1 Basics of claim reserving
1.1 Introduction and motivation
1.2 Basic terms and definitions
1.3 Literature and software

2 Chain-Ladder-Method (CLM)
2.1 How does the Chain-Ladder method work
2.2 Future development
2.3 Validation and examples (part 1 of 3)
2.4 Ultimate uncertainty
2.5 Validation and examples (part 2 of 3)
2.6 Solvency uncertainty
2.7 Validation and examples (part 3 of 3)
2.8 Literature

3 Other classical reserving methods
3.1 Complementary-Loss-Ration method (CLRM)
3.2 Bornhuetter-Ferguson method (BFM)
3.3 Benktander-Hovinen method (BHM)
3.4 Cape-Cod method
3.5 Extended-Complementary-Loss-Ration method (ECLRM)
3.6 Other methods
3.7 Literature
Table of contents

1 Basics of claim reserving
1.1 Introduction and motivation
1.2 Basic terms and definitions
1.3 Literature and software

2 Chain-Ladder-Method (CLM)
2.1 How does the Chain-Ladder method work
2.2 Future development
2.3 Validation and examples (part 1 of 3)
2.4 Ultimate uncertainty
2.5 Validation and examples (part 2 of 3)
2.6 Solvency uncertainty
2.7 Validation and examples (part 3 of 3)
2.8 Literature

3 Other classical reserving methods
3.1 Complementary-Loss-Ration method (CLRM)
3.2 Bornhuetter-Ferguson method (BFM)
3.3 Benktander-Hovinen method (BHM)
3.4 Cape-Cod method
3.5 Extended-Complementary-Loss-Ration method (ECLRM)
3.6 Other methods
3.7 Literature
4 Linear-Stochastic-Reserving methods

4.1 How do Linear-Stochastic-Reserving methods (LSRM) work
4.2 Future development
4.3 Ultimate uncertainty
4.4 Solvency uncertainty
4.5 Examples
4.6 Estimation of correlation of reserving Risks
4.7 Literature

5 Separation of small and large claims

5.1 What is the problem with large claims
5.2 How to separate small from large claims
5.3 Estimation methods for small and large claims
5.4 Modelling the transition from small to large
5.5 Literature

6 Poisson-Model

6.1 Modelling the number of reported claims
6.2 Projection of the future outcome
6.3 Ultimate uncertainty of the Poisson-Model
6.4 Generalised linear models and reserving
6.5 Literature
4 Linear Stochastic Reserving methods
4.1 How do Linear Stochastic Reserving methods (LSRM) work
4.2 Future development
4.3 Ultimate uncertainty
4.4 Solvency uncertainty
4.5 Examples
4.6 Estimation of correlation of reserving risks
4.7 Literature

5 Separation of small and large claims
5.1 What is the problem with large claims
5.2 How to separate small from large claims
5.3 Estimation methods for small and large claims
5.4 Modelling the transition from small to large
5.5 Literature

6 Poisson-Model
6.1 Modelling the number of reported claims
6.2 Projection of the future outcome
6.3 Ultimate uncertainty of the Poisson-Model
6.4 Generalised linear models and reserving
6.5 Literature
7 Bootstrap for CLM
7.1 Motivation
7.2 Chain-Ladder method and bootstrapping, variant 1
7.3 Bootstrapping Chain-Ladder step by step, variant 1
7.4 Chain-Ladder method and bootstrapping, variant 2
7.5 Bootstrapping Chain-Ladder step by step, variant 2
7.6 Possible problems with bootstrapping
7.7 Parametric vs. non-parametric bootstrap
7.8 Literature

8 CLM: Bayesian & credibility approach
8.1 A Bayesian approach to the Chain-Ladder method
8.2 A credibility approach to the Chain-Ladder method
8.3 Example
8.4 Literature
7 Bootstrap for CLM
7.1 Motivation
7.2 Chain-Ladder method and bootstrapping, variant 1
7.3 Bootstrapping Chain-Ladder step by step, variant 1
7.4 Chain-Ladder method and bootstrapping, variant 2
7.5 Bootstrapping Chain-Ladder step by step, variant 2
7.6 Possible problems with bootstrapping
7.7 Parametric vs. non-parametric bootstrap
7.8 Literature

8 CLM: Bayesian & credibility approach
8.1 A Bayesian approach to the Chain-Ladder method
8.2 A credibility approach to the Chain-Ladder method
8.3 Example
8.4 Literature
9 Mid year reserving

