)
61	DL = 2	0.56	(0.39)	-1.36	(0.39)	-0.89	(0.39)	-2.08	(0.39)	-2.05	(0.39)	-0.68	(0.39)
62	DL = 3	0.58	(0.39)	-1.75	(0.40)	-1.53	(0.39)	-2.79	(0.40)	-3.11	(0.39)	-1.05	(0.40)
63	DL = 4	-0.16	(0.40)	-1.72	(0.40)	-1.32	(0.39)	-2.46	(0.39)	-3.43	(0.40)	-1.03	(0.40)
64	DL = 5	-0.87	(0.40)	-1.81	(0.41)	-1.55	(0.40)	-2.79	(0.40)	-4.14	(0.41)	-2.76	(0.44)
65	DL = 6	-1.33	(0.41)	-2.02	(0.41)	-2.00	(0.40)	-3.01	(0.40)	-4.71	(0.42)	-2.34	(0.42)
66	DL = 7	-1.65	(0.41)	-2.20	(0.42)	-3.15	(0.42)	-4.35	(0.44)	-4.91	(0.43)	-2.19	(0.42)
67	DL = 8	-2.47	(0.43)	-2.46	(0.43)	-3.53	(0.44)	-3.92	(0.43)	-4.68	(0.43)	-1.87	(0.42)
68	DL = 9	-3.23	(0.46)	-1.76	(0.42)	-3.36	(0.43)	-3.44	(0.42)	-4.05	(0.42)	-1.19	(0.42)
69	DL = 10	-2.88	(0.45)	-1.72	(0.43)	-2.86	(0.43)	-2.53	(0.42)	-3.26	(0.42)	-1.00	(0.42)
70	DL = 11	-2.21	(0.44)	-1.54	(0.43)	-2.72	(0.43)	-2.40	(0.43)	-2.94	(0.42)	-0.73	(0.43)
71	DL = 12	-1.50	(0.44)	-0.55	(0.43)	-2.52	(0.44)	-2.07	(0.43)	-2.46	(0.43)	-0.38	(0.43)
72	DL = 13	-0.77	(0.44)	-0.35	(0.44)	-2.03	(0.44)	-1.72	(0.44)	-2.56	(0.44)	-0.50	(0.44)
73	DL = 14	-0.43	(0.45)	-0.48	(0.45)	-1.58	(0.45)	-1.40	(0.45)	-2.04	(0.44)	0.00	(0.45)
74	DL = 15	-0.11	(0.46)	-0.66	(0.46)	-1.83	(0.46)	-1.29	(0.45)	-2.01	(0.45)	0.28	(0.46)
75	DL = 16	0.36	(0.47)	-0.23	(0.47)	-1.89	(0.47)	-0.98	(0.46)	-1.93	(0.46)	-0.07	(0.47)
76	DL = 17	0.04	(0.48)	0.13	(0.48)	-1.45	(0.48)	-0.89	(0.48)	-1.87	(0.48)	-0.20	(0.48)
77	DL = 18	-0.08	(0.50)	0.64	(0.49)	-1.26	(0.49)	-1.53	(0.49)	-1.82	(0.49)	-0.47	(0.49)
78	DL = 19	0.22	(0.51)	0.43	(0.51)	-1.46	(0.51)	-0.97	(0.50)	-2.27	(0.50)	-0.33	(0.51)
79	DL = 20	-0.03	(0.53)	0.81	(0.52)	-1.50	(0.53)	-0.66	(0.52)	-2.42	(0.52)	0.06	(0.53)
80	DL = 21	0.48	(0.55)	0.16	(0.54)	-1.05	(0.55)	-0.65	(0.54)	-2.51	(0.54)	0.27	(0.55)
81	DL = 22	0.47	(0.57)	0.49	(0.57)	-0.70	(0.57)	-0.36	(0.57)	-2.41	(0.57)	0.82	(0.57)
82	DL = 23	1.06	(0.60)	0.66	(0.60)	-0.50	(0.60)	-0.50	(0.60)	-2.30	(0.60)	0.95	(0.60)
83	DL = 24	1.25	(0.64)	0.98	(0.64)	-0.24	(0.64)	-0.23	(0.64)	-1.45	(0.64)	0.64	(0.64)
84	DL = 25	0.80	(0.69)	0.51	(0.69)	-0.46	(0.69)	-0.07	(0.69)	-1.43	(0.69)	0.25	(0.69)
85	DL = 26	1.39	(0.76)	0.01	(0.76)	-1.05	(0.76)	0.08	(0.76)	-1.92	(0.76)	0.66	(0.76)
86	DL = 27	1.33	(0.86)	-1.81	(0.86)	-1.27	(0.86)	0.60	(0.86)	-2.21	(0.86)	-0.95	(0.86)
87	DL = 28	0.27	(1.04)	-3.19	(1.04)	-1.56	(1.04)	-0.97	(1.04)	-2.57	(1.04)	-0.64	(1.04)
88	DL = 29	0.71	(1.44)	-3.10	(1.45)	-4.10	(1.47)	-2.46	(1.44)	-5.21	(1.45)	0.22	(1.44)
89	DL = 30	1.80	(6.33)	-8.96	(6.33)	-8.67	(6.33)	0.02	(6.33)	-9.00	(6.33)	-2.17	(6.33)

