TECHNISCHE UNIVERSITÄT WIEN
Vienna University of Technology

DIPLOMARBEIT

Incorporating higher order moments into the claims reserving process

ausgeführt am Institut für
Wirtschaftsmathematik
der Technischen Universität Wien

unter Anleitung von
Ao.Univ.Prof. Dipl.-Ing. Dr.techn. Friedrich Hubalek

durch

Philip Gruber

Ernst-Melchior-Gasse 11/1/27
1020 Wien

Kontakt: philip_gruber@hotmail.com

Datum	Unterschrift

Abstract

The aim of this thesis is introducing higher order statistics into the claims reserving process. After establishing the standard chain ladder model from [WM08], estimators for the skewness and kurtosis of the reserve risk distribution are developed. The derivation of these estimators is built on the work in [Mor12] and [Mor13]. The estimators are then used in a model framework to simulate the whole reserve risk distribution, which allows for the application of other statistics such as the Value at Risk.

Keywords: *skewness, kurtosis, chain ladder, claims reserving, Value at Risk, volatility*

Contents

Chapter 1

Introduction

Claims reserving is one of the most important tasks for every non-life insurance company. There are many different ways to tackle the claims reserving problem, with the most famous and probably most used one being the chain ladder model.
While this model was initially only an algorithm to get a best estimate for the reserves, the introduction of stochastic models by Mack and later Wüthrich went one step further to measure the volatility of these estimators. With the concept of the MSEP (mean squared error of prediction) applied to the claims reserving model we now have a good sense of the first two moments of the reserve risk distribution.
But knowing only these two moments leaves a lot to be desired since two distributions can coincide in both but still bear wildly different risks, especially where the tails of the distribution are concerned.
In the modern actuarial world, where these tails of the distributions gain more and more importance (the measurement of the Value at Risk being the prime example) it is important to get finer measurements for the volatility of our data.
Building on two papers by Dal Moro we will find estimators for the skewness and the kurtosis of individual elements of the claims triangle and use them in a stochastic simulation to simulate the ultimate claims.
In doing so we are not only able to estimate higher order statistics of the claims reserves from the simulation but can also compute other key figures fo the reserve risk distribution, such as the Value at Risk or the Expected Shortfall.

Chapter 2

The stochastic model [1]

2.1 General notation and the definition of the model

In this section we will design a stochastic model that allows us to describe the stochastic claims reserving process in a formal way. It is taken from [WM08], where it is presented as a theoretical foundation for the claims reserving problem. While this formalization is not absolutely necessary for practical calculations, it still serves as a good mathematical background for the following chapters.
We assume $N \in \mathbb{N}$ claims with accompanying $T_1, \ldots, T_N$ reporting dates, where $T_i \leq T_{i+1}$, $\forall i = 1, \ldots, n$. Each reporting date T_i initiates a time-series process $(T_{i,j})_{j \geq 0}$.
For i fixed but arbitrary, we define $T_i = T_{i,0} \leq T_{i,1} \leq \cdots \leq T_{i,N_i}$ as the time points, where an action for the i-th claim is observed. This can be anything, from a payment or a new reserve estimation to any other new information concerning claim i. T_{i,N_i} denotes the settlement date of the i-th claim. We set $T_{i,N_i+k} = \infty$ für $k \geq 1$.

In our stochastic model we have two stochastic processes associated with the time series $(T_{i,j})_{j \geq 0}$. At each time point $T_{i,j}$ we have:

$$X_{i,j} = \begin{cases} \text{Payment at time } T_{i,j} \text{ for the } i\text{-th claim} \\ 0, \text{ if there is no payment at time } T_{i,j} \end{cases}$$

$$I_{i,j} = \begin{cases} \text{new information at time } T_{i,j} \text{ for the } i\text{-th claim} \\ \emptyset, \text{ if there is no new information at time } T_{i,j} \end{cases}$$

For $T_{i,j} = \infty$ we set $X_{i,j} = 0$ and $I_{i,j} = \emptyset$.

[1] This chapter follows [WM08].

Now we can define the stochastic processes relevant for the claims reserving process.

- **The payment process of the i-th claim:**
  Based on $(T_{i,j}, X_{i,j})_{j\geq 0}$ we define the cumulative payment process $C_i(t)$ by

$$C_i(t) := \sum_{j\in\{k;T_{i,k}\leq t\}} X_{i,j} \tag{2.1}$$

  where $C_i(t) = 0$ for $t < T_i$. The absolute claim amount is then given by

$$C_i(\infty) := C_i(T_{i,N_i}) = \sum_{j=0}^{\infty} X_{i,j}$$

  The reserve for future payments for the i-th claim is calculated as

$$R_i(t) = C_i(\infty) - C_i(t) = \sum_{j\in\{k;T_{i,k}>t\}} X_{i,j} \tag{2.2}$$

  2.1.1 Remark. We note that $R_i(t)$ is a random variable, which we will later try to predict from the data available at time t. One possible method is to use the conditional expectation as an estimator. This estimator is called claims reserve for outstanding liabilities and its calculation (and measuring its volatility) will be the aim of the following chapters.

- **The information process of the i-th claim:**
  The information process is given by $(T_{i,j}, I_{i,j})_{j\geq 0}$.

- **The settlement process of the i-th claim:**
  The settlement process is given by $(T_{i,j}, I_{i,j}, X_{i,j})_{j\geq 0}$.

We denote the aggregated processes of all claims by

$$C(t) = \sum_{i=1}^{N} C_i(t), \text{ and} \tag{2.3}$$

$$R(t) = \sum_{i=1}^{N} R_i(t) \tag{2.4}$$

$C(t)$ denotes all payments for the N claims up to time t and $R(t)$ the sum of all future payments at time t.
Now we will consider the reserving problem as an estimation problem. First we define an information σ-algebra, which is generated by the settlement process.

$$\mathcal{F}_t^N = \sigma(\{(T_{i,j}, I_{i,j}, X_{i,j}) : 1 \leq i \leq N, j \geq 0, T_{i,j} \leq t\})$$

Quite often there is also external information, which has to be taken into account (for example a-priori estimates of the responsible actuary, some sort of "expert judgement" or external economic indices like inflation and legislative changes). We combine these external factors into one filtration $(\mathcal{E}_t)_{t\geq 0}$. The entire information available to the insurance company at time t is then given by

$$\mathcal{F}_t = \sigma(\mathcal{F}_t^N \times \mathcal{E})$$

Our estimation problem consists of estimating the conditional distribution

$$\mathbb{P}[C(\infty) \in \cdot|\mathcal{F}_t].$$

To achieve this we will find estimators for the first four (central) moments of each cumulative claim $C_i(s)$ conditional on F_{s-1} for each $s \geq 0$. Then we will use a simulation method to obtain an estimate for the conditional (on F_t) cumulative distribution function for the ultimate claim. We define the first four conditional central moments as

$$\begin{aligned}
\mu_t^{(1)}\,(C(\infty)|\mathcal{F}_t) &= \mathbb{E}[C(\infty)|\mathcal{F}_t],\\
\mu_t^{(2)}\,(C(\infty)|\mathcal{F}_t) &= \mathbb{E}[(C(\infty)-\mu_t^{(1)})^2|\mathcal{F}_t] = \mathbb{V}[C(\infty)|\mathcal{F}_t],\\
\mu_t^{(3)}\,(C(\infty)|\mathcal{F}_t) &= \mathbb{E}[(C(\infty)-\mu_t^{(1)})^3|\mathcal{F}_t],\\
\mu_t^{(4)}\,(C(\infty)|\mathcal{F}_t) &= \mathbb{E}[(C(\infty)-\mu_t^{(1)})^4|\mathcal{F}_t].
\end{aligned}$$

2.1.2 Definition. The conditional skewness of a random variable is defined as

$$\text{Skew}\,(C(\infty)|\mathcal{F}_t) = \frac{\mu_t^{(3)}}{\sigma_t^3} = \frac{\mu_t^{(3)}}{\left(\mu_t^{(2)}\right)^{\frac{3}{2}}}$$

where $\sigma_t = \sqrt{\mu_t^{(2)}}$ is the standard deviation.

2.1.3 Definition. The conditional kurtosis of a random variable is defined as

$$\text{Kurt}\,(C(\infty)|\mathcal{F}_t) = \frac{\mu_t^{(4)}}{\sigma_t^4} = \frac{\mu_t^{(4)}}{\left(\mu_t^{(2)}\right)^{2}}$$

where $\sigma_t = \sqrt{\mu_t^{(2)}}$ is again the standard deviation.

Before we will calculate the above estimators we will introduce the standard notation of the claims reserving problem.

2.2 Notation of the claims reserving problem

Past claims and outstanding liabilities can be depicted in a so called claims triangle, which applies the claims on a two-dimensional axis. See table 2.1 for an illustration.
The y-axis represents the accident years and the x-axis the development years. The most recent accident year is denoted by I and the last development year by J. Therefore we have $i \in \{0, \ldots, I\}$ and $j \in \{0, \ldots, J\}$. We will denote the entries of the triangle matrix

AY/DY	0	1	...	j	...	J
0	$C_{0,0}$	$C_{0,1}$	...	...	...	$C_{0,J}$
1	$C_{1,0}$	$\iddots$	$\iddots$	$\iddots$	$C_{1,J-1}$	
$\vdots$	$\vdots$	$\iddots$	$\iddots$	$\iddots$		
i	$\vdots$	$\iddots$	$\iddots$			
$\vdots$	$\vdots$	$C_{I-1,1}$				
I	$C_{I,0}$					

Table 2.1: Claims triangle

Range	**Explanation**
upper triangle	observed data
lower triangle	to be predicted

Table 2.2: Color table for table 2.1

by $X_{i,j}$. Possible interpretations for $X_{i,j}$ can be found in table 2.3. Under standard annotation we consider $X_{i,j}$ as incremental data. An often used alternative is to use cumulative data, which is then given by

$$C_{i,j} = \sum_{k=0}^{j} X_{i,k}$$

Incremental claims		**Cumulative claims**
$X_{i,j}$ incremental payments	$\Leftrightarrow$	$C_{i,j}$ cumulative payments
$X_{i,j}$ number of reported claims with AY i and DY j	$\Leftrightarrow$	$C_{i,j}$ total number of reported claims for AY i up to DY j
$X_{i,j}$ change of reported claim amount	$\Leftrightarrow$	$C_{i,j}$ claims incurred

Table 2.3: different interpretations for $X_{i,j}$ and $C_{i,j}$

2.2.1 Stipulation. For ease of notation and interpretation we set $X_{i,j}$ as incremental payments and $C_{i,j}$ as cumulative payments for the rest of this paper. All results apply to all other interpretations as well.

Typically the claims triangle is separated into two parts at time I. The upper triangle (or trapezoid if $I > J$, which means that we have more accident years than development years) contains observed values $X_{i,j}$, where $i + j \leq I$. On the contrary the lower triangle is empty a priori and has to be filled with our estimates for the future values $X_{i,j}$, where $i + j > I$.

Formally we have the following two sets

$$\mathcal{D}_I = \{X_{i,j} : i+j \leq I, 0 \leq i \leq I, 0 \leq j \leq J\} \tag{2.5}$$
$$\mathcal{D}_I^c = \{X_{i,j} : i+j > I, 0 \leq i \leq I, 0 \leq j \leq J\} \tag{2.6}$$

where $\mathcal{D}_I$ is the set of observations (the upper triangle in table 2.1) and $\mathcal{D}_I^c$ the unknown future set of data (the lower triangle in table 2.1) that has to be estimated from $\mathcal{D}_I$.
The accounting years are found on the diagonals $i+j=k, k \geq 0$. The accompanying incremental payments for accounting year k are then given by

$$X_k = \sum_{i+j=k} X_{i,j}, \quad k = 1, \ldots, I.$$

They are situated on the $(k+1)$-th diagonal of the development triangle.
Outstanding liabilities for the accident year i at time j are then calculated by

$$R_{i,j} = \sum_{k=j+1}^{J} X_{i,k} = C_{i,J} - C_{i,j}, \quad i = 0, \ldots, I.$$

AY/DY	0	1	...	j	...	J
0	$C_{0,0}$	$C_{0,1}$	...	...	...	$C_{0,J}$
1	$C_{1,0}$	⋰	⋰	⋰	$C_{1,J-1}$	$\widehat{C}_{1,J}^{CL}$
⋮	⋮	⋰	⋰	⋰	⋰	⋮
i	⋮	⋰	⋰	⋰	⋰	⋮
⋮	⋮	$C_{I-1,1}$	$\widehat{C}_{I-1,2}^{CL}$	⋰	⋰	$\widehat{C}_{I-1,J}^{CL}$
I	$C_{I,0}$	$\widehat{C}_{I,1}^{CL}$	...	...	$\widehat{C}_{I,J-1}^{CL}$	$\widehat{C}_{I,J}^{CL}$

Table 2.4: Claims triangle

Range		**Explanation**
red diagonal		current accounting year
upper triangle	(incl. red diagonal)	$\mathcal{D}_I$
lower triangle	(excl. red diagonal)	$\mathcal{D}_I^c$

Table 2.5: Color table for table 2.4

2.2.2 Model assumptions. We will assume for the rest of this paper that

$$I = J$$

and with that $X_{i,j} = 0, \ \forall j > J$.
This assumption has no impact on any results (which all also hold true in the general case), but simplifies the notation in the following chapters.
Note that we will still use both I and J in the formulas and proofs to make the thought process behind the derivation of the formulas clearer and to make it easier for the interested reader to adapt the estimators and the simulation model introduced in chapter 5 for the more general case, where $I \neq J$.

Chapter 3

The chain-ladder method [1]

3.1 Framework

The chain ladder method is probably the most famous claims reserving method. Together with the Bornhuetter-Ferguson method it is the most commonly used method in insurance practice, as it is easy to implement and understand. Both methods can be seen solely as an algorithm to arrive at an estimate for the claims reserves. But if we want more than just a best estimate (say we want to quantify the volatility of our prediction), we need to define an underlying stochastic model.
Hence in this chapter we will first define this model and then use that framework to estimate the second moment of the claims reserving distribution. Thereupon we will derive an estimator for the mean square error of prediction in section 3.3, which means that we not only try to estimate the variance of the underlying process $(C_{i,j})_{i,j\geq 0}$ but also of the estimators $(\widehat{C}_{i,j}^{CL})_{i,j\in D_I^c}$. The derivation of skewness and kurtosis estimators is then the aim of chapter 4.

3.2 The stochastic model

There are many different ways to arrive at the chain ladder estimators. We want to follow the distribution-free method presented in [WM08], which was first introduced by Thomas Mack in [Mac93].
In the distribution-free method cumulative claims are linked together by so called development factors. Our aim will be to estimate the ultimate claims amount $C_{i,J}$ and with it the outstanding claims reserve

$$R_i = C_{i,J} - C_{i,I-i} \tag{3.1}$$

for $i = 1, \ldots, I$ (observe that $R_0 = 0$ because of model assumptions 2.2.2).

[1] The structure and notation of this chapter follow [WM08], with the proofs being newly performed.

3.2.1 Model assumptions (distribution free chain ladder model).

- Cumulative claims $C_{i,j}$ of different accident years i are independent. (3.2)
- There exist development factors $f_0, \dots, f_{J-1} > 0$, so that

$$\forall\ 0 \le i \le I \ \text{ and } \ \forall\ 1 \le j \le J \ \text{ we have}$$
$$\mathbb{E}\left[C_{i,j}|C_{i,0},\dots,C_{i,j-1}\right] = \mathbb{E}\left[C_{i,j}|C_{i,j-1}\right] = f_{j-1}C_{i,j-1}. \tag{3.3}$$

3.2.2 Remark.

- Model assumptions 3.2.1 describe the basic assumptions necessary to use the chain ladder algorithm. These assumptions are sufficient to find an estimator for the conditional expectation, but we will have to expand them step by step for the calculation of the higher-order moments.
- Independence between claims of different accident years is a basic assumption of nearly all claims reserving methods. Because of this assumption, effects stemming from different accounting years should be eliminated from the underlying data.
- The factors f_i have varying names in technical literature like development factors, chain-ladder factors, link ratios and so forth. Our main goal will be to first estimate these factors and then quantify their volatility.

Let $\mathcal{D}_I = \{C_{i,j} : i + j \le I, 0 \le I, 0 \le j \le J\}$ denote the set of observations at time I (see (2.5)).

3.2.3 Lemma. *Under model assumptions 3.2.1 we have*

$$\mathbb{E}\left[C_{i,J}|\mathcal{D}_I\right] = \mathbb{E}\left[C_{i,J}|C_{i,I-i}\right] = C_{i,I-i}\ f_{I-i}\ \cdots\ f_{J-1} \tag{3.4}$$

for $i = 1, \dots, I$.

Proof. The proof consists of an iterative application of (3.3) and the attributes of the conditional expectation. $\mathcal{D}_i$ is an increasing sequence of sets (i.e. $\mathcal{D}_i \subseteq \mathcal{D}_j,\ \forall i \le j$), which means that the sequence $\left(\sigma\left(\mathcal{D}_i\right)\right)_{0\le i\le I}$ is an filtration and we can use the tower property. It follows that

$$\begin{aligned}
\mathbb{E}\left[C_{i,J}|\mathcal{D}_I\right] &= \mathbb{E}\left[\mathbb{E}\left[C_{i,J}|\mathcal{D}_{J-1}\right]|\mathcal{D}_I\right] \\
&\overset{(3.2)}{=} \mathbb{E}\left[\mathbb{E}\left[C_{i,J}|C_{i,0},\dots,C_{i,J-1}\right]|\mathcal{D}_I\right] \\
&\overset{(3.3)}{=} \mathbb{E}\left[\mathbb{E}\left[C_{i,J}|C_{i,J-1}\right]|\mathcal{D}_I\right] \\
&\overset{(3.3)}{=} f_{J-1}\ \mathbb{E}\left[C_{i,J-1}|\mathcal{D}_I\right] \\
&= \dots \\
&= f_{I-i}\ \cdots\ f_{J-1}\ \mathbb{E}\left[C_{i,I-i}|\mathcal{D}_I\right] \\
&\overset{(C_{i,I-i}\subseteq\mathcal{D}_I)}{=} f_{I-i}\ \cdots\ f_{J-1}\ C_{i,I-i}
\end{aligned}$$

With the same argument we get $\mathbb{E}\left[C_{i,J}|\mathcal{D}_I\right] \overset{(3.2)+(3.3)}{=} \mathbb{E}\left[C_{i,J}|C_{i,I-i}\right]$. ❑

With lemma 3.2.3 we get an algorithm to estimate the ultimate loss $C_{i,J}$ of an accident year, provided we know the development factors f_i. The claims reserve for outstanding liabilities is then calculated by

$$\widehat{R}_i = \mathbb{E}\left[R_i|\mathcal{D}_I\right] = \mathbb{E}\left[C_{i,J}|\mathcal{D}_I\right] - C_{i,I-i} = C_{i,I-i}\ \left(f_{I-i}\ \cdots\ f_{J-1} - 1\right) \tag{3.5}$$

This estimator is generally called the best estimate for the claims reserve of accident year i based on information $\mathcal{D}_I$.
In practice the development factors f_j are unknown and therefore have to be estimated from the data too.

3.2.4 Estimator (Development Factors)**.** We will estimate the factors $f_j,\ j = 0, \dots, J-1$ by (see lemma 3.2.6 for the reasoning behind this).

$$\widehat{f}_j := \frac{\sum_{i=0}^{I-j-1} C_{i,j+1}}{\sum_{i=0}^{I-j-1} C_{i,j}} = \sum_{i=0}^{I-j-1} \frac{C_{i,j}}{\sum_{k=0}^{I-j-1} C_{k,j}} \frac{C_{i,j+1}}{C_{i,j}} \tag{3.6}$$

Observe that the chain ladder factors are a weighed average of the individual development factors

$$F_{i,j+1} := \frac{C_{i,j+1}}{C_{i,j}}. \tag{3.7}$$

3.2.5 Definition (Chain Ladder Estimator)**.** For $i = 1, \dots, I$ and $j = I - i + 1, \dots, J$ the chain ladder estimator for $\mathbb{E}\left[C_{i,j}|\mathcal{D}_I\right]$ is then given by

$$\widehat{C}_{i,j}^{CL} := \widehat{\mathbb{E}}\left[C_{i,j}|\mathcal{D}_I\right] = C_{i,I-i}\ \widehat{f}_{I-i}\ \cdots\ \widehat{f}_{j-1} \tag{3.8}$$

We will now define a new set of observations as

$$\mathcal{B}_k := \{C_{i,j} : i + j \leq I, 0 \leq j \leq k\} \subseteq \mathcal{D}_I \tag{3.9}$$

We have $\mathcal{B}_J = \mathcal{D}_I$, the set of all observations at time I. $\mathcal{B}_k$ represents the set of information up to development year k (see table 3.1).

3.2.6 Lemma. *Under model assumptions 3.2.1 we have*

(i) $\mathbb{E}\left[\widehat{f}_j\middle|\mathcal{B}_j\right] = f_j$, *i.e.* $\widehat{f}_j$ *is an unbiased estimator for* f_j *given* $\mathcal{B}_j$,

(ii) $\mathbb{E}\left[\widehat{f}_j\right] = f_j$, *i.e.* $\widehat{f}_j$ *is an (unconditionally) unbiased estimator for* f_j,

(iii) $\mathbb{E}\left[\widehat{f}_0\ \cdots\ \widehat{f}_j\right] = \mathbb{E}\left[\widehat{f}_0\right]\ \cdots\ \mathbb{E}\left[\widehat{f}_j\right]$, $j = 0, \dots, J-1$, *i.e. the* $\widehat{f}_j$ *are uncorrelated,*

(iv) $\mathbb{E}\left[\widehat{C}_{i,J}^{CL}\middle|C_{i,I-i}\right] = \mathbb{E}\left[C_{i,J}\middle|\mathcal{D}_I\right]$, *i.e.* $\widehat{C}_{i,J}^{CL}$ *is an unbiased estimator for* $\mathbb{E}\left[C_{i,J}\middle|\mathcal{D}_I\right]$ $\left(= \mathbb{E}\left[C_{i,J}\middle|C_{i,I-i}\right]\right)$ *given* $C_{i,I-i}$,

(v) $\mathbb{E}\left[\widehat{C}_{i,J}^{CL}\right] = \mathbb{E}\left[C_{i,J}\right]$, *i.e.* $\widehat{C}_{i,J}^{CL}$ *is an (unconditionally) unbiased estimator for* $\mathbb{E}\left[C_{i,J}\right]$.

AY/DY	0	1	...	k	...	...	J
0	$C_{0,0}$	$C_{0,1}$	...	$C_{0,k}$	...	$\iddots$	$C_{0,J}$
1	$C_{1,0}$	$\iddots$	$\iddots$	$\vdots$	$\vdots$	$C_{1,J-1}$	
$\vdots$	$\vdots$	$\iddots$	$\iddots$	$\vdots$	$\iddots$		
i	$\vdots$	$\iddots$	$\iddots$	$C_{i,k}$			
$\vdots$	$\vdots$	$\iddots$	$\iddots$				
$\vdots$	$\vdots$	$C_{I-1,2}$					
I	$C_{I,1}$						

Table 3.1: Claims triangle with $\mathcal{B}_k$ in blue

Proof.

(i) $C_{i,j}$ is measurable with respect to (w.r.t) $\mathcal{B}_j$ $\forall i = 0, \ldots, I-j-1$, which leads to

$$\mathbb{E}\left[\widehat{f}_j \middle| \mathcal{B}_j\right] = \mathbb{E}\left[\frac{\sum_{i=0}^{I-j-1} C_{i,j+1}}{\sum_{i=0}^{I-j-1} C_{i,j}} \middle| \mathcal{B}_j\right] = \frac{\sum_{i=0}^{I-j-1} \mathbb{E}\left[C_{i,j+1} | \mathcal{B}_j\right]}{\sum_{i=0}^{I-j-1} C_{i,j}} \overset{(3.3)}{=} \frac{\sum_{i=0}^{I-j-1} C_{i,j}\ f_j}{\sum_{i=0}^{I-j-1} C_{i,j}} = f_j$$

(ii) $\mathbb{E}\left[\widehat{f}_j\right] = \mathbb{E}\left[\mathbb{E}\left[\widehat{f}_j \middle| \mathcal{B}_j\right]\right] \overset{(i)}{=} f_j$

(iii) We know that $\widehat{f}_i$ is measurable w.r.t. $\mathcal{B}_j$ for $i < j$ $(*)$. If we use the attributes of the conditional expectation iteratively we get

$$\begin{aligned}\mathbb{E}\left[\widehat{f}_0 \cdots \widehat{f}_j\right] &= \mathbb{E}\left[\mathbb{E}\left[\widehat{f}_0 \cdots \widehat{f}_j \middle| \mathcal{B}_j\right]\right] \overset{(*)}{=} \mathbb{E}\left[\widehat{f}_0 \cdots \widehat{f}_{j-1} \mathbb{E}\left[\widehat{f}_j \middle| \mathcal{B}_j\right]\right] \\ &\overset{(i)}{=} \mathbb{E}\left[\widehat{f}_0 \cdots \widehat{f}_{j-1}\right]\ f_j \overset{(ii)}{=} \mathbb{E}\left[\widehat{f}_0 \cdots \widehat{f}_{j-1}\right]\ \mathbb{E}\left[\widehat{f}_j\right] \\ &= \ldots = \mathbb{E}\left[\widehat{f}_0\right] \cdots \mathbb{E}\left[\widehat{f}_j\right]\end{aligned}$$

(iv) With the same argument as in lemma 3.2.3 we have

$$\begin{aligned}\mathbb{E}\left[\widehat{C}_{i,J}^{CL} \middle| C_{i,I-i}\right] &\overset{(3.8)}{=} \mathbb{E}\left[C_{i,I-i}\ \widehat{f}_{I-i} \cdots \widehat{f}_{J-1} \middle| C_{i,I-i}\right] \\ &= \mathbb{E}\left[C_{i,I-i}\ \widehat{f}_{I-i} \cdots \widehat{f}_{J-2}\ \mathbb{E}\left[\widehat{f}_{J-1} \middle| \mathcal{B}_{J-1} \middle| C_{i,I-i}\right]\right] \\ &\overset{(i)}{=} f_{J-1}\ \mathbb{E}\left[\widehat{C}_{i,J-1}^{CL} \middle| C_{i,I-i}\right] \\ &= \ldots \\ &= C_{i,I-i}\ f_{I-i} \cdots f_{J-1} = \mathbb{E}\left[C_{i,j} | \mathcal{D}_I\right]\end{aligned}$$

(v) $\mathbb{E}\left[\widehat{C}_{i,J}^{CL}\right] = \mathbb{E}\left[\mathbb{E}\left[\widehat{C}_{i,J}^{CL}\middle|C_{i,I-i}\right]\right] \overset{(iv)}{=} \mathbb{E}\left[C_{i,J}\right]$

❏

3.2.7 Remark.

- We have shown the uncorrelatedness of the estimators $\widehat{f}_j$, but it has to be emphasized that this does not implicate independence of the factors. In fact it can be shown that the squares of two consecutive estimators $\widehat{f}_j$ and $\widehat{f}_{j+1}$ are negatively correlated (for further information see [WM08]).
- Lemma 3.2.6 shows that our estimators $\widehat{f}_j$ are unbiased estimators of the development factors and thereby justifies our choice in estimator 3.2.4. In [WM08] it is further shown, that under certain conditions, the choice (3.6) fulfills an optimality condition amongst all unbiased estimators.

3.3 Mean square error

In the previous section we have established best estimates for the claims reserves, but we are also interested in their volatility. The aim of this chapter is to estimate the second moment of these estimators by introducing the concept of the mean square error of prediction. Building on this foundation we will then also calculate higher moments and estimate the full reserve risk distribution in the subsequent chapters.

3.3.1 Definition (conditional MSEP)**.** The conditional mean square error of prediction of an estimator $\widehat{X}$ for X is defined by

$$\mathbf{msep}_{X|\mathcal{D}}\left(\widehat{X}\right) = \mathbb{E}\left[\left(\widehat{X} - X\right)^2\middle|\mathcal{D}\right]$$

For a $\mathcal{D}$ measurable estimator $\widehat{X}$ we have

$$\mathbf{msep}_{X|\mathcal{D}}\left(\widehat{X}\right) = \mathbb{V}\left[X|\mathcal{D}\right] + \left(\widehat{X} - \mathbb{E}\left[X|\mathcal{D}\right]\right)^2 \tag{3.10}$$

3.3.2 Remark. We can now interprete the two terms on the right hand side of eq. (3.10) separately

- The first term is generally called conditional process variance (stochastic error) and describes the volatility within the stochastic model. This means that this part cannot be eliminated by further refinement of our estimator.
- The second term is the parameter estimation error. It reflects the uncertainty in the estimation of the parameters and the estimation of the conditional expectation. This factor should generally decrease as the number of observations increases, but it should be noted that in many practical situations it remains positive and never disappears entirely.

- We can see our first problem here very clearly. To calculate the parameter estimation error we would have to know the exact value of $\mathbb{E}[X|\mathcal{D}]$, which we do not have (we have estimated it by $\widehat{X}$). Therefore we cannot use a straightforward approach for this calculation. One way to proceed is to study the possible fluctuations of $\widehat{X}$ around $\mathbb{E}[X|\mathcal{D}]$.

We will now expand our model assumptions to be able to find estimators for the second moment.

3.3.3 Model assumptions.

- Cumulative claims $C_{i,j}$ of different accident years i are independent.
- $(C_{i,j})_{j\geq 0}$ form a Markov chain. There exist factors $f_0, \dots, f_{J-1} > 0$ and variance parameters $\sigma_0^2, \dots, \sigma_{J-1}^2 > 0$ such that

$$\forall 0 \leq i \leq I \quad \text{and} \quad \forall 1 \leq j \leq J \quad \text{we have}$$
$$\mathbb{E}[C_{i,j}|C_{i,j-1}] = f_{j-1}C_{i,j-1}, \tag{3.11}$$
$$\mathbb{V}[C_{i,j}|C_{i,j-1}] = \sigma_{j-1}^2 C_{i,j-1}. \tag{3.12}$$

3.3.4 Estimator. We define the estimators for the σ_j^2 for $0 \leq j \leq J-2$ by

$$\widehat{\sigma}_j^2 = \frac{1}{I-j-1}\sum_{i=0}^{I-j-1} C_{i,j}\left(\frac{C_{i,j+1}}{C_{i,j}} - \widehat{f}_j\right)^2 \tag{3.13}$$

3.3.5 Remark. Note that we do not have enough data to estimate $\widehat{\sigma}_{J-1}$, which would only be possible with the above formula if $I > J$. So a different approach has to be used to estimate the last variance factor. We will choose the extrapolation introduced in [Mac93] which states

$$\widehat{\sigma}_{J-1}^2 = \min\left\{\frac{\left(\widehat{\sigma}_{J-2}^2\right)^2}{\widehat{\sigma}_{J-3}^2}, \widehat{\sigma}_{J-3}^2, \widehat{\sigma}_{J-2}^2\right\} \tag{3.14}$$

3.3.6 Lemma. *Under model assumptions 3.3.3 we have*

(i) $\mathbb{E}\left[\widehat{\sigma}_j^2|\mathcal{B}_j\right] = \sigma_j^2$, *i.e.* $\widehat{\sigma}_j^2$ *is an unbiased estimator for* σ_j^2 *given* $\mathcal{B}_j$,

(ii) $\mathbb{E}\left[\widehat{\sigma}_j^2\right] = \sigma_j^2$, *i.e.* $\widehat{\sigma}_j^2$ *is an (unconditionally) unbiased estimator for* σ_j^2,

Proof. (i) We have

$$\begin{aligned}\mathbb{E}\left[\widehat{\sigma}_j^2|\mathcal{B}_j\right] &= \mathbb{E}\left[\frac{1}{I-j-1}\sum_{i=0}^{I-j-1} C_{i,j}\left(\frac{C_{i,j+1}}{C_{i,j}} - \widehat{f}_j\right)^2\middle|\mathcal{B}_j\right]\\ &= \frac{1}{I-j-1}\sum_{i=0}^{I-j-1} C_{i,j}\mathbb{E}\left[\left(\frac{C_{i,j+1}}{C_{i,j}} - \widehat{f}_j\right)^2\middle|\mathcal{B}_j\right]\end{aligned}$$

hence we will calculate the conditional expectations in the sum for $i = 0, \dots, I-j-1$. We have

$$\mathbb{E}\left[\left(\frac{C_{i,j+1}}{C_{i,j}} - \widehat{f}_j\right)^2 \middle| \mathcal{B}_j\right] = \mathbb{E}\left[\left(\frac{C_{i,j+1}}{C_{i,j}} - f_j - \left(\widehat{f}_j - f_j\right)\right)^2 \middle| \mathcal{B}_j\right]$$
$$= \mathbb{E}\left[\left(\frac{C_{i,j+1}}{C_{i,j}} - f_j\right)^2 \middle| \mathcal{B}_j\right] - 2\,\mathbb{E}\left[\left(\frac{C_{i,j+1}}{C_{i,j}} - f_j\right)\left(\widehat{f}_j - f_j\right) \middle| \mathcal{B}_j\right] + \mathbb{E}\left[\left(\widehat{f}_j - f_j\right)^2 \middle| \mathcal{B}_j\right]$$

We will now calculate each of these terms separately. Using model assumptions 3.3.3 we have

$$\mathbb{E}\left[\left(\frac{C_{i,j+1}}{C_{i,j}} - f_j\right)^2 \middle| \mathcal{B}_j\right] \overset{(3.11)}{=} \mathbb{V}\left[\frac{C_{i,j+1}}{C_{i,j}} \middle| \mathcal{B}_j\right] \overset{(3.12)}{=} \frac{\sigma_j^2}{C_{i,j}} \tag{3.15}$$

The independence of different accident years yields

$$\begin{aligned}
\mathbb{E}\left[\left(\frac{C_{i,j+1}}{C_{i,j}} - f_j\right)\left(\widehat{f}_j - f_j\right) \middle| \mathcal{B}_j\right] &= \mathrm{Cov}\left(\frac{C_{i,j+1}}{C_{i,j}}, \widehat{f}_j \middle| \mathcal{B}_j\right) \\
&\overset{(3.6)}{=} \mathrm{Cov}\left(\frac{C_{i,j+1}}{C_{i,j}}, \sum_{n=0}^{I-j-1} \frac{C_{n,j}}{\sum_{k=0}^{I-j-1} C_{k,j}} \frac{C_{n,j+1}}{C_{n,j}} \middle| \mathcal{B}_j\right) \\
&= \sum_{n=0}^{I-j-1} \mathrm{Cov}\left(\frac{C_{i,j+1}}{C_{i,j}}, \frac{C_{n,j}}{\sum_{k=0}^{I-j-1} C_{k,j}} \frac{C_{n,j+1}}{C_{n,j}} \middle| \mathcal{B}_j\right) \\
&\overset{(3.2)}{=} \frac{C_{i,j}}{\sum_{k=0}^{I-j-1} C_{k,j}} \mathrm{Cov}\left(\frac{C_{i,j+1}}{C_{i,j}}, \frac{C_{i,j+1}}{C_{i,j}} \middle| \mathcal{B}_j\right) \\
&= \frac{C_{i,j}}{\sum_{k=0}^{I-j-1} C_{k,j}} \mathbb{V}\left[\frac{C_{i,j+1}}{C_{i,j}} \middle| \mathcal{B}_j\right] \\
&\overset{(3.15)}{=} \frac{\sigma_j^2}{\sum_{k=0}^{I-j-1} C_{k,j}}
\end{aligned} \tag{3.16}$$

where Cov denotes the covariance which is defined in eq. (5.6). For the last term we have

$$\begin{aligned}
\mathbb{E}\left[\left(\widehat{f}_j - f_j\right)^2 \middle| \mathcal{B}_j\right] &= \mathbb{E}\left[\left(\widehat{f}_j - \mathbb{E}\left[\widehat{f}_j\right]\right)^2 \middle| \mathcal{B}_j\right] \\
&= \mathbb{V}\left[\widehat{f}_j \middle| \mathcal{B}_j\right] \\
&\overset{(3.6)}{=} \mathbb{V}\left[\frac{\sum_{k=0}^{I-j-1} C_{k,j+1}}{\sum_{k=0}^{I-j-1} C_{k,j}} \middle| \mathcal{B}_j\right] \\
&= \left(\frac{1}{\sum_{k=0}^{I-j-1} C_{k,j}}\right)^2 \mathbb{V}\left[\sum_{k=0}^{I-j-1} C_{k,j+1} \middle| \mathcal{B}_j\right] \\
&\overset{(3.2)}{=} \left(\frac{1}{\sum_{k=0}^{I-j-1} C_{k,j}}\right)^2 \sum_{k=0}^{I-j-1} \mathbb{V}\left[C_{k,j+1} | \mathcal{B}_j\right]
\end{aligned}$$

$$\overset{(3.12)}{=} \left(\frac{1}{\sum_{k=0}^{I-j-1} C_{k,j}}\right)^2 \sigma_j^2 \left(\sum_{k=0}^{I-j-1} C_{k,j}\right)$$
$$= \frac{\sigma_j^2}{\sum_{k=0}^{I-j-1} C_{k,j}} \tag{3.17}$$

Putting all of this together we have

$$\mathbb{E}\left[\left(\frac{C_{i,j+1}}{C_{i,j}} - \widehat{f}_j\right)^2 \middle| \mathcal{B}_j\right] \overset{(3.15)+(3.16)+(3.17)}{=} \sigma_j^2 \left(\frac{1}{C_{i,j}} - \frac{1}{\sum_{k=0}^{I-j-1} C_{k,j}}\right) \tag{3.18}$$

Thus we can conclude

$$\begin{aligned}
\mathbb{E}\left[\widehat{\sigma}_j^2 \middle| \mathcal{B}_j\right] &= \frac{1}{I-j-1} \sum_{i=0}^{I-j-1} C_{i,j} \mathbb{E}\left[\left(\frac{C_{i,j+1}}{C_{i,j}} - \widehat{f}_j\right]^2 \middle| \mathcal{B}_j\right] \\
&\overset{(3.18)}{=} \frac{1}{I-j-1} \sum_{i=0}^{I-j-1} C_{i,j}\, \sigma_j^2 \left(\frac{1}{C_{i,j}} - \frac{1}{\sum_{k=0}^{I-j-1} C_{k,j}}\right) \\
&= \frac{\sigma_j^2}{I-j-1} \sum_{i=0}^{I-j-1} \left(1 - \frac{C_{i,j}}{\sum_{k=0}^{I-j-1} C_{k,j}}\right) \\
&= \frac{\sigma_j^2}{I-j-1} \left(I - j - \sum_{i=0}^{I-j-1} \frac{C_{i,j}}{\sum_{k=0}^{I-j-1} C_{k,j}}\right) \\
&= \frac{\sigma_j^2}{I-j-1} \left(I - j - \frac{\sum_{i=0}^{I-j-1} C_{i,j}}{\sum_{k=0}^{I-j-1} C_{k,j}}\right) \\
&= \frac{\sigma_j^2}{I-j-1} (I-j-1) \\
&= \sigma_j^2
\end{aligned}$$

(ii) Using item (i) and the tower property of the conditional expectation we have

$$\mathbb{E}\left[\widehat{\sigma}_j^2\right] = \mathbb{E}\left[\mathbb{E}\left[\widehat{\sigma}_j^2 \middle| \mathcal{B}_j\right]\right] = \sigma_j^2.$$

❑

From the above proof we can deduce the following equation

$$\mathbb{E}\left[\widehat{f}_j^2 \middle| \mathcal{B}_j\right] = \mathbb{V}\left[\widehat{f}_j^2 \middle| \mathcal{B}_j\right] + \left(\mathbb{E}\left[\widehat{f}_j \middle| \mathcal{B}_j\right]\right)^2 \overset{(3.17)}{=} \frac{\sigma_j^2}{\sum_{i=0}^{I-j-1} C_{i,j}} + f_j^2 \tag{3.19}$$

We will need this equation later on in the derivation of the conditional estimation error and the derivations of the skewness and kurtosis estimators.
Our aim for the rest of this chapter is now to find an estimator for the MSEP of $\widehat{C}_{i,J}^{CL}$, where $i = 1, \ldots, I$.