9.1 Problem of mid-year reserving
9.2 Methods for mid-year reserving
9.3 Conclusion
9.4 Literature
Stochastic Reserving

9 Mid year reserving

9.1 Problem of mid-year reserving
9.2 Methods for mid-year reserving
9.3 Conclusion
9.4 Literature
Stochastic Reserving

Lecture 1

Introduction

René Dahms

ETH Zurich, Spring 2019

20 February 2019

(Last update: 18 February 2019)
1 Basics of claim reserving

1.1 Introduction and motivation
1.1.1 General insurance
1.1.2 Claim reserves
1.1.3 Relevance of claim reserves
1.1.4 Purposes of (stochastic) loss reserving

1.2 Basic terms and definitions
1.2.1 Terminology
1.2.2 Triangles (trapezoids)
1.2.3 Stochastic reserving and Best Estimate

1.3 Literature and software
Lecture 1: Table of contents

1 Basics of claim reserving
  1.1 Introduction and motivation
  1.2 Basic terms and definitions
  1.3 Literature and software
All starts with:
An insured (policyholder) pays some premium to an insurer in order to transfer the (more or less directly related) significant monetary consequences (loss) of a randomly incurring future event (risk).

Examples 1.1

<table>
<thead>
<tr>
<th>Insurance</th>
<th>Insured Loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Motor Liability (MTPL)</td>
<td>loss to a 3rd person caused by a self-inflicted car accident</td>
</tr>
<tr>
<td>General Liability (GL)</td>
<td>loss to a 3rd person caused by the policyholder, except car accidents</td>
</tr>
<tr>
<td>Fire (Property)</td>
<td>policyholders loss to household and property caused by fire</td>
</tr>
<tr>
<td>Health</td>
<td>policyholders loss caused by illness</td>
</tr>
<tr>
<td>Pension</td>
<td>policyholders loss, because of a long life</td>
</tr>
<tr>
<td>Life</td>
<td>‘another persons loss’ caused by the death of the insured</td>
</tr>
</tbody>
</table>

Life insurance
The insured risk depends directly on the life of the insured.

General (or Non-Life) insurance
The insured risk does not depends directly on the life of the insured.
All starts with:
An insured (policyholder) pays some premium to an insurer in order to transfer the (more or less directly related) significant monetary consequences (loss) of a randomly occurring future event (risk).

Examples 1.1

<table>
<thead>
<tr>
<th>Insurance</th>
<th>Insured loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Motor Liability (MTPL)</td>
<td>loss to a 3rd person caused by a self-inflicted car accident, except car accidents</td>
</tr>
<tr>
<td>General Liability (GL)</td>
<td>loss to a 3rd person caused by the policyholder, except car accidents</td>
</tr>
<tr>
<td>Fire (Property)</td>
<td>policyholders loss to household and property caused by fire</td>
</tr>
<tr>
<td>Health</td>
<td>policyholders loss caused by illness</td>
</tr>
<tr>
<td>Pension</td>
<td>policyholders loss, because of a long life</td>
</tr>
<tr>
<td>Life</td>
<td>‘another persons loss‘ caused by the death of the insured</td>
</tr>
</tbody>
</table>

Life insurance
The insured risk depends directly on the life of the insured.

General (or Non-Life) insurance
The insured risk does not depend directly on the life of the insured.

Important words of the definition:

- **transfer**: therefore no self-insurance
- **random future**: not (completely) known, random in timing or amount
- **loss**: no lotteries and no betting
- **significant loss**: therefore no service contract
Reinsurance, Health and Accident

There are types of insurances which have components of both, Life and General insurance. The classification depends on the regulator, the company and the accounting standard.

Switzerland

Life (and Pensions), Non-Life (General insurance or P&C), Health and Reinsurance

IFRS 17

An insurance contract is ‘a contract under which one party (the issuer) accepts significant insurance risk from another party (the policyholder) by agreeing to compensate the policyholder if a specified uncertain future event (the insured event) adversely affects the policyholder’
Reinsurance, Health and Accident

There are types of insurances which have components of both, Life and General insurance. The classification depends on the regulator, the company and the accounting standard.

Switzerland
Life (and Pensions), Non-Life (General insurance or P&C), Health and Reinsurance

IFRS 17
An insurance contract is ‘a contract under which one party (the insurer) accepts significant insurance risk from another party (the policyholder) by agreeing to compensate the policyholder if a specified uncertain future event (the insured event) adversely affects the policyholder’
Problem 1.2

At the end of a business year the insurer usually knows all its contracts but not all the corresponding claims and their ultimate losses. Reasons may be:

1. Not yet materialised or detected claims. For instance, product liability insurance.

2. Not yet reported claims. For instance, time delay, because of holidays.

3. Unknown future payments for not yet finally settled claims.
Problem 1.2
At the end of a business year the insurer usually knows all its contracts but not all the corresponding claims and their ultimate losses. Reasons may be:

1. Not yet materialised or detected claims. For instance, product liability insurance.
2. Not yet reported claims. For instance, time delay because of holidays.
3. Unknown future payments for not yet finally settled claims.