Table A.3: Dispersion submodel - Development lag effects.

Correlation	PA ON	CA ON	PA AB	CA AB	PA ATL	CA ATL
$\rho$	0.80	0.67	0.72	0.68	0.75	0.69

Table A.4: Correlation parameter  $\rho$  estimated using GEE.

Index parameter	PA ON	CA ON	PA AB	CA AB	PA ATL	CA ATL
$p$	1.900	1.200	1.500	1.500	1.215	1.200

Table A.5: Index parameter  $p$  for the Tweedie distribution.

# Appendix B

## Other copula families

### B.1 Archimedean Copulas

One of the main advantages of this copula family is that Archimedean copulas can be written in closed-form. In contrast with elliptical copulas, Archimedean copulas are able to capture lower and upper tail dependencies and not just radial symmetry [McNeil et al. \(2005\)](#). They allow for positive or negative association, one of the primary reasons for its popularity in applications in insurance, finance and medical statistics [Shi and Frees \(2011\)](#). Clayton, Frank and Gumbel-Hougaard copulas are famous examples of Archimedean copulas.

**Definition B.1.1.** *Archimedean copulas [Nelsen \(2006\)](#). Let the copula generator  $\varphi$  be a continuous, strictly decreasing function  $\varphi : [0, 1] \rightarrow [0, \infty]$ , such that  $\varphi(1) = 0$ , then  $C$  is an Archimedean copula if and only if  $\varphi$  is convex. The copula  $C$  has the form,*

$$C(u, v) = \varphi^{-1}(\varphi(u) + \varphi(v)). \tag{B.1.1}$$

With help from equation [\(B.1.1\)](#) we can summarize the most common Archimedean copulas in a bivariate setting with their copula generator  $\varphi$ .

Copula family	Generator $\varphi_\theta(t)$	Parameter domain	Independence
Clayton	$\frac{1}{\theta}(t^{-\theta} - 1)$	$\theta \in [-1, \infty)$	$\theta \rightarrow 0$
Frank	$-\ln \left( \frac{e^{-\theta t} - 1}{e^{-\theta} - 1} \right)$	$\theta \in (-\infty, \infty)$	$\theta \rightarrow 0$
Gumbel-Hougaard	$(-\ln t)^\theta$	$\theta \in [1, \infty]$	$\theta = 1$

Table B.1: Common Archimedean copula families and some of their properties.

Table [B.1](#) is constructed from [McNeil et al. \(2005\)](#).

## B.2 Plackett copula

**Definition B.2.1.** For a random vector  $(X, Y)$ , the Plackett copula, see [Plackett \(1965\)](#), with dependence parameter  $\theta > 0$ , has the form,

$$C(u, v) = \frac{2\theta uv}{B + \sqrt{B^2 - 4uv\theta(\theta - 1)}}, \quad (\text{B.2.1})$$

where  $B = 1 + (\theta - 1)(u + v)$ .

The Plackett copula in [\(B.2.1\)](#) only exists in a bivariate setting. It was proposed by [Plackett \(1965\)](#) but the closed-form was found by [Mardia \(1967\)](#). The Plackett copula was constructed to satisfy the following properties,

- If  $\theta \rightarrow 0$ , then  $C \rightarrow W$ , the counter-monotonic copula.
- If  $\theta = 1$ , then  $C = \Pi$ , the product copula and thus, stochastic independence.
- If  $\theta \rightarrow \infty$ , then  $C \rightarrow M$ , the comonotonic copula.