We have

$$\mathbf{msep}_{C_{i,j}|\mathcal{D}_I}\left(\widehat{C}_{i,J}^{CL}\right) = \mathbb{E}\left[\left(\widehat{C}_{i,J}^{CL} - C_{i,J}\right)^2 \middle| \mathcal{D}_I\right]$$
$$\overset{(3.10)}{=} \mathbb{V}\left[C_{i,J}|\mathcal{D}_I\right] + \left(\widehat{C}_{i,J}^{CL} - \mathbb{E}\left[C_{i,J}|\mathcal{D}_I\right]\right) \tag{3.20}$$

for single accident years and

$$\mathbf{msep}_{\sum_{i=1}^I C_{i,j}|\mathcal{D}_I}\left(\sum_{i=1}^I \widehat{C}_{i,J}^{CL}\right) = \mathbb{E}\left[\left(\sum_{i=1}^I \widehat{C}_{i,J}^{CL} - \sum_{i=1}^I C_{i,J}\right)^2 \middle| \mathcal{D}_I\right] \tag{3.21}$$

for aggregated accident years.

3.3.1 Conditional Process Variance

3.3.7 Lemma. *Under model assumptions 3.3.3 the conditional process variance for the ultimate claim amount $C_{i,J}$ of a single accident year i with $i = 1, \ldots, I$, is given by*

$$\mathbb{V}\left[C_{i,J}|\mathcal{D}_I\right] = \left(\mathbb{E}\left[C_{i,J}|C_{i,I-i}\right]\right)^2 \sum_{j=I-i}^{J-1} \frac{\sigma_j^2}{f_j^2 \; \mathbb{E}\left[C_{i,j}|C_{i,I-i}\right]} \tag{3.22}$$

Proof.

$$\begin{aligned}
\mathbb{V}\left[C_{i,J}|\mathcal{D}_I\right] &= \mathbb{V}\left[C_{i,J}|C_{i,I-i}\right] \\
&= \mathbb{E}\left[\mathbb{V}\left[C_{i,J}|C_{i,J-1}\right]|C_{i,I-i}\right] + \mathbb{V}\left[\mathbb{E}\left[C_{i,J}|C_{i,J-1}\right]|C_{i,I-i}\right] \\
&= \sigma_{J-1}^2 \; \mathbb{E}\left[C_{i,J-1}|C_{i,I-i}\right] + f_{J-1}^2 \; \mathbb{V}\left[C_{i,J-1}|C_{i,I-i}\right] \\
&= \sigma_{J-1}^2 \; C_{i,I-i} \prod_{m=I-i}^{J-2} f_m + f_{J-1}^2 \; \mathbb{V}\left[C_{i,J-1}|C_{i,I-i}\right].
\end{aligned} \tag{3.23}$$

We can now do the same procedure for $\mathbb{V}\left[C_{i,J-1}|C_{i,I-i}\right]$ to get

$$\mathbb{V}\left[C_{i,J-1}|C_{i,I-i}\right] = \sigma_{J-2}^2 \; C_{i,I-i} \prod_{m=I-i}^{J-3} f_m + f_{J-2}^2 \; \mathbb{V}\left[C_{i,J-2}|C_{i,I-i}\right].$$

Iterating this until we get $\mathbb{V}\left[C_{i,I-i}|C_{i,I-i}\right] (= 0)$ on the right hand side of the above equation we get

$$\begin{aligned}
\mathbb{V}\left[C_{i,J}|\mathcal{D}_I\right] &= C_{i,I-i} \sum_{j=I-i}^{J-1} \prod_{n=j+1}^{J-1} f_n^2 \; \sigma_j^2 \prod_{m=I-i}^{j-1} f_m \\
&= \sum_{j=I-i}^{J-1} \prod_{n=j+1}^{J-1} f_n^2 \; \sigma_j^2 \mathbb{E}\left[C_{i,j}|C_{i,I-i}\right] \\
&= \left(\mathbb{E}\left[C_{i,J}|C_{i,I-i}\right]\right)^2 \sum_{j=I-i}^{J-1} \frac{\sigma_j^2}{f_j^2 \; \mathbb{E}\left[C_{i,j}|C_{i,I-i}\right]}.
\end{aligned}$$

❑

3.3.8 Estimator. Our estimator for the conditional process variance is then

$$\widehat{\mathbb{V}}\left[C_{i,J}|\mathcal{D}_I\right] = \widehat{\mathbb{E}}\left[\left(C_{i,J} - \mathbb{E}\left[C_{i,J}|\mathcal{D}_I\right]\right)^2\Big|\mathcal{D}_I\right]$$
$$\overset{(3.22)}{=} \left(\widehat{C}_{i,J}^{CL}\right)^2 \sum_{j=I-i}^{J-1} \frac{\widehat{\sigma}_j^2}{\widehat{f}_j^2\ \widehat{C}_{i,j}^{CL}}. \tag{3.24}$$

where for ease of notation in the formula $\widehat{C}_{i,I-i}^{CL} := C_{i,I-i}$.
We can also use eq. (3.23) to get the recursive formula

$$\widehat{\mathbb{V}}\left[C_{i,j}|\mathcal{D}_I\right] = \widehat{\mathbb{V}}\left[C_{i,j-1}|\mathcal{D}_I\right]\widehat{f}_{j-1}^2 + \widehat{\sigma}_{j-1}^2\ \widehat{C}_{i,j-1}^{CL},$$
$$\forall j = I-i+1,\ldots,J,$$

where $\widehat{\mathbb{V}}\left[C_{i,I-i}|\mathcal{D}_I\right] = 0$ and again $\widehat{C}_{i,I-i}^{CL} = C_{i,I-i}$.

Under model assumptions 3.3.3 different accident years are independent, which leads to

$$\mathbb{V}\left[\sum_{i=1}^{I} C_{i,J}\middle|\mathcal{D}_I\right] = \sum_{i=1}^{I}\mathbb{V}\left[C_{i,J}|\mathcal{D}_I\right].$$

Therefore we estimate the conditional process variance for aggregated accident years by

$$\widehat{\mathbb{V}}\left[\sum_{i=1}^{I} C_{i,J}\middle|\mathcal{D}_I\right] = \sum_{i=1}^{I}\widehat{\mathbb{V}}\left[C_{i,J}|\mathcal{D}_I\right].$$

3.3.2 The parameter estimation error for single accident years

To calculate the MSEP we now need to find an estimator for the second term of eq. (3.20). This estimator will tell us about the accuracy of the chain ladder estimators $\widehat{f}_j$. We begin by calculating

$$\left(\widehat{C}_{i,J}^{CL} - \mathbb{E}\left[C_{i,J}|\mathcal{D}_I\right]\right)^2 \overset{(3.4)+(3.8)}{=} C_{i,I-i}^2\left(\widehat{f}_{I-i}\cdots\widehat{f}_{J-1} - f_{I-i}\cdots f_{J-1}\right)^2$$
$$= C_{i,I-i}^2\left(\prod_{J=I-i}^{J-1}\widehat{f}_j^2 + \prod_{J=I-i}^{J-1} f_j^2 - 2\prod_{J=I-i}^{J-1}\widehat{f}_j\ f_j\right) \tag{3.25}$$

Observe that while the factors $\widehat{f}_j$ are known at time I we do not know the factors f_j, $j = 0,\ldots,J-1$ (that is why we had to estimate them in the first place), hence we cannot compute eq. (3.25) directly. We will now use an analytic resampling approach, introduced in [WM08]. In this approach we will analyse the extent to which the estimators $\widehat{f}_j$ fluctuate around the true values f_j.
We start by focusing on resampling the following products

$$\widehat{f}_{I-i}^2\cdots\widehat{f}_{J-1}^2, \quad i = 1,\ldots,I \tag{3.26}$$

AY/DY	0	1	$\dots$	$I-i$	$I-i+1$	$\dots$	J
0	$C_{0,0}$	$C_{0,1}$	$\dots$	$C_{0,I-i}$	$C_{0,I-i+1}$	$\iddots$	$C_{0,J}$
1	$C_{1,0}$	$\iddots$	$\iddots$	$\vdots$	$\vdots$	$\iddots$	
$\vdots$	$\vdots$	$\iddots$	$\iddots$	$\vdots$	$\iddots$		
i	$\vdots$	$\iddots$	$\iddots$	$C_{i,I-i}$			
$\vdots$	$\vdots$	$\iddots$	$\iddots$				
$\vdots$	$\vdots$	$C_{I-1,2}$					
I	$C_{I,1}$						

Table 3.2: Claims triangle with $\mathcal{D}^O_{I,i}$ in green

First we again define a new subset of $\mathcal{D}_i$ (see table 3.2 for a visualization)

$$\mathcal{D}^O_{I,i} := \{C_{k,j} \in D_I : j > I-i\} \subseteq \mathcal{D}_I. \tag{3.27}$$

Observe that $\widehat{f}_j$ is $\mathcal{B}_{j+1}$ measurable. We will now conditionally resample in $\mathcal{D}^O_{I,i}$, which formally means that we will sample under a different probability measure. For this we will calculate the value of

$$\mathbb{E}\left[\widehat{f}^2_{I-i}\middle|\mathcal{B}_{I-i}\right]\ \mathbb{E}\left[\widehat{f}^2_{I-i+1}\middle|\mathcal{B}_{I-i+1}\right]\ \cdots\ \mathbb{E}\left[\widehat{f}^2_{J-1}\middle|\mathcal{B}_{J-1}\right] \tag{3.28}$$

This means we are averaging the factors $\widehat{f}_j$ at every position $j = I-i, \dots, J-1$ on the conditional structure. Note that eq. (3.28) depends on the observations in $\mathcal{D}^O_{I,i}$, because we have that $\mathcal{D}^O_{I,i} \cap \mathcal{B}_j \neq \emptyset$ if $j > I-i$.
To formalize things we note that the estimated CL factors $\widehat{f}_j$ are in fact functions of $(C_{k,j+1})_{k=0,\dots,I-j-1}$ and $(C_{k,j})_{k=0,\dots,I-j-1}$ so we can write

$$\widehat{f}_j = \widehat{f}_j\left((C_{k,j+1})_{k=0,\dots,I-j-1}, (C_{k,j})_{k=0,\dots,I-j-1}\right) = \frac{\sum_{k=0}^{I-j-1} C_{k,j+1}}{\sum_{k=0}^{I-j-1} C_{k,j}}$$

According to model assumptions 3.3.3 $(C_{k,j})_{j\geq 0}$ is a Markov chain, so we can write its probability distribution as

$$\begin{aligned}\mathrm{dP}_k(x_0,\dots,x_J) &= \mathrm{K}^{(k)}_0(\mathrm{d}x_0)\ \mathrm{K}^{(k)}_1(x_o,\mathrm{d}x_1)\ \mathrm{K}^{(k)}_2(x_o,x_1,\mathrm{d}x_2)\ \cdots\ \mathrm{K}^{(k)}_J(x_o,\dots,x_{J-1}\mathrm{d}x_J)\\ &= \mathrm{K}^{(k)}_0(\mathrm{d}x_0)\ \mathrm{K}^{(k)}_1(x_o,\mathrm{d}x_1)\ \mathrm{K}^{(k)}_2(x_1,\mathrm{d}x_2)\ \cdots\ \mathrm{K}^{(k)}_J(x_{J-1}\mathrm{d}x_J)\end{aligned}$$

In this resampling process we will now always keep the set of actual observations $C_{k,j}$ fixed and resample the next step in the time series. This means that, given $\mathcal{D}_I$, we consider the following probability measures

$$\mathrm{dP}^*_k\left((x_{k,j})_{k+j\leq I}\right) = \prod_{k=0}^{I-1} \mathrm{K}^{(k)}_1\left(C_{k,0},\mathrm{d}x_{k,1}\right)\ \cdots\ \mathrm{K}^{(k)}_{I-k}\left(C_{k,I-k-1},\mathrm{d}x_{k,I-k}\right)$$

for the resampling of

$$\prod_{j=I-i}^{J-1} \widehat{f}_j = \prod_{j=I-i}^{J-1} \widehat{f}_j \left((x_{k,j+1})_{k=0,\dots,I-j-1}, (C_{k,j})_{k=0,\dots,I-j-1} \right) = \prod_{j=I-i}^{J-1} \frac{\sum_{k=0}^{I-j-1} x_{k,j+1}}{\sum_{k=0}^{I-j-1} C_{k,j}}$$

The values $C_{k,j}$ serve as a fixed volume measure for the resampled values of $x_{k,j+1}$.

3.3.9 Remark. [WM08] also introduces other approaches to tackle the resampling problem, but we will only focus on this one as it leads to a closed analytical formula for the MSEP of the reserves.

To accomplish the resampling of the CL factors $\widehat{f}_j$ we introduce a time series assumption to our model. Our new stronger model assumptions can than be defined as follows

3.3.10 Model assumptions (time series model).

- Cumulative claims $C_{i,j}$ of different accident years are independent
- There exist constants $f_j > 0,\ \sigma_j > 0$ and random variables $\varepsilon_{i,j+1}$ such that

$$C_{i,j+1} = f_j\ C_{i,j} + \sigma_j\ \sqrt{C_{i,j}}\ \varepsilon_{i,j+1} \qquad (3.29)$$
$$\forall\ i \in \{0, \dots, I\}\,, \text{ and } \forall\ j \in \{0, \dots, J-1\}$$

  where $\varepsilon_{i,j+1}$ are conditionally independent given $\mathcal{B}_0$, with $\mathbb{E}\left[\varepsilon_{i,j+1}|\mathcal{B}_0\right] = 0$, $\mathbb{E}\left[\varepsilon_{i,j+1}^2\middle|\mathcal{B}_0\right] = 1$ and $\mathbb{P}\left[C_{i,j+1} > 0|\mathcal{B}_0\right] = 1\ \forall i \in \{0, \dots, I\}$, and $\forall j \in \{0, \dots, J-1\}$.

3.3.11 Remark.

- We can see that eq. (3.29) defines an autoregressive process. We will use it in our derivation of the estimation error.
- The random variables $\varepsilon_{i,j+1}$ are defined via the conditional probability measure $\mathbb{P}[\cdot|\mathcal{B}_0]$. All subsequent calculations are then done under this probability measure.
- Observe that we have $C_{i,j+1} > 0$ almost surely under $\mathbb{P}[\cdot|\mathcal{B}_0]$ for our cumulative claims. Note that a similar assumption also underlies model assumptions 3.3.3 to make sense of the variance condition (eq. (3.12)).
- It can easily be verified that model assumptions 3.3.10 imply model assumptions 3.3.3.

We will now resample the observations for $\widehat{f}_{I-i}, \dots, \widehat{f}_{J-1}$, given the upper triangle $\mathcal{D}_I$. To do this we generate a new set of observations $\widetilde{C}_{i,j+1}$, given $\mathcal{D}_I$, $i = 0, \dots, I$ and $j = 0, \dots, J-1$ using the formula

$$\widetilde{C}_{i,j+1} = f_j\ C_{i,j} + \sigma_j\ \sqrt{C_{i,j}}\ \tilde{\varepsilon}_{i,j+1} \qquad (3.30)$$

with $\sigma_j > 0$ and $\tilde{\varepsilon}_{i,j+1}$ being an independent and identically distributed copy of $\varepsilon_{i,j+1}$, given $\mathcal{B}_0$.

Before we procede we define the following short notation

$$^{[k]}S_j = \sum_{i=0}^{k} C_{i,j} \qquad (3.31)$$

Now we can calculate our development factors under model assumptions 3.3.10.

3.3.12 Estimator.

$$\widehat{f}_j^* = \frac{\sum_{i=0}^{I-j-1} \widetilde{C}_{i,j+1}}{\sum_{i=0}^{I-j-1} C_{i,j}} \overset{(3.30)}{=} f_j + \frac{\sigma_j}{{}^{[I-j-1]}S_j} \sum_{i=0}^{I-j-1} \sqrt{C_{i,j}}\ \tilde{\varepsilon}_{i,j+1} \tag{3.32}$$

We denote the probability measure of the resampled factors by $\mathbb{P}^*_{\mathcal{D}_I}$. With this we have

3.3.13 Lemma.

(i) the resampled estimates $\widehat{f}^*_{I-i}, \ldots, \widehat{f}^*_{J-1}$ are independent w.r.t. to $\mathbb{P}^*_{\mathcal{D}_I}$,

(ii) $\mathbb{E}^*_{\mathcal{D}_I}\left[\widehat{f}^*_j\right] = f_j$ for $j = 0, \ldots, J-1$, and

(iii) $\mathbb{E}^*_{\mathcal{D}_I}\left[\left(\widehat{f}^*_j\right)^2\right] = f_j^2 + \frac{\sigma_j^2}{{}^{[I-j-1]}S_j}$ for $j = 0, \ldots, J-1$.

Proof. (i) We have that the $\tilde{\varepsilon}_{i,j}$ are independent, given $\mathcal{B}_0 \subseteq \mathcal{D}_I$. From eq. (3.32) we can see that the $\widehat{f}^*_j$ are real-valued functions of the $\tilde{\varepsilon}_{i,j}$ and deterministic variables, which immediately implies that (i) holds true.

(ii) For $j = 0, \ldots, J-1$ we have

$$\mathbb{E}^*_{\mathcal{D}_I}\left[\widehat{f}^*_j\right] = f_j + \frac{\sigma_j}{{}^{[I-j-1]}S_j} \sum_{i=0}^{I-j-1} \sqrt{C_{i,j}}\ \mathbb{E}^*_{\mathcal{D}_I}\left[\tilde{\varepsilon}_{i,j+1}\right] = f_j,$$

since $\mathbb{E}\left[\varepsilon_{i,j+1}|\mathcal{B}_0\right] = 0$ $(\triangle)$.

(iii) For $j = 0, \ldots, J-1$ and using the conditional independence of $\tilde{\varepsilon}_{i,j}(\circ)$ and $\mathbb{E}\left[\varepsilon^2_{i,j+1}|\mathcal{B}_0\right] = 1$ $(*)$ we have

$$\begin{aligned}
&\mathbb{E}^*_{\mathcal{D}_I}\left[\left(\widehat{f}^*_j\right)^2\right] = \\
&= f_j^2 + \mathbb{E}^*_{\mathcal{D}_I}\left[\left(\frac{\sigma_j}{{}^{[I-j-1]}S_j} \sum_{i=0}^{I-j-1} \sqrt{C_{i,j}}\ \tilde{\varepsilon}_{i,j+1}\right)^2\right] + 2\ \mathbb{E}^*_{\mathcal{D}_I}\left[\frac{f_j\ \sigma_j}{{}^{[I-j-1]}S_j} \sum_{i=0}^{I-j-1} \sqrt{C_{i,j}}\ \tilde{\varepsilon}_{i,j+1}\right] \\
&= f_j^2 + \frac{\sigma_j^2}{\left({}^{[I-j-1]}S_j\right)^2}\ \mathbb{E}^*_{\mathcal{D}_I}\left[\left(\sum_{i=0}^{I-j-1} \sqrt{C_{i,j}}\ \tilde{\varepsilon}_{i,j+1}\right)^2\right] + \frac{2\ f_j\ \sigma_j}{{}^{[I-j-1]}S_j} \sum_{i=0}^{I-j-1} \sqrt{C_{i,j}}\ \mathbb{E}^*_{\mathcal{D}_I}\left[\tilde{\varepsilon}_{i,j+1}\right] \\
&\overset{(\triangle)}{=} f_j^2 + \frac{\sigma_j^2}{\left({}^{[I-j-1]}S_j\right)^2}\ \mathbb{V}^*_{\mathcal{D}_I}\left[\sum_{i=0}^{I-j-1} \sqrt{C_{i,j}}\ \tilde{\varepsilon}_{i,j+1}\right] \\
&\overset{(\circ)}{=} f_j^2 + \frac{\sigma_j^2}{\left({}^{[I-j-1]}S_j\right)^2} \sum_{i=0}^{I-j-1} C_{i,j}\ \mathbb{V}^*_{\mathcal{D}_I}[\tilde{\varepsilon}_{i,j+1}] \\
&\overset{(*)}{=} f_j^2 + \frac{\sigma_j^2}{\left({}^{[I-j-1]}S_j\right)^2}\ {}^{[I-j-1]}S_j
\end{aligned}$$

$$= f_j^2 + \frac{\sigma_j^2}{{}^{[I-j-1]}S_j}$$

❑

Using this and the fact that for independent random variables $X_1, \ldots, X_n$ we have

$$\mathbb{V}\left[\prod_{j=1}^{n} X_j\right] = \prod_{j=1}^{n} \mathbb{E}\left[X_j^2\right] - \prod_{j=1}^{n} \left(\mathbb{E}\left[X_j\right]\right)^2 \tag{3.33}$$

we can resample the estimation error

$$\left(\widehat{C}_{i,J}^{CL} - \mathbb{E}\left[C_{i,J}|\mathcal{D}_I\right]\right)^2 = C_{i,I-i}^2 \left(\widehat{f}_{I-i} \cdots \widehat{f}_{J-1} - f_{I-i} \cdots f_{J-1}\right)^2$$

by calculating

$$\begin{aligned}
&C_{i,I-i}^2 \mathbb{E}_{\mathcal{D}_I}^* \left[\left(\widehat{f}_{I-i}^* \cdots \widehat{f}_{J-1}^* - f_{I-i} \cdots f_{J-1}\right)^2\right] = \\
&= C_{i,I-i}^2 \mathbb{V}_{\mathcal{D}_I}^* \left[\widehat{f}_{I-i}^* \cdots \widehat{f}_{J-1}^*\right] \\
&\overset{(3.33)}{=} C_{i,I-i}^2 \left(\prod_{j=I-i}^{J-1} \mathbb{E}_{\mathcal{D}_I}^* \left[\left(\widehat{f}_j^*\right)^2\right] - \prod_{j=I-i}^{J-1} f_j^2\right) \\
&= C_{i,I-i}^2 \left(\prod_{j=I-i}^{J-1} \left(f_j^2 + \frac{\sigma_j^2}{{}^{[I-j-1]}S_j}\right) - \prod_{j=I-i}^{J-1} f_j^2\right)
\end{aligned} \tag{3.34}$$

By replacing the unknown parameters by their estimates we can now define an estimator for the conditional estimation error.

3.3.14 Estimator. We estimate the conditional estimation error for accident year $i = 1, \ldots, I$ by

$$\begin{aligned}
\widehat{\mathbb{V}}\left[\widehat{C}_{i,J}^{CL}\middle|\mathcal{D}_I\right] &= \mathbb{E}_{\mathcal{D}_I}^* \left[\left(\widehat{C}_{i,J}^{CL} - \mathbb{E}\left[C_{i,J}|\mathcal{D}_I\right]\right)\right] \\
&\overset{(3.34)}{=} C_{i,I-i}^2 \left(\prod_{j=I-i}^{J-1} \left(\widehat{f}_j^2 + \frac{\widehat{\sigma}_j^2}{{}^{[I-j-1]}S_j}\right) - \prod_{j=I-i}^{J-1} \widehat{f}_j^2\right)
\end{aligned} \tag{3.35}$$

This equation can be rewritten in a recursive form. We obtain for $j = I-i+1, \ldots, J$

$$\begin{aligned}
\widehat{\mathbb{V}}\left[\widehat{C}_{i,j}^{CL}\middle|\mathcal{D}_I\right] &= \widehat{\mathbb{V}}\left[\widehat{C}_{i,j-1}^{CL}\middle|\mathcal{D}_I\right] \ \widehat{f}_{j-1}^2 + C_{i,I-i}^2 \ \frac{\widehat{\sigma}_{j-1}^2}{{}^{[I-j-1]}S_{j-1}} \prod_{m=I-i}^{j-2} \left(f_m^2 + \frac{\widehat{\sigma}_m^2}{{}^{[I-m-1]}S_m}\right) \\
&= \widehat{\mathbb{V}}\left[\widehat{C}_{i,j-1}^{CL}\middle|\mathcal{D}_I\right] \left(\widehat{f}_{j-1}^2 + \frac{\widehat{\sigma}_{j-1}^2}{{}^{[I-j-1]}S_{j-1}}\right) + C_{i,I-i}^2 \ \frac{\widehat{\sigma}_{j-1}^2}{{}^{[I-j-1]}S_{j-1}} \prod_{m=I-i}^{j-2} f_m^2
\end{aligned} \tag{3.36}$$

3.3.15 Estimator (Mean square error of prediction for single accident years)**.** Under model assumptions 3.3.10 we have the following estimator for the conditional MSEP of the ultimate claim for a single accident year $i = 1, \dots, I$

$$\widehat{\mathbf{msep}}_{C_{i,J}|\mathcal{D}_I}\left(\widehat{C}_{i,J}^{CL}\right) = \mathbb{E}\left[\left(\widehat{C}_{i,J}^{CL} - C_{i,J}\right)^2 \middle| \mathcal{D}_I\right]$$

$$\overset{(3.20)+(3.24)+(3.35)}{=} \left(\widehat{C}_{i,J}^{CL}\right)^2 \sum_{j=I-i}^{J-1} \frac{\widehat{\sigma}_j^2}{\widehat{f}_j^2\, \widehat{C}_{i,j}^{CL}} + C_{i,I-i}^2 \left(\prod_{j=I-i}^{J-1}\left(\widehat{f}_j^2 + \frac{\widehat{\sigma}_j^2}{{}^{[I-j-1]}S_j}\right) - \prod_{j=I-i}^{J-1} \widehat{f}_j^2\right) \tag{3.37}$$

By observing $\left(\widehat{C}_{i,J}^{CL}\right)^2 = C_{i,I-i}^2 \prod_{j=I-i}^{J-1} \widehat{f}_j^2$ we can rewrite eq. (3.37) as

$$\widehat{\mathbf{msep}}_{C_{i,J}|\mathcal{D}_I}\left(\widehat{C}_{i,J}^{CL}\right) = \left(\widehat{C}_{i,J}^{CL}\right)^2 \left(\sum_{j=I-i}^{J-1} \frac{\widehat{\sigma}_j^2}{\widehat{f}_j^2\, \widehat{C}_{i,j}^{CL}} + \prod_{j=I-i}^{J-1}\left(1 + \frac{\widehat{\sigma}_j^2}{{}^{[I-j-1]}S_j}\right) - 1\right) \tag{3.38}$$

To get the Mack formula which we will need later on we do a linear approximation from below. We have

$$\begin{aligned}
\prod_{j=I-i}^{J-1}\left(\widehat{f}_j^2 + \frac{\widehat{\sigma}_j^2}{{}^{[I-j-1]}S_j}\right) - \prod_{j=I-i}^{J-1} \widehat{f}_j^2 &= \prod_{j=I-i}^{J-1} \widehat{f}_j^2\left(1 + \frac{\widehat{\sigma}_j^2}{\widehat{f}_j^2\, {}^{[I-j-1]}S_j}\right) - \prod_{j=I-i}^{J-1} \widehat{f}_j^2 \\
&= \prod_{j=I-i}^{J-1} \widehat{f}_j^2 \prod_{j=I-i}^{J-1}\left(1 + \frac{\widehat{\sigma}_j^2}{\widehat{f}_j^2\, {}^{[I-j-1]}S_j}\right) - \prod_{j=I-i}^{J-1} \widehat{f}_j^2
\end{aligned}$$

If we drop all cross products from $\prod_{j=I-i}^{J-1}\left(1 + \frac{\widehat{\sigma}_j^2}{\widehat{f}_j^2\, {}^{[I-j-1]}S_j}\right)$, which actually means developing the Taylor series around point $a = 0$, we have

$$\prod_{j=I-i}^{J-1}\left(1 + \frac{\widehat{\sigma}_j^2}{\widehat{f}_j^2\, {}^{[I-j-1]}S_j}\right) \approx 1 + \sum_{j=I-i}^{J-1} \frac{\widehat{\sigma}_j^2}{\widehat{f}_j^2\ {}^{[I-j-1]}S_j} \tag{3.39}$$

So we can conclude that

$$\begin{aligned}
&C_{i,I-i}^2\left(\prod_{j=I-i}^{J-1}\left(\widehat{f}_j^2 + \frac{\widehat{\sigma}_j^2}{{}^{[I-j-1]}S_j}\right) - \prod_{j=I-i}^{J-1} \widehat{f}_j^2\right) \\
&\approx C_{i,I-i}^2\left(\prod_{j=I-i}^{J-1} \widehat{f}_j^2\left(1 + \sum_{j=I-i}^{J-1} \frac{\widehat{\sigma}_j^2}{\widehat{f}_j^2\ {}^{[I-j-1]}S_j} - 1\right)\right) \\
&= C_{i,I-i}^2 \prod_{j=I-i}^{J-1} \widehat{f}_j^2 \sum_{j=I-i}^{J-1} \frac{\widehat{\sigma}_j^2}{\widehat{f}_j^2\ {}^{[I-j-1]}S_j}
\end{aligned}$$

which leads us to the following estimator.

3.3.16 Estimator (Mack Estimator for the MSEP of single accident years)**.** Under model assumptions 3.3.10 we have the following Mack estimator for the conditional MSEP of the ultimate claim for a single accident year $i = 1, \ldots, I$

$$\widehat{\mathbf{msep}}_{C_{i,J}|\mathcal{D}_I}\left(\widehat{C}^{CL}_{i,J}\right) := \left(\widehat{C}^{CL}_{i,J}\right)^2 \sum_{j=I-i}^{J-1} \frac{\widehat{\sigma}^2_j}{\widehat{f}^2_j}\left(\frac{1}{\widehat{C}^{CL}_{i,j}} + \frac{1}{{}^{[I-j-1]}S_j}\right) \tag{3.40}$$

3.3.17 Remark. Note that the $\approx$ in eq. (3.39) is in fact a $\geq$, so $\widehat{\mathbf{msep}}_{C_{i,J}|\mathcal{D}_I}\left(\widehat{C}^{CL}_{i,J}\right)$ is a lower bound of $\widehat{\mathbf{msep}}_{C_{i,J}|\mathcal{D}_I}\left(\widehat{C}^{CL}_{i,J}\right)$.