- Strictly taken: From the point in time where the insurance contract is in force (or the insurance company has send a binding offer), you have to account for all potential claims. The precise rules for this depend on the accounting standard.
Payment pattern

depend strongly on the underlying risk (exposure). Therefore, in practice an actuary not only have to look at number based statistics, but also have to understand the type of the underlying exposure.
• red may be Motor Hull
• blue is typical for Garantie Decannale in France
• gray may be mandatory accident insurance in Switzerland
Claim reserves are often the most important part of the balance sheet of a general insurer. Moreover, a slightly false estimate of claim reserves may make the difference between an annual profit or loss.

Some examples*:

<table>
<thead>
<tr>
<th>insurer</th>
<th>equity</th>
<th>gain</th>
<th>reserves</th>
<th>gain reserves</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zurich</td>
<td>$ 21.0</td>
<td>$ 3.0</td>
<td>$ 82.7</td>
<td>3.6%</td>
</tr>
<tr>
<td>Allianz</td>
<td>€ 31.4</td>
<td>€ 3.5</td>
<td>€ 78.0</td>
<td>4.5%</td>
</tr>
<tr>
<td>Swiss Re</td>
<td>$ 11.7</td>
<td>$ 3.0</td>
<td>$ 49.5</td>
<td>6.1%</td>
</tr>
<tr>
<td>Munich Re</td>
<td>€ 14.1</td>
<td>€ 2.6</td>
<td>€ 45.0</td>
<td>5.8%</td>
</tr>
</tbody>
</table>

*Amounts (in billion) representing only the general insurance part of the company and are taken from the annual reports of 2012. The amounts are not entirely comparable, because the separation of the general insurance business from the other parts may be different from company to company.
Claim reserves are often the most important part of the balance sheet of a general insurer. Moreover, a slightly false estimate of claim reserves may make the difference between an annual profit or loss.

Some examples:

<table>
<thead>
<tr>
<th>Insurer</th>
<th>Equity</th>
<th>Gain</th>
<th>Reserves</th>
<th>Gain Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zurich</td>
<td>€ 31.0</td>
<td>€ 3.5</td>
<td>€ 82.7</td>
<td>9.6%</td>
</tr>
<tr>
<td>Allianz</td>
<td>€ 31.4</td>
<td>€ 3.5</td>
<td>€ 78.0</td>
<td>4.5%</td>
</tr>
<tr>
<td>Swiss Re</td>
<td>$ 17.7</td>
<td>$ 3.0</td>
<td>$ 49.5</td>
<td>6.1%</td>
</tr>
<tr>
<td>Munich Re</td>
<td>€ 14.1</td>
<td>€ 2.6</td>
<td>€ 45.0</td>
<td>5.8%</td>
</tr>
</tbody>
</table>

*Amounts (in billion) representing only the general insurance part of the company and are taken from the annual reports of 2012. The amounts are not entirely comparable, because the separation of the general insurance business from other parts may differ from company to company.
Example: Converium AG

Converium AG was one of the largest reinsurers worldwide. At 20\textsuperscript{th} July 2004 the company issued a profit warning caused by a strengthening of the claim reserves of the US general liability portfolio by $400 million.

Consequences:

- loss of 35\% of equity
- an immediate deep plunge of over 50\% (about 70\% until October 2004) of the stock price
- rating downgrade from A to BBB+ by Standard & Poors
- unfriendly takeover by SCOR in 2007 (although Converium did make profit again and got its A rating back)
Example: Converium AG
Converium AG was one of the largest reinsurers worldwide. At 20th July 2004 the company issued a profit warning caused by a strengthening of the claim reserves of the US general liability portfolio by $400 million.
Consequences:
• loss of 35% of equity
• an immediate deep plunge of over 50% (about 70% until October 2004) of the stock price
• rating downgrade from A to BBB+ by Standard & Poors
• unfriendly takeover by SCOR in 2007 (although Converium did make profit again and got its A rating back)
Loss reserving

is an integral part of many processes. For instance:

- annual closings
- pricing
- forecasts
- measuring risks, like under IFRS 17, Solvency II and Swiss Solvency Test (SST)
- modelling the value of customers
- ...