3.3.3 MSEP for aggregated accident years

While model assumptions 3.3.10 state that different accident years are independent we still have to take dependencies into account when aggregating the factors $\widehat{C}^{CL}_{i,j}$, since they are estimated from the same set of data ($\widehat{f}_j$ and $\widehat{\sigma}_j$ do not depend on i). We start with the case of two aggregated variables and compute for $0 \leq i \neq k \leq I$

$$\begin{aligned}\widehat{\mathbf{msep}}_{C_{i,J}+C_{k,J}|\mathcal{D}_I}\left(\widehat{C}^{CL}_{i,j} + \widehat{C}^{CL}_{k,j}\right) &= \mathbb{E}\left[\left(\widehat{C}^{CL}_{i,j} + \widehat{C}^{CL}_{k,j} - (C_{i,J} + C_{k,j})\right)^2 \middle| \mathcal{D}_I\right] \\ &\overset{(3.10)}{=} \mathbb{V}\left[C_{i,J} + C_{k,j}|\mathcal{D}_I\right] + \left(\widehat{C}^{CL}_{i,j} + \widehat{C}^{CL}_{k,j} - \mathbb{E}\left[C_{i,J} + C_{k,j}|\mathcal{D}_I\right]\right)^2\end{aligned}$$

Using the independence of different accident years the first term is simply

$$\mathbb{V}\left[C_{i,J} + C_{k,j}|\mathcal{D}_I\right] = \mathbb{V}\left[C_{i,J}|\mathcal{D}_I\right] + \mathbb{V}\left[+C_{k,j}|\mathcal{D}_I\right]$$

where we already have estimators for the individual terms. For the second term we get

$$\begin{aligned}\left(\widehat{C}^{CL}_{i,j} + \widehat{C}^{CL}_{k,j} - \mathbb{E}\left[C_{i,J} + C_{k,j}|\mathcal{D}_I\right]\right)^2 &= \left(\widehat{C}^{CL}_{i,j} - \mathbb{E}\left[C_{i,J}|\mathcal{D}_I\right]\right)^2 + \left(\widehat{C}^{CL}_{k,j} - \mathbb{E}\left[C_{k,j}|\mathcal{D}_I\right]\right)^2 \\ &\quad + 2\left(\widehat{C}^{CL}_{i,j} - \mathbb{E}\left[C_{i,J}|\mathcal{D}_I\right]\right)\left(\widehat{C}^{CL}_{k,j} - \mathbb{E}\left[C_{k,j}|\mathcal{D}_I\right]\right)\end{aligned}$$

Combining these two equations yields

$$\begin{aligned}\widehat{\mathbf{msep}}_{C_{i,J}+C_{k,J}|\mathcal{D}_I}\left(\widehat{C}^{CL}_{i,j} + \widehat{C}^{CL}_{k,j}\right) &= \mathbb{E}\left[\left(\widehat{C}^{CL}_{i,j} - C_{i,J}\right)^2 \middle| \mathcal{D}_I\right] + \mathbb{E}\left[\left(\widehat{C}^{CL}_{k,j} - C_{k,j}\right)^2 \middle| \mathcal{D}_I\right] \\ &\quad + 2\left(\widehat{C}^{CL}_{i,j} - \mathbb{E}\left[C_{i,J}|\mathcal{D}_I\right]\right)\left(\widehat{C}^{CL}_{k,j} - \mathbb{E}\left[C_{k,j}|\mathcal{D}_I\right]\right) \\ &= \widehat{\mathbf{msep}}_{C_{i,J}|\mathcal{D}_I}\left(\widehat{C}^{CL}_{i,j}\right) + \widehat{\mathbf{msep}}_{C_{k,J}|\mathcal{D}_I}\left(+\widehat{C}^{CL}_{k,j}\right) \\ &\quad + 2\left(\widehat{C}^{CL}_{i,j} - \mathbb{E}\left[C_{i,J}|\mathcal{D}_I\right]\right)\left(\widehat{C}^{CL}_{k,j} - \mathbb{E}\left[C_{k,j}|\mathcal{D}_I\right]\right)\end{aligned}$$

We will now resample the development factors to get an analogous expression to eq. (3.34).

Under $\mathbb{P}^*_{\mathcal{D}_I}$ we resample the terms

$$\left(\widehat{C}^{CL}_{i,J} - \mathbb{E}\left[C_{i,J}|\mathcal{D}_I\right]\right)\left(\widehat{C}^{CL}_{k,j} - \mathbb{E}\left[C_{k,j}|\mathcal{D}_I\right]\right) = C_{i,I-i}\left(\widehat{f}_{I-i}\cdots\widehat{f}_{J-1} - f_{I-i}\cdots f_{J-1}\right)$$
$$C_{k,I-k}\left(\widehat{f}_{I-k}\cdots\widehat{f}_{J-1} - f_{I-k}\cdots f_{J-1}\right)$$

by calculating (without loss of generality we set $i < k$)

$$\begin{aligned}
&C_{i,I-i}\, C_{k,I-k}\, \mathbb{E}^*_{\mathcal{D}_I}\left[\left(\prod_{j=I-i}^{J-1}\widehat{f}^*_j - \prod_{j=I-i}^{J-1} f_j\right)\left(\prod_{j=I-k}^{J-1}\widehat{f}^*_j - \prod_{j=I-k}^{J-1} f_j\right)\right]\\
&= C_{i,I-i}\, C_{k,I-k}\, \mathbb{E}^*_{\mathcal{D}_I}\Bigg[\prod_{j=I-i}^{J-1}\widehat{f}^*_j\prod_{j=I-k}^{J-1}\widehat{f}^*_j - \prod_{j=I-i}^{J-1} f_j\prod_{j=I-k}^{J-1}\widehat{f}^*_j\\
&\qquad - \prod_{j=I-i}^{J-1}\widehat{f}^*_j\prod_{j=I-k}^{J-1} f_j + \prod_{j=I-i}^{J-1} f_j\prod_{j=I-k}^{J-1} f_j\Bigg]\\
&= C_{i,I-i}\, C_{k,I-k}\Bigg[\prod_{j=I-k}^{I-i-1}\mathbb{E}\left[\widehat{f}^*_j\right]\prod_{j=I-i}^{J-1}\mathbb{E}\left[\left(\widehat{f}^*_j\right)^2\right] - \prod_{j=I-i}^{J-1} f_j\prod_{j=I-k}^{J-1}\mathbb{E}\left[\widehat{f}^*_j\right]\\
&\qquad - \prod_{j=I-i}^{J-1}\mathbb{E}\left[\widehat{f}^*_j\right]\prod_{j=I-k}^{J-1} f_j + \prod_{j=I-i}^{J-1} f_j\prod_{j=I-k}^{J-1} f_j\Bigg]\\
&= C_{i,I-i}\, C_{k,I-k}\prod_{j=I-k}^{I-i-1} f_j\Bigg[\prod_{j=I-i}^{J-1}\mathbb{E}\left[\left(\widehat{f}^*_j\right)^2\right] - \prod_{j=I-i}^{J-1} f^2_j - \prod_{j=I-i}^{J-1} f^2_j + \prod_{j=I-i}^{J-1} f^2_j\Bigg]\\
&= C_{i,I-i}\mathbb{E}\left[C_{k,I-i}|D_I\right]\left(\prod_{j=I-i}^{J-1}\left(f^2_j + \frac{\sigma^2_j}{{}^{[I-j-1]}S_j}\right) - \prod_{j=I-i}^{J-1} f^2_j\right)
\end{aligned}\tag{3.41}$$

Plugging in our estimators we get

3.3.18 Estimator (Mean square error of prediction for aggregated accident years)**.** Under model assumptions 3.3.10 we have the following estimator for the conditional MSEP of the ultimate claim for aggregated accident years $i = 1, \dots, I$

$$\begin{aligned}
\widehat{\mathbf{msep}}_{\sum_i C_{i,J}|\mathcal{D}_I}\left(\sum_{i=1}^{I}\widehat{C}^{CL}_{i,J}\right) &= \mathbb{E}\left[\left(\sum_{i=1}^{I}\widehat{C}^{CL}_{i,J} - \sum_{i=1}^{I} C_{i,J}\right)^2\middle|\mathcal{D}_I\right]\\
&\overset{(3.41)}{=} \sum_{i=1}^{I}\widehat{\mathbf{msep}}_{C_{i,J}|\mathcal{D}_I}\left(\widehat{C}^{CL}_{i,J}\right)\\
&\quad + 2\sum_{1\le i<k\le I} C_{i,I-i}\,\widehat{C}^{CL}_{k,I-i}\left(\prod_{j=I-i}^{J-1}\left(\widehat{f}^2_j + \frac{\widehat{\sigma}^2_j}{{}^{[I-j-1]}S_j}\right) - \prod_{j=I-i}^{J-1}\widehat{f}^2_j\right)
\end{aligned}\tag{3.42}$$

Analogously as in the derivation of eq. (3.39) we can do a linear approximation from below to get the Mack Estimator presented in [Mac93].

3.3.19 Estimator (Mack MSEP for aggregated accident years)**.** Under model assumptions 3.3.10 we have the following estimator for the conditional MSEP of the ultimate claim for aggregated accident years $i = 1, \ldots, I$

$$\widehat{\mathbf{msep}}_{\sum_i C_{i,J}|\mathcal{D}_I}\left(\sum_{i=1}^{I} \widehat{C}_{i,J}^{CL}\right) = \sum_{i=1}^{I} \widehat{\mathbf{msep}}_{C_{i,J}|\mathcal{D}_I}\left(\widehat{C}_{i,J}^{CL}\right)$$
$$+ 2 \sum_{1 \le i < k \le I} \left(\widehat{C}_{i,J}^{CL}\ \widehat{C}_{k,J}^{CL} \sum_{j=I-i}^{J-1} \frac{\widehat{\sigma}_j^2}{\widehat{f}_j^{2\,[I-j-1]} S_j} \right) \tag{3.43}$$

Chapter 4

Estimation of skewness and kurtosis [1]

Having established estimators for the first two moments in the previous chapter we now want to estimate the skewness and the kurtosis of our claim reserves. Therefore we will first expand the model assumptions and then find an appropriate estimator. Higher-order moments often give us a more intricate look at the distribution of our random variable. Especially when it comes to the tails of the distribution (which are often needed for Value at Risk calculations), we can find better distributional fits than with just the first two moments.

4.1 Estimation of the Skewness

4.1.1 Definition (Conditional Skewness)**.** Following the notation from chapter 2 we define the conditional third moment of a random variable A, given B as

$$\mu^{(3)}(A|B) = \mathbb{E}\left[\left(A - \mathbb{E}[A|B]\right)^3 \middle| B\right] \tag{4.1}$$

Then the conditional skewness is given by

$$\text{Skew}(A|B) = \frac{\mu^{(3)}(A|B)}{\left(\mathbb{V}[A|B]\right)^{\frac{3}{2}}} \tag{4.2}$$

Next we will expand our model assumptions to fit higher order moments.

4.1.2 Model assumptions.

- Cumulative claims $C_{i,j}$ of different accident years i are independent.

[1] The structure and notation of this chapter follow [Mor12] and [Mor13], with the proofs being newly performed and resulting in slightly different estimators.

- $(C_{i,j})_{j\geq 0}$ form a Markov chain. There exist factors $f_0,\dots,f_{J-1}>0$ and variance parameters $\sigma_0^2,\dots,\sigma_{J-1}^2>0$ such that

$$\begin{aligned} &\forall\ 0\leq i\leq I \text{ and } \forall\ 1\leq j\leq J \text{ we have}\\ &\mathbb{E}\left[C_{i,j}|C_{i,j-1}\right]=f_{j-1}C_{i,j-1}, &(4.3)\\ &\mathbb{V}\left[C_{i,j}|C_{i,j-1}\right]=\sigma_{j-1}^2C_{i,j-1}. &(4.4)\end{aligned}$$

- Skew $(C_{i,j+1}|\mathcal{D}_I)$ depends on j but does not depend on i.

From the last point in model assumptions 4.1.2 we can deduce

$$\begin{aligned} \forall\ j\in\{0,\dots,J-1\}\ \exists\ \gamma_j\ \forall\ i\in\{0,\dots,I\}:&\\ \gamma_j &= \text{Skew}\left(C_{i,j+1}|\mathcal{D}_I\right)\\ &= \frac{\mu^{(3)}\left(C_{i,j+1}|\mathcal{D}_I\right)}{\left(\mathbb{V}\left[C_{i,j+1}|\mathcal{D}_I\right]\right)^{\frac{3}{2}}}\\ &\overset{(4.4)}{=} \frac{\mu^{(3)}\left(C_{i,j+1}|\mathcal{D}_I\right)}{\left(\sigma_j^2\ C_{i,j}\right)^{\frac{3}{2}}}\end{aligned}$$

This can be rearranged to

$$\mu^{(3)}\left(C_{i,j+1}|\mathcal{D}_I\right)=\gamma_j\left(\sigma_j^2\ C_{i,j}\right)^{\frac{3}{2}}, \tag{4.5}$$

which means that the third moment of $C_{i,j+1}$ is proportional to $(C_{i,j})^{\frac{3}{2}}$.
We will now use the special form of the above equation to define a modification of the skewness in which we will incorporate the variance factors σ_j^2. From eq. (4.5) we have that there exist factors γ_j^* so that

$$\begin{aligned}\mu^{(3)}\left(C_{i,j+1}|\mathcal{D}_I\right)&=\gamma_j^*\ \left(C_{i,j}\right)^{\frac{3}{2}} &(4.6)\\ \forall i=0,\dots,I \text{ and } &\forall j=0,\dots,I-1\end{aligned}$$

4.1.3 Remark. Note that

$$\gamma_j=\frac{\gamma_j^*}{\left(\sigma_j^2\right)^{\frac{3}{2}}},\qquad j=0,\dots,J-1 \tag{4.7}$$

Before the next formula we want to generalize the short notation from eq. (3.31) to simplify the notation in the proof.

4.1.4 Definition. We set

$$^{[k]}S_j^{[p]}:=\sum_{i=0}^{k}C_{i,j}^p \tag{4.8}$$

and in accordance with eq. (3.31) we set

$$^{[k]}S_j={}^{[k]}S_j^{[1]}=\sum_{i=0}^{k}C_{i,j} \tag{4.9}$$

4.1.5 Estimator. We will estimate γ_j^* for $j = 0, \dots, J-3$ by

$$\widehat{\gamma}_j^* = \frac{\sum_{i=0}^{I-j-1} C_{i,j}^{3/2}\left(\frac{C_{i,j+1}}{C_{i,j}} - \widehat{f}_j\right)^3}{\sum_{i=0}^{I-j-1}\left(1 - \frac{C_{i,j}}{\sum_{k=0}^{I-j-1} C_{k,j}}\right)^3 + \frac{\sum_{k=0}^{I-j-1} C_{k,j}^3 - \left(\sum_{k=0}^{I-j-1} C_{k,j}^{3/2}\right)^2}{\left(\sum_{k=0}^{I-j-1} C_{k,j}\right)^3}} \tag{4.10}$$

4.1.6 Remark. Note that estimator 4.1.5 differs from the one given in [Mor12], which is

$$\widehat{\gamma}_j^{\mathrm{DM}} = \frac{1}{I - j - \frac{\left(\sum_{k=0}^{I-j-1} C_{k,j}^{3/2}\right)^2}{\left(\sum_{k=0}^{I-j-1} C_{k,j}\right)^3}} \sum_{i=0}^{I-j-1} C_{i,j}^{3/2}\left(\frac{C_{i,j+1}}{C_{i,j}} - \widehat{f}_j\right)^3$$

The form of eq. (4.10) results from the proof of lemma 4.1.7 and is very similar in form to eq. (4.26). Therefore we will use estimator 4.1.5 from here on out.

4.1.7 Lemma. *The estimator $\widehat{\gamma}^*$ is a conditionally (and unconditionally) unbiased estimator for γ^*, which means that*

(i) $\mathbb{E}\left[\widehat{\gamma}_j^* \middle| \mathcal{B}_j\right] = \gamma^*$ *for* $j = 0, \dots, J-3$

(ii) $\mathbb{E}\left[\widehat{\gamma}_j^*\right] = \gamma^*$ *for* $j = 0, \dots, J-3$

where $\mathcal{B}_j = \{C_{i,k} : i + k \leq I, 0 \leq k \leq j\} \subseteq \mathcal{D}_I$ *(see eq. (3.9)).*

Proof. To simplify the notation we define

$$\mathbb{E}_j\left[\cdot\right] \coloneqq \mathbb{E}\left[\cdot|\mathcal{B}_j\right] \quad \text{and} \quad \mathbb{V}_j\left[\cdot\right] \coloneqq \mathbb{V}\left[\cdot|\mathcal{B}_j\right] \tag{4.11}$$

which means that all of the expectations are taken on the conditional probability measures $\mathbb{P}\left(\cdot|\mathcal{B}_j\right)$, where $j = 0\dots, J-3$.

(i) For $j = 0, \dots, J-3$ arbitrary but fixed we start by calculating

$$\mathbb{E}_j\left[\sum_{i=0}^{I-j-1} C_{i,j}^{3/2}\left(\frac{C_{i,j+1}}{C_{i,j}} - \widehat{f}_j\right)^3\right] = \underbrace{\sum_{i=0}^{I-j-1} \frac{1}{C_{i,j}^{3/2}}\mathbb{E}_j\left[C_{i,j+1}^3\right]}_{(a)} - \underbrace{3\sum_{i=0}^{I-j-1} \frac{1}{\sqrt{C_{i,j}}}\mathbb{E}_j\left[\widehat{f}_j C_{i,j+1}^2\right]}_{(b)}$$
$$+ \underbrace{3\sum_{i=0}^{I-j-1} \sqrt{C_{i,j}}\mathbb{E}_j\left[\widehat{f}_j^2 C_{i,j+1}\right]}_{(c)} - \underbrace{\sum_{i=0}^{I-j-1} C_{i,j}^{3/2}\mathbb{E}_j\left[\widehat{f}_j^3\right]}_{(d)}$$

We will now calculate each element of the above equation individually:

(a) We have

$$
\begin{aligned}
\gamma_j^* \, C_{i,j}^{3/2} &\overset{(4.6)}{=} \mu^{(3)}\left(C_{i,j+1}|\mathcal{D}_I\right) \\
&= \mathbb{E}_j\left[C_{i,j+1}^3\right] - 3\mathbb{E}_j\left[C_{i,j+1}^2\right]\mathbb{E}_j\left[C_{i,j+1}\right] + 3\mathbb{E}_j\left[C_{i,j+1}\right]\mathbb{E}_j\left[C_{i,j+1}\right]^2 - \mathbb{E}_j\left[C_{i,j+1}\right]^3 \\
&= \mathbb{E}_j\left[C_{i,j+1}^3\right] - 3\left(\mathbb{V}_j\left[C_{i,j+1}\right] + \mathbb{E}_j\left[C_{i,j+1}\right]^2\right)\mathbb{E}_j\left[C_{i,j+1}\right] + 2\mathbb{E}_j\left[C_{i,j+1}\right]^3 \\
&\overset{(4.3)+(4.4)}{=} \mathbb{E}_j\left[C_{i,j+1}^3\right] - 3\left(\sigma_j^2 C_{i,j} + f_j^2 C_{i,j}^2\right) f_j C_{i,j} + 2 f_j^3 C_{i,j}^3 \\
&= \mathbb{E}_j\left[C_{i,j+1}^3\right] - 3 f_j \sigma_j^2 C_{i,j}^2 - f_j^3 C_{i,j}^3
\end{aligned}
$$

which yields

$$
\textbf{(a)} = \sum_{i=0}^{I-j-1} \frac{1}{C_{i,j}^{3/2}} \left(\gamma_j^* \, C_{i,j}^{3/2} + 3 f_j \sigma_j^2 C_{i,j}^2 + f_j^3 C_{i,j}^3\right) \tag{4.12}
$$

(b) By inserting the definition of $\widehat{f}_j$ and using the independence between different accident years $(*)$ we have

$$
\begin{aligned}
\mathbb{E}_j\left[\widehat{f}_j C_{i,j+1}^2\right] &\overset{(3.6)}{=} \underbrace{\frac{1}{\sum_{k=0}^{I-j-1} C_{k,j}}}_{=:L} \mathbb{E}_j\left[\sum_{\substack{k=0\\k\neq i}}^{I-j-1} C_{k,j+1} C_{i,j+1}^2 + C_{i,j+1}^3\right] \\
&\overset{(*)}{=} L\left(\mathbb{E}_j\left[C_{i,j+1}^2\right]\sum_{\substack{k=0\\k\neq i}}^{I-j-1} \mathbb{E}_j\left[C_{k,j+1}\right] + \mathbb{E}_j\left[C_{i,j+1}^3\right]\right) \\
&\overset{(4.12)+(4.3)+(4.4)}{=} L\left(\left(\sigma_j^2 C_{i,j} + f_j^2 C_{i,j}^2\right)\sum_{\substack{k=0\\k\neq i}}^{I-j-1}\left(f_j C_{k,j}\right) + \gamma_j^* \, C_{i,j}^{3/2} + 3 f_j \sigma_j^2 C_{i,j}^2 + f_j^3 C_{i,j}^3\right) \\
&= L\left(f_j \sigma_j^2 C_{i,j}\sum_{\substack{k=0\\k\neq i}}^{I-j-1} C_{k,j} + f_j^3 C_{i,j}^2 \sum_{\substack{k=0\\k\neq i}}^{I-j-1} C_{k,j} + \gamma_j^* \, C_{i,j}^{3/2} + 3 f_j \sigma_j^2 C_{i,j}^2 + f_j^3 C_{i,j}^3\right) \\
&= L\left(f_j \sigma_j^2 C_{i,j}\sum_{k=0}^{I-j-1} C_{k,j} + f_j^3 C_{i,j}^2 \sum_{k=0}^{I-j-1} C_{k,j} + \gamma_j^* \, C_{i,j}^{3/2} + 2 f_j \sigma_j^2 C_{i,j}^2\right) \\
&= f_j \sigma_j^2 C_{i,j} + f_j^3 C_{i,j}^2 + L\left(\gamma_j^* \, C_{i,j}^{3/2} + 2 f_j \sigma_j^2 C_{i,j}^2\right)
\end{aligned}
$$

so we have

$$
\textbf{(b)} = 3 \sum_{i=0}^{I-j-1} \frac{1}{\sqrt{C_{i,j}}} \left(f_j \sigma_j^2 C_{i,j} + f_j^3 C_{i,j}^2 + \frac{\gamma_j^* \, C_{i,j}^{3/2}}{\sum_{k=0}^{I-j-1} C_{k,j}} + \frac{2 f_j \sigma_j^2 C_{i,j}^2}{\sum_{k=0}^{I-j-1} C_{k,j}}\right) \tag{4.13}
$$

(c) We again insert the definition of $\widehat{f}_j$ and use the independence between different accident years $(*)$ to get

$$
\begin{aligned}
\mathbb{E}_j\left[\widehat{f}_j^2 C_{i,j+1}\right] &\overset{(3.6)}{=} \underbrace{\frac{1}{\left(\sum_{k=0}^{I-j-1} C_{k,j}\right)^2}}_{=:M} \mathbb{E}_j\left[\left(\sum_{k=0}^{I-j-1} C_{k,j+1}\right)^2 C_{i,j+1}\right] \\
&= M\ \mathbb{E}_j\left[C_{i,j+1}\left(\sum_{k=0}^{I-j-1} C_{k,j+1}^2 + 2\sum_{k=1}^{I-j-1}\sum_{n=0}^{k-1} C_{k,j+1}\ C_{n,j+1}\right)\right] \\
&= M\ \mathbb{E}_j\Bigg[C_{i,j+1}^3 + \sum_{\substack{k=0\\k\neq i}}^{I-j-1}\left(C_{k,j+1}^2 C_{i,j+1}\right) + 2\sum_{\substack{k=1\\k\neq i}}^{I-j-1}\sum_{n=0}^{k-1}\left(C_{k,j+1}\ C_{n,j+1}\ C_{i,j+1}\right) \\
&\qquad + 2\sum_{n=0}^{i-1}\left(C_{n,j+1}\ C_{i,j+1}^2\right)\Bigg] \\
&= M\ \mathbb{E}_j\Bigg[C_{i,j+1}^3 + \sum_{\substack{k=0\\k\neq i}}^{I-j-1}\left(C_{k,j+1}^2 C_{i,j+1}\right) + 2\sum_{\substack{k=1\\k\neq i}}^{I-j-1}\sum_{\substack{n=0\\n\neq i}}^{k-1}\left(C_{k,j+1}\ C_{n,j+1}\ C_{i,j+1}\right) \\
&\qquad + 2\sum_{n=0}^{i-1}\left(C_{n,j+1}\ C_{i,j+1}^2\right) + 2\sum_{k=i+1}^{I-j-1}\left(C_{k,j+1}\ C_{i,j+1}^2\right)\Bigg] \\
&= M\ \mathbb{E}_j\Bigg[C_{i,j+1}^3 + \sum_{\substack{k=0\\k\neq i}}^{I-j-1}\left(C_{k,j+1}^2 C_{i,j+1}\right) + 2\sum_{\substack{k=1\\k\neq i}}^{I-j-1}\sum_{\substack{n=0\\n\neq i}}^{k-1}\left(C_{k,j+1}\ C_{n,j+1}\ C_{i,j+1}\right) \\
&\qquad + 2\sum_{\substack{k=0\\k\neq i}}^{I-j-1}\left(C_{k,j+1}\ C_{i,j+1}^2\right)\Bigg] \\
&\overset{(*)}{=} M\Bigg(\mathbb{E}_j\left[C_{i,j+1}^3\right] + \sum_{\substack{k=0\\k\neq i}}^{I-j-1}\left(\mathbb{E}_j\left[C_{k,j+1}^2\right]\mathbb{E}_j\left[C_{i,j+1}\right]\right) \\
&\qquad + 2\sum_{\substack{k=1\\k\neq i}}^{I-j-1}\sum_{\substack{n=0\\n\neq i}}^{k-1}\left(\mathbb{E}_j\left[C_{k,j+1}\right]\mathbb{E}_j\left[C_{n,j+1}\right]\mathbb{E}_j\left[C_{i,j+1}\right]\right) \\
&\qquad + 2\sum_{\substack{k=0\\k\neq i}}^{I-j-1}\left(\mathbb{E}_j\left[C_{k,j+1}\right]\mathbb{E}_j\left[C_{i,j+1}^2\right]\right)\Bigg) \\
&= M\Bigg(\gamma_j^*\ C_{i,j}^{3/2} + 3f_j\sigma_j^2 C_{i,j}^2 + f_j^3 C_{i,j}^3 + f_j C_{i,j}\sum_{\substack{k=0\\k\neq i}}^{I-j-1}\left(\sigma_j^2 C_{k,j} + f_j^2 C_{k,j}^2\right) \\
&\qquad + 2f_j^3\sum_{\substack{k=1\\k\neq i}}^{I-j-1}\sum_{\substack{n=0\\n\neq i}}^{k-1}\left(C_{k,j}\ C_{n,j}\ C_{i,j}\right) + 2\left(\sigma_j^2 C_{i,j} + f_j^2 C_{i,j}^2\right) f_j\sum_{\substack{k=0\\k\neq i}}^{I-j-1} C_{k,j}\Bigg) \\
&= M\Bigg(\gamma_j^*\ C_{i,j}^{3/2} + f_j\sigma_j^2 C_{i,j}\Bigg(3\ C_{i,j} + \sum_{\substack{k=0\\k\neq i}}^{I-j-1} C_{k,j} + 2\sum_{\substack{k=0\\k\neq i}}^{I-j-1} C_{k,j}\Bigg)
\end{aligned}
$$

$$+ f_j^3 C_{i,j} \left(C_{i,j}^2 + \sum_{\substack{k=0\\k\neq i}}^{I-j-1} C_{k,j}^2 + 2 \sum_{\substack{k=1\\k\neq i}}^{I-j-1} \sum_{\substack{n=0\\n\neq i}}^{k-1} (C_{k,j}\, C_{n,j}) + C_{i,j} \sum_{\substack{k=0\\k\neq i}}^{I-j-1} C_{k,j} \right)\Bigg)$$

$$= M \left(\gamma_j^*\, C_{i,j}^{3/2} + 3 f_j \sigma_j^2 C_{i,j} \sum_{k=0}^{I-j-1} C_{k,j} + f_j^3 C_{i,j} \left(\sum_{k=0}^{I-j-1} C_{k,j}^2 + 2 \sum_{k=1}^{I-j-1} \sum_{n=0}^{k-1} (C_{k,j}\, C_{n,j}) \right) \right)$$

$$= M \left(\gamma_j^*\, C_{i,j}^{3/2} + 3 f_j \sigma_j^2 C_{i,j} \sum_{k=0}^{I-j-1} C_{k,j} + f_j^3 C_{i,j} \left(\sum_{k=0}^{I-j-1} C_{k,j} \right)^2 \right)$$

which yields

$$\textbf{(c)} = 3 \sum_{i=0}^{I-j-1} \frac{\sqrt{C_{i,j}}}{\left(\sum_{k=0}^{I-j-1} C_{k,j} \right)^2} \left(\gamma_j^*\, C_{i,j}^{3/2} + 3 f_j \sigma_j^2 C_{i,j} \sum_{k=0}^{I-j-1} C_{k,j} + f_j^3 C_{i,j} \left(\sum_{k=0}^{I-j-1} C_{k,j} \right)^2 \right) \tag{4.14}$$

(d) For the last element we will use eq. (3.19) which states

$$\mathbb{E}_j \left[\widehat{f}_j^2 \middle| \mathcal{B}_j \right] = \frac{\sigma_j^2}{\sum_{i=0}^{I-j-1} C_{i,j}} + f_j^2.$$

Additionally we will use the fact that for independent random variables $X_1, \ldots, X_n$ with expected value of zero and each with finite first three moments we have

$$\mathbb{E}_j \left[\left(\sum_{i=1}^{n} X_i \right)^3 \right] = \sum_{i=1}^{n} \mathbb{E}_j \left[X_i^3 \right] \tag{4.15}$$

This follows from the fact that all cross products contain a factor of the form $\mathbb{E}_j [X_i] (= 0)$ and are therefore zero.
We use this theorem on the centralized random variables $C_{i,j+1} - \mathbb{E}_j [C_{i,j+1}]$ to prove

$$\mathbb{E}_j \left[\left(\widehat{f}_j - \mathbb{E}_j \left[\widehat{f}_j \right] \right)^3 \right] \overset{(4.3)}{=} \mathbb{E}_j \left[\left(\frac{\sum_{i=0}^{I-j-1} C_{i,j+1}}{\sum_{i=0}^{I-j-1} C_{i,j}} - \mathbb{E}_j \left[\frac{\sum_{i=0}^{I-j-1} C_{i,j+1}}{\sum_{i=0}^{I-j-1} C_{i,j}} \right] \right)^3 \right]$$

$$= \frac{1}{\left(\sum_{i=0}^{I-j-1} C_{i,j} \right)^3} \mathbb{E}_j \left[\left(\sum_{i=0}^{I-j-1} C_{i,j+1} - \sum_{i=0}^{I-j-1} \mathbb{E}_j [C_{i,j+1}] \right)^3 \right]$$

$$= \frac{1}{\left(\sum_{i=0}^{I-j-1} C_{i,j} \right)^3} \mathbb{E}_j \left[\left(\sum_{i=0}^{I-j-1} C_{i,j+1} - \mathbb{E}_j [C_{i,j+1}] \right)^3 \right]$$

$$\overset{(4.15)}{=} \frac{1}{\left(\sum_{i=0}^{I-j-1} C_{i,j} \right)^3} \sum_{i=0}^{I-j-1} \mathbb{E}_j \left[(C_{i,j+1} - \mathbb{E}_j [C_{i,j+1}])^3 \right]$$

$$\overset{(4.6)}{=} \frac{\gamma^*}{\left(\sum_{i=0}^{I-j-1} C_{i,j} \right)^3} \sum_{i=0}^{I-j-1} C_{i,j}^{3/2} \tag{4.16}$$

Using all of the above we have

$$\begin{aligned}
\mathbb{E}_j\left[\widehat{f}_j^3\right] &= \mathbb{E}_j\left[\left(\widehat{f}_j - \mathbb{E}_j\left[\widehat{f}_j\right]\right)^3\right] + 3\mathbb{E}_j\left[\widehat{f}_j^2\right]\mathbb{E}_j\left[\widehat{f}_j\right] - 3\mathbb{E}_j\left[\widehat{f}_j\right]\mathbb{E}_j\left[\widehat{f}_j\right]^2 + \mathbb{E}_j\left[\widehat{f}_j\right]^3 \\
&= \mathbb{E}_j\left[\left(\widehat{f}_j - \mathbb{E}_j\left[\widehat{f}_j\right]\right)^3\right] + 3\mathbb{E}_j\left[\widehat{f}_j^2\right]\mathbb{E}_j\left[\widehat{f}_j\right] - 2\mathbb{E}_j\left[\widehat{f}_j\right]^3 \\
&\overset{(4.16)+(3.19)}{=} \frac{\gamma^*}{\left(\sum_{i=0}^{I-j-1} C_{i,j}\right)^3}\sum_{i=0}^{I-j-1} C_{i,j}^{3/2} + 3\left(\frac{\sigma_j^2}{\sum_{i=0}^{I-j-1} C_{i,j}} + f_j^2\right) f_j - 2f_j^3 \\
&= \frac{\gamma^*}{\left(\sum_{i=0}^{I-j-1} C_{i,j}\right)^3}\sum_{i=0}^{I-j-1} C_{i,j}^{3/2} + 3\frac{f_j\sigma_j^2}{\sum_{i=0}^{I-j-1} C_{i,j}} + f_j^3
\end{aligned}$$

which yields

$$\textbf{(d)} = \sum_{i=0}^{I-j-1} C_{i,j}^{3/2}\left(\frac{\gamma^*}{\left(\sum_{k=0}^{I-j-1} C_{k,j}\right)^3}\sum_{k=0}^{I-j-1} C_{k,j}^{3/2} + 3\frac{f_j\sigma_j^2}{\sum_{k=0}^{I-j-1} C_{k,j}} + f_j^3\right) \tag{4.17}$$

Putting everything together we have

$$\begin{aligned}
&\mathbb{E}_j\left[\sum_{i=0}^{I-j-1} C_{i,j}^{3/2}\left(\frac{C_{i,j+1}}{C_{i,j}} - \widehat{f}_j\right)^3\right] = (a) - (b) + (c) - (d) \\
&= \sum_{i=0}^{I-j-1}\Bigg[\frac{1}{C_{i,j}^{3/2}}\left(\gamma_j^*\, C_{i,j}^{3/2} + 3f_j\sigma_j^2 C_{i,j}^2 + f_j^3 C_{i,j}^3\right) \\
&\quad - \frac{3}{\sqrt{C_{i,j}}}\left(f_j\sigma_j^2 C_{i,j} + f_j^3 C_{i,j}^2 + \frac{\gamma_j^*\, C_{i,j}^{3/2}}{{}^{[I-j-1]}S_j} + \frac{2f_j\sigma_j^2 C_{i,j}^2}{{}^{[I-j-1]}S_j}\right) \\
&\quad + \frac{3\sqrt{C_{i,j}}}{\left({}^{[I-j-1]}S_j\right)^2}\left(\gamma_j^*\, C_{i,j}^{3/2} + 3f_j\sigma_j^2 C_{i,j}\,{}^{[I-j-1]}S_j + f_j^3 C_{i,j}\left({}^{[I-j-1]}S_j\right)^2\right) \\
&\quad - C_{i,j}^{3/2}\left(\frac{\gamma_j^*}{\left({}^{[I-j-1]}S_j\right)^3}\sum_{k=0}^{I-j-1} C_{k,j}^{3/2} + 3\frac{f_j\sigma_j^2}{{}^{[I-j-1]}S_j} + f_j^3\right)\Bigg] \\
&= \sum_{i=0}^{I-j-1}\Bigg[\gamma_j^* + 3f_j\sigma_j^2\sqrt{C_{i,j}} + f_j^3 C_{i,j}^{3/2} \\
&\quad - 3\left(f_j\sigma_j^2\sqrt{C_{i,j}} + f_j^3 C_{i,j}^{3/2} + \frac{\gamma_j^*\, C_{i,j}}{{}^{[I-j-1]}S_j} + \frac{2f_j\sigma_j^2 C_{i,j}^{3/2}}{{}^{[I-j-1]}S_j}\right)
\end{aligned}$$

$$+\frac{3}{\left({}^{[I-j-1]}S_j\right)^2}\left(\gamma_j^*\ C_{i,j}^2+3f_j\sigma_j^2 C_{i,j}^{3/2}\,{}^{[I-j-1]}S_j\right)+3f_j^3 C_{i,j}^{3/2}$$

$$\left.-C_{i,j}^{3/2}\left(\frac{\gamma_j^*}{\left({}^{[I-j-1]}S_j\right)^3}\sum_{k=0}^{I-j-1}C_{k,j}^{3/2}+3\frac{f_j\sigma_j^2}{{}^{[I-j-1]}S_j}+f_j^3\right)\right]$$

$$=\sum_{i=0}^{I-j-1}\gamma_j^*\left[1-\frac{3\ C_{i,j}}{{}^{[I-j-1]}S_j}+\frac{3\ C_{i,j}^2}{\left({}^{[I-j-1]}S_j\right)^2}-\frac{C_{i,j}^{3/2}}{\left({}^{[I-j-1]}S_j\right)^3}\sum_{k=0}^{I-j-1}C_{k,j}^{3/2}\right]$$

$$=\gamma_j^*\left[I-j-3+3\frac{\sum_{i=0}^{I-j-1}C_{i,j}^2}{\left({}^{[I-j-1]}S_j\right)^2}-\frac{\left(\sum_{k=0}^{I-j-1}C_{k,j}^{3/2}\right)^2}{\left({}^{[I-j-1]}S_j\right)^3}\right]$$

$$=\gamma_j^*\left[\sum_{i=0}^{I-j-1}\left(1-\frac{C_{i,j}}{\sum_{k=0}^{I-j-1}C_{k,j}}\right)^3+\frac{\sum_{k=0}^{I-j-1}C_{k,j}^3-\left(\sum_{k=0}^{I-j-1}C_{k,j}^{3/2}\right)^2}{\left(\sum_{k=0}^{I-j-1}C_{k,j}\right)^3}\right]$$

which proves $\mathbb{E}_j\left[\widehat{\gamma}_j^*\middle|\mathcal{B}_j\right]=\gamma_j^*$ and where we used the following equation for the last equality

$$\sum_{i=0}^{I-j-1}\left(1-\frac{C_{i,j}}{{}^{[I-j-1]}S_j}\right)^3=I-j-3+3\frac{\sum_{i=0}^{I-j-1}C_{i,j}^2}{\left({}^{[I-j-1]}S_j\right)^2}-\frac{\sum_{k=0}^{I-j-1}C_{k,j}^3}{\left({}^{[I-j-1]}S_j\right)^3}$$

(*ii*) Using (*i*) we have $\mathbb{E}\left[\widehat{\gamma}_j^*\right]=\mathbb{E}\left[\mathbb{E}\left[\widehat{\gamma}_j^*\middle|\mathcal{B}_j\right]\right]=\gamma^*$ for $j=0,\ldots,J-3$.

❑

4.1.8 Estimator (Skewness estimator)**.** By inserting our estimators into eq. (4.7) we can estimate the skewness of development year $j+1$ by

$$\widehat{\gamma}_j=\frac{\widehat{\gamma}_j^*}{\left(\widehat{\sigma}_j^2\right)^{\frac{3}{2}}},\quad \forall j=0,\ldots,J-3 \tag{4.18}$$

Because the estimation is unstable if there are too few data points (which for the estimation of the skewness is normally the case for the last two development years), we make the simplification to estimate $\widehat{\gamma}_{J-1}$ and $\widehat{\gamma}_{J-2}$ by

$$\widehat{\gamma}_{J-1}=\widehat{\gamma}_{J-2}=0,$$

which assumes an underlying normal distribution. In the case of more accident years than development years, i.e. $I>J$, it is possible to use eq. (4.18) for $j\geq J-2$, because the proof of lemma 4.1.7 is not dependent on $j<J-2$. For each step that I is greater than J we can estimate one more factor $\widehat{\gamma}_j$ with eq. (4.18).

4.1.9 Remark. Note that while our estimators $\widehat{\gamma}_j^*$ and $\widehat{\sigma}_j^2$ are unbiased, this does not have to be the case for $\widehat{\gamma}_j$.

4.2 Estimation of the Kurtosis

In this section we will cover the same steps as in the previous one but now we will adapt them to the estimation of the kurtosis of a distribution.

4.2.1 Definition (Conditional Kurtosis)**.** Following the notation from chapter 2 we define the conditional fourth moment of a random variable A given B as

$$\mu^{(4)}(A|B) = \mathbb{E}\left[(A - \mathbb{E}[A|B])^4 \Big| B\right] \tag{4.19}$$

The conditional kurtosis, as defined by Karl Pearson, is then given by

$$\text{Kurt}(A|B) = \frac{\mu^{(4)}(A|B)}{\left(\mu^{(2)}(A|B)\right)^2} = \frac{\mu^{(4)}(A|B)}{(\mathbb{V}[A|B])^2} \tag{4.20}$$

Next we will once more expand the model assumptions to allow us to find an estimator for the kurtosis.

4.2.2 Model assumptions.

- Cumulative claims $C_{i,j}$ of different accident years i are independent.
- $(C_{i,j})_{j\geq 0}$ form a Markov chain. There exist factors $f_0, \ldots, f_{J-1} > 0$ and variance parameters $\sigma_0^2, \ldots, \sigma_{J-1}^2 > 0$ such that

$$\forall 0 \leq i \leq I \quad \text{and} \quad \forall 1 \leq j \leq J \quad \text{we have}$$

$$\mathbb{E}[C_{i,j}|C_{i,j-1}] = f_{j-1} C_{i,j-1}, \tag{4.21}$$

$$\mathbb{V}[C_{i,j}|C_{i,j-1}] = \sigma_{j-1}^2 C_{i,j-1}. \tag{4.22}$$

- $\text{Skew}(C_{i,j+1}|\mathcal{D}_I)$ depends on j but does not depend on i.
- $\text{Kurt}(C_{i,j+1}|\mathcal{D}_I)$ depends on j but does not depend on i.

Again we can use the last point to conclude

$$\forall\, j \in \{0, \ldots, J-1\}\ \exists\, \kappa_j\ \forall\, i \in \{0, \ldots, I\}:$$

$$\begin{aligned}\kappa_j &= \text{Kurt}(C_{i,j+1}|\mathcal{D}_I) \\ &= \frac{\mu^{(4)}(A|B)}{(\mathbb{V}[A|B])^2} \\ &\overset{(4.22)}{=} \frac{\mu^{(4)}(A|B)}{\left(\sigma_j^2\, C_{i,j}\right)^2}\end{aligned}$$

This can be rearranged to

$$\mu^{(4)}(C_{i,j+1}|\mathcal{D}_I) = \kappa_j \left(\sigma_j^2\, C_{i,j}\right)^2, \tag{4.23}$$

which means that the fourth moment of $C_{i,j+1}$ is proportional to $(C_{i,j})^2$.
To find an estimator for the kurtosis we will now go the same way as in the last section. First we use eq. (4.23) to define a modification of the kurtosis, appropriate for this particular case. We deduce that there exist factors κ_j^* so that

$$\mu^{(4)}\left(C_{i,j+1}|\mathcal{D}_I\right) = \kappa_j^*\left(C_{i,j}\right)^2 \qquad (4.24)$$
$$\forall i = 0,\ldots,I \text{ and } \forall j = 0,\ldots,I-1$$

4.2.3 Remark. Note that

$$\kappa_j = \frac{\kappa_j^*}{\left(\sigma_j^2\right)^2} \qquad (4.25)$$

We will have to work a bit differently than in section 4.1 because we do not have an expression for

$$\mathbb{E}\left[\left(\widehat{\sigma}_j^2\right)^2\middle|\mathcal{B}_j\right] = \mathbb{V}\left[\widehat{\sigma}_j^2|\mathcal{B}_j\right] + \left(\mathbb{E}\left[\widehat{\sigma}_j^2|\mathcal{B}_j\right]\right)^2$$

Trying to calculate $\mathbb{V}\left[\widehat{\sigma}_j^2|\mathcal{B}_j\right]$ becomes tricky very fast, because terms of the form

$$\mathbb{V}\left[\sum_{i=0}^{I-j.1} C_{i,j}\left(\frac{C_{i,j+1}}{C_{i,j}} - \widehat{f}_j\right)^2\middle|\mathcal{B}_j\right]$$

arise. Because the inner terms are not (proven to be) uncorrelated we cannot interchange the sum and the variance and thus cannot proceed as in eq. (3.17). For this reason we first have to assume to know the parameters σ_j for $j = 0,\ldots,J-1$ and will then define our estimator in estimator 4.2.7

4.2.4 Estimator. We estimate κ_j^* for $j = 0,\ldots,J-4$ by

$$\widehat{\kappa}_j^* = \frac{\sum_{i=0}^{I-j-1} C_{i,j}^2\left(\frac{C_{i,j+1}}{C_{i,j}} - \widehat{f}_j\right)^4}{\sum_{i=0}^{I-j-1}\left(1 - \frac{C_{i,j}}{\sum_{k=0}^{I-j-1} C_{k,j}}\right)^4 + \frac{\left(\sum_{i=0}^{I-j-1} C_{i,j}^2\right)^2 - \sum_{i=0}^{I-j-1} C_{i,j}^4}{\left(\sum_{i=0}^{I-j-1} C_{i,j}\right)^4}}$$
$$- 3\left(\sigma_j^2\right)^2 \frac{\left(2 - 6\frac{\sum_{i=0}^{I-j-1} C_{i,j}^2}{\left(\sum_{i=0}^{I-j-1} C_{i,j}\right)^2} + 4\frac{\sum_{i=0}^{I-j-1} C_{i,j}^3}{\left(\sum_{i=0}^{I-j-1} C_{i,j}\right)^3} + 2\frac{\sum_{i=0}^{I-j-1} C_{i,j}^2 \sum_{k=1}^{I-j-1}\sum_{n=0}^{k-1} C_{k,j}C_{n,j}}{\left(\sum_{i=0}^{I-j-1} C_{i,j}\right)^4}\right)}{\sum_{i=0}^{I-j-1}\left(1 - \frac{C_{i,j}}{\sum_{k=0}^{I-j-1} C_{k,j}}\right)^4 + \frac{\left(\sum_{i=0}^{I-j-1} C_{i,j}^2\right)^2 - \sum_{i=0}^{I-j-1} C_{i,j}^4}{\left(\sum_{i=0}^{I-j-1} C_{i,j}\right)^4}} \qquad (4.26)$$

4.2.5 Remark. Note that estimator 4.2.4 differs from the one given in [Mor13], which is

$$\widehat{\kappa}_j^{DM} = \frac{\sum_{i=0}^{I-j-1} C_{i,j}^2 \left(\frac{C_{i,j+1}}{C_{i,j}} - \widehat{f}_j\right)^4 - 3\ \left(\widehat{\sigma}_j^2\right)^2 \left(2 - 6\ \frac{\sum_{i=0}^{I-j-1} C_{i,j}^2}{\left(\sum_{i=0}^{I-j-1} C_{i,j}\right)^2} + 4\ \frac{\sum_{i=0}^{I-j-1} C_{i,j}^3}{\left(\sum_{i=0}^{I-j-1} C_{i,j}\right)^3}\right)}{\sum_{i=0}^{I-j-1} \left(1 - \frac{C_{i,j}}{\sum_{k=0}^{I-j-1} C_{k,j}}\right)^4 + \frac{\left(\sum_{i=0}^{I-j-1} C_{i,j}^2\right)^2 - \sum_{i=0}^{I-j-1} C_{i,j}^4}{\left(\sum_{i=0}^{I-j-1} C_{i,j}\right)^4}}$$

The form of eq. (4.26) results from the proof of lemma 4.2.6 below and we will therefore use estimator 4.1.5 from here on out.

4.2.6 Lemma. *The estimator $\widehat{\kappa}^*$ is a conditionally (and unconditionally) unbiased estimator for κ^*, which means that*

(i) $\mathbb{E}\left[\widehat{\kappa}_j^* \middle| \mathcal{D}_I\right] = \kappa^*$ *for* $j = 0, \ldots, J-4$

(ii) $\mathbb{E}\left[\widehat{\kappa}_j^*\right] = \kappa^*$ *for* $j = 0, \ldots, J-4$

Proof. We will again use the short notations $\mathbb{E}_j$ and $\mathbb{V}_j$ which are defined as (see eq. (4.11))

$$\mathbb{E}_j\left[\cdot\right] \coloneqq \mathbb{E}\left[\cdot|\mathcal{B}_j\right] \quad \text{and} \quad \mathbb{V}_j\left[\cdot\right] \coloneqq \mathbb{V}\left[\cdot|\mathcal{B}_j\right]$$

(i) For $j = 0, \ldots, J-4$ arbitrary but fixed we start by calculating

$$\begin{aligned}\mathbb{E}_j\left[\sum_{i=0}^{I-j-1} C_{i,j}^2 \left(\frac{C_{i,j+1}}{C_{i,j}} - \widehat{f}_j\right)^4\right] &= \underbrace{\sum_{i=0}^{I-j-1} \frac{1}{C_{i,j}^2} \mathbb{E}_j\left[C_{i,j+1}^4\right]}_{=:(a)} - 4 \underbrace{\sum_{i=0}^{I-j-1} \frac{1}{C_{i,j}} \mathbb{E}_j\left[C_{i,j+1}^3 \widehat{f}_j\right]}_{=:(b)} \\ &+ 6 \underbrace{\sum_{i=0}^{I-j-1} \mathbb{E}_j\left[C_{i,j+1}^2 \widehat{f}_j^2\right]}_{=:(c)} - 4 \underbrace{\sum_{i=0}^{I-j-1} C_{i,j} \mathbb{E}_j\left[C_{i,j+1} \widehat{f}_j^3\right]}_{=:(d)} + \underbrace{\sum_{i=0}^{I-j-1} C_{i,j}^2 \mathbb{E}_j\left[\widehat{f}_j^4\right]}_{=:(e)}\end{aligned}$$

We will now calculate each element of the above equation individually:

(a) Using part (a) of the proof of lemma 4.1.7 we have

$$\begin{aligned}\kappa_j^*\ C_{i,j}^2 &\overset{(4.24)}{=} \mu^{(4)}\left(C_{i,j+1}|\mathcal{D}_I\right) \\ &= \mathbb{E}_j\left[C_{i,j+1}^4\right] - 4\mathbb{E}_j\left[C_{i,j+1}^3\right]\mathbb{E}_j\left[C_{i,j+1}\right] + 6\mathbb{E}_j\left[C_{i,j+1}^2\right]\mathbb{E}_j\left[C_{i,j+1}\right]^2 - 3\mathbb{E}_j\left[C_{i,j+1}\right]^4 \\ &= \mathbb{E}_j\left[C_{i,j+1}^4\right] - 4\left(\gamma_j^* C_{i,j}^{3/2} + 3 f_j \sigma_j^2 C_{i,j}^2 + f_j^3 C_{i,j}^3\right) f_j C_{i,j} \\ &\qquad + 6\left(\sigma_j^2 C_{i,j} + f_j^2 C_{i,j}^2\right) f_j^2 C_{i,j}^2 - 3 f_j^4 C_{i,j}^4 \\ &= \mathbb{E}_j\left[C_{i,j+1}^4\right] - 4\gamma_j^* f_j C_{i,j}^{5/2} - 12 f_j^2 \sigma_j^2 C_{i,j}^3 - 4 f_j^4 C_{i,j}^4 + 6 f_j^2 \sigma_j^2 C_{i,j}^3 + 6 f_j^4 C_{i,j}^4 - 3 f_j^4 C_{i,j}^4 \\ &= \mathbb{E}_j\left[C_{i,j+1}^4\right] - 4\gamma_j^* f_j C_{i,j}^{5/2} - 6 f_j^2 \sigma_j^2 C_{i,j}^3 - f_j^4 C_{i,j}^4\end{aligned}$$

which yields

$$\begin{aligned}
\textbf{(a)} &= \sum_{i=0}^{I-j-1} \frac{1}{C_{i,j}^2} \left(\kappa_j^* C_{i,j}^2 + 4\gamma_j^* f_j C_{i,j}^{5/2} + 6 f_j^2 \sigma_j^2 C_{i,j}^3 + f_j^4 C_{i,j}^4 \right) \\
&= \sum_{i=0}^{I-j-1} \left(\kappa_j^* + 4\gamma_j^* f_j \sqrt{C_{i,j}} + 6 f_j^2 \sigma_j^2 C_{i,j} + f_j^4 C_{i,j}^2 \right) \qquad (4.27)
\end{aligned}$$

(b) By inserting the definition of $\widehat{f}_j$ and using the independence between different accident years $(*)$ we have

$$\begin{aligned}
\mathbb{E}_j \left[\widehat{f}_j C_{i,j+1}^3 \right] &\overset{(3.6)}{=} \underbrace{\frac{1}{\sum_{k=0}^{I-j-1} C_{k,j}}}_{=:L} \mathbb{E}_j \left[\sum_{\substack{k=0 \\ k \neq i}}^{I-j-1} C_{k,j+1} C_{i,j+1}^3 + C_{i,j+1}^4 \right] \\
&\overset{(*)}{=} L \left(\mathbb{E}_j \left[C_{i,j+1}^3 \right] \sum_{\substack{k=0 \\ k \neq i}}^{I-j-1} \mathbb{E}_j \left[C_{k,j+1} \right] + \mathbb{E}_j \left[C_{i,j+1}^4 \right] \right) \\
&\overset{(4.27)}{=} L \Bigg(\left(\gamma_j^* \; C_{i,j}^{3/2} + 3 f_j \sigma_j^2 C_{i,j}^2 + f_j^3 C_{i,j}^3 \right) \sum_{\substack{k=0 \\ k \neq i}}^{I-j-1} (f_j C_{k,j}) \\
&\qquad + \kappa_j^* C_{i,j}^2 + 4\gamma_j^* f_j C_{i,j}^{5/2} + 6 f_j^2 \sigma_j^2 C_{i,j}^3 + f_j^4 C_{i,j}^4 \Bigg) \\
&= L \Bigg(\gamma_j^* f_j C_{i,j}^{3/2} \sum_{k=0}^{I-j-1} C_{k,j} + 3 f_j^2 \sigma_j^2 C_{i,j}^2 \sum_{k=0}^{I-j-1} C_{k,j} + f_j^4 C_{i,j}^3 \sum_{k=0}^{I-j-1} C_{k,j} \\
&\qquad + \kappa_j^* C_{i,j}^2 + 3\gamma_j^* f_j C_{i,j}^{5/2} + 3 f_j^2 \sigma_j^2 C_{i,j}^3 \Bigg) \\
&= \gamma_j^* f_j C_{i,j}^{3/2} + 3 f_j^2 \sigma_j^2 C_{i,j}^2 + f_j^4 C_{i,j}^3 + L \left(\kappa_j^* C_{i,j}^2 + 3\gamma_j^* f_j C_{i,j}^{5/2} + 3 f_j^2 \sigma_j^2 C_{i,j}^3 \right)
\end{aligned}$$

We conclude

$$\begin{aligned}
\textbf{(b)} &= \sum_{i=0}^{I-j-1} \frac{4}{C_{i,j}} \left(\gamma_j^* f_j C_{i,j}^{3/2} + 3 f_j^2 \sigma_j^2 C_{i,j}^2 + f_j^4 C_{i,j}^3 + \frac{\kappa_j^* C_{i,j}^2 + 3\gamma_j^* f_j C_{i,j}^{5/2} + 3 f_j^2 \sigma_j^2 C_{i,j}^3}{\sum_{k=0}^{I-j-1} C_{k,j}} \right) \\
&= \sum_{i=0}^{I-j-1} 4 \left(\gamma_j^* f_j \sqrt{C_{i,j}} + 3 f_j^2 \sigma_j^2 C_{i,j} + f_j^4 C_{i,j}^2 + \frac{\kappa_j^* C_{i,j} + 3\gamma_j^* f_j C_{i,j}^{3/2} + 3 f_j^2 \sigma_j^2 C_{i,j}^2}{\sum_{k=0}^{I-j-1} C_{k,j}} \right) \qquad (4.28)
\end{aligned}$$

(c) Analogously to above we have

$$\mathbb{E}_j\left[C_{i,j+1}^2\widehat{f}_j^2\right] \overset{(3.6)}{=} \underbrace{\frac{1}{\left(\sum_{k=0}^{I-j-1} C_{k,j}\right)^2}}_{=:M} \mathbb{E}_j\left[\left(\sum_{k=0}^{I-j-1} C_{k,j+1}\right)^2 C_{i,j+1}^2\right]$$

$$= M\,\mathbb{E}_j\left[C_{i,j+1}^2\left(\sum_{k=0}^{I-j-1} C_{k,j+1}^2 + 2\sum_{k=1}^{I-j-1}\sum_{n=0}^{k-1} C_{k,j+1}\,C_{n,j+1}\right)\right]$$

$$= M\,\mathbb{E}_j\Bigg[C_{i,j+1}^4 + \sum_{\substack{k=0\\k\neq i}}^{I-j-1}\left(C_{k,j+1}^2 C_{i,j+1}^2\right) + 2\sum_{\substack{k=1\\k\neq i}}^{I-j-1}\sum_{n=0}^{k-1}\left(C_{k,j+1}\,C_{n,j+1}\,C_{i,j+1}^2\right)$$
$$+ 2\sum_{n=0}^{i-1}\left(C_{n,j+1}\,C_{i,j+1}^3\right)\Bigg]$$

$$= M\,\mathbb{E}_j\Bigg[C_{i,j+1}^4 + \sum_{\substack{k=0\\k\neq i}}^{I-j-1}\left(C_{k,j+1}^2 C_{i,j+1}^2\right) + 2\sum_{\substack{k=1\\k\neq i}}^{I-j-1}\sum_{\substack{n=0\\n\neq i}}^{k-1}\left(C_{k,j+1}\,C_{n,j+1}\,C_{i,j+1}^2\right)$$
$$+ 2\sum_{n=0}^{i-1}\left(C_{n,j+1}\,C_{i,j+1}^3\right) + 2\sum_{k=i+1}^{I-j-1}\left(C_{k,j+1}\,C_{i,j+1}^3\right)\Bigg]$$

$$= M\,\mathbb{E}_j\Bigg[C_{i,j+1}^4 + \sum_{\substack{k=0\\k\neq i}}^{I-j-1}\left(C_{k,j+1}^2 C_{i,j+1}^2\right) + 2\sum_{\substack{k=1\\k\neq i}}^{I-j-1}\sum_{\substack{n=0\\n\neq i}}^{k-1}\left(C_{k,j+1}\,C_{n,j+1}\,C_{i,j+1}^2\right)$$
$$+ 2\sum_{\substack{k=0\\k\neq i}}^{I-j-1}\left(C_{k,j+1}\,C_{i,j+1}^3\right)\Bigg]$$

$$\overset{(*)}{=} M\Bigg[\mathbb{E}_j\left[C_{i,j+1}^4\right] + \sum_{\substack{k=0\\k\neq i}}^{I-j-1}\left(\mathbb{E}_j\left[C_{k,j+1}^2\right]\mathbb{E}_j\left[C_{i,j+1}^2\right]\right)$$
$$+ 2\sum_{\substack{k=1\\k\neq i}}^{I-j-1}\sum_{\substack{n=0\\n\neq i}}^{k-1}\left(\mathbb{E}_j\left[C_{k,j+1}\right]\mathbb{E}_j\left[C_{n,j+1}\right]\mathbb{E}_j\left[C_{i,j+1}^2\right]\right)$$
$$+ 2\sum_{\substack{k=0\\k\neq i}}^{I-j-1}\left(\mathbb{E}_j\left[C_{k,j+1}\right]\mathbb{E}_j\left[C_{i,j+1}^3\right]\right)\Bigg]$$

$$= M\Bigg[\kappa_j^* C_{i,j}^2 + 4\gamma_j^* f_j C_{i,j}^{5/2} + 6 f_j^2\sigma_j^2 C_{i,j}^3 + f_j^4 C_{i,j}^4$$
$$+ \left(\sigma_j^2 C_{i,j} + f_j^2 C_{i,j}^2\right)\sum_{\substack{k=0\\k\neq i}}^{I-j-1}\left(\sigma_j^2 C_{k,j} + f_j^2 C_{k,j}^2\right)$$
$$+ 2\left(\sigma_j^2 C_{i,j} + f_j^2 C_{i,j}^2\right) f_j^2 \sum_{\substack{k=1\\k\neq i}}^{I-j-1}\sum_{\substack{n=0\\n\neq i}}^{k-1}\left(C_{k,j} C_{n,j}\right)$$

$$+ 2\left(\gamma_j^* \; C_{i,j}^{3/2} + 3 f_j \sigma_j^2 C_{i,j}^2 + f_j^3 C_{i,j}^3\right) f_j \sum_{\substack{k=0 \\ k \neq i}}^{I-j-1} C_{k,j} \Bigg]$$

$$= M\Bigg[\kappa_j^* C_{i,j}^2 + \gamma_j^* \left(4 f_j C_{i,j}^{5/2} + 2 C_{i,j}^{3/2} f_j \sum_{\substack{k=0 \\ k \neq i}}^{I-j-1} C_{k,j}\right) + \left(\sigma_j^2\right)^2 C_{i,j} \left(\sum_{\substack{k=0 \\ k \neq i}}^{I-j-1} C_{k,j}\right)$$

$$+ f_j^4 C_{i,j}^2 \left(C_{i,j}^2 + \sum_{\substack{k=0 \\ k \neq i}}^{I-j-1} C_{k,j}^2 + 2 \sum_{\substack{k=1 \\ k \neq i}}^{I-j-1} \sum_{\substack{n=0 \\ n \neq i}}^{k-1} (C_{k,j} C_{n,j}) + 2 C_{i,j} \sum_{\substack{k=0 \\ k \neq i}}^{I-j-1} C_{k,j}\right)$$

$$+ f_j^2 \sigma_j^2 C_{i,j} \Bigg(6 C_{i,j}^2 + \sum_{\substack{k=0 \\ k \neq i}}^{I-j-1} C_{k,j}^2 + C_{i,j} \sum_{\substack{k=0 \\ k \neq i}}^{I-j-1} C_{k,j}$$

$$+ 2 \sum_{\substack{k=1 \\ k \neq i}}^{I-j-1} \sum_{\substack{n=0 \\ n \neq i}}^{k-1} (C_{k,j} C_{n,j}) + 6 C_{i,j} \sum_{\substack{k=0 \\ k \neq i}}^{I-j-1} C_{k,j}\Bigg) c\Bigg]$$

$$= M\Bigg[\kappa_j^* C_{i,j}^2 + \gamma_j^* \left(2 f_j C_{i,j}^{5/2} + 2 C_{i,j}^{3/2} f_j \sum_{k=0}^{I-j-1} C_{k,j}\right)$$

$$+ f_j^4 C_{i,j}^2 \left(\sum_{k=0}^{I-j-1} C_{k,j}^2 + 2 \sum_{k=1}^{I-j-1} \sum_{n=0}^{k-1} (C_{k,j} C_{n,j})\right) + \left(\sigma_j^2\right)^2 C_{i,j} \left(\sum_{\substack{k=0 \\ k \neq i}}^{I-j-1} C_{k,j}\right)$$

$$+ f_j^2 \sigma_j^2 C_{i,j} \left(5 C_{i,j}^2 + \sum_{k=0}^{I-j-1} C_{k,j}^2 + 2 \sum_{k=1}^{I-j-1} \sum_{n=0}^{k-1} (C_{k,j} C_{n,j}) + 5 C_{i,j} \sum_{\substack{k=0 \\ k \neq i}}^{I-j-1} C_{k,j}\right)\Bigg]$$

$$= M\Bigg[\kappa_j^* C_{i,j}^2 + \gamma_j^* \left(2 f_j C_{i,j}^{5/2} + 2 f_j C_{i,j}^{3/2} \sum_{k=0}^{I-j-1} C_{k,j}\right) + f_j^4 C_{i,j}^2 \left(\sum_{k=0}^{I-j-1} C_{k,j}\right)^2$$

$$+ \left(\sigma_j^2\right)^2 C_{i,j} \left(\sum_{\substack{k=0 \\ k \neq i}}^{I-j-1} C_{k,j}\right) + f_j^2 \sigma_j^2 C_{i,j} \left(\sum_{k=0}^{I-j-1} C_{k,j}\right)^2 + 5 f_j^2 \sigma_j^2 C_{i,j}^2 \sum_{k=0}^{I-j-1} C_{k,j}\Bigg]$$

which yields

$$\textbf{(c)} = 6 \sum_{i=0}^{I-j-1} \Bigg[\frac{\kappa_j^* C_{i,j}^2}{\left({}^{[I-j-1]}S_j\right)^2} + \frac{2 \gamma_j^* f_j C_{i,j}^{5/2}}{\left({}^{[I-j-1]}S_j\right)^2} + \frac{2 \gamma_j^* f_j C_{i,j}^{3/2}}{{}^{[I-j-1]}S_j} + \frac{5 f_j^2 \sigma_j^2 C_{i,j}^2}{{}^{[I-j-1]}S_j}$$

$$+ f_j^4 C_{i,j}^2 + f_j^2 \sigma_j^2 C_{i,j} + \frac{\left(\sigma_j^2\right)^2 C_{i,j} \sum_{\substack{k=0 \\ k \neq i}}^{I-j-1} C_{k,j}}{\left({}^{[I-j-1]}S_j\right)^2}\Bigg] \tag{4.29}$$

(d) For the next element we will use the results from the proof of lemma 4.1.7. We have

$$
\begin{aligned}
\mathbb{E}_j\left[C_{i,j+1}\widehat{f}_j^3\right] &\overset{(3.6)}{=} \underbrace{\frac{1}{\left(\sum_{k=0}^{I-j-1} C_{k,j}\right)^3}}_{=:Z} \mathbb{E}_j\left[C_{i,j+1}\left(\sum_{k=0}^{I-j-1} C_{k,j+1}\right)^3\right] \\
&= Z\,\mathbb{E}_j\Bigg[C_{i,j+1}\Bigg(\sum_{k=0}^{I-j-1} C_{k,j+1}^3 + 3\sum_{k=0}^{I-j-1}\sum_{\substack{n=0\\n\neq k}}^{I-j-1} C_{k,j+1}^2\, C_{n,j+1} \\
&\qquad + 6\sum_{k=2}^{I-j-1}\sum_{n=1}^{k-1}\sum_{m=0}^{n-1} C_{k,j+1}\, C_{n,j+1}\, C_{m,j+1}\Bigg)\Bigg] \\
&= Z\,\mathbb{E}_j\Bigg[C_{i,j+1}^4 + \sum_{\substack{k=0\\k\neq i}}^{I-j-1} C_{k,j+1}^3 C_{i,j+1} + 3\sum_{\substack{n=0\\n\neq i}}^{I-j-1} C_{i,j+1}^3\, C_{n,j+1} + 3\sum_{\substack{k=0\\k\neq i}}^{I-j-1} C_{k,j+1}^2\, C_{i,j+1}^2 \\
&\qquad + 3\sum_{\substack{k=0\\k\neq i}}^{I-j-1}\sum_{\substack{n=0\\n\neq k\\n\neq i}}^{I-j-1} C_{k,j+1}^2\, C_{n,j+1}\, C_{i,j+1} + 6\sum_{\substack{k=1\\k\neq i}}^{I-j-1}\sum_{\substack{n=0\\n\neq i}}^{k-1} C_{i,j+1}^2\, C_{k,j+1}\, C_{n,j+1} \\
&\qquad + 6\sum_{\substack{k=2\\k\neq i}}^{I-j-1}\sum_{\substack{n=1\\n\neq i}}^{k-1}\sum_{\substack{m=0\\m\neq i}}^{n-1} C_{k,j+1}\, C_{n,j+1}\, C_{m,j+1}\, C_{i,j+1}\Bigg] \\
&\overset{(*)}{=} Z\,\Bigg[\mathbb{E}_j\left[C_{i,j+1}^4\right] + \sum_{\substack{k=0\\k\neq i}}^{I-j-1} \mathbb{E}_j\left[C_{k,j+1}^3\right]\mathbb{E}_j\left[C_{i,j+1}\right] + 3\sum_{\substack{n=0\\n\neq i}}^{I-j-1} \mathbb{E}_j\left[C_{i,j+1}^3\right]\mathbb{E}_j\left[C_{n,j+1}\right] \\
&\qquad + 3\sum_{\substack{k=0\\k\neq i}}^{I-j-1} \mathbb{E}_j\left[C_{k,j+1}^2\right]\mathbb{E}_j\left[C_{i,j+1}^2\right] \\
&\qquad + 3\sum_{\substack{k=0\\k\neq i}}^{I-j-1}\sum_{\substack{n=0\\n\neq k\\n\neq i}}^{I-j-1} \mathbb{E}_j\left[C_{k,j+1}^2\right]\mathbb{E}_j\left[C_{n,j+1}\right]\mathbb{E}_j\left[C_{i,j+1}\right] \\
&\qquad + 6\sum_{\substack{k=1\\k\neq i}}^{I-j-1}\sum_{\substack{n=0\\n\neq i}}^{k-1} \mathbb{E}_j\left[C_{i,j+1}^2\right]\mathbb{E}_j\left[C_{k,j+1}\right]\mathbb{E}_j\left[C_{n,j+1}\right] \\
&\qquad + 6\sum_{\substack{k=2\\k\neq i}}^{I-j-1}\sum_{\substack{n=1\\n\neq i}}^{k-1}\sum_{\substack{m=0\\m\neq i}}^{n-1} \mathbb{E}_j\left[C_{k,j+1}\right]\mathbb{E}_j\left[C_{n,j+1}\right]\mathbb{E}_j\left[C_{m,j+1}\right]\mathbb{E}_j\left[C_{i,j+1}\right]\Bigg] \\
&= Z\,\Bigg[\kappa_j^* C_{i,j}^2 + 4\gamma_j^* f_j C_{i,j}^{5/2} + 6 f_j^2\sigma_j^2 C_{i,j}^3 + f_j^4 C_{i,j}^4 \\
&\qquad + f_j C_{i,j}\sum_{\substack{k=0\\k\neq i}}^{I-j-1}\left(\gamma_j^*\, C_{k,j}^{3/2} + 3 f_j\sigma_j^2 C_{k,j}^2 + f_j^3 C_{k,j}^3\right)
\end{aligned}
$$

$$
\begin{aligned}
&+3\left(\gamma_j^* \ C_{i,j}^{3/2}+3f_j\sigma_j^2C_{i,j}^2+f_j^3C_{i,j}^3\right)\sum_{\substack{n=0\\n\neq i}}^{I-j-1} f_jC_{n,j}\\
&+3\left(\sigma_j^2C_{i,j}+f_j^2C_{i,j}^2\right)\sum_{\substack{k=0\\k\neq i}}^{I-j-1}\left(\sigma_j^2C_{k,j}+f_j^2C_{k,j}^2\right)\\
&+3\sum_{\substack{k=0\\k\neq i}}^{I-j-1}\sum_{\substack{n=0\\n\neq k\\n\neq i}}^{I-j-1}\left(\sigma_j^2C_{k,j}+f_j^2C_{k,j}^2\right)f_j^2C_{n,j}C_{i,j}\\
&+6\left(\sigma_j^2C_{i,j}+f_j^2C_{i,j}^2\right)\sum_{\substack{k=1\\k\neq i}}^{I-j-1}\sum_{\substack{n=0\\n\neq i}}^{k-1}f_j^2C_{k,j}C_{n,j}\\
&+6f_j^4C_{i,j}\sum_{\substack{k=2\\k\neq i}}^{I-j-1}\sum_{\substack{n=1\\n\neq i}}^{k-1}\sum_{\substack{m=0\\m\neq i}}^{n-1}C_{k,j}C_{m,j}C_{m,j}\Bigg]\\
=Z&\Bigg[\kappa_j^*C_{i,j}^2+\gamma_j^*f_j\left(4C_{i,j}^{5/2}+C_{i,j}\sum_{\substack{k=0\\k\neq i}}^{I-j-1}C_{k,j}^{3/2}+3C_{i,j}^{3/2}\sum_{\substack{k=0\\k\neq i}}^{I-j-1}C_{k,j}\right)\\
&+3\left(\sigma_j^2\right)^2C_{i,j}\sum_{\substack{k=0\\k\neq i}}^{I-j-1}C_{k,j}+f_j^2\sigma_j^2C_{i,j}\Bigg(6C_{i,j}^2+6\sum_{\substack{k=0\\k\neq i}}^{I-j-1}C_{k,j}^2+12C_{i,j}\sum_{\substack{k=0\\k\neq i}}^{I-j-1}C_{k,j}\\
&\qquad+6\sum_{\substack{k=1\\k\neq i}}^{I-j-1}\sum_{\substack{n=0\\n\neq i}}^{k-1}C_{k,j}C_{n,j}+3\sum_{\substack{k=0\\k\neq i}}^{I-j-1}\sum_{\substack{n=0\\n\neq k\\n\neq i}}^{I-j-1}C_{k,j}C_{n,j}\Bigg)\\
&+f_j^4C_{i,j}\Bigg(C_{i,j}^3+\sum_{\substack{k=0\\k\neq i}}^{I-j-1}C_{k,j}^3+3C_{i,j}^2\sum_{\substack{k=0\\k\neq i}}^{I-j-1}C_{k,j}\\
&\qquad+3C_{i,j}\sum_{\substack{k=0\\k\neq i}}^{I-j-1}C_{k,j}^2+3\sum_{\substack{k=0\\k\neq i}}^{I-j-1}\sum_{\substack{n=0\\n\neq k\\n\neq i}}^{I-j-1}C_{k,j}^2C_{n,j}\\
&\qquad+6C_{i,j}\sum_{\substack{k=1\\k\neq i}}^{I-j-1}\sum_{\substack{n=0\\n\neq i}}^{k-1}C_{k,j}C_{n,j}+6\sum_{\substack{k=2\\k\neq i}}^{I-j-1}\sum_{\substack{n=1\\n\neq i}}^{k-1}\sum_{\substack{m=0\\m\neq i}}^{n-1}C_{k,j}C_{m,j}C_{m,j}\Bigg)\Bigg]\\
=Z&\Bigg[\kappa_j^*C_{i,j}^2+\gamma_j^*f_j\left(C_{i,j}\sum_{k=0}^{I-j-1}C_{k,j}^{3/2}+3C_{i,j}^{3/2}\sum_{k=0}^{I-j-1}C_{k,j}\right)\\
&+3\left(\sigma_j^2\right)^2C_{i,j}\sum_{\substack{k=0\\k\neq i}}^{I-j-1}C_{k,j}+f_j^2\sigma_j^2C_{i,j}\left(6\sum_{k=0}^{I-j-1}C_{k,j}^2+12\sum_{k=1}^{I-j-1}\sum_{n=0}^{k-1}C_{k,j}C_{n,j}\right)
\end{aligned}
$$

$$+ f_j^4 C_{i,j} \left(\sum_{k=0}^{I-j-1} C_{k,j}^3 + 3 \sum_{k=0}^{I-j-1} \sum_{\substack{n=0 \\ n \neq k}}^{I-j-1} C_{k,j}^2 C_{n,j} + 6 \sum_{k=2}^{I-j-1} \sum_{n=1}^{k-1} \sum_{m=0}^{n-1} C_{k,j} C_{m,j} C_{m,j} \right) \Bigg]$$

$$= Z \left[\kappa_j^* C_{i,j}^2 + \gamma_j^* f_j \left(C_{i,j} \sum_{k=0}^{I-j-1} C_{k,j}^{3/2} + 3 C_{i,j}^{3/2} \sum_{k=0}^{I-j-1} C_{k,j} \right) + 3 \left(\sigma_j^2\right)^2 C_{i,j} \sum_{\substack{k=0 \\ k \neq i}}^{I-j-1} C_{k,j} \right.$$

$$\left. + 6 f_j^2 \sigma_j^2 C_{i,j} \left(\sum_{k=0}^{I-j-1} C_{k,j} \right)^2 + f_j^4 C_{i,j} \left(\sum_{k=0}^{I-j-1} C_{k,j} \right)^3 \right]$$