The resulting estimates for claim reserves depend on the purpose. For instance, loss reserving in the context of annual closings deals with the past, whereas in the context of pricing we are interested in the future. Moreover, in pricing one usually looks at a more detailed split in subportfolios than during closings.
Loss reserving is an integral part of many processes. For instance:
- annual closings
- pricing
- forecasts
- measuring risks, like under IFRS 17, Solvency II and Swiss Solvency Test (SST)
- modelling the value of customers
- ...

The resulting estimates for claim reserves depend on the purpose. For instance, loss reserving in the context of annual closings deals with the past, whereas in the context of pricing we are interested in the future. Moreover, in pricing one usually looks at a more detailed split in subportfolios than during closings.
Definition 1.3 (Case reserves or outstanding)

**Case reserves** are estimates of the (undiscounted) sum of all future payments made by claim managers on a claim by claim basis.

Definition 1.4 (Claim reserves or (technical) provisions)

**Claim reserves** are the estimates of the (undiscounted) sum of all future payments for claims (of a portfolio) that have already happened.

\[
\text{claim reserves} = \text{case reserves} + \text{IBNR}
\]

Definition 1.5 (Incurred but not yet reported (IBNyR) reserves)

**IBNyR reserves** are the part of the claim reserves that corresponds to not yet reported claims.

Definition 1.6 (Incurred but not enough reserved (IBNeR) reserves)

**IBNeR reserves** are the difference between the claim reserves for claims known to the insurer and the corresponding case reserves.

Definition 1.7 (IBNR or IBN(e/y)R)

\[
\text{IBNR reserves} = \text{IBNeR} + \text{IBNyR}
\]
Provided we take a positive sign for claim reserves IBNyR are non-negative, whereas IBNeR may be positive or negative.

Usually, we will not look at discounted reserves, because discounting (and inflation) disturbs the development of claims and is dealt with separately, i.e. first get undiscounted figures and corresponding payment patterns and then discount.
Definition 1.8 (Incurred (losses) or reported amounts)
\[
\text{incurred} = \text{payments} + \text{case reserves}
\]

Definition 1.9 (Ultimate)
\[
\text{ultimate} = \text{payments} + \text{claim reserves} = \text{incurred} + \text{IBNR}
\]

Remarks 1.10
- Payments are often called paid (losses).
- The naming is not consistent within the actuarial world. For instance, actuaries often understand under IBNR only the IBNyR part.
- The definitions depend on the accounting standard. For instance, under IFRS 17 one has to discount the cash flows and one has to take the inception date (or the begin of the coverage period) instead of the accident date.
Definition 1.8 (Incurred (losses) or reported amounts)
\[ \text{incurred} = \text{payments} + \text{case reserves} \]

Definition 1.9 (Ultimate)
\[ \text{ultimate} = \text{payments} + \text{claim reserves} = \text{incurred} + \text{IBNR} \]
\[ \text{ultimate} = \text{claim reserves} + \text{payments} = \text{IBNR} + \text{case reserves} + \text{payments} = \text{IBNR} + \text{incurred} \]

Remarks 1.10
- Payments are often called paid (losses).
- The naming is not consistent within the actuarial world. For instance, actuaries often understand under IBNR only the IBNyR part.
- The definitions depend on the accounting standard. For instance, under IFRS 17 one has to discount the cash flows and one has to take the inception date (or the beginning of the coverage period) instead of the accident date.
Main objects

of reserving are claim development triangles (trapezoids), containing the development of payments (or other claim properties) per accident period for a whole portfolio.

- We assume that \( I \geq J \). If \( I = J \) we have a triangle and otherwise a trapezoid, but for simplicity we will call it triangle anyway.
- rows = accident (or origin) periods
- columns = development periods
- diagonals = calendar periods
- \( S_{i,k} \) are the payments during development period \( k \) for claims happened in accident period \( i \).

If more than one portfolio is involved we add an additional upper index \( m \) to indicate the triangle.
- Payments could be replaced by other claim properties like
  * changes of reported amounts (\( = \) incremental incurred)
  * number of newly reported claims
  * payments on just getting large claims
  * ...
Main objects of reserving are claim development triangles (trapezoids), containing the development of payments (or other claim properties) per accident period for a whole portfolio.

- rows = accident (or origin) periods
- columns = development periods
- diagonals = calendar periods
- \( S_{i,k} \) are the payments during development period \( k \) for claims happened in accident period \( i \).
- Payments could be replaced by other claim properties like:
  - changes of reported amounts (\( \Delta \): incremental incurred)
  - number of newly reported claims
  - payments on just getting large claims

We assume that \( I \geq J \). If \( I = J \) we have a triangle and otherwise a trapezoid, but for simplicity we will call it triangle anyway.

Some actuaries look at those numbers from a different angel:

- accident periods or development periods decreasing instead of increasing
- permutation of accident, development and calendar periods

Moreover, the different kinds of periods have not to be based on the same single unit, like month, quarter or year. For instance, sometimes one looks at accident years and development months.
Reserving means to project the future of the triangles in order to get full rectangles.

- $\mathcal{D}^n$ is the $\sigma$-algebra of all information up to calendar period $n$:

$$\mathcal{D}^n := \sigma(S_{i,k}^m : 0 \leq m \leq M, 0 \leq i \leq I, 0 \leq k \leq J \land (n - i))$$

- $\mathcal{D}^I$ is called the past of the triangles.
- The future of the triangles is:

$$\{ S_{i,k}^m : 0 \leq m \leq M, 0 < k \leq J, I - k < i \leq I \}$$

We assume that there is no development after development period $J$. That means we assume that there is no tail development.

- Ultimate of accident period $i$

$$\text{ultimate of accident period } i = \sum_{k=0}^{J} S_{i,k}^m$$

- Claim reserves of accident period $i$

$$\text{claim reserves of accident period } i = \sum_{k=I+1-i}^{J} S_{i,k}^m$$
On a diagonal $n$ we have for all accident and development periods $i$ and $k$: \[ n = i + k, \]

in particular on the last known diagonal $I$ we have $I = k + i$. 

$\sigma$-algebra of all information up to calendar period $n$: 
\[ D_n := \sigma(S_{m,i,k} : 0 \leq m \leq M, 0 \leq i \leq I, 0 \leq k \leq \lfloor n-i \rfloor) \]

The future of the triangles is:
\[ \{S_{m,i,k} : 0 \leq m \leq M, 0 \leq k \leq J, I - k < i \leq I - k + 1 \} \]

We assume that there is no development after development period $J$. That means we assume that there is no tail development.

ultimate of accident period $i$ = \[ \sum_{k=0}^{I} S_{m,i,k} \]

claim reserves of accident period $i$ = \[ \sum_{k=I+1}^{J} S_{m,i,k} \]
Definition 1.11 (Stochastic loss reserving)

We call a reserving method a stochastic reserving method if it is based on a stochastic model.

Remark 1.12

- Some actuaries call reserving methods that are based on simulations stochastic, even if they are not based on a stochastic model.
- Since we have a stochastic model, we usually expect beside the estimate of claim reserves some estimate of the corresponding uncertainties.

Types of stochastic reserving methods

We differentiate between

- **distribution based reserving methods**, which make explicit assumptions on the distribution of claim properties $S_{i,k}^m$ or related objects.
- **distribution free reserving methods**, which only makes assumptions on moments of the distribution of claim properties $S_{i,k}^m$ or related objects.
Definition 1.11 (Stochastic loss reserving)
We call a reserving method a stochastic reserving method if it is based on a stochastic model.

Remark 1.12
• Some actuaries call reserving methods that are based on simulations stochastic, even if they are not based on a stochastic model.
• Since we have a stochastic model, we usually expect beside the estimate of claim reserves some estimate of the corresponding uncertainties.

Types of stochastic reserving methods
We differentiate between
• distribution based reserving methods, which make explicit assumptions on the distribution of claim properties $S_{m_i,k}$ or related objects.
• distribution free reserving methods, which only make assumptions on moments of the distribution of claim properties $S_{m_i,k}$ or related objects.
Definition 1.13 (Best Estimate)

The Swiss regulator defines (translation)

...Best Estimate reserves are the conditional unbiased estimator of the conditional expectation of all future (undiscounted) cash flows based on all at the time of estimation available information ...

FINMA Rundschreiben 2008/42 Rückstellungen Schadenversicherung

Mathematically that could be interpreted like:

\[
\hat{E} \left[ \hat{E} \left[ \sum_{k=0}^{J} S_{i,k}^m \mid D^I \right] - \hat{E} \left[ \sum_{k=0}^{J} S_{i,k}^m \mid D^{I+1} \right] \right] \mid D^I \right] = 0.
\]

estimated claims development result

estimated at time \( I \)
A definition of Best Estimate reserves is not easily to find. We will look at the one of the Swiss regulator.

At the first look this definition looks promising. But if you try to translate the phrase ‘conditional unbiased estimator of a conditional expectation’ into formulas you will get problems.

One possibility is the following:

First we do not look at future cash flows (or reserves) but at the ultimate payments. Since we know the already paid amounts, both views are equivalent, but ultimates are