We conclude

$$\textbf{(d)} = 4 \sum_{i=0}^{I-j-1} C_{i,j} \left[\frac{\kappa_j^* C_{i,j}^2}{\left({}^{[I-j-1]}S_j\right)^3} + \frac{\gamma_j^* f_j \left(C_{i,j} \sum_{k=0}^{I-j-1} C_{k,j}^{3/2} + 3 C_{i,j}^{3/2} \sum_{k=0}^{I-j-1} C_{k,j} \right)}{\left({}^{[I-j-1]}S_j\right)^3} \right.$$
$$\left. + \frac{3 \left(\sigma_j^2\right)^2 C_{i,j} \sum_{\substack{k=0 \\ k \neq i}}^{I-j-1} C_{k,j}}{\left({}^{[I-j-1]}S_j\right)^3} + \frac{6 f_j^2 \sigma_j^2 C_{i,j}}{{}^{[I-j-1]}S_j} + f_j^4 C_{i,j} \right]$$
$$= 4 \sum_{i=0}^{I-j-1} \left[\frac{\kappa_j^* C_{i,j}^3}{\left({}^{[I-j-1]}S_j\right)^3} + \frac{\gamma_j^* f_j \left(C_{i,j}^2 \sum_{k=0}^{I-j-1} C_{k,j}^{3/2} + 3 C_{i,j}^{5/2} \sum_{k=0}^{I-j-1} C_{k,j} \right)}{\left({}^{[I-j-1]}S_j\right)^3} \right.$$
$$\left. + \frac{3 \left(\sigma_j^2\right)^2 C_{i,j}^2 \sum_{\substack{k=0 \\ k \neq i}}^{I-j-1} C_{k,j}}{\left({}^{[I-j-1]}S_j\right)^3} + \frac{6 f_j^2 \sigma_j^2 C_{i,j}^2}{{}^{[I-j-1]}S_j} + f_j^4 C_{i,j}^2 \right] \tag{4.30}$$

(e) For the last element we will use eqs. (3.19) and (4.17) which state

$$\mathbb{E}_j \left[\widehat{f}_j^2 \middle| \mathcal{B}_j \right] = \frac{\sigma_j^2}{\sum_{i=0}^{I-j-1} C_{i,j}} + f_j^2,$$

$$\mathbb{E}_j \left[\widehat{f}_j^3 \middle| \mathcal{B}_j \right] = \frac{\gamma^*}{\left(\sum_{i=0}^{I-j-1} C_{i,j} \right)^3} \sum_{i=0}^{I-j-1} C_{i,j}^{3/2} + 3 \frac{f_j \sigma_j^2}{\sum_{i=0}^{I-j-1} C_{i,j}} + f_j^3.$$

Additionally we will use the fact that for independent random variables $X_1, \ldots, X_n$ with expected value of zero and each with finite first four moments we have

$$\mathbb{E}_j \left[\left(\sum_{i=1}^{n} X_i \right)^4 \right] = \sum_{i=1}^{n} \mathbb{E}_j \left[X_i^4 \right] + 6 \sum_{i=2}^{n} \sum_{j=1}^{i-1} \mathbb{E}_j \left[X_i^2 \right] \mathbb{E}_j \left[X_j^2 \right] \tag{4.31}$$

This follows from the fact that all other cross products contain a factor of the form $\mathbb{E}_j \left[X_i \right] (= 0)$ and are therefore zero.

We use this theorem on the centralized random variables $C_{i,j+1} - \mathbb{E}_j[C_{i,j+1}]$ to prove

$$\begin{aligned}
\mathbb{E}_j\left[\left(\widehat{f}_j - \mathbb{E}_j\left[\widehat{f}_j\right]\right)^4\right] &\overset{(4.21)}{=} \mathbb{E}_j\left[\left(\frac{\sum_{i=0}^{I-j-1} C_{i,j+1}}{\sum_{i=0}^{I-j-1} C_{i,j}} - \mathbb{E}_j\left[\frac{\sum_{i=0}^{I-j-1} C_{i,j+1}}{\sum_{i=0}^{I-j-1} C_{i,j}}\right]\right)^4\right] \\
&= \frac{1}{\left(\sum_{i=0}^{I-j-1} C_{i,j}\right)^4}\mathbb{E}_j\left[\left(\sum_{i=0}^{I-j-1} C_{i,j+1} - \sum_{i=0}^{I-j-1} \mathbb{E}_j[C_{i,j+1}]\right)^4\right] \\
&= \frac{1}{\left(\sum_{i=0}^{I-j-1} C_{i,j}\right)^4}\mathbb{E}_j\left[\left(\sum_{i=0}^{I-j-1} C_{i,j+1} - \mathbb{E}_j[C_{i,j+1}]\right)^4\right] \\
&\overset{(4.31)}{=} \frac{1}{\left(\sum_{i=0}^{I-j-1} C_{i,j}\right)^4}\left(\sum_{i=0}^{I-j-1} \mathbb{E}_j\left[(C_{i,j+1} - \mathbb{E}_j[C_{i,j+1}])^4\right]\right. \\
&\qquad \left. + 6\sum_{i=1}^{I-j-1}\sum_{n=0}^{i-1} \mathbb{V}_j[C_{i,j+1}]\,\mathbb{V}_j[C_{n,j+1}]\right) \\
&\overset{(4.24)}{=} \frac{1}{\left(\sum_{i=0}^{I-j-1} C_{i,j}\right)^4}\left(\kappa^* \sum_{i=0}^{I-j-1} C_{i,j}^2 + 6\left(\sigma_j^2\right)^2 \sum_{i=1}^{I-j-1}\sum_{n=0}^{i-1} C_{i,j}C_{n,j}\right)
\end{aligned} \tag{4.32}$$

Using all of the above we have

$$\begin{aligned}
\mathbb{E}_j\left[\widehat{f}_j^4\right] &= \mathbb{E}_j\left[\left(\widehat{f}_j - \mathbb{E}_j\left[\widehat{f}_j\right]\right)^4\right] + 4\mathbb{E}_j\left[\widehat{f}_j^3\right]\mathbb{E}_j\left[\widehat{f}_j\right] - 6\mathbb{E}_j\left[\widehat{f}_j^2\right]\mathbb{E}_j\left[\widehat{f}_j\right]^2 + 3\mathbb{E}_j\left[\widehat{f}_j\right]^4 \\
&= \kappa^* \frac{\sum_{i=0}^{I-j-1} C_{i,j}^2}{\left(\sum_{i=0}^{I-j-1} C_{i,j}\right)^4} + 4f_j\left(\gamma^* \frac{\sum_{i=0}^{I-j-1} C_{i,j}^{3/2}}{\left(\sum_{i=0}^{I-j-1} C_{i,j}\right)^3} + 3\frac{f_j\sigma_j^2}{\sum_{i=0}^{I-j-1} C_{i,j}} + f_j^3\right) \\
&\quad + 6\frac{\left(\sigma_j^2\right)^2 \sum_{i=1}^{I-j-1}\sum_{n=0}^{i-1} C_{i,j}C_{n,j}}{\left(\sum_{i=0}^{I-j-1} C_{i,j}\right)^4} - 6f_j^2\left(\frac{\sigma_j^2}{\sum_{i=0}^{I-j-1} C_{i,j}} + f_j^2\right) + 3f_j^4 \\
&= \kappa^* \frac{\sum_{i=0}^{I-j-1} C_{i,j}^2}{\left(\sum_{i=0}^{I-j-1} C_{i,j}\right)^4} + 4f_j\gamma^* \frac{\sum_{i=0}^{I-j-1} C_{i,j}^{3/2}}{\left(\sum_{i=0}^{I-j-1} C_{i,j}\right)^3} + 6\frac{f_j^2\sigma_j^2}{\sum_{i=0}^{I-j-1} C_{i,j}} + f_j^4 \\
&\quad + 6\frac{\left(\sigma_j^2\right)^2 \sum_{i=1}^{I-j-1}\sum_{n=0}^{i-1} C_{i,j}C_{n,j}}{\left(\sum_{i=0}^{I-j-1} C_{i,j}\right)^4}
\end{aligned}$$

which yields

$$\begin{aligned}
\textbf{(e)} = \sum_{i=0}^{I-j-1} C_{i,j}^2 &\left(\kappa^* \frac{\sum_{k=0}^{I-j-1} C_{k,j}^2}{\left({}^{[I-j-1]}S_j\right)^4} + 4f_j\gamma_j^* \frac{\sum_{k=0}^{I-j-1} C_{k,j}^{3/2}}{\left({}^{[I-j-1]}S_j\right)^3}\right. \\
&\left. + 6\frac{f_j^2\sigma_j^2}{{}^{[I-j-1]}S_j} + f_j^4 + 6\frac{\left(\sigma_j^2\right)^2 \sum_{k=1}^{I-j-1}\sum_{n=0}^{k-1} C_{k,j}C_{n,j}}{\left({}^{[I-j-1]}S_j\right)^4}\right)
\end{aligned} \tag{4.33}$$

Putting everything together nearly all terms cancel each other out and we have

$$
\begin{aligned}
&\mathbb{E}_j\left[\sum_{i=0}^{I-j-1} C_{i,j}^2\left(\frac{C_{i,j+1}}{C_{i,j}}-\widehat{f}_j\right)^4\right]=(a)-(b)+(c)-(d)+(e)\\
&=\sum_{i=0}^{I-j-1}\Bigg(\kappa_j^*+4\gamma_j^* f_j\sqrt{C_{i,j}}+6f_j^2\sigma_j^2 C_{i,j}+f_j^4C_{i,j}^2\\
&\quad-4\left[\gamma_j^* f_j\sqrt{C_{i,j}}+3f_j^2\sigma_j^2C_{i,j}+f_j^4C_{i,j}^2+\frac{\kappa_j^*C_{i,j}+3\gamma_j^*f_jC_{i,j}^{3/2}+3f_j^2\sigma_j^2C_{i,j}^2}{\sum_{k=0}^{I-j-1}C_{k,j}}\right]\\
&\quad+6\left[\frac{\kappa_j^*C_{i,j}^2}{\left({}^{[I-j-1]}S_j\right)^2}+\frac{2\gamma_j^*f_jC_{i,j}^{5/2}}{\left({}^{[I-j-1]}S_j\right)^2}+\frac{2\gamma_j^*f_jC_{i,j}^{3/2}}{{}^{[I-j-1]}S_j}+\frac{5f_j^2\sigma_j^2C_{i,j}^2}{{}^{[I-j-1]}S_j}\right.\\
&\qquad\left.+f_j^4C_{i,j}^2+f_j^2\sigma_j^2C_{i,j}+\frac{\left(\sigma_j^2\right)^2C_{i,j}\sum_{\substack{k=0\\k\neq i}}^{I-j-1}C_{k,j}}{\left({}^{[I-j-1]}S_j\right)^2}\right]\\
&\quad-4\left[\frac{\kappa_j^*C_{i,j}^3}{\left({}^{[I-j-1]}S_j\right)^3}+\frac{\gamma_j^*f_j\left(C_{i,j}^2\sum_{k=0}^{I-j-1}C_{k,j}^{3/2}+3C_{i,j}^{5/2}\sum_{k=0}^{I-j-1}C_{k,j}\right)}{\left({}^{[I-j-1]}S_j\right)^3}\right.\\
&\qquad\left.+\frac{3\left(\sigma_j^2\right)^2C_{i,j}^2\sum_{\substack{k=0\\k\neq i}}^{I-j-1}C_{k,j}}{\left({}^{[I-j-1]}S_j\right)^3}+\frac{6f_j^2\sigma_j^2C_{i,j}^2}{{}^{[I-j-1]}S_j}+f_j^4C_{i,j}^2\right]\\
&\quad+C_{i,j}^2\left[\kappa^*\frac{\sum_{k=0}^{I-j-1}C_{k,j}^2}{\left({}^{[I-j-1]}S_j\right)^4}+4f_j\gamma_j^*\frac{\sum_{k=0}^{I-j-1}C_{k,j}^{3/2}}{\left({}^{[I-j-1]}S_j\right)^3}+6\frac{f_j^2\sigma_j^2}{{}^{[I-j-1]}S_j}\right.\\
&\qquad\left.+f_j^4+6\frac{\left(\sigma_j^2\right)^2\sum_{k=1}^{I-j-1}\sum_{n=0}^{k-1}C_{k,j}C_{n,j}}{\left({}^{[I-j-1]}S_j\right)^4}\right]\Bigg)\\
&=\sum_{i=0}^{I-j-1}\left[\kappa_j^*\left(1-\frac{4\ C_{i,j}}{{}^{[I-j-1]}S_j}+\frac{6\ C_{i,j}^2}{\left({}^{[I-j-1]}S_j\right)^2}-\frac{4\ C_{i,j}^3}{\left({}^{[I-j-1]}S_j\right)^3}+\frac{C_{i,j}^2\sum_{k=0}^{I-j-1}C_{k,j}^2}{\left({}^{[I-j-1]}S_j\right)^4}\right)\right.\\
&\qquad+\left(\sigma_j^2\right)^2\left(\frac{6\ C_{i,j}\left(\sum_{k=0}^{I-j-1}C_{k,j}-C_{i,j}\right)}{\left({}^{[I-j-1]}S_j\right)^2}-\frac{12C_{i,j}^2\left(\sum_{k=0}^{I-j-1}C_{k,j}-C_{i,j}\right)}{\left({}^{[I-j-1]}S_j\right)^3}\right.\\
&\qquad\left.\left.+\frac{6C_{i,j}^2\sum_{k=1}^{I-j-1}\sum_{n=0}^{k-1}C_{k,j}C_{n,j}}{\left({}^{[I-j-1]}S_j\right)^4}\right)\right]\\
&=\kappa_j^*\left(I-j-4+\frac{6\sum_{i=0}^{I-j-1}C_{i,j}^2}{\left({}^{[I-j-1]}S_j\right)^2}-\frac{4\sum_{i=0}^{I-j-1}C_{i,j}^3}{\left({}^{[I-j-1]}S_j\right)^3}+\frac{\left(\sum_{i=0}^{I-j-1}C_{i,j}^2\right)^2}{\left({}^{[I-j-1]}S_j\right)^4}\right)
\end{aligned}
$$

$$
\begin{aligned}
&+\sum_{i=0}^{I-j-1} 3\left(\sigma_j^2\right)^2\left(\frac{2\,C_{i,j}}{{}^{[I-j-1]}S_j}-\frac{2\,C_{i,j}^2}{\left({}^{[I-j-1]}S_j\right)^2}-\frac{4C_{i,j}^2}{\left({}^{[I-j-1]}S_j\right)^2}\right.\\
&\left.+\frac{4C_{i,j}^3}{\left({}^{[I-j-1]}S_j\right)^3}+\frac{2C_{i,j}^2\sum_{k=1}^{I-j-1}\sum_{n=0}^{k-1}C_{k,j}C_{n,j}}{\left({}^{[I-j-1]}S_j\right)^4}\right)\\
=\kappa_j^*&\left(I-j-4+\frac{6\sum_{i=0}^{I-j-1}C_{i,j}^2}{\left({}^{[I-j-1]}S_j\right)^2}-\frac{4\sum_{i=0}^{I-j-1}C_{i,j}^3}{\left({}^{[I-j-1]}S_j\right)^3}+\frac{\left(\sum_{i=0}^{I-j-1}C_{i,j}^2\right)^2}{\left({}^{[I-j-1]}S_j\right)^4}\right)\\
&+3\left(\sigma_j^2\right)^2\left(2-\frac{2\sum_{i=0}^{I-j-1}C_{i,j}^2}{\left({}^{[I-j-1]}S_j\right)^2}-\frac{4\sum_{i=0}^{I-j-1}C_{i,j}^2}{\left({}^{[I-j-1]}S_j\right)^2}\right.\\
&\left.+\frac{4\sum_{i=0}^{I-j-1}C_{i,j}^3}{\left({}^{[I-j-1]}S_j\right)^3}+\frac{2\sum_{i=0}^{I-j-1}C_{i,j}^2\sum_{k=1}^{I-j-1}\sum_{n=0}^{k-1}C_{k,j}C_{n,j}}{\left({}^{[I-j-1]}S_j\right)^4}\right)\\
=\kappa_j^*&\left(\sum_{i=0}^{I-j-1}\left[1-\frac{C_{i,j}}{\sum_{i=0}^{I-j-1}C_{i,j}}\right]^4+\frac{\left(\sum_{i=0}^{I-j-1}C_{i,j}^2\right)^2-\sum_{i=0}^{I-j-1}C_{i,j}^4}{\left(\sum_{i=0}^{I-j-1}C_{i,j}\right)^4}\right)\\
&+3\left(\sigma_j^2\right)^2\left(2-6\frac{\sum_{i=0}^{I-j-1}C_{i,j}^2}{\left({}^{[I-j-1]}S_j\right)^2}+4\frac{\sum_{i=0}^{I-j-1}C_{i,j}^3}{\left({}^{[I-j-1]}S_j\right)^3}\right.\\
&\left.+2\frac{\sum_{i=0}^{I-j-1}C_{i,j}^2\sum_{k=1}^{I-j-1}\sum_{n=0}^{k-1}C_{k,j}C_{n,j}}{\left({}^{[I-j-1]}S_j\right)^4}\right)
\end{aligned}
$$

where we used the following equation for the last equality

$$
\sum_{i=0}^{I-j-1}\left[1-\frac{C_{i,j}}{\sum_{i=0}^{I-j-1}C_{i,j}}\right]^4=I-j-4+\frac{6\sum_{i=0}^{I-j-1}C_{i,j}^2}{\left({}^{[I-j-1]}S_j\right)^2}-\frac{4\sum_{i=0}^{I-j-1}C_{i,j}^3}{\left({}^{[I-j-1]}S_j\right)^3}+\frac{\sum_{i=0}^{I-j-1}C_{i,j}^4}{\left({}^{[I-j-1]}S_j\right)^4}
$$

and which shows that

$$
\mathbb{E}_j\left[\widehat{\kappa}_j^*\middle|\mathcal{B}_j\right]=\kappa^*.
$$

(ii) Using (i) we have $\mathbb{E}\left[\widehat{\kappa}_j^*\right]=\mathbb{E}\left[\mathbb{E}\left[\widehat{\kappa}_j^*\middle|\mathcal{B}_j\right]\right]=\kappa^*$ for $j=0,\ldots,J-4$.

❑

4.2.7 Estimator (Kurtosis estimator)**.** By inserting our estimators into eqs. (4.25) and (4.26) we can estimate the kurtosis of development year $j+1$ by

$$\widehat{\kappa}_j = \frac{1}{\left(\widehat{\sigma}_j^2\right)^2} \left[\frac{\sum\limits_{i=0}^{I-j-1} C_{i,j}^2 \left(\frac{C_{i,j+1}}{C_{i,j}} - \widehat{f}_j \right)^4}{\sum\limits_{i=0}^{I-j-1} \left(1 - \frac{C_{i,j}}{\sum_{k=0}^{I-j-1} C_{k,j}} \right)^4 + \frac{\left(\sum_{i=0}^{I-j-1} C_{i,j}^2\right)^2 - \sum_{i=0}^{I-j-1} C_{i,j}^4}{\left(\sum_{i=0}^{I-j-1} C_{i,j}\right)^4}} \right.$$
$$\left. - 3\left(\widehat{\sigma}_j^2\right)^2 \frac{\left(2 - 6 \frac{\sum\limits_{i=0}^{I-j-1} C_{i,j}^2}{\left(\sum\limits_{i=0}^{I-j-1} C_{i,j}\right)^2} + 4 \frac{\sum\limits_{i=0}^{I-j-1} C_{i,j}^3}{\left(\sum\limits_{i=0}^{I-j-1} C_{i,j}\right)^3} + 2 \frac{\sum\limits_{i=0}^{I-j-1} C_{i,j}^2 \sum\limits_{k=1}^{I-j-1} \sum\limits_{n=0}^{k-1} C_{k,j} C_{n,j}}{\left(\sum\limits_{i=0}^{I-j-1} C_{i,j}\right)^4} \right)}{\sum\limits_{i=0}^{I-j-1} \left(1 - \frac{C_{i,j}}{\sum_{k=0}^{I-j-1} C_{k,j}} \right)^4 + \frac{\left(\sum_{i=0}^{I-j-1} C_{i,j}^2\right)^2 - \sum_{i=0}^{I-j-1} C_{i,j}^4}{\left(\sum_{i=0}^{I-j-1} C_{i,j}\right)^4}} \right] \tag{4.34}$$

for $j = 0, \dots, J-4$.
Because the estimation is unstable if there are too few data points (which for the estimation of the kurtosis is normally the case for the last three development years), we make the simplified approach of estimating $\widehat{\kappa}_{J-1}$, $\widehat{\kappa}_{J-2}$ and $\widehat{\kappa}_{J-3}$ by

$$\widehat{\kappa}_{J-1} = \widehat{\kappa}_{J-2} = \widehat{\kappa}_{J-3} = 3$$

This assumes an underlying normal distribution.
In the case of more accident years than development years, i.e. $I > J$, it is possible to use eq. (4.34) for $j \geq J-3$, because the proof of lemma 4.2.6 is not dependent on $j < J-3$. For each step that I is greater than J we can estimate one more factor $\widehat{\kappa}_j$ with eq. (4.34).

4.2.8 Remark. Note that while our estimators $\widehat{\kappa}_j^*$ and $\widehat{\sigma}_j^2$ are unbiased, this does not have to be the case for $\widehat{\kappa}_j$.

4.3 Skewness and kurtosis of the prediction error

Looking at our estimators for skewness and kurtosis from the previous sections we observe that we do not have any estimators for the last development year J, which is why we use a simplified normal distribution approach. This stems from the fact that we only have one observation so it is not possible to quantify the volatility of the individual development factors $F_{i,j+1}$ (see eq. (3.7)).
This means that we cannot estimate the higher-order moments of the ultimate claims and with that of the outstanding reserves. Additionally we do not have any estimators on the higher order moments of the prediction error, which would be desirable since we do not know the real values f_j and σ_j for $J = 0, \dots, J-1$ used in the estimators. Ideally we would want something like the MSEP but for the third and fourth moment. However trying to calculate this directly, one runs into analytical problems quickly.

For this reason we will go another route and use a simulation approach to estimate the distribution of the ultimate claims. This lets us then use empirical statistics to estimate higher moments and also other functions like the quantile of the distribution. We can even make histograms and compare them with fitted density curves of often used distributions to compare the shape of the probability density. We will expand on this approach in the next section.

Chapter 5

The simulation process [1]

The goal is now to predict the distribution of the ultimate claims $C_{i,J}$, or equivalently of the reserves R_i, where $i = 1, \dots, I$. Following the last chapter we will use a stochastic model to estimate the distribution of the ultimate claims.
In our stochastic model we will simulate the unknown cumulative claims $C_{i,j}$ for $i + j > I$ by sampling them from a given distribution using moment matching and our estimators for the first four moments. This way we will then have the full triangle of the cumulative claims, where the lower half was simulated according to our model. We can then compute statistics like the 99% quantile from the results of the simulation.
The idea behind this is that if the simulation number is high enough to get stable results, the moments for each single $\widehat{C}_{i,j}^{Sim}$ should follow our estimated moments, no matter which distribution is used. This holds true because we used moment matching to estimate the parameters for the distribution.
Let N be the number of simulations used. After running the simulation we will have N simulated results for each ultimate claim $C_{i,J}$, with $i = 1, \dots, I$. We can then use them to get a measure of the volatility of the end reserves. We will discuss this in more detail in section 5.2, but first we define the model specifications.

5.1 The model

If we want to use the MSEP calculations from section 3.3 in conjunction with the skewness and kurtosis estimators from chapter 4 we need to integrate the respective model assumptions into one combined assumption.
The following assumptions are our final expansion of the chain ladder model and will be used for the simulation models described in sections 5.2, 5.4 and 6.4.

[1] The simulation model is based on the model introduced in [Mor12] and expanded upon in [Mor13]. It is adapted slightly to fit the Wüthrich model.

5.1.1 Model assumptions (Simulation Model).

- Cumulative claims $C_{i,j}$ of different accident years are independent
- There exist constants $f_j > 0$, $\sigma_j > 0, \gamma_j \in \mathbb{R}, \zeta_j > 0$ and random variables $\varepsilon_{i,j+1}$ such that

$$C_{i,j+1} = f_j \ C_{i,j} + \sigma_j \ \sqrt{C_{i,j}} \ \varepsilon_{i,j+1} \qquad (5.1)$$
$$\forall \ i \in \{0, \dots, I\}, \text{ and } \forall \ j \in \{0, \dots, J-1\}$$

  where $\varepsilon_{i,j+1}$ are conditionally independent given $\mathcal{B}_0$, with $\mathbb{E}\left[\varepsilon_{i,j+1}|\mathcal{B}_0\right] = 0$, $\mathbb{E}\left[\varepsilon^2_{i,j+1}\middle|\mathcal{B}_0\right] = 1$, $\mathbb{E}\left[\varepsilon^3_{i,j+1}\middle|\mathcal{B}_0\right] = \gamma_j$, $\mathbb{E}\left[\varepsilon^4_{i,j+1}\middle|\mathcal{B}_0\right] = \kappa_j$ and $\mathbb{P}\left[C_{i,j+1} > 0|\mathcal{B}_0\right] = 1 \ \forall i \in \{0, \dots, I\}$, and $\forall j \in \{0, \dots, J-1\}$.

Model assumptions 5.1.1 imply model assumptions 3.3.10 and the assumptions on the higher-order moments of the $\varepsilon_{i,j+1}$ do not interfere with the MSEP calculations since only the first two moments are needed there, so all results still apply. For the skewness and kurtosis estimators we observe that eq. (5.1) yields

$$\begin{aligned}\mathbb{E}\left[\left(C_{i,j+1} - \mathbb{E}\left[C_{i,j+1}|\mathcal{D}_I\right]\right)^3\middle|D_I\right] &= \left(\sigma_j^2 C_{i,j}\right)^{\frac{3}{2}} \mathbb{E}\left[\varepsilon^3_{i,j+1}\middle|D_I\right] \\ &= \left(\sigma_j^2 C_{i,j}\right)^{\frac{3}{2}} \gamma_j\end{aligned}$$

Similarly we conclude

$$\begin{aligned}\mathbb{E}\left[\left(C_{i,j+1} - \mathbb{E}\left[C_{i,j+1}|\mathcal{D}_I\right]\right)^4\middle|D_I\right] &= \left(\sigma_j^2 C_{i,j}\right)^{\frac{4}{2}} \mathbb{E}\left[\varepsilon^4_{i,j+1}\middle|D_I\right] \\ &= \left(\sigma_j^2 C_{i,j}\right)^{\frac{3}{2}} \kappa_j\end{aligned}$$

Comparing these equations to eqs. (4.5) and (4.23) we can conclude that model assumptions 5.1.1 imply model assumptions 4.2.2 and that we can use the estimators from chapter 4.

We will denote the cumulative values simulated by our model by $\widehat{C}^{Sim}_{i,j}, (i,j) \in \mathcal{D}^c_I$. We can interpret them in two different ways, which influences the type of variance estimator we use in the simulation.

- First we can interprete the $\widehat{C}^{Sim}_{i,j}$ as realizations of the cumulative values $C_{i,j}$. Then we set

$$\mathbb{V}\left[\widehat{C}^{Sim}_{i,j}\middle|\mathcal{D}_I\right] = \widehat{C}_{i,j-1} \ \widehat{\sigma}_{j-1},$$

  for $j = 1, \dots, J$ and $i = I - j + 1, \dots, I$ and where for ease of notation

$$\widehat{C}_{i,j} = \begin{cases} C_{i,j} & 0 \le i \le I-j, j = 0, \dots, J \\ \widehat{C}^{Sim}_{i,j} & I-j < i \le I, j = 1, \dots, J \end{cases} \qquad (5.2)$$

  The problem with this approach is that it does not take the estimation error, arising from not knowing the real values f_j and s_j, into account.

- This is why we will use the second apprach from now on, where we interprete the $\widehat{C}^{Sim}_{i,j}$ as realizations of the estimated chain ladder values $\widehat{C}^{CL}_{i,j}$. As an estimator for the variance we derive a single step form of the MSEP estimator from section 3.3. Generalizing eq. (3.37) we have (see also eq. (5.11))

$$\widehat{\mathbf{msep}}_{C_{i,j}|\mathcal{D}_I}\left(\widehat{C}^{CL}_{i,j}\right) = \mathbb{E}\left[\left(\widehat{C}^{CL}_{i,j} - C_{i,j}\right)^2 \middle| \mathcal{D}_I\right]$$
$$= \left(\widehat{C}^{CL}_{i,j}\right)^2 \sum_{m=I-i}^{j-1} \frac{\widehat{\sigma}^2_m}{\widehat{f}^2_m\ \widehat{C}^{CL}_{i,j}} + C^2_{i,I-i}\left(\prod_{m=I-i}^{j-1}\left(\widehat{f}^2_m + \frac{\widehat{\sigma}^2_m}{{}^{[I-m-1]}S_m}\right) - \prod_{m=I-i}^{j-1}\widehat{f}^2_m\right)$$

for $j = 1, \ldots, J$ and $i = I - j + 1, \ldots, I$.
The estimator $\widehat{\mathbf{msep}}_{C_{i,j}|\mathcal{D}_I}$ for $\widehat{C}^{CL}_{i,j}$ takes the prediction error at each timestep $j = I - i + 1, \ldots, j$ into account. In our simulation model we simulate the claims triangle column after column so at step j all values $\widehat{C}_{\cdot,j-1}$ are known. So for the variance expression we only use a single step form of the above equation. First we define

$$\mathcal{E}_{I,j} = \{C_{i,k} : i + k \leq I, 0 \leq i \leq I, 0 \leq k \leq J\} \cup \{\widehat{C}_{i,k} : 0 \leq i \leq I, 0 \leq k \leq J\}, \quad (5.3)$$

which is visualized in table 5.1.

AY/DY	0	1	...	$j-1$	j	...	...	J
0	$C_{0,0}$	$C_{0,1}$	...	...	$C_{0,j}$	...	$\cdot^{\cdot^{\cdot}}$	$C_{0,J}$
1	$C_{1,0}$	$\cdot^{\cdot^{\cdot}}$	$\cdot^{\cdot^{\cdot}}$	$\cdot^{\cdot^{\cdot}}$	$\cdot^{\cdot^{\cdot}}$	$\cdot^{\cdot^{\cdot}}$	$C_{1,J-1}$	
$\vdots$	$\vdots$	$\cdot^{\cdot^{\cdot}}$	$\cdot^{\cdot^{\cdot}}$	$\cdot^{\cdot^{\cdot}}$	$\cdot^{\cdot^{\cdot}}$	$\cdot^{\cdot^{\cdot}}$		
$I-j$	$\vdots$	$\cdot^{\cdot^{\cdot}}$	$\cdot^{\cdot^{\cdot}}$	$\cdot^{\cdot^{\cdot}}$	$C_{I-j,j}$			
$I-j+1$	$\vdots$	$\cdot^{\cdot^{\cdot}}$	$\cdot^{\cdot^{\cdot}}$	$C_{I-j+,j-1}$	$\widehat{C}^{Sim}_{I-j+1,j}$			
$\vdots$	$\vdots$	$\cdot^{\cdot^{\cdot}}$	$\cdot^{\cdot^{\cdot}}$	$\cdot^{\cdot^{\cdot}}$	$\vdots$			
$\vdots$	$\vdots$	$C_{I-1,1}$	$\cdot^{\cdot^{\cdot}}$	$\cdot^{\cdot^{\cdot}}$	$\vdots$			
I	$C_{I,0}$	$\widehat{C}^{Sim}_{I,1}$	...	$\widehat{C}^{Sim}_{I,j-1}$	$\widehat{C}^{Sim}_{I,j}$			

Table 5.1: $\mathcal{E}_{I,j-1}$ in green

For the values on the second diagonal we have

$$\widehat{\mathbf{msep}}_{C_{I-j+1,j}|\mathcal{D}_I}\left(\widehat{C}^{CL}_{I-j+1,j}\right) = \mathbb{E}\left[\left(\widehat{C}^{CL}_{I-j+1,j} - C_{I-j+1,j}\right)^2 \middle| \mathcal{D}_I\right]$$
$$= \left(\widehat{C}^{CL}_{I-j+1,j}\right)^2 \frac{\widehat{\sigma}^2_{j-1}}{\widehat{f}^2_{j-1}\ C_{I-j+1,j-1}} + \left(C_{I-j+1,j-1}\right)^2\left(\widehat{f}^2_{j-1} + \frac{\widehat{\sigma}^2_{j-1}}{{}^{[I-j]}S_{j-1}} - \widehat{f}^2_{j-1}\right)$$
$$\stackrel{(3.3)}{=} \left(C_{I-j+1,j-1}\right)^2 \frac{\widehat{\sigma}^2_{j-1}}{C_{I-j+1,j-1}} + \left(C_{I-j+1,j-1}\right)^2 \frac{\widehat{\sigma}^2_{j-1}}{{}^{[I-j]}S_{j-1}}$$

$$= C_{I-j+1,j-1}\ \widehat{\sigma}^2_{j-1} + (C_{I-j+1,j-1})^2\ \frac{\widehat{\sigma}^2_{j-1}}{{}^{[I-j]}S_{j-1}}$$

$$= C_{I-j+1,j-1}\ \widehat{\sigma}^2_{j-1} \left(1 + \frac{C_{I-j+1,j-1}}{{}^{[I-j]}S_{j-1}}\right)$$

Based on this equation we define the following estimator for the variance of the $\widehat{C}^{Sim}_{i,j}$

$$\mathbb{V}\left[\widehat{C}^{Sim}_{i,j}\middle|\mathcal{E}_{I,j-1}\right] = \widehat{\mathbf{msep}}_{C_{i,j}|\mathcal{E}_{I,j-1}}\left(\widehat{C}^{CL}_{i,j}\right) := C_{i,j-1}\ \widehat{\sigma}^2_{j-1}\left(1 + \frac{C_{i,j-1}}{\sum_{k=0}^{I-j} C_{k,j-1}}\right) \quad (5.4)$$

for $j = 1, \ldots, J$ and $i = I - j + 1, \ldots, I$. From here on forth we will use this expression as the estimator of the variance to match the chain ladder MSEP in the overall result.

5.1.2 Remark. Note that we do not have estimators for the skewness and the kurtosis of the chain ladder estimators $\widehat{C}^{CL}_{i,j}$ so we will take the estimators γ_j and κ_j, $j = 1, \ldots, J$, from the original time series process $C_{i,j}$.

5.2 The simulation algorithm

We will now go through the simulation process step for step. See table 5.2 for the color scheme used in this chapter.

Color	**Explanation**
blue	deterministic data
red	data to be simulated at current step
green	already simulated data

Table 5.2: Color table for this chapter

1. The first step of the simulation process is computing the values

   $$\widehat{f}_j,\ \widehat{\sigma}_j,\ \widehat{\gamma}_j,\ \widehat{\kappa}_j \text{ for } j = 0, \ldots, J-1.$$

   Our claims triangle at this point looks like table 5.3.

   We will now simulate each development year $j = 1, \ldots, J$ to fill up the claims triangle. For each development year j we have constant skewness and kurtosis (see model assumptions 4.2.2), which are given by the estimators $\widehat{\gamma}_{j-1}$ and $\widehat{\kappa}_{j-1}$. The expected value and the variance differ for each accident year i and depend also on the realizations at step $j - 1$.

AY/DY	0	1	...	k	...	...	J
0	$C_{0,0}$	$C_{0,1}$	...	$C_{0,k}$	...	$\iddots$	$C_{0,J}$
1	$C_{1,0}$	$\iddots$	$\iddots$	$\iddots$	$\iddots$	$C_{1,J-1}$	
$\vdots$	$\vdots$	$\iddots$	$\iddots$	$\iddots$	$\iddots$		
i	$\vdots$	$\iddots$	$\iddots$	$C_{i,k}$			
$\vdots$	$\vdots$	$\iddots$	$\iddots$				
$\vdots$	$\vdots$	$C_{I-1,1}$					
I	$C_{I,0}$						

Table 5.3: Simulation step 1

So this means that at each step $j = 1, \dots, J$ we have

$$\mathbb{E}\left[\widehat{C}^{Sim}_{i,j}\middle|\mathcal{E}_{I,j-1}\right] = \widehat{C}_{i,j-1}\ \widehat{f}_{j-1},$$
$$\mathbb{V}\left[\widehat{C}^{Sim}_{i,j}\middle|\mathcal{E}_{I,j-1}\right] = \widehat{C}_{i,j-1}\ \widehat{\sigma}^2_{j-1}\left(1 + \frac{C_{i,j-1}}{\sum_{k=0}^{I-j} C_{k,j-1}}\right),$$
$$\text{Skew}\left(\widehat{C}^{Sim}_{i,j}\middle|\mathcal{E}_{I,j-1}\right) = \widehat{\gamma}_{j-1},$$
$$\text{Kurt}\left(\widehat{C}^{Sim}_{i,j}\middle|\mathcal{E}_{I,j-1}\right) = \widehat{\kappa}_{j-1},$$

for $i = I - j + 1, \dots, I$ and where the $\widehat{C}_{i,j}$ are defined as in eq. (5.2) and the $\mathcal{E}_{I,j-1}$ as in eq. (5.3). With these equations we can now simulate the whole triangle development year after development year.

2. This means that in the second step we simulate $C_{I,2}$ by computing its mean, variance, skewness and kurtosis and then sampling from a fitted distribution (for more specifics on the sampling see section 5.4 and section 6.4). This leaves us with table 5.4.

3. At step k we have table 5.5. Note that we do not sample each $\widehat{C}^{Sim}_{m,k}$, where $m = i+1, \dots, I$, individually but together, because of implicit correlations in the chain ladder model, which we will discuss in section 5.3.

4. At the last point of our simulation we have table 5.6, where

$$\widehat{R}_i = \widehat{C}_{i,J} - \widehat{C}_{i,I-i} \text{ for } i = 1, \dots, I.$$

5. Now we can compute the total reserve by taking

$$\widehat{\mathbf{TR}} = \sum_{i=1}^{I} \widehat{R}_i \tag{5.5}$$

AY/DY	0	1	...	k	...	...	J
0	$C_{0,0}$	$C_{0,1}$	...	$C_{0,k}$	...	⋰	$C_{0,J}$
1	$C_{1,0}$	⋰	⋰	⋰	⋰	$C_{1,J-1}$	
⋮	⋮	⋰	⋰	⋰	⋰		
i	⋮	⋰	⋰	$C_{i,k}$			
⋮	⋮	⋰	⋰				
⋮	⋮	$C_{I-1,1}$					
I	$C_{I,0}$	$\widehat{C}_{I,1}^{Sim}$					

Table 5.4: Simulation step 2

AY/DY	0	1	...	k	...	...	J
0	$C_{0,0}$	$C_{0,1}$	...	$C_{0,k}$	...	⋰	$C_{0,J}$
1	$C_{1,0}$	⋰	⋰	⋰	⋰	$C_{1,J-1}$	
⋮	⋮	⋰	⋰	⋰	⋰		
i	⋮	⋰	⋰	$C_{i,k}$			
⋮	⋮	⋰	⋰	$\widehat{C}_{i+1,k}^{Sim}$			
⋮	⋮	$C_{I-1,1}$	⋰	⋮			
I	$C_{I,0}$	$\widehat{C}_{I,1}^{Sim}$	...	$\widehat{C}_{I,k}^{Sim}$			

Table 5.5: Simulation step k

For comparison purposes we also define $\widehat{\mathbf{TR}}^{CL} := \sum_{i=1}^{I} \widehat{R}_i^{CL}$ as the total chain ladder reserve.

If we repeat this simulation process N times we get N different reserves. Since we used moment matching to simulate the data we have

$$\begin{aligned}
\mathbb{E}\left[\widehat{\mathbf{TR}}\right] &= \mathbb{E}\left[\sum_{i=1}^{I} \widehat{R}_i\right] \\
&= \mathbb{E}\left[\sum_{i=1}^{I} \left(\widehat{C}_{i,J} - \widehat{C}_{i,I-i}\right)\right] \\
&= \sum_{i=1}^{I} \left(\mathbb{E}\left[\widehat{C}_{i,J}^{Sim}\right] - C_{i,I-i}\right)
\end{aligned}$$

$$= \sum_{i=1}^{I} \left(\mathbb{E}\left[\widehat{C}_{i,J}^{CL} \right] - C_{i,I-i} \right)$$
$$= \mathbb{E}\left[\widehat{\mathbf{TR}}^{CL} \right]$$

So our outstanding reserves (and with it the total reserve) have the same expectation as the standard chain ladder estimator. Now we can use empirical estimators on our $\widehat{\mathbf{TR}}$ to also compute higher moments or other statistics like a quantile (making Value at Risk calculations possible).

AY/DY	0	1	...	k	...	J	$\widehat{R}$
0	$C_{0,0}$	$C_{0,1}$	...	$C_{0,k}$	$\iddots$	$C_{0,J}$	0
1	$C_{1,0}$	$\iddots$	$\iddots$	$\iddots$	$C_{1,J-1}$	$\widehat{C}_{1,J}^{Sim}$	$\widehat{R}_1$
$\vdots$	$\vdots$	$\iddots$	$\iddots$	$\iddots$	$\iddots$	$\vdots$	$\vdots$
i	$\vdots$	$\iddots$	$\iddots$	$C_{i,k}$	$\iddots$	$\vdots$	$\widehat{R}_i$
$\vdots$	$\vdots$	$\iddots$	$\iddots$	$\widehat{C}_{i+1,k}^{Sim}$	$\iddots$	$\vdots$	$\vdots$
$\vdots$	$\vdots$	$C_{I-1,1}$	$\iddots$	$\vdots$	$\iddots$	$\vdots$	$\vdots$
I	$C_{I,0}$	$\widehat{C}_{I,1}^{Sim}$	...	$\widehat{C}_{I,k}^{Sim}$	...	$\widehat{C}_{I,J}^{Sim}$	$\widehat{R}_J$

Table 5.6: Simulation step J

5.2.1 Remark. It must be emphasized that while it is true that for the first four moments of the ultimate claims the type of distribution used to simulate the $\widehat{C}_{i,j}^{Sim}$ is not of any importance (because of the use of moment matching) the same cannot be said for other measures (like higher moments or even the whole empirical distribution function), as the type of distribution used does influence these results.

We will now discuss two different ways to realize the above simulation model. The first one was introduced in [Mor12] (and further enhanced in [Mor13]) and uses the Generalized Pareto distribution (see appendix A.1.3) as the underlying distribution. One big shortcoming of this method is that since GPD only has three parameters we have to decide which moments we want to fit the parameters to. For this reason we will then introduce the Pearson System in section 6.1 and use this set of distributions in section 6.4 for a simulation model incorporating all four of our moment estimators.

5.3 Correlation

According to model assumptions 4.2.2 claims of different accident years i are independent of each other. While this holds true for the known claims $C_{i,j}$ we do have a dependency

structure for our estimators $\widehat{C}_{i,j}^{CL}$. This can be attributed to the fact that their estimation depends on the same set of data for each development year j (the factors $\widehat{f}_j$ do not depend on i).
If follows that we cannot simulate each $\widehat{C}_{i,j}^{CL}$ independently but instead have to take the dependency structure into account and simulate each column of the claims triangle together. We will follow Dal Moro's idea in [Mor12] to use the formula of the MSEP for aggregated accident years to compute the linear correlation matrix. But instead of Mack's formula we will use the Wüthrich formula presented in estimator 3.3.18.
This approach will be explained in detail in section 5.3.2, but first we make some general definitions.

5.3.1 General definitions

5.3.1 Definition. Let $X_1, \dots, X_n$ be n random variables. We define for $i, k = 1, \dots, n$

$$\sigma_{X_i} := \sqrt{\mathbb{V}[X_i]},$$

as the standard deviation of the variable X_i,

$$\operatorname{Cov}(X_i, X_k) := \mathbb{E}\left[(X_i - \mathbb{E}[X_i])\ (X_k - \mathbb{E}[X_k])\right] \tag{5.6}$$

as the covariance between two variables and

$$\operatorname{Corr}(X_i, X_k) := \frac{\operatorname{Cov}(X_i, X_k)}{\sigma_{X_i}\sigma_{X_k}}$$

as the linear correlation between the random variables X_i and X_k.
The linear correlation matrix ρ is then the $n \times n$ matrix with the entries

$$\rho_{i,k} = \operatorname{Corr}(X_i, X_k).$$

The aim will now be to estimate the correlation matrices for each column $j = 2, \dots, J$ of the claims triangle (at time 1 we only simulate one value, hence we need no correlation matrix). This means that at each point j we have to estimate the correlation for the red area in table 5.7.
Observe that at each step $j = 2, \dots, J$ we have to correlate the set of random variables

$$\left\{\widehat{C}_{I-j+1,j}^{CL}, \dots, \widehat{C}_{I,j}^{CL}\right\}$$

which has j entries.

5.3.2 Definition (Chain Ladder Correlation Matrices)**.** The above means we are searching for $J-1$ matrices $\rho^{\langle j\rangle}$, where $j = 2, \dots, J$ with the properties

- $\rho^{\langle j\rangle} \in [0,1]^{j\times j}$
- For each $i, k = 1, \dots, j$ we have

$$\rho_{i,k}^{\langle j\rangle} = \operatorname{Corr}\left(\widehat{C}_{I-j+i,j}^{CL}, \widehat{C}_{I-j+k,j}^{CL}\right)$$

AY/DY	0	1	$\dots$	j	$\dots$	$\dots$	J
0	$C_{0,0}$	$C_{0,1}$	$\dots$	$C_{0,j}$	$\dots$	$\iddots$	$C_{0,J}$
1	$C_{1,0}$	$\iddots$	$\iddots$	$\iddots$	$\iddots$	$C_{1,J-1}$	
$\vdots$	$\vdots$	$\iddots$	$\iddots$	$\iddots$	$\iddots$		
$I-j$	$\vdots$	$\iddots$	$\iddots$	$C_{I-j,j}$			
$\vdots$	$\vdots$	$\iddots$	$\iddots$	$\widehat{C}^{CL}_{I-j+1,j}$			
$\vdots$	$\vdots$	$C_{I-1,1}$		$\vdots$			
I	$C_{I,0}$			$\widehat{C}^{CL}_{I,j}$			

Table 5.7: Simulation step j

To make the notation a little easier we make an index shift and define the proxy variables $\mathrm{r}^{\langle j \rangle}_{i,k}$ where

$$\mathrm{r}^{\langle j \rangle}_{i,k} = \mathrm{Corr}\left(\widehat{C}^{CL}_{i,j}, \widehat{C}^{CL}_{k,j}\right) \quad \left(= \rho^{\langle j \rangle}_{i-I+j,k-I+j}\right) \tag{5.7}$$

for $j = 2, \dots, J$ and $i, k = I - j + 1, \dots, I$.

5.3.2 The estimation method

Now that we have all the formalizations we can start with the estimation.
First note that the variance of the sum of n correlated random variables X_i, where $i = 1, \dots, n$ is given by

$$\begin{aligned} \mathbb{V}\left(\sum_{i=1}^{n} X_i\right) &= \sum_{i=1}^{n} \mathbb{V}(X_i) + 2 \sum_{1 \le i < k \le n} \mathrm{Cov}\left(X_i, X_k\right) \\ &= \sum_{i=1}^{n} \mathbb{V}(X_i) + 2 \sum_{1 \le i < k \le n} \rho\left(X_i, X_k\right) \sigma_{X_i} \sigma_{X_k} \end{aligned} \tag{5.8}$$

If we take another look at eq. (3.21) we have

$$\begin{aligned} \widehat{\mathbf{msep}}_{\sum_i C_{i,J} | \mathcal{D}_I}\left(\sum_{i=1}^{I} \widehat{C}^{CL}_{i,J}\right) &= \sum_{i=1}^{I} \widehat{\mathbf{msep}}_{C_{i,J} | \mathcal{D}_I}\left(\widehat{C}^{CL}_{i,J}\right) \\ &+ 2 \sum_{1 \le i < k \le I} C_{i,I-i}\, \widehat{C}^{CL}_{k,I-i} \left(\prod_{j=I-i}^{J-1} \left(\widehat{f}_j^2 + \frac{\widehat{\sigma}_j^2}{{}^{[I-j-1]}S_j}\right) - \prod_{j=I-i}^{J-1} \widehat{f}_j^2\right). \end{aligned} \tag{5.9}$$

If we compare eqs. (5.8) and (5.9) we can observe that they are of similar form. If we identify the MSEP with the variance of our factors $\widehat{C}_{i,J}^{CL}$ we can compare the coefficients of the sums to get linear correlation estimators (for now only for time J). It follows that

$$\widehat{\mathrm{r}}_{i,k}^{\langle J\rangle} = \frac{C_{i,I-i}\ \widehat{C}_{k,I-i}^{CL}\left(\prod_{j=I-i}^{J-1}\left(\widehat{f}_j^2 + \frac{\widehat{\sigma}_j^2}{{}^{[I-j-1]}S_j}\right) - \prod_{j=I-i}^{J-1}\widehat{f}_j^2\right)}{\sqrt{\widehat{\mathbf{msep}}_{C_{i,J}|\mathcal{D}_I}\ \widehat{\mathbf{msep}}_{C_{k,J}|\mathcal{D}_I}}} \tag{5.10}$$

We now have an estimator for the correlation matrix $\mathrm{r}^{\langle J\rangle}$ at time J. To get $\mathrm{r}^{\langle j\rangle}$ for $j < J$ we simply have to generalize eq. (5.9) for every time-point j. This can easily be done because of the multiplicative structure of the chain ladder method. Thus we have (remember that we assume $I = J$)

$$\begin{aligned}\widehat{\mathbf{msep}}_{C_{i,j}|\mathcal{D}_I}\left(\widehat{C}_{i,j}^{CL}\right) &= \mathbb{E}\left[\left(\widehat{C}_{i,j}^{CL} - C_{i,j}\right)^2\middle|\mathcal{D}_I\right]\\ &= \left(\widehat{C}_{i,j}^{CL}\right)^2\sum_{m=I-i}^{j-1}\frac{\widehat{\sigma}_m^2}{\widehat{f}_m^2\ \widehat{C}_{i,j}^{CL}} + C_{i,I-i}^2\left(\prod_{m=I-i}^{j-1}\left(\widehat{f}_m^2 + \frac{\widehat{\sigma}_m^2}{{}^{[I-m-1]}S_m}\right) - \prod_{m=I-i}^{j-1}\widehat{f}_m^2\right)\end{aligned} \tag{5.11}$$

for $i = I - j + 1, \ldots, I$ and

$$\begin{aligned}\widehat{\mathbf{msep}}_{\sum_i C_{i,j}|\mathcal{D}_I}\left(\sum_{i=I-j+1}^{I}\widehat{C}_{i,j}^{CL}\right) &= \sum_{i=I-j+1}^{I}\widehat{\mathbf{msep}}_{C_{i,j}|\mathcal{D}_I}\left(\widehat{C}_{i,j}^{CL}\right)\\ &+ 2\sum_{I-j+1\le i<k\le I}C_{i,I-i}\ \widehat{C}_{k,I-i}^{CL}\left(\prod_{m=I-i}^{j-1}\left(\widehat{f}_m^2 + \frac{\widehat{\sigma}_m^2}{{}^{[I-m-1]}S_m}\right) - \prod_{m=I-i}^{j-1}\widehat{f}_m^2\right).\end{aligned} \tag{5.12}$$

Similar to above we conclude

5.3.3 Estimator (Correlation Estimator)**.**

$$\widehat{\mathrm{r}}_{i,k}^{\langle j\rangle} = \frac{C_{i,I-i}\ \widehat{C}_{k,I-i}^{CL}\left(\prod_{m=I-i}^{j-1}\left(\widehat{f}_m^2 + \frac{\widehat{\sigma}_m^2}{{}^{[I-m-1]}S_m}\right) - \prod_{m=I-i}^{j-1}\widehat{f}_m^2\right)}{\sqrt{\widehat{\mathbf{msep}}_{C_{i,j}|\mathcal{D}_I}\ \widehat{\mathbf{msep}}_{C_{k,j}|\mathcal{D}_I}}} \tag{5.13}$$

Substituting back according to eq. (5.7) yields the estimators of the matrices $\rho^{\langle j\rangle}$.

5.3.4 Estimator. For $j = 2, \ldots, J$ we estimate the entries of the correlation matrix $\widehat{\rho}^{\langle j\rangle}$ by

$$\widehat{\rho}_{i,k}^{\langle j\rangle} = \widehat{\mathrm{r}}_{I-j+i,I-j+k}^{\langle j\rangle}, \tag{5.14}$$

where $i, k = 1, \ldots, j$.

Having estimated the correlation matrices $\rho^{\langle j\rangle}$ for $j = 2, \ldots, J$ we will use them to sample correlated variables in our simulation process. We accomplish this by using a Gaussian copula in combination with inversion sampling (see appendix A.2.1 and definition A.2.7).

5.4 The GPD simulation

In this section we will introduce the GPD approach presented in [Mor13]. We follow the general algorithm presented in section 5.2 and use the generalized Pareto distribution to sample the $\widehat{C}_{i,j}^{Sim}$. Since this distribution only has three parameters it is not possible to use all our estimated moments for moment-matching. In [Mor13] it is suggested to use the first two moments in any case and then choose which of the third or fourth moment to use for the last equation.
For our samples we will fit the simulation to the skewness estimator, for different approaches consult the original paper.

5.4.1 Remark. In the original paper two models are introduced. One where the data is fitted to a generalized Pareto distribution and one where it is fitted to a generalized extreme value distribution. We will only discuss the first one here, since both implementations follow the same structure. For further information consult [Mor12] or [Mor13].

The simulation follows the algorithm presented in section 5.2 but we will discuss in detail what happens at each step $j = 1, \dots, J$. As a reminder at step j we have the situation of table 5.8.

AY/DY	0	1	…	j	…	…	J
0	$C_{0,0}$	$C_{0,1}$	…	$C_{0,j}$	…	⋰	$C_{0,J}$
1	$C_{1,0}$	⋰	⋰	⋰	⋰	$C_{1,J-1}$	
⋮	⋮	⋰	⋰	⋰	⋰		
$I-j$	⋮	⋰	⋰	$C_{I-j,j}$			
⋮	⋮	⋰	⋰	$\widehat{C}_{I-j+1,j}^{Sim}$			
⋮	⋮	$C_{I-1,1}$	⋰	⋮			
I	$C_{I,0}$	$\widehat{C}_{I,1}^{Sim}$	…	$\widehat{C}_{I,j}^{Sim}$			

Table 5.8: Simulation step j

At step j all parameters $\widehat{f}_j$, $\widehat{\sigma}_j$ and $\widehat{\gamma}_j$, as well as the set $\mathcal{E}_{I,j-1}$ (see eq. (5.3)) are known. So we can calculate the three moments of the $\widehat{C}_{i,j}^{Sim}$ (the kurtosis is not needed here). We have

$$\begin{aligned}
\mathbb{E}\left[\widehat{C}_{i,j}^{Sim}\middle|\mathcal{E}_{I,j-1}\right] &= \widehat{C}_{i,j-1}\ \widehat{f}_{j-1} \\
\mathbb{V}\left[\widehat{C}_{i,j}^{Sim}\middle|\mathcal{E}_{I,j-1}\right] &= \widehat{C}_{i,j-1}\ \widehat{\sigma}_{j-1}^2\left(1 + \frac{C_{i,j-1}}{\sum_{k=0}^{I-j} C_{k,j-1}}\right), \\
\text{Skew}\left(\widehat{C}_{i,j}^{Sim}\middle|\mathcal{E}_{I,j-1}\right) &= \widehat{\gamma}_{j-1},
\end{aligned}$$

for $j = 1, \dots, J$ and $i = I - j + 1, \dots, I$.
We will leave out the conditional expectation from the following formulas as to not overload the notation, but remember that all calculations are done on the conditional probability measure $\mathbb{P}(\cdot|\mathcal{E}_{I,j-1})$. In this model we suppose that $\widehat{C}_{i,j}^{Sim} \sim \text{GPD}(\mu_{i,j}; s_{i,j}; \xi_j)$. We use the method of moments to estimate the parameters from the above equations. We have

$$\mathbb{E}\left[\widehat{C}_{i,j}^{Sim}\right] = \widehat{C}_{i,j-1}\ \widehat{f}_{j-1} \stackrel{!}{=} \mu_{i,j} + \frac{s_{i,j}}{1-\xi_j}\ , \tag{5.15a}$$

$$\mathbb{V}\left[\widehat{C}_{i,j}^{Sim}\right] = \widehat{C}_{i,j-1}\ \widehat{\sigma}_{j-1}^2 \left(1 + \frac{C_{i,j-1}}{\sum_{k=0}^{I-j} C_{k,j-1}}\right) \stackrel{!}{=} \frac{s_{i,j}^2}{(1-\xi_j)(1-2\xi_j)}\ , \tag{5.15b}$$

$$\text{Skew}\left(\widehat{C}_{i,j}^{Sim}\right) = \widehat{\gamma}_{j-1} \stackrel{!}{=} \frac{2(1+\xi_j)\sqrt{1-2\xi_j}}{1-3\xi_j}\ . \tag{5.15c}$$

We can then use eqs. (5.15a) to (5.15c) to calculate the parameters $\mu_{i,j}, s_{i,j}$ and ξ_j. Then we use inversion sampling to sample the random variable $\widehat{C}_{i,j}^{Sim}$. If $U_{i,j}$ is a sample from a uniform distribution on $(0, 1)$ we have that

$$\widehat{C}_{i,j}^{Sim} = \mu_{i,j} + \frac{s_{i,j}\left(U_{i,j}^{-\xi_j} - 1\right)}{\xi_j} \tag{5.16}$$

is a sample from a random variable $X \sim \text{GPD}(\mu_{i,j}; s_{i,j}; \xi_j)$. To sample the variable $U_{i,j}$ for $i = I - j + 1, \dots, I$ we follow a similar route as described in section 5.3. To estimate the correlation matrix $\rho^{\langle j \rangle}$ Dal Moro uses the same coefficient comparison but with the Mack formula for the MSEP instead of our approach with the Wüthrich formula. We will state only the results here. Dal Moro estimates the correlation for each development year $j = 1, \dots, J$ and $i = I - j + 1, \dots, I$ with

$$\widehat{\text{Corr}}\left(\widehat{C}_{i,j}^{Sim}, \widehat{C}_{k,j}^{Sim}\right) := \frac{\widehat{C}_{i,j}^{Sim}\ \widehat{C}_{k,j}^{Sim} \sum_{m=I-i}^{j-1} \frac{\widehat{\sigma}_m^2}{\widehat{f}_m^2 \sum_{n=0}^{I-m-1} C_{n,m}}}{\sqrt{\widehat{\mathbf{msep}}_{C_{i,j}|\mathcal{D}_I}\left(\widehat{C}_{i,j}^{CL}\right)}\sqrt{\widehat{\mathbf{msep}}_{C_{i,j}|\mathcal{D}_I}\left(\widehat{C}_{i,j}^{CL}\right)}}$$

where

$$\widehat{\mathbf{msep}}_{C_{i,j}|\mathcal{D}_I}\left(\widehat{C}_{i,j}^{CL}\right) = \left(\widehat{C}_{i,j}^{CL}\right)^2 \sum_{k=I-i}^{j-1} \frac{\widehat{\sigma}_k^2}{\widehat{f}_k^2}\left(\frac{1}{\widehat{C}_{i,k}^{CL}} + \frac{1}{\sum_{n=0}^{I-k-1} C_{n,k}}\right).$$

We then use a Gaussian copula to sample the uniformly distributed random numbers $\{U_{I-j+1,j}, \dots, U_{I,j}\}$ and use eq. (5.16) to calculate the set $\left\{\widehat{C}_{I-j+1,j}^{Sim}, \dots, \widehat{C}_{I,j}^{Sim}\right\}$.

Chapter 6

Simulation with four moments

In this chapter we will define a set of distributions called the Pearson system, named after the English mathematician Karl Pearson. One reason this system is a natural fit for our simulation is that we have assumed constant skewness and kurtosis for each development year $j,\ j = 1, \ldots, J$ (see model assumptions 4.2.2). This fits naturally with the way the Pearson coefficients are computed. In the Pearson system one can use the skewness and kurtosis of a random variable to identify the type of Pearson distribution (see table 6.1). Since both are constant for one development year j we have one type of distribution per development year. The resulting variables then get shifted with the expected value and the variance,which are also dependent on the accident year $i = 1, \ldots, I$, to match all four moments (cf. section 6.3).

6.1 The Pearson System [1]

The Pearson system is a system of distributions, where for every member the probability density function (pdf) $f(x)$ satisfies the differential equation

$$\frac{1}{f}\frac{\mathrm{d}f}{\mathrm{d}x} = -\frac{a + x}{c_0 + c_1 x + c_2 x^2} \tag{6.1}$$

This means that the shape of the distribution depends on the parameters a, c_0, c_1, and c_2. We have that

- $\frac{\mathrm{d}f}{\mathrm{d}x} = 0$ when $x = -a$ or when $f \equiv 0$
- f is finite when $-a$ is not a root of $c_0 + c_1 x + c_2 x^2 = 0$

Since we are not just looking for any functions $f : \mathbb{R} \to \mathbb{R}$ but for probability distribution functions we have the extra conditions that

$$f(x) \geq 0,\ \forall x \in \mathbb{R} \tag{6.2}$$

[1] This derivation of the Pearson system is taken from [JKB94a]

and

$$\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = 1. \tag{6.3}$$

From this we can deduce that $\lim_{x\to\infty} f(x) = 0$ and that $\lim_{x\to\infty} \frac{\mathrm{d}f}{\mathrm{d}x} = 0$. Pearson classifies the distributions that solve eq. (6.1) into different types depending on the values a, c_0, c_1, and c_2. The solutions of eq. (6.1) depend on the nature of the roots of

$$c_0 + c_1 x + c_2 x^2 = 0. \tag{6.4}$$

We will now look at the different possible ranges of the above parameters to classify the distribution into different types.

6.1.1 Pearson Type 0 - the normal distribution

If $c_1 = c_2 = 0$ then eq. (6.1) simplifies itself to

$$\frac{\mathrm{d}\log f(x)}{\mathrm{d}x} = -\frac{a+x}{c_0}$$

which is solved by

$$f(x) = K \exp\left(-\frac{(x+a)^2}{2c_0}\right)$$

where K is the constant chosen so that f fulfills eq. (6.3). It also follows that $c_0 > 0$ and that $K = \sqrt{2\pi c_0}$. This means that our distribution is a normal distribution with expected value $-a$ and variance c_0. For further details on the normal distribution see appendix A.1.1.

For the derivation of the other types we will suppose that the origin of the scale of X has been chosen so that $\mathbb{E}[X] = 0$. This assumption does not limit our derivations since each random variable (with finite expectation) can be shifted, so that its shifted expected value is zero.

6.1.2 Pearson Type 1 (akin to the beta distribution)

Let the roots of eq. (6.4) be denoted by a_1 and a_2. We say that f is of type 1 if

$$a_1, a_2 \in \mathbb{R} \text{ and } a_1 < 0 < a_2$$

This means that

$$c_0 + c_1 x + c_2 x^2 = -c_2 (x - a_1)(a_2 - x)$$

and we can rewrite eq. (6.1) to

$$\frac{1}{f}\frac{\mathrm{d}f}{\mathrm{d}x} = -\frac{a+x}{(x-a_1)(a_2-x)}$$

The solution of which is

$$f(x) = K\,(x-a_1)^{m_1}\,(a_2-x)^{m_2} \tag{6.5}$$

with

$$m_1 = \frac{a + a_1}{c_2(a_2 - a_1)}$$
$$m_2 = -\frac{a + a_2}{c_2(a_2 - a_1)}$$

To satisfy eq. (6.2) we limit the range of the variable x to $a_1 < x < a_2$ so that both $x - a_1$ and $a_2 - x$ are positive. Equation (6.5) represents a proper pdf if $m_1, m_2 > -1$ and corresponds to a generalized form of the beta distribution (see appendix A.1.4).

6.1.3 Pearson Type 2 (akin to the symmetrical beta distribution)

The second type is now just a special case of the first one. We say that f is of type 2 if f follows the form (6.5) with $m_1 = m_2$, which corresponds to a symmetrical generalized beta distribution.

6.1.4 Pearson Type 3 (akin to the gamma distribution)

Type 3 corresponds to the case $c_2 = 0$ (and $c_1 \neq 0$). Then eq. (6.1) transforms to

$$\frac{\mathrm{d} \log f(x)}{\mathrm{d}x} = -\frac{a + x}{c_0 + c_1 x} = -\frac{1}{c_1} - \frac{a - \frac{c_0}{c_1}}{c_0 + c_1 x},$$

which is solved by

$$f(x) = K\left(c_0 + c_1 x\right)^m \exp\left(\frac{-x}{c_1}\right), \tag{6.6}$$

where

$$m = \frac{\frac{c_0}{c_1} - a}{c_1}.$$

For the range of x we set the boundaries

$$\begin{cases} x > -\frac{c_0}{c_1} & \text{if } c_1 > 0, \\ x < -\frac{c_0}{c_1} & \text{if } c_1 < 0. \end{cases}$$

Type 3 corresponds to a gamma distribution, which is defined in appendix A.1.5.

6.1.5 Pearson Type 4

The type 4 distribution which does not correspond to any standard distribution denotes the case where eq. (6.4) does not have any real roots. We use the identity

$$c_0 + c_1 x + c_2 x^2 = C_0 + c_2\left(x + C_1\right)^2$$

with $C_0 = c_0 - \frac{1}{4}\frac{c_1^2}{c_2}$ and $C_1 = \frac{1}{2}\frac{c_1}{c_2}$. Then we can write eq. (6.1) as

$$\frac{\mathrm{d} \log f(x)}{\mathrm{d}x} = -\frac{-\left(x + C_1\right) - \left(a - C_1\right)}{C_0 + c_2\left(x + C_1\right)^2}$$

From this we can conclude that

$$f(x) = K\left[C_0 + c_2\left(x + C_1\right)^2\right]^{-\frac{1}{2c_2}} \exp\left(-\frac{a - C_1}{\sqrt{c_2 C_0}} \tan^{-1}\left(\frac{x + C_1}{\sqrt{\frac{C_0}{c_2}}}\right)\right) \tag{6.7}$$

6.1.6 Pearson Type 5 (akin to the inverse gamma location-scale distribution)

For this type we consider the case where $c_0 + c_1 x + c_2 x^2$ is a perfect square, which means that $c_1^2 = 4c_0 c_2$. Then we can rewrite eq. (6.1) to

$$\begin{aligned}\frac{\mathrm{d}\log f(x)}{\mathrm{d}x} &= -\frac{x + a}{c_2\left(x + C_1\right)^2} \\ &= -\frac{1}{c_2\left(x + C_1\right)} - \frac{a - C_1}{c_2\left(x + C_1\right)^2}\end{aligned}$$

which leads to the solution

$$f(x) = K\left(x + C_1\right)^{-\frac{1}{c_2}} \exp\left(\frac{a - C_1}{c_2\left(x + C_1\right)}\right) \tag{6.8}$$

For the range of x we set the boundaries

$$\begin{cases} x > -C_1 & \text{if } \frac{a - C_1}{c_2} < 0, \\ x < -C_1 & \text{if } \frac{a - C_1}{c_2} > 0. \end{cases}$$

If $a = C_1$ and $|c_2| < 1$ then we have the special case

$$f(x) = K\left(x + C_1\right)^{-\frac{1}{c_2}}.$$

Equation (6.8) corresponds to an inverse gamma location-scale distribution (observe that if we denote by X the random variable with pdf f then $(X + C_1)^{-1}$ has a type 3 distribution).

6.1.7 Pearson Type 6 (akin to the F location-scale distribution)

Type 6 corresponds to the case where the roots of eq. (6.4) are real and of the same sign. If both of them are negative, which means that we can write without loss of generality that $a_1 < a_2 < 0$, we can do the same analysis as we did for Type 1 (see eq. (6.5) to end at

$$f(x) = K\left(x - a_1\right)^{m_1}\left(x - a_2\right)^{m_2} \tag{6.9}$$

where again

$$\begin{aligned} m_1 &= \frac{a + a_1}{c_2(a_2 - a_1)}, \\ m_2 &= -\frac{a + a_2}{c_2(a_2 - a_1)}. \end{aligned}$$

We have $x > a_2$ as the expected value is greater than a_2. For eq. (6.9) to represent a pdf we also need that $m_2 < -1$ and $m_1 + m_2 < 0$.

6.1.8 Pearson Type 7 (akin to the Student's t location-scale distribution)

Lastly type 7 corresponds to the case where $c_1 = a = 0$, and $c_2 > 0$. This transforms eq. (6.1) to

$$\frac{\mathrm{d}\log f(x)}{\mathrm{d}x} = -\frac{x}{c_0 + c_2 x^2}$$

which leads to

$$f(x) = K\left(c_0 + c_2 x^2\right)^{-\frac{1}{2c_2}}. \tag{6.10}$$

An important distribution belonging to this family is the Student's t distribution (see appendix A.1.7). We can derive eq. (6.10) by a multiplicative transformation from a t distribution with degrees of freedom $\frac{1}{c_2} - 1$.

6.2 Identifying the types according to the moments [2]

The Pearson system is well-suited for our simulation model because the parameters a, c_0, c_1, and c_2 can be expressed in terms of the moments of the distribution. We begin by making some notational definitions. Let X be a random variable and denote its moments by $\mu_r := \mu_r(X) = \mathbb{E}[X^r]$. Similarly we denote the centralized moments by $\mu'_r := \mu'_r(X) = \mathbb{E}\left[(X - \mathbb{E}[X])^r\right]$. Additionally we define

$$\beta_1 := [\mathrm{Skew}(X)]^2 \tag{6.11}$$

$$\beta_2 := \mathrm{Kurt}(X) \tag{6.12}$$

We start our calculations by multiplying both sides of eq. (6.1) with x^r and rewriting it to

$$x^r\left(c_0 + c_1 x + c_2 x^2\right)\frac{\mathrm{d}f(x)}{\mathrm{d}x} + x^r(a + x)f(x) = 0 \tag{6.13}$$

By assuming that

$$x^r f(x) \xrightarrow{x\to\pm\infty} 0 \text{ for } r \leq 5$$

and integrating both sides of eq. (6.13) from $-\infty$ to $+\infty$ we obtain

$$-rc_0\mu'_{r-1} + \left[-(r+1)c_1 + a\right]\mu'_r + \left[-(r+2)c_2 + 1\right]\mu'_{r+1} = 0 \tag{6.14}$$

First we note that $\mu'_0 = 1$ and that (in this context) $\mu'_{-1} = 1$. By then putting $r = 0, 1, 2, 3$ in eq. (6.14) we obtain four linear equation for our arguments a, c_0, c_1, and c_2 with coefficients that are functions of μ'_1, μ'_2, μ'_3 and μ'_4. We can shift the variable X so that its expected value is zero, which leads to $\mu'_1 = 0$ and $\mu'_r = \mu_r$ for $r \geq 2$. Then we arrive at the following formulas to calculate the coefficients from the moments

$$c_0 = \frac{4\beta_2 - 3\beta_1}{10\beta_2 - 12\beta_1 - 18}\,\mu_2 \tag{6.15}$$

[2] following [JKB94a]

$$a = c_1 = \frac{\sqrt{\beta_1}(\beta_2 + 3)}{10\beta_2 - 12\beta_1 - 18}\sqrt{\mu_2} \tag{6.16}$$

$$c_2 = \frac{2\mu_4\mu_2 - 3\mu_3^2 - 6\mu_2^2}{10\mu_4\mu_2 - 12\mu_3^2 - 18\mu_2^3}$$

$$= \frac{2\beta_2 - 3\beta_1 - 6}{10\beta_2 - 12\beta_1 - 18} \tag{6.17}$$

We can use eqs. (6.15) to (6.17) to calculate the Pearson coefficients (and identify the type) given the moments of a random variable.

Most importantly for our simulation, it is now possible to classify the different types depending only on the skewness and kurtosis of the underlying random variable. For this we first define

$$\kappa = \frac{1}{4}\frac{c_1^2}{c_0 c_2} = \frac{1}{4}\frac{\beta_1(\beta_2 + 3)^2}{(4\beta_2 - 3\beta_1)(2\beta_2 - 3\beta_1 - 6)}$$

and then classify each type in table 6.1.

Type	**Conditions on β_1 and β_2**
0	$\beta_1 = 0 \wedge \beta_2 = 3$
1	$\kappa < 0$
2	$\beta_1 = 0 \wedge \beta_2 < 3$
3	$2\beta_2 - 3\beta_1 - 6 = 0$
4	$0 < \kappa < 1$
5	$\kappa = 1$
6	$\kappa > 1$
7	$\beta_1 = 0 \wedge \beta_2 > 3$

Table 6.1: Pearson classification

6.3 Generating correlated samples from the Pearson system in Matlab

Before we continue with the simulation model of the claims triangle, we will take a look at the technical implementation used to create the outputs in chapter 7.

In the sampling process a modified version of the Matlab built-in *pearsrnd*[3] function was used. This function generates random samples from the Pearson system given the expected value, the standard deviation, the skewness and the kurtosis of the underlying random variable.

To compute the Pearson coefficients and identify the type, *pearsrnd* proceeds similarly to section 6.2, including the identification of the types after table 6.1. One very useful trait

[3]The *pearsrnd* function is part of the Statistics Toolbox 8.1

of the Pearson system which will help us in the sampling process is that all Pearson types except type 4 can be seen as transformations of standard distributions (see table 6.2).

Type	**Associated distribution**
0	Normal distribution
1	Four-parameter beta
2	Symmetric four-parameter beta
3	Three-parameter gamma
4	Not related to any standard distribution
5	Inverse gamma location-scale
6	F location-scale
7	Student's t location-scale

Table 6.2: Pearson distributions

Before we state the algorithms to create Pearson samples we delve into the theory behind them. Our goal is to create samples from a random variable X given the expected value (denoted by μ), the standard deviation (denoted by σ), the skewness (denoted by γ) and the kurtosis (denoted by κ).
First define the random variable ε by

$$\varepsilon = \frac{X - \mu}{\sigma}. \tag{6.18}$$

This means that

$$\begin{aligned} \mathbb{E}[\varepsilon] &= 0, \\ \mathbb{V}[\varepsilon] &= 1, \\ \mathrm{Skew}(\varepsilon) &= \mathrm{Skew}(X) = \gamma, \\ \mathrm{Kurt}(\varepsilon) &= \mathrm{Kurt}(X) = \kappa. \end{aligned}$$

Observe that both ε and X have the same Pearson type since only the skewness and kurtosis are used in the identification (see table 6.1). This means we can sample ε and transform the output according to eq. (6.18) to get samples of X.

6.3.1 Algorithm (Pearson sampling - basic version). To create a uncorrelated Pearson sample we have
Input: The expected value (denoted by μ), the standard deviation (denoted by σ), the skewness (denoted by γ) and the kurtosis (denoted by κ) of a random variable X and the number of simulations N. The desired output is a sample vector $\left(\widehat{X}_1, \ldots, \widehat{X}_N\right)$.

1. First calculate $\beta_1 = (\gamma)^2$, $\beta_2 = \kappa$ and the parameters a, c_0, c_1, and c_2 to identify the Pearson type of ε.

2. Then calculate the roots a_1 and a_2 of eq. (6.4).

3. Afterwards calculate the necessary parameters (m_1,m_2 or C_1), sample from the associated distribution (in our case with the related built-in function) and transform

the output accordingly. If Pearson type 4 is identified, a rejection algorithm is implemented to sample from the distribution.
At this step we get n samples $(\widehat{\varepsilon}_1, \ldots, \widehat{\varepsilon}_N)$ of ε.

4. In a final step we transform the random vector $(\widehat{\varepsilon}_1, \ldots, \widehat{\varepsilon}_N)$ by setting

$$\widehat{X}_i = \sigma \widehat{\varepsilon}_i + \mu,$$

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

Output: The output is then a random vector $(\widehat{X}_1, \ldots, \widehat{X}_N)$ sampled from a random variable matching all input moments of X.

6.3.2 Remark.

- Reorganizing eq. (6.18) to

$$X = \mu + \sigma\varepsilon$$

and comparing that to eq. (5.1) we can see that this way of sampling is perfectly in line with model assumptions 5.1.1.

- For further information on which transformations to use to relate each Pearson type with their associated distribution consult [MI75], the Matlab documentation or [Dev86]. [MI75] is the most detailed (in terms of sampling) of those and gives algorithms to sample from each type of the Pearson system.

For the sampling of correlated variables some steps have to be modified. In contrast to the above all correlated random variables have to be sampled together (otherwise there would be no correlation). The new algorithm takes the linear correlation matrix ρ as an extra input. It is also tailored to the special case that skewness and kurtosis of all random variables are the same (which is the case in our simulation model - cf. model assumptions 5.1.1).
In addition to the above it should also be stated that (let M be the number of random variables) the Pearson parameters a, c_0, c_1, and c_2 do not depend on $j = 1, \ldots, M$ since for the ε_j all four input parameters are the same (cf. eqs. (6.15) to (6.17)).

6.3.3 Algorithm (Pearson sampling - expanded version). To create correlated Pearson samples for the random variables $(X_1, \ldots, X_M)$ we have
Input: The number of simulations N, the expected values (denoted by μ_j), the standard deviations (denoted by σ_j), the skewness and kurtosis (denoted by γ and κ respectively and both not dependent on j) of M random variables X_j, $j = 1, \ldots, M$. Additionally we have the linear correlation matrix ρ, where

$$\rho_{i,j} = \text{Corr}\,(X_i, X_j)$$

as an extra input. The desired output is N samples of each of the X_j.

1. After calculating β_1, β_2 and the parameters a, c_0, c_1, and c_2, identify the Pearson type (which is the same for all ε_j and with that all X_j)..

2. Then calculate the roots a_1 and a_2 of eq. (6.4).

3. In the expanded version the sampling is done a bit differently from the base version (except for Pearson type 4 where correlation is not implemented and the standard built-in function from above is used). After calculating the necessary parameters the sampling is done by using a gaussian copula and the algorithm presented in appendix A.2.2. This means that for each $i = 1, \dots, N$ we follow the following procedure:
   (a) First we use ρ to sample from a uniform distribution via a gaussian copula approach (see appendix A.2.1). This yields a vector $(U_1, \dots, U_M)$, where $U_i \sim U(0,1)$ and $\mathrm{Corr}\,(U_i, U_j) \approx \rho_{i,j}$, $i, j = 1, \dots, M$.
   (b) Then we use the inversion method (see definition A.2.7) to get samples from our associated distribution (with its cdf denoted by F_j), which we will then transform accordingly to fit the Pearson type. We will denote the transformation function by h_{F_j}. It is dependent on the Pearson type and its relation to its associated distribution F_j (cf. remark 6.3.2).
   That means we receive a sample $(\widehat{\varepsilon}_1, \dots, \widehat{\varepsilon}_M)$ where for $i = 1, \dots, M$ we have that
   $$\widehat{\varepsilon}_i = h_j\left(F_j^{-1}(U_i)\right).$$
   (c) If we do this N times (in fact computationally we can do it all at once) we have N sets of vectors $\left(\widehat{\varepsilon}_1^{\langle n\rangle}, \dots, \widehat{\varepsilon}_M^{\langle n\rangle}\right)$ $(n = 1 \dots, N)$.

4. In a final step we transform the sample vectors $\left(\widehat{\varepsilon}_1^{\langle n\rangle}, \dots, \widehat{\varepsilon}_M^{\langle n\rangle}\right)$ by setting
   $$\widehat{X}_j^{\langle n\rangle} = \sigma_j \widehat{\varepsilon}_j^{\langle n\rangle} + \mu_j.$$
   for $n = 1, \dots, N$ and $j = 1, \dots, M$.

Output: The output consists of N sample vectors $\left(\widehat{X}_1^{\langle n\rangle}, \dots, \widehat{X}_M^{\langle n\rangle}\right)$, where $n = 1, \dots, N$, sampled from random variables matching all input moments of each X_j and where each sample vector is generated using the input correlation matrix .

6.4 Simulation with the Pearson system

We will now go through the simulation step by step .

1. The first step of the simulation process is computing the values
   $$\widehat{f}_j,\ \widehat{\sigma}_j,\ \widehat{\gamma}_j,\ \text{and}\ \widehat{\kappa}_j$$
   for $j = 0, \dots, J-1$.

2. Then we simulate the missing cumulative claims for each development year $j = 1, \dots, J$ to fill up the claims triangle. For each development year j we have constant skewness and kurtosis (see model assumptions 4.2.2), which are given by the estimators $\widehat{\gamma}_{j-1}$ and $\widehat{\kappa}_{j-1}$. The expected value and the variance differ for each accident year i.

So this means that at each step $j = 1, \dots, J$ we have

$$\begin{aligned}
\mathbb{E}\left[\widehat{C}_{i,j}^{Sim}\middle|\mathcal{E}_{I,j-1}\right] &= \widehat{C}_{i,j-1}\ \widehat{f}_{j-1} && =: a_{i,j},\\
\mathbb{V}\left[\widehat{C}_{i,j}^{Sim}\middle|\mathcal{E}_{I,j-1}\right] &= \widehat{C}_{i,j-1}\ \widehat{\sigma}_{j-1}^2\left(1+\frac{C_{i,j-1}}{\sum_{k=0}^{I-j} C_{k,j-1}}\right) && =: b_{i,j},\\
\text{Skew}\left(\widehat{C}_{i,j}^{Sim}\middle|\mathcal{E}_{I,j-1}\right) &= \widehat{\gamma}_{j-1} && =: c_j,\\
\text{Kurt}\left(\widehat{C}_{i,j}^{Sim}\middle|\mathcal{E}_{I,j-1}\right) &= \widehat{\kappa}_{j-1} && =: d_j,
\end{aligned}$$

for $i = I - j + 1, \dots, I$ and where for ease of notation

$$\widehat{C}_{i,j} = \begin{cases} C_{i,j} & 0 \le i \le I-j, j = 0, \dots, J \\ \widehat{C}_{i,j}^{Sim} & I-j < i \le I, j = 1, \dots, J \end{cases}$$

and $\mathcal{E}_{I,j-1}$ as in eq. (5.3). Now we can use c_j and d_j to calculate the Pearson parameters β_1 and β_2 (see eqs. (6.11) and (6.12)). With these we identify the Pearson type and its associated distribution.

3. The claims triangle at step j looks like table 6.3. Our aim is to simulate the random variables $\left(\widehat{C}_{I-j+1,j}^{Sim}, \dots, \widehat{C}_{I,j}^{Sim}\right)$. For this we input the parameters $a_{i,j}, b_{i,j}, c_j, d_j$ and the linear correlation matrix $\widehat{\rho}^{\langle j\rangle}$ into algorithm 6.3.3 which identifies the Pearson type and creates the samples. The output is then N sets of vectors $\left(\widehat{X}_1^{\langle n\rangle}, \dots, \widehat{X}_j^{\langle n\rangle}\right)$ $(n = 1 \dots, N)$ which we combine into one sample matrix $\Theta^{\langle j\rangle}$ by setting

$$\Theta_{k,n}^{\langle j\rangle} = \widehat{X}_k^{\langle n\rangle} \quad \forall\ k = 1, \dots, j \text{ and } n = 1, \dots, N$$

The matrix $\Theta^{\langle j\rangle}$ is a matrix of random variables with entries $\Theta_{k,n}^{\langle j\rangle}$, where $k = 1, \dots, j$ and $n = 1, \dots, N$ and where each column $\Theta_{\cdot,n}^{\langle j\rangle}$ is a sample from $\left(\widehat{C}_{I-j+1,j}^{Sim}, \dots, \widehat{C}_{I,j}^{Sim}\right)$. This means that if we fix the simulation number n we have

$$^{\langle n\rangle}\widehat{C}_{I-j+i,j}^{Sim} = \Theta_{i,n}^{\langle j\rangle} \tag{6.19}$$

for each $j = 1, \dots, J$, $i = 1, \dots, j$ and $n = 1, \dots, N$.

4. For the ultimate claims this gives us N simulations of the values $\widehat{C}_{i,J}^{Sim}$, $i = 1, \dots, I$. If we want to compute a statistic of them we can do this directly with the simulated values saved in $\Theta_{i,n}^{\langle J\rangle}$. As a first example we want to show the computation of the best estimate reserves. We have

$$\widehat{\mathbb{E}}\left[\widehat{C}_{i,J}\right] = \frac{1}{N}\sum_{n=1}^{N}\Theta_{i,n}^{\langle J\rangle}$$

and with that we can also calculate

$$\widehat{\mathbb{E}}\left[\widehat{R}_i\right] = \widehat{\mathbb{E}}\left[\widehat{C}_{i,J}\right] - C_{i,I-i} \text{ for } i = 1, \dots, I.$$

AY/DY	0	1	...	j	...	...	J
0	$C_{0,0}$	$C_{0,1}$	...	$C_{0,j}$	...	$\iddots$	$C_{0,J}$
1	$C_{1,0}$	$\iddots$	$\iddots$	$\iddots$	$\iddots$	$C_{1,J-1}$	
$\vdots$	$\vdots$	$\iddots$	$\iddots$	$\iddots$	$\iddots$		
$I-j$	$\vdots$	$\iddots$	$\iddots$	$C_{I-j,j}$			
$\vdots$	$\vdots$	$\iddots$	$\iddots$	$\widehat{C}^{Sim}_{I-j+1,j}$			
$\vdots$	$\vdots$	$C_{I-1,1}$	$\iddots$	$\vdots$			
I	$C_{I,0}$	$\widehat{C}^{Sim}_{I,1}$	...	$\widehat{C}^{Sim}_{I,j}$			

Table 6.3: Simulation step j

For ease of notation we will introduce another matrix. We will call it the reserve matrix $\Psi^{\langle J\rangle}$ and define it by

$$\Psi^{\langle J\rangle}_{i,n} = \Theta^{\langle J\rangle}_{i,n} - C_{i,I-i},$$
$$\text{for } i = 1,\dots,I \text{ and } n = 1,\dots,N.$$

Each column of the matrix $\Psi^{\langle J\rangle}$ represents one simulation of the outstanding reserve for each accident year i.

6.5 Statistical Estimators

6.5.1 Estimator (Sample Estimators)**.** We will now define the statistics used in chapter 7 to showcase the simulation. We have

$$\widehat{\mathbb{E}}\left[\widehat{R}_i\right] = \frac{1}{N}\sum_{n=1}^{N}\Psi^{\langle J\rangle}_{i,n}, \tag{6.20}$$

$$\widehat{\mathbb{V}}\left[\widehat{R}_i\right] = \frac{1}{N-1}\sum_{n=1}^{N}\left(\Psi^{\langle J\rangle}_{i,n} - \widehat{\mathbb{E}}\left[\widehat{R}_i\right]\right)^2, \tag{6.21}$$

$$\widehat{\mathrm{Std}}\left(\widehat{R}_i\right) = \sqrt{\widehat{\mathbb{V}}\left[\widehat{R}_i\right]}, \tag{6.22}$$

$$\widehat{\mathrm{Skew}}\left(\widehat{R}_i\right) = \frac{\sqrt{N\left(N-1\right)}}{N-2}\frac{\frac{1}{N}\sum_{n=1}^{N}\left(\Psi^{\langle J\rangle}_{i,n} - \widehat{\mathbb{E}}\left[\widehat{R}_i\right]\right)^3}{\left(\sqrt{\frac{1}{N}\sum_{n=1}^{N}\left(\Psi^{\langle J\rangle}_{i,n} - \widehat{\mathbb{E}}\left[\widehat{R}_i\right]\right)^2}\right)^3}, \tag{6.23}$$

$$k_i^{tmp} = \frac{\frac{1}{N}\sum_{n=1}^{N}\left(\Psi_{i,n}^{\langle J\rangle} - \widehat{\mathbb{E}}\left[\widehat{R}_i\right]\right)^4}{\left(\frac{1}{N}\sum_{n=1}^{N}\left(\Psi_{i,n}^{\langle J\rangle} - \widehat{\mathbb{E}}\left[\widehat{R}_i\right]\right)^2\right)^3},$$

$$\widehat{\text{Kurt}}\left(\widehat{R}_i\right) = \frac{N-1}{(N-2)(N-3)}\left((N+1)\,k_i^{tmp} - 3(N-1)\right) + 3, \tag{6.24}$$

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

These are the bias corrected empirical estimators used by Matlab. Additionally we will calculate the Value at Risk of the outstanding reserve.

6.5.2 Definition (Value at Risk)**.** The Value at Risk at significance level α of a random variable X with cumulative distribution function F_X is defined as

$$\text{VaR}_\alpha(X) = \inf\{q \in \mathbb{R} : F_X(q) \geq \alpha\} \tag{6.25}$$

6.5.3 Remark. The VaR_α is in fact just the α-quantile of the distribution F_X.

Since we do not have an analytical expression of the cdf F_X we will use a quantile estimator to calculate the Value at Risk of our sample.

6.5.4 Algorithm (Matlab quantile function)*.* The Matlab function *quantile* works the following way

(i) Let x be an n-element vector of samples from the random variable X. Then let y be the sorted vector of x.

(ii) The sorted values y then correspond to the $({}^{0.5}/_{n}), ({}^{1.5}/_{n}), \ldots, ({}^{(n-0.5)}/_{n})$ sample quantiles of X.We now have three different scenarios.

If ${}^{0.5}/_{n} \leq p \leq {}^{(n-0.5)}/_{n}$**:**

(a) If there exists an $i \in \{0, 1, \ldots, n-1\}$ such that

$$p = {}^{(n-0.5-i)}/_{n}$$

then $quantile(x, p) = y(n-i)$.

(b) Otherwise there exists a $j \in \{0, 1, \ldots, n-1\}$ such that

$${}^{(n-0.5-(j+1))}/_{n} < p < {}^{(n-0.5-j)}/_{n}$$

In this case *quantile* uses linear interpolation between the above values. We have

$$quantile(x,p) = y(n-(j+1)) + [p - {}^{(n-0.5-(j+1))}/_{n}]\,[y(n-j) - y(n-(j+1))]$$

Else:

(c) For the quantiles corresponding to probabilities outside of the above range, *quantile* assigns the minimum or maximum values of X. This means that

$$quantile(x,p) = \begin{cases} y(1) & \text{if } p < {}^{0.5}/_{n} \\ y(n) & \text{if } p > {}^{(n-0.5)}/_{n}. \end{cases}$$

6.5.5 Estimator (VaR Estimator)**.** With the above definition of the function *quantile* we then have

$$\widehat{\mathrm{VaR}}_\alpha\left(\widehat{R}_i\right) = quantile\left(\Psi_{i,\cdot}^{\langle J\rangle}, \alpha\right) \tag{6.26}$$

6.5.6 Remark. The sampling process described in this section and in algorithm 6.3.3 can be interpreted in two different ways.
We can see algorithm 6.3.3 as a means to simulate the cumulative claims $\widehat{C}_{i,j}^{Sim}$. In this case the variables ε_k are only temporary variables used in the simulation process.
However we can also take eq. (5.1) from model assumptions 5.1.1 as the basis of our simulation. We have assumed that

$$C_{i,j+1} = f_j\ C_{i,j} + \sigma_j\ \sqrt{C_{i,j}}\ \varepsilon_{i,j+1}$$
$$\forall\ i \in \{0,\ldots,I\}\text{, and } \forall\ j \in \{0,\ldots,J-1\}$$

This means that we can also simulate the $\varepsilon_{i,j+1}$ and then calculate $\widehat{C}_{i,j+1}^{Sim}$ from them. We have

$$\begin{aligned}\varepsilon_{i,j+1} &= \frac{C_{i,j+1} - f_j\ C_{i,j}}{\sigma_j\ \sqrt{C_{i,j}}} \\ &= \frac{C_{i,j+1} - \mathbb{E}\left[C_{i,j+1}|\mathcal{D}_I\right]}{\sqrt{\mathbb{V}\left[C_{i,j+1}|\mathcal{D}_I\right]}}\end{aligned}$$

Comparing this to eq. (6.18) we can see that for a fixed $j \in \{0,\ldots,J-1\}$ we can interpret the ε_k $(k = 1,\ldots,j+1)$ from algorithm 6.3.3 as samples of the $\varepsilon_{i,j+1}$ $(i = I-j,\ldots,I)$ and conclude that both interpretations yield the same result.

Chapter 7

Examples

We will now utilize both methods from the previous sections on different claims triangles and compare the results against each other. Additionally we will also compare the results to the standard chain ladder estimator and the MSEP estimator introduced in [WM08].

7.1 Application of the simulation

For each of the following triangles we will first calculate all parameters and then run the simulation models described in sections 5.4 and 6.4. The results of this run will then be displayed in a table and interpreted subsequently. Additionally a histogram of the total reserve is shown, featuring overlaid density curves of standard distribution functions.
The normal, lognormal and gamma distributions used below are fitted to the chain ladder estimators via moment matching. The generalized extreme value distribution, which is also displayed in the figure, is fitted to the data generated by the simulation via the maximum likelihood method.
The simulation was realized in Matlab R2012b with the Statistics Toolbox 8.1 installed and the number of simulations set to one million. A special thanks has to go to Nico Schlömer for creating and sharing his Matlab script *matlab2tikz*, which was used in the creation of the output histograms [1].
To get an idea about the duration of one simulation table 7.1 depicts the average runtime per triangle for different simulation numbers. The second column denotes the time the Pearson simulation from section 6.4 took, while the third column denotes the runtime of both simulations (Pearson and GPD) plus the runtime for the creation of the output tables and figures.

[1] It can be downloaded from the Matlab file exchange - see `http://www.mathworks.com/matlabcentral/fileexchange/22022-matlab2tikz`

Nr. of simulations	**Pearson sim.**	**Whole sim.**
10,000	1.39	3.06
100,000	8.40	11.82
1,000,000	77.96	101.58

Table 7.1: Mean runtimes of simulation models in seconds

7.1.1 Example 1: Wüthrich Triangle

The first claims triangle is taken from [WM08]. It is given by fig. 7.1.

AY/DY	0	1	2	3	4	5	6	7	8	9
0	5,946,975	9,668,212	10,563,929	10,771,690	10,978,394	11,040,518	11,106,331	11,121,181	11,132,310	11,148,124
1	6,346,756	9,593,162	10,316,383	10,468,180	10,536,004	10,572,608	10,625,360	10,636,546	10,648,192	
2	6,269,090	9,245,313	10,092,366	10,355,134	10,507,837	10,573,282	10,626,827	10,635,751		
3	5,863,015	8,546,239	9,268,771	9,459,424	9,592,399	9,680,740	9,724,068			
4	5,778,885	8,524,114	9,178,009	9,451,404	9,681,692	9,786,916				
5	6,184,793	9,013,132	9,585,897	9,830,796	9,935,753					
6	5,600,184	8,493,391	9,056,505	9,282,022						
7	5,288,066	7,728,169	8,256,211							
8	5,290,793	7,648,729								
9	5,675,568									
$\widehat{f}_j$	1.4925	1.0778	1.0229	1.0148	1.007	1.0051	1.0011	1.001	1.0014	
$\widehat{\sigma}_j$	135.253	33.8029	15.7596	19.8467	9.3362	2.0011	0.8232	0.2196	0.0586	
$\widehat{\gamma}_j$	2.0465	0.2386	-0.5048	0.2291	0.2806	0.8868	0.383	0	0	
$\widehat{\kappa}_j$	5.2748	1.1998	1.9027	1.6792	0.9745	1.3181	3	3	3	

Figure 7.1: Example 1: Wüthrich Triangle

A summary of the key figures can be found in fig. 7.2 while fig. 7.3 depicts a histogram based on the simulation of the total reserve in comparison to fitted standard density curves. As can be seen from the output, non of the standard distributions offer a good fit to this distribution. The bimodal look of the histogram can probably be attributed to most columns having Pearson type 1, which is akin to the beta distribution and has bimodal properties when the parameters are low.

	Pearson Simulation			**Chain Ladder Estimators**		**GPD Sim.**
	$\mathbb{E}\left[\widehat{R}_i\right]$	$\sqrt{\mathbb{V}\left[\widehat{R}_i\right]}$	$\widehat{\mathrm{VaR}}_{99\%}\left(\widehat{R}_i\right)$	$\widehat{R}_i^{CL}$	$\sqrt{\widehat{\mathbf{msep}}_{R_i \mid \mathcal{D}_I}}$	$\widehat{\mathrm{VaR}}_{99\%}\left(\widehat{R}_i^{GPD}\right)$
1	15,126.54	270.46	15,756.56	15,126.29	267.51	15,580.26
2	26,258.40	918.07	28,394.48	26,257.45	915.24	27,999.92
3	34,543.18	3,063.16	42,312.91	34,538.47	3,058.74	41,371.50
4	85,303.14	7,626.96	102,341.29	85,301.62	7,628.15	105,991.36
5	156,473.23	33,340.02	211,659.83	156,494.25	33,341.22	226,987.61
6	286,217.67	73,443.48	436,572.98	286,121.02	73,466.90	448,429.77
7	449,045.30	85,383.18	633,065.90	449,166.98	85,398.21	636,457.12
8	1,043,246.40	134,181.85	1,331,729.34	1,043,242.44	134,336.55	1,349,585.88
9	3,950,621.30	410,783.86	5,146,276.79	3,950,815.25	410,817.59	5,377,532.99
Total	6,046,835.15	452,739.29	7,368,874.51	6,047,063.77	462,960.58	7,522,123.05

Figure 7.2: Example 1: Summary

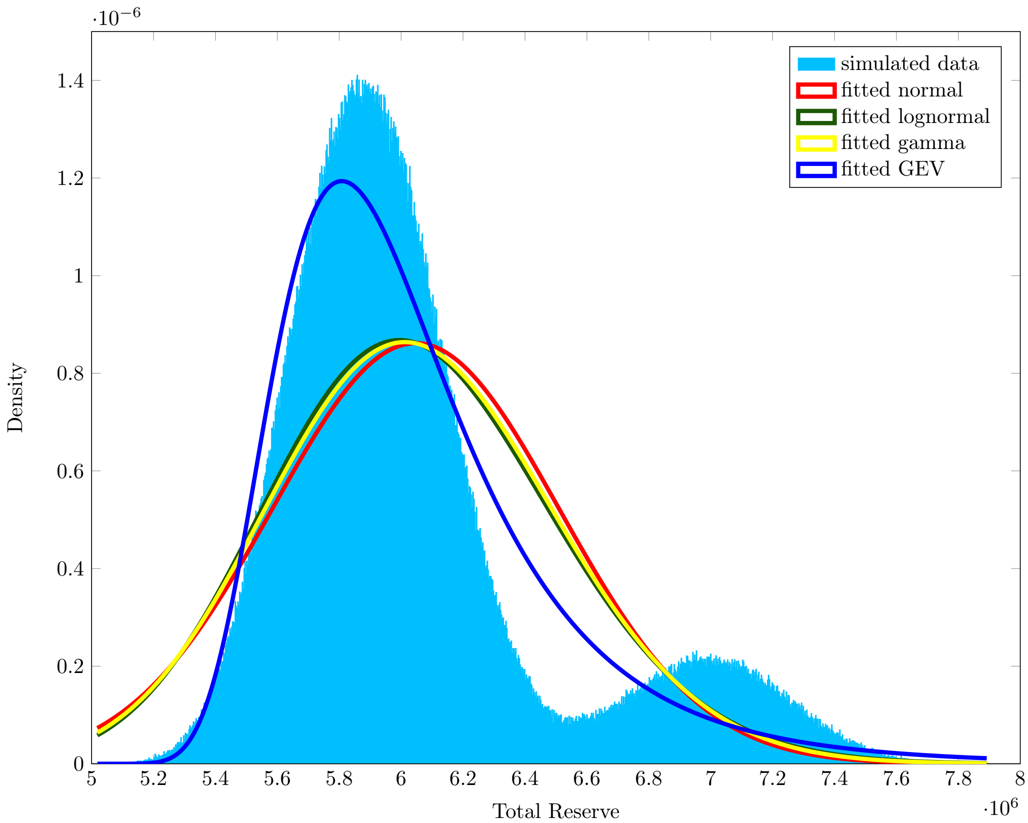

Figure 7.3: Example 1: Simulation Output

7.1.2 Example 2: Mack Triangle

The next triangle can be found in [Mac93] and is depicted in fig. 7.4.

AY/DY	0	1	2	3	4	5	6	7	8	9
0	357,848	1,124,788	1,735,330	2,218,270	2,745,596	3,319,994	3,466,336	3,606,286	3,833,515	3,901,463
1	352,118	1,236,139	2,170,033	3,353,322	3,799,067	4,120,063	4,647,867	4,914,039	5,339,085	
2	290,507	1,292,306	2,218,525	3,235,179	3,985,995	4,132,918	4,628,910	4,909,315		
3	310,608	1,418,858	2,195,047	3,757,447	4,029,929	4,381,982	4,588,268			
4	443,160	1,136,350	2,128,333	2,897,821	3,402,672	3,873,311				
5	396,132	1,333,217	2,180,715	2,985,752	3,691,712					
6	440,832	1,288,463	2,419,861	3,483,130						
7	359,480	1,421,128	2,864,498							
8	376,686	1,363,294								
9	344,014									
$\widehat{f}_j$	3.4906	1.7473	1.4574	1.1739	1.1038	1.0863	1.0539	1.0766	1.0177	
$\widehat{\sigma}_j$	400.3503	194.2598	204.8541	123.2189	117.1807	90.4753	21.1333	33.8728	21.1333	
$\widehat{\gamma}_j$	0.1961	0.3229	1.0196	-0.7557	0.8008	-0.0641	-1.948	0	0	
$\widehat{\kappa}_j$	1.7958	1.6328	2.559	1.4845	1.6243	-0.3701	3	3	3	

Figure 7.4: Example 2: Mack Triangle

A summary of the key figures can be found in fig. 7.5 while fig. 7.6 depicts a histogram based on the simulation of the total reserve in comparison to fitted standard density curves. From the numbers it seems that this is a very regular triangle, where each of the distributions offers a reasonable fit.

	Pearson Simulation			**Chain Ladder Estimators**		**GPD Sim.**
	$\mathbb{E}\left[\widehat{R}_i\right]$	$\sqrt{\mathbb{V}\left[\widehat{R}_i\right]}$	$\widehat{\text{VaR}}_{99\%}\left(\widehat{R}_i\right)$	$\widehat{R}_i^{CL}$	$\sqrt{\widehat{\mathbf{msep}}_{R_i \mid \mathcal{D}_I}}$	$\widehat{\text{VaR}}_{99\%}\left(\widehat{R}_i^{GPD}\right)$
1	94,702.93	85,408.35	293,298.54	94,633.81	75,535.04	222,820.83
2	469,576.68	126,500.24	766,267.54	469,511.29	121,700.12	725,990.37
3	709,713.31	127,151.05	993,012.65	709,637.82	133,550.98	987,533.95
4	984,690.66	261,476.32	1,473,709.82	984,888.64	261,412.47	1,520,655.97
5	1,419,403.27	411,042.26	2,353,664.47	1,419,459.46	411,027.80	2,475,222.83
6	2,177,406.23	557,738.56	3,408,171.04	2,177,640.62	558,355.88	3,510,842.65
7	3,918,819.33	874,928.43	6,245,350.39	3,920,301.01	875,429.58	6,288,008.20
8	4,278,631.24	972,439.81	6,839,308.64	4,278,972.26	971,385.37	6,858,733.77
9	4,626,061.88	1,365,121.40	8,133,420.81	4,625,810.69	1,363,384.66	8,140,636.09
Total	18,679,005.54	2,173,908.23	24,085,707.70	18,680,855.61	2,447,618.31	24,498,884.79

Figure 7.5: Example 2: Summary

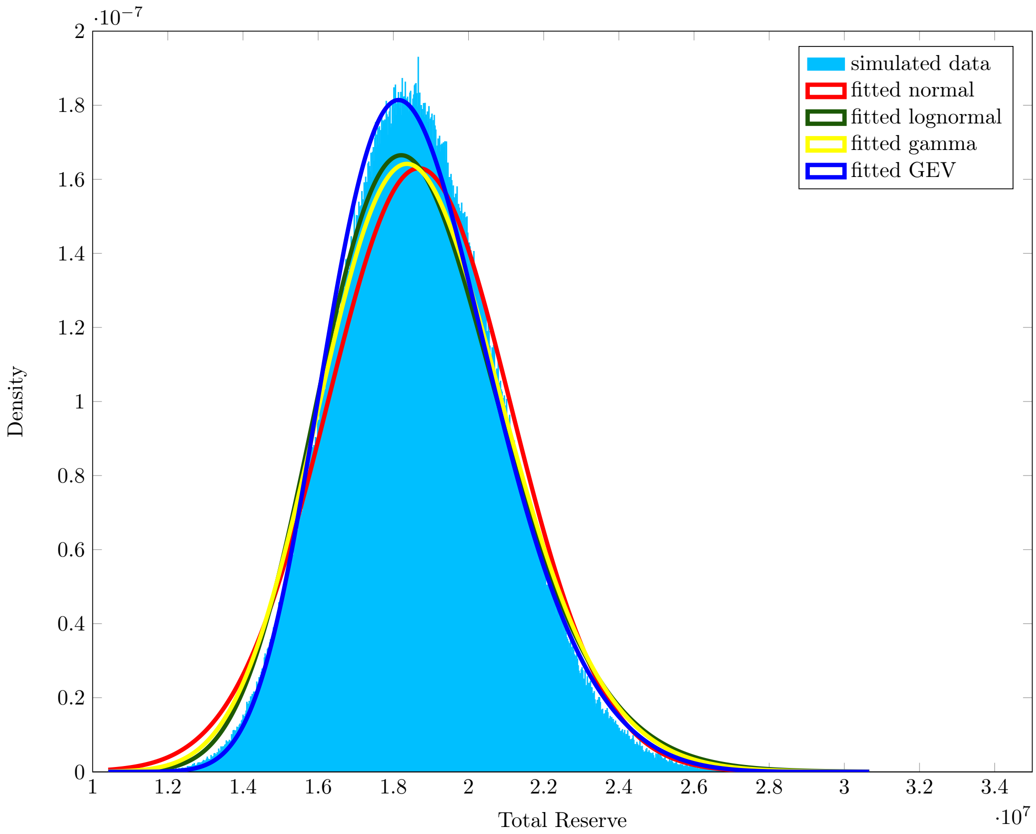

Figure 7.6: Example 2: Simulation Output

7.1.3 Example 3: Dal Moro Triangle

This triangle is taken from [Mor13]. It is given by fig. 7.7.

AY/DY	0	1	2	3	4	5	6	7	8	9
0	10,798	11,313	15,110	15,163	14,232	14,063	12,050	12,163	11,624	12,942
1	11,595	13,743	15,143	15,253	14,999	15,468	12,907	13,086	13,122	
2	11,724	13,621	15,401	14,577	14,932	15,052	13,156	12,847		
3	11,820	13,666	14,915	14,269	14,933	15,263	13,016			
4	11,746	13,352	14,998	14,456	14,915	15,042				
5	11,641	13,182	14,858	14,721	14,788					
6	11,557	13,186	14,811	14,898						
7	11,552	13,159	14,887							
8	11,525	13,061								
9	11,522									
$\widehat{f}_j$	1.1378	1.1416	0.982	1.0041	1.0118	0.8543	0.9996	0.9801	1.1134	
$\widehat{\sigma}_j$	3.9568	8.4241	3.1661	4.7823	1.976	2.008	2.3229	3.7369	2.3229	
$\widehat{\gamma}_j$	-1.7175	2.6942	-0.3731	-0.9838	-0.4009	-0.0872	-1.5418	0	0	
$\widehat{\kappa}_j$	5.318	6.8202	0.787	2.0494	1.4971	1.2178	3	3	3	

Figure 7.7: Example 3: Dal Moro Triangle

A summary of the key figures can be found in fig. 7.8 while fig. 7.9 depicts a histogram based on the simulation of the total reserve in comparison to fitted standard density curves. In this case we can see one of the major shortcomings of using a distribution like the lognormal or the gamma distribution. Since they cannot take negative values into account they are very bad fits for triangles where there are a lot of regresses.

	Pearson Simulation			**Chain Ladder Estimators**		**GPD Sim.**
	$\mathbb{E}\left[\widehat{R}_i\right]$	$\sqrt{\mathbb{V}\left[\widehat{R}_i\right]}$	$\widehat{\text{VaR}}_{99\%}\left(\widehat{R}_i\right)$	$\widehat{R}_i^{CL}$	$\sqrt{\widehat{\textbf{msep}}_{R_i\mid\mathcal{D}_I}}$	$\widehat{\text{VaR}}_{99\%}\left(\widehat{R}_i^{GPD}\right)$
1	1,488.23	453.06	2,540.15	1,487.85	388.25	2,146.75
2	1,172.13	792.71	3,049.27	1,171.72	690.74	2,623.90
3	1,180.31	793.63	2,996.01	1,180.80	773.18	2,808.48
4	-1,026.07	808.94	828.92	-1,025.14	823.27	775.24
5	-844.76	824.88	1,047.82	-844.54	856.26	1,057.70
6	-794.43	1,028.11	1,436.55	-793.65	1,048.07	1,471.49
7	-1,045.39	1,085.55	1,340.88	-1,047.39	1,106.83	1,388.41
8	800.59	1,449.33	5,254.27	800.58	1,458.81	5,034.94
9	2,389.95	1,527.24	6,979.04	2,391.25	1,537.31	6,749.34
Total	3,320.56	3,945.01	12,829.78	3,321.48	4,889.25	13,827.66

Figure 7.8: Example 3: Summary

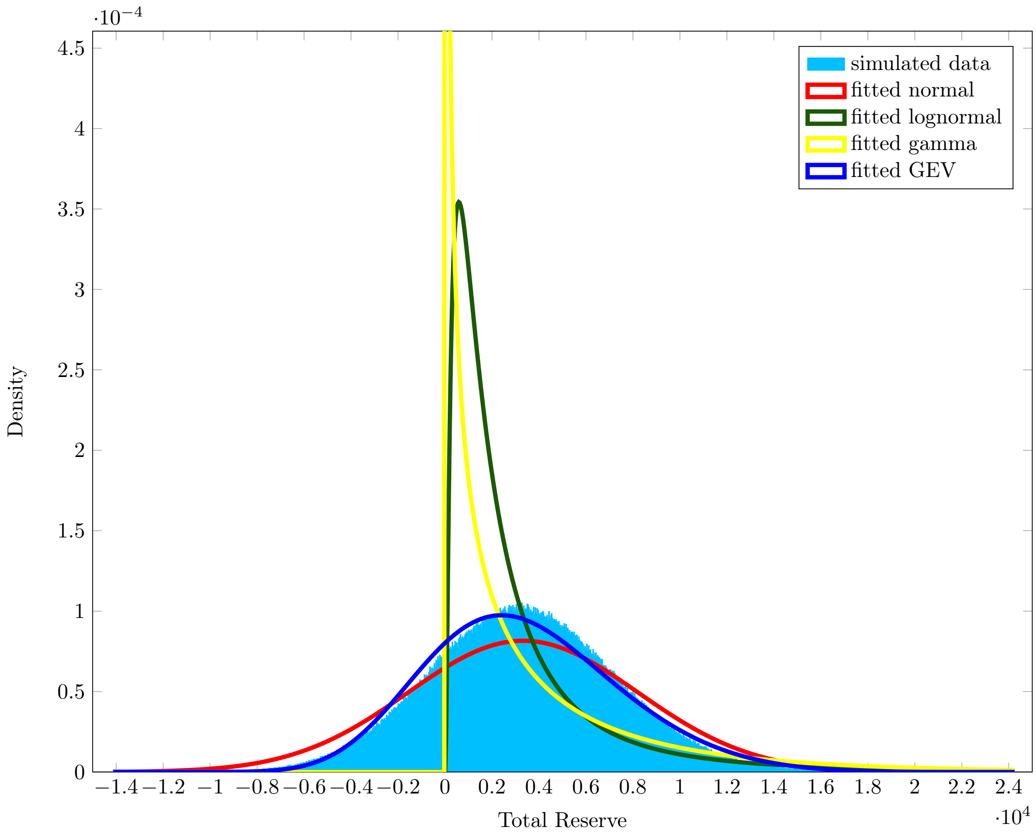

Figure 7.9: Example 3: Simulation Output

7.1.4 Example 4: Product Liability Triangle

This triangle is also taken from [Mor13] and it is given by fig. 7.10.

AY/DY	0	1	2	3	4	5	6	7	8	9
0	199	246	360	404	401	321	328	340	332	342
1	1,312	1,181	1,347	1,390	1,511	1,686	1,860	1,689	1,680	
2	493	700	817	740	912	903	897	902		
3	391	487	1,472	1,640	611	593	597			
4	586	741	1,251	1,509	1,864	2,040				
5	892	1,285	1,556	1,720	1,828					
6	654	1,644	2,060	2,270						
7	379	906	1,255							
8	705	950								
9	384									
$\widehat{f}_j$	1.4507	1.4072	1.0914	0.9627	1.046	1.0511	0.9501	0.9916	1.0301	
$\widehat{\sigma}_j$	13.809	14.849	2.9171	12.4593	3.2492	1.7349	2.0616	0.3062	0.0455	
$\widehat{\gamma}_j$	1.0925	2.5375	-0.8514	-1.8146	-1.1706	0.3586	0.3009	0	0	
$\widehat{\kappa}_j$	3.1839	6.1203	2.881	3.5381	1.3946	0.7722	3	3	3	

Figure 7.10: Example 4: Product Liability Triangle

A summary of the key figures can be found in fig. 7.11 while fig. 7.12 depicts a histogram based on the simulation of the total reserve in comparison to fitted standard density curves. Similarly to Example 3 the lognormal distribution offers a very bad fit because some of the reserves are very low, or even negative in the case of accident year 3. Additionally the total reserve seems to be negatively skewed, which means that even the normal distribution overestimates the right side of the tail (cf. fig. 7.13).

	Pearson Simulation			**Chain Ladder Estimators**		**GPD Sim.**
	$\mathbb{E}\left[\widehat{R}_i\right]$	$\sqrt{\mathbb{V}\left[\widehat{R}_i\right]}$	$\widehat{\mathrm{VaR}}_{99\%}\left(\widehat{R}_i\right)$	$\widehat{R}_i^{CL}$	$\sqrt{\widehat{\mathbf{msep}}_{R_i \mid \mathcal{D}_I}}$	$\widehat{\mathrm{VaR}}_{99\%}\left(\widehat{R}_i^{GPD}\right)$
1	50.61	4.66	61.42	50.60	4.59	58.39
2	19.38	11.70	46.66	19.38	11.68	41.02
3	-17.57	56.89	125.66	-17.61	56.88	102.33
4	41.13	160.27	439.00	40.98	160.35	427.76
5	122.60	225.05	592.09	122.59	224.78	596.54
6	62.06	768.90	957.99	61.92	767.40	1,055.10
7	151.53	579.46	886.71	152.06	577.27	955.28
8	549.12	816.31	3,001.36	548.85	811.08	3,025.75
9	513.74	731.78	3,168.94	494.92	746.51	2,926.51
Total	1,492.60	1,588.91	5,344.89	1,473.70	1,786.39	5,523.79

Figure 7.11: Example 4: Summary

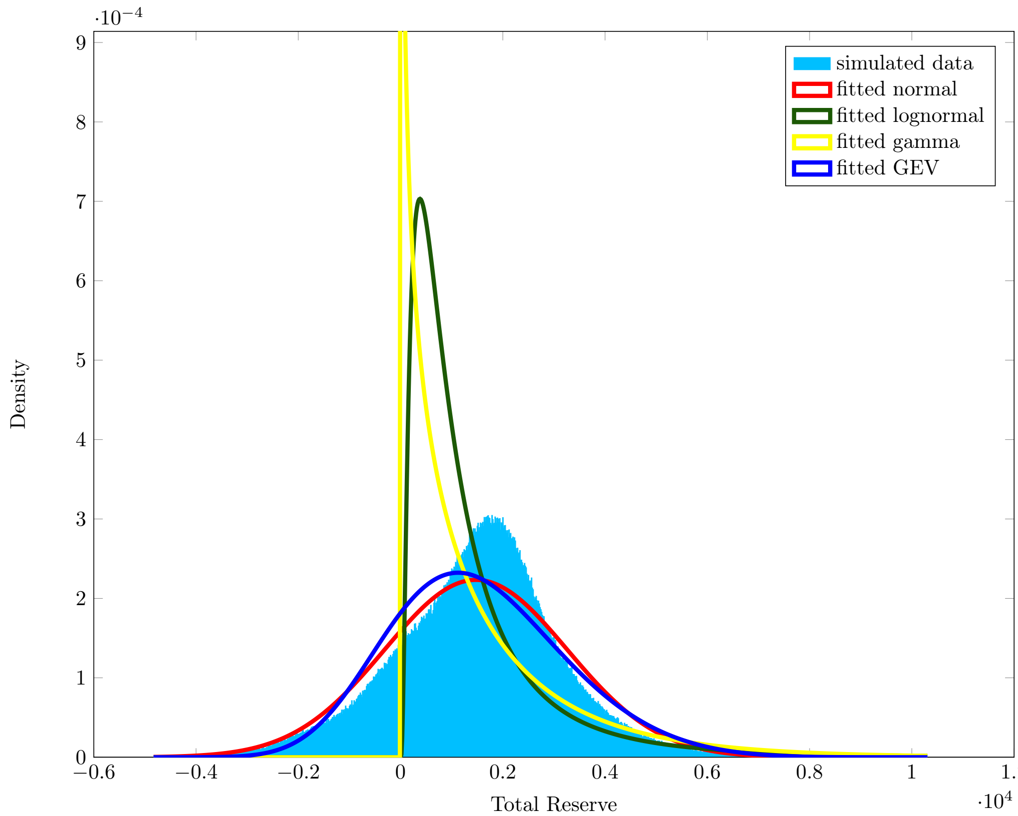

Figure 7.12: Example 4: Simulation Output

7.1.5 Comparing the methods

In fig. 7.13 we compare the different methods on the basis of the 99% Value at Risk. As can be seen in the table it seems like there can be no one distribution which fits all claims triangles. Especially the factor of negative reserves plays an important role in the choice of the distribution.

The normal, lognormal and gamma distribution are chosen because they are the most popular distributions used to estimate the claims reserve distribution in today's insurance practice. All three of them can be fitted to two moments. This is not true for the generalized extreme value distribution, which has three parameters and was therefore omitted from this comparison. The parameter estimation was done via moment matching and using the chain ladder best estimate and MSEP developed in chapter 3.

Example:	Pearson Simulation	GPD Sim.	fitted Normal	fitted Lognormal	fitted Gamma
1	7,368,874.51	7,522,123.05	7,124,071.13	7,202,979.24	7,175,890.75
2	24,085,707.70	24,498,884.79	24,374,867.27	25,090,696.54	24,841,480.35
3	12,829.78	13,827.66	14,695.59	22,685.67	23,025.05
4	5,344.89	5,523.79	5,629.45	8,564.42	8,282.12

Figure 7.13: VaR_α of total reserve with $\alpha = 99\%$

Figures 7.14 and 7.15 compare the skewness and kurtosis of the above distributions. As predicted the Pearson and the GPD model result in a similar skewness for the total reserves. This does not hold true for the kurtosis, since the GPD model was not fitted to the kurtosis estimators. The fitted distributions offer varying results, in which especially the skewness and the kurtosis of the lognormal and gamma distribution for examples 3 and 4 stand out. These very high values can be explained by the fact that both distributions have origin zero and are therefore not suitable for these claims triangles.

Example:	Pearson Simulation	GPD Sim.	fitted Normal	fitted Lognormal	fitted Gamma
1	1.2864	1.2687	0	0.2301	0.1531
2	0.2395	0.2486	0	0.3953	0.262
3	0.0976	0.0964	0	7.6056	2.944
4	-0.1062	-0.0393	0	5.4177	2.4244

Figure 7.14: Skewness of total reserve

Example:	Pearson Simulation	GPD Sim.	fitted Normal	fitted Lognormal	fitted Gamma
1	4.2242	6.3493	3	3.0943	3.0352
2	2.9897	3.0433	3	3.2791	3.103
3	3.0668	3.03	3	191.1785	16.0009
4	3.5123	3.7607	3	82.5934	11.8163

Figure 7.15: Kurtosis of total reserve

Overall the simulation model seems to give a good impression of the shape of the reserve risk distribution and can help in finding better estimators and distributional fits in the claims reserving process.

Appendix A

Mathematical Background

A.1 Distributions [1]

A.1.1 The Normal Distribution

A random variable X is said to be normally distributed, i.e. $X \sim \mathcal{N}(\mu, \sigma^2)$, if it has the probability density function (pdf)

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$$

We have

$$\begin{aligned} \mathbb{E}[X] &= \mu, \\ \mathbb{V}[X] &= \sigma^2, \\ \text{Skew}(X) &= 0, \\ \text{Kurt}(X) &= 3. \end{aligned}$$

A.1.2 The Lognormal Distribution

We say a random variable X is lognormally distributed if $Z = \log(X)$ is normally distributed. We denote $X \sim \log\mathcal{N}(\mu, \sigma)$ and for the pdf it follows that

$$f(x) = \frac{1}{x\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2}\frac{[\log(x) - \mu]^2}{\sigma^2}\right)$$

For the moments we have

$$\mathbb{E}[X^r] = \exp\left(r\mu + \frac{1}{2}r^2\sigma^2\right)$$

[1] definitions and results taken from [JKB94a] and [JKB94b]

which leads to

$$\begin{aligned}
\mathbb{E}[X] &= \exp\left(\mu + \frac{1}{2}\sigma^2\right), \\
\mathbb{V}[X] &= \exp\left(2\mu + \sigma^2\right)\left(\exp\left(\sigma^2\right) - 1\right), \\
\text{Skew}(X) &= \sqrt{\exp\left(\sigma^2\right) - 1}\left(\exp\left(\sigma^2\right) + 2\right), \\
\text{Kurt}(X) &= \exp\left(4\sigma^2\right) + 2\exp\left(3\sigma^2\right) + 3\exp\left(2\sigma^2\right) - 3.
\end{aligned}$$

A.1.3 Generalized Pareto Distribution [2]

There are many different forms of the Pareto distribution. We will use a form conform with [Mor13] to be able to use the moment estimators from the paper.
We say a random variable X follows a generalized Pareto distribution ($X \sim \text{GPD}(\mu, s, \xi)$) if it has the cumulative distribution function

$$F(x) = 1 - \left(1 + \frac{\xi(x-\mu)}{s}\right)^{-1/\xi}$$

for $x > \mu \in \mathbb{R}$, $s > 0$ and $\xi \in \mathbb{R}$. For the moments we have

$$\begin{aligned}
\mathbb{E}[X] &= \mu + \frac{s}{1-\xi}, \\
\mathbb{V}[X] &= \frac{s^2}{(1-\xi)(1-2\xi)}, \\
\text{Skew}(X) &= \frac{2(1+\xi)\sqrt{1-2\xi}}{1-3\xi}, \\
\text{Kurt}(X) &= \frac{3\left(3-5\xi-4\xi^3\right)}{1-7\xi+12\xi^2}.
\end{aligned}$$

A.1.4 Beta Distribution

We say a random variable X follows a beta distribution, noted $X \sim \text{Beta}(p, q)$, if its pdf is of the form

$$f(x) = \frac{\text{B}(p,q)}{x}^{p-1}(1-x)^{q-1}, \quad 0 \le x \le 1$$

where $\text{B}(p, q)$ is the beta function $\text{B}(p, q) = \int_0^1 t^{p-1}(1-t)^{q-1}\text{d}t$. For more information on the beta function see for example [AS65]. We have

$$\begin{aligned}
\mathbb{E}[X] &= \frac{p}{p+q}, \\
\mathbb{V}[X] &= \frac{pq}{(p+q)^2(p+q+1)}, \\
\text{Skew}(X) &= \frac{2(q-p)\sqrt{p^{-1}+q^{-1}+(pq)^{-1}}}{p+q+2},
\end{aligned}$$

[2] to match [Mor13] the definition is taken from [EKM03]

$$\text{Kurt}\,(X) = \frac{3\,(p+q+1)\left[2(p+q)^2 + pq(p+q-6)\right]}{pq(p+q+2)(p+q+3)}.$$

A.1.5 Gamma Distribution

A random variable Y follows a generalized gamma distribution, noted $Y \sim \text{gGam}\,(\alpha, \beta, \gamma)$ if its distribution function is

$$f_Y(x) = \frac{(x-\gamma)^{\alpha-1} \exp\left(-\frac{x-\gamma}{\beta}\right)}{\beta^\alpha \Gamma(\alpha)}$$

where $\alpha, \beta > 0$ and $x > \gamma$ and where $\Gamma(\alpha)$ is the gamma function $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} \mathrm{d}t$. For more information on the gamma function see for example [AS65]. For the standard form of the the distribution we set $\gamma = 0$ and then have $X \sim \text{Gam}\,(\alpha, \beta)$ with the pdf

$$f_X(x) = \frac{x^{\alpha-1} \exp\left(-\frac{x}{\beta}\right)}{\beta^\alpha \Gamma(\alpha)}, \quad \alpha, \beta > 0.$$

We then have

$$\begin{aligned}
\mathbb{E}\,[X] &= \alpha\beta, \\
\mathbb{V}\,[X] &= \alpha\beta^2, \\
\text{Skew}\,(X) &= \frac{2}{\sqrt{\alpha}}, \\
\text{Kurt}\,(X) &= 3 + \frac{6}{\alpha}.
\end{aligned}$$

A.1.6 F Distribution

If X_1, X_2 are independent chi-square variables with degrees of freedom ν_1 and ν_2, i.e. $X_i \sim \chi^2_{\nu_i}$ for $I = 1, 2$, then the distribution of

$$X = \left(\frac{X_1}{\nu_1}\right) \left(\frac{X_2}{\nu_2}\right)^{-1}$$

is the F distribution with ν_1, ν_2 degrees of freedom. Its pdf is given by

$$f(x) = \frac{\left(\dfrac{\nu_1}{\nu_2}\right)^{\nu_1/2}}{\mathrm{B}\left(\dfrac{1}{2}\nu_1, \dfrac{1}{2}\nu_2\right)} \frac{x^{(\nu_1/2)-1}}{\left(1 + \dfrac{\nu_1}{\nu_2} x\right)^{(\nu_1+\nu_2)/2}}, \quad x > 0.$$

Its moments are given by

$$\mathbb{E}\,[X] = \frac{\nu_2}{\nu_2 - 2}, \quad \nu_2 > 2,$$

$$\mathbb{V}[X] = \frac{2\nu_2^2(\nu_1+\nu_2-2)}{\nu_1(\nu_2-2)^2(\nu_2-4)}, \quad \nu_2 > 4,$$

$$\mathrm{Skew}(X) = \sqrt{\frac{8(\nu_2-4)}{(\nu_1+\nu_2-2)\nu_1}}\frac{2\nu_1+\nu_2-2}{\nu_2-6}, \quad \nu_2 > 6,$$

$$\mathrm{Kurt}(X) = \frac{3\left[\nu_2-4+\frac{1}{2}(\nu_2-6)[\mathrm{Skew}(X)]^2\right]}{\nu_2-8}, \quad \nu_2 > 8.$$

A.1.1 Remark. For the definition of the chi-square distribution consult [JKB94a].

A.1.7 Student's t Distribution

Let $U \sim U(0,1)$ and $V \sim \chi_\nu^2$ then we have that

$$X = U\frac{\sqrt{\nu}}{V}$$

follows a student t distribution with ν degrees of freedom. It has the pdf

$$f(x) = \frac{\Gamma\left(\frac{1}{2}(\nu+1)\right)}{\sqrt{\pi\nu}\,\Gamma\left(\frac{1}{2}\nu\right)}\frac{1}{\left(1+\frac{x^2}{\nu}\right)^{\frac{\nu+1}{2}}}$$

All odd moments of X are zero. We have

$$\mathbb{E}[X] = 0,$$

$$\mathbb{V}[X] = \frac{\nu}{\nu-2}, \quad \nu \geq 2,$$

$$\mathrm{Skew}(X) = 0,$$

$$\mathrm{Kurt}(X) = \frac{3(\nu-2)}{\nu-4}, \quad \nu \geq 4.$$

A.1.8 Uniform Distribution on $[0,1]$

A random variable X follows a uniform distribution on the interval $[0,1]$, noted $X \sim U(0,1)$ if its distribution function is

$$f(x) = \mathbf{1}_{[0,1]}$$

where $\mathbf{1}_{[0,1]}$ is the indicator function on $[0,1]$. We have

$$\mathbb{E}[X] = \frac{1}{2},$$

$$\mathbb{V}[X] = \frac{1}{12},$$

$$\mathrm{Skew}(X) = 0,$$

$$\mathrm{Kurt}(X) = \frac{9}{5}.$$

A.2 Correlation

A.2.1 Copula [3]

Copulas are used in mathematical statistics to describe dependencies between random variables. In general it is not possible to determine the cumulative distribution function (cdf) F of a d dimensional random vector X from just its margin F_i, $i = 1, \dots, d$. They have to be coupled in some way, which is where copulas come into place. First some general definitions:

A.2.1 Definition. Let $X = (X_1, \dots, X_d)$ be a d-dimensional random vector defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Its properties are defined by its cumulative distribution function

$$F(x_1, \dots, x_d) \coloneqq \mathbb{P}\left[X_1 \leq x_1, \dots, X_d \leq x_d\right], \quad x_1, \dots, x_d \in \mathbb{R}.$$

Then for $i = 1, \dots, d$ the distribution functions F_i of X_i are called the marginal distributions of F and can be calculated from F via

$$F_i(x_i) = \lim_{x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_d \to \infty} F(x_1, \dots, x_d),$$

which is sometimes also written as

$$F_i(x_i) = F(\infty, \dots, \infty, x_i, \infty, \dots, \infty).$$

A.2.2 Definition (Copula)**.**

- A function $C : [0,1]^d \to [0,1]$ is called a (d-dimensional) copula, if there is a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ supporting a random vector $(U_1, \dots, U_d)$ such that $U_k \sim U(0,1)$ (the uniform distribution) for all $k = 1, \dots, d$ and

$$C\left(u_1, \dots, U_d\right) = \mathbb{P}\left[U_1 \leq u_1, \dots, U_d \leq u_d\right], \quad u_1, \dots, u_d \in \mathbb{R}$$

- On a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ let $(U_1, \dots, U_d)$ be a random vector on $[0,1]^d$ whose joint distribution function (restricted to $[0,1]^d$) is a copula $C : [0,1]^d \to [0,1]$. For $i = 2, \dots, d$ and indices $1 \leq j_i < \dots < j_i \leq d$ the notation $C_{j_1, \dots, j_i} : [0,1]^i \to [0,1]$ is introduced for the joint distribution function of the random subvector $(U_{j_1}, \dots, U_{j_i})$. It is itself a copula and called an i-margin of C.

A.2.3 Remark. Note that for a random vector $(U_1, \dots, U_d) \in [0,1]^d$ the values of its cdf on $\mathbb{R}^d \setminus [0,1]^d$ are completely determined by its values on $[0,1]^d$. For this reason copulas are only defined on the d-dimensional unit cube.

The theoretical foundation for the use of copulas is given by the Sklar's Theorem. In short it states that it is always possible to decouple a multivariate probability distribution into its marginal distributions and a copula. Conversely it is possible to build a multivariate probability distribution by combining given margins with a copula. These two elements are then often easier to handle than the law of the joint probability distribution.

[3] following [MS12]

A.2.4 Theorem (Sklar's Theorem)**.** *Let F be a d-dimensional distribution function with margins $F_1, \dots, F_d$. Then there exists a d-dimensional copula C such that for all $(x_1, \dots, x_d) \in \mathbb{R}^d$ it holds that*

$$F(x_1, \dots, x_d) = C\left(F_1(x_1), \dots, F_d(x_d)\right) \tag{A.1}$$

If $F_1, \dots, F_d$ are continuous then C is unique. Conversely if C is a d-dimensional copula and $F_1, \dots, F_d$ are univariate distribution functions then the function F defined via eq. (A.1) is a d-dimensional distribution function.

Let us now define the copula that we will use in our simulation

A.2.5 Definition (The Gaussian copula)**.** On a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ let $(X_1, \dots, X_d)$ be a normally distributed random vector with joint distribution function

$$F(x_1, \dots, x_d) = \int\limits_{\times_{i=1}^{d}(-\infty, x_i]} (2\pi)^{-\frac{d}{2}} \det(\Sigma)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(\boldsymbol{s} - \boldsymbol{\mu})\Sigma^{-1}(\boldsymbol{s} - \boldsymbol{\mu})^\intercal\right) \mathrm{d}s$$

for a symmetric, positive-definite matrix Σ and a mean vector $\boldsymbol{\mu} = (\mu_1, \dots, \mu_d) \in \mathbb{R}^d$, $\boldsymbol{s} := (s_1, \dots, s_d)$ and $\det(\Sigma)$ is the determinant of Σ. Denoting by $\sigma_1^2 := \Sigma_{11}, \dots, \sigma_d^2 := \Sigma_{dd} > 0$ the diagonal entries of Σ, the marginal law F_i of X_i is a normal distribution with mean μ_i and variance σ_i, $I = 1, \dots, d$. The copula C_Σ^{Gauss} of $(X_1, \dots, X_d)$ is called the Gaussian copula and is given by

$$C_\Sigma^{Gauss}(u_1, \dots, u_d) := F\left(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d)\right). \tag{A.2}$$

Lastly we will state a general theorem that will help us in the next subsection.

A.2.6 Theorem. *Let X be a random variable with a continuous cdf F_X. Then we have*

- *The random variable $Y := F_X(X)$ is uniformly distributed, i.e. $Y \sim U(0, 1)$.*
- *Let $U \sim U(0, 1)$, then $Z := F_X^{-1}(U)$ follows the same distribution as X.*

For the proof of this theorem we refer to [MS12] (where the more general case of a non continuous cdf F_X is treated too). The second point of Theorem A.2.6 is the basis for an often used algorithm to create pseudo random samples named inversion sampling.

A.2.7 Definition (Inversion Sampling)**.** The inversion sampling method lets us generate random samples of a random variable X which follows a cumulative distribution function F_X. It requires a method to generate uniform samples.
To sample from a given random variable X via inversion sampling we have to follow these steps

1. Create a sample u from a uniform distribution on the intervall $[0, 1]$
2. Use the (generalized) inverse function F_X^{-1} on u to get the value $x = F_X^{-1}(u)$. We have that
$$F_X(x) = F_X\left(F_X^{-1}(u)\right) = u$$
3. Then x can be taken as a random number drawn from the distribution given by F_X.

A.2.2 Generation of random variables with the Gaussian copula

One of the practical applications of the above is, that it is possible to generate pseudo random samples of a multivariate probability distribution with given marginals and a copula.
We use the following algorithm to generate correlated random variables:

A.2.8 Algorithm.
Inputs: A set of correlated risks $(X_1, \ldots, X_n)$ with marginal cumulative distribution Functions F_i, $i = 1, \ldots, n$ and a linear correlation matrix $\rho \in [0,1]^{n \times n}$.

1. If ρ is positive definite the Choleski decomposition is used to find a matrix $\mathbf{C}$ such that $\rho = \mathbf{C}^\intercal \mathbf{C}$. In the case that ρ is only positive semidefinite calculate the spectral decomposition of ρ, i.e. $\rho = \mathbf{U} \Lambda \mathbf{U}^\intercal$, where $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ is the diagonal matrix consisting of the eigenvalues $\lambda_1, \ldots, \lambda_n$ of ρ and $\mathbf{U}$ is an orthogonal matrix, and set $\mathbf{C} := \Lambda^{\frac{1}{2}} \mathbf{U}^\intercal$.

2. Generate n standard normal variables $\mathbf{Y} = (Y_1, \ldots, Y_n)$.

3. Set $\mathbf{Z} = \mathbf{YC}$.

4. Set $u_i = \Phi(Z_i)$ for $i = 1, \ldots, n$, where Φ is the cdf of the standard normal distribution.

5. Set $\widehat{X}_i = F_i^{-1}(u_i)$.

Output: The vector $(\widehat{X}_1, \ldots, \widehat{X}_n)$ forms a sample from a multivariate distribution which was generated using the correlations $(\rho_{ij})_{i,j=1,\ldots,n}$ and marginals F_i, $i = 1, \ldots, n$.

A.2.9 Remark. Notes to some of the steps

- Item 2 can be achieved multiple ways. We will use the built-in Matlab functions for the creation of normal random variables. Generally the most used form is the Box-Muller transformation method (see for example [Dev86]).

- Items 3 and 4 represent the copula part of the algorithm and generate a sample vector $\boldsymbol{U} = (u_1, \ldots, u_n)$, where each u_i is a sample of a uniformly distributed random variable $U_i \sim U(0,1)$, $i = 1, \ldots, n$ (cf. Theorem A.2.6) generated by using the input correlation ρ.

- Note that the linear correlation of the output is not exactly matching the input correlation but is only an approximation. The application of the normal cumulative distribution function in item 4 and the inverse cdf in item 5 lead to distortions.

- Item 5 is the inversion method and yields an output as described (cf. definition A.2.7).

Bibliography

[AS65] Milton Abramowitz and Irene A. Stegun. *Handbook of mathematical functions with formulas, graphs and mathematical tables.* 9th ed. Dover Publications, 1965.

[Dev86] Luc Devroye. *Non-uniform random variate generation.* 1st ed. Springer, New York, 1986.

[EKM03] Paul Embrechts, Claudia Klüppelberg and Thomas Mikosch. *Modelling Extremal Events for Insurance and Finance.* 4th ed. Springer, 2003.

[JKB94a] Norman L. Johnson, Samuel Kotz and N. Balakrishnan. *Continuous Univariate Distributions.* 2nd ed. Vol. 1. John Wiley & Sons Ltd, USA, 1994.

[JKB94b] Norman L. Johnson, Samuel Kotz and N. Balakrishnan. *Continuous Univariate Distributions.* 2nd ed. Vol. 2. John Wiley & Sons Ltd, USA, 1994.

[Mac93] Thomas Mack. 'Distribution-free calculation of the standard error of chain ladder reserve stimates'. In: *Astin Bulletin* 23.2 (1993).

[MI75] Elgie J. McGrath and David C. Irving. 'Random number generation for selected probability distributions'. In: *Techniques for efficient Monte Carlo sampling.* Vol. 2. 1975. URL: http://web.ornl.gov/info/reports/1975/3445605715394.pdf.

[Mor12] Eric Dal Moro. *Application of Skewness to Non-Life Reserving.* ASTIN Colloquium 2012. Available at SSRN: http://ssrn.com/abstract=2281878. 2012.

[Mor13] Eric Dal Moro. *Kurtosis and skewness estimation for non-life reserve risk distribution.* ASTIN Colloquium 2013. Available at SSRN: http://ssrn.com/abstract=2281902. 2013.

[MS12] Jan-Frederik Mai and Matthias Scherer. *Simulating Copulas: Stochastic Models, Sampling Algorithms, and Applications.* Imperial College Press, 2012.

[WM08] Mario V. Wüthrich and Michael Merz. *Stochastic claims reserving methods in insurance.* John Wiley & Sons Ltd, Sussex, England, 2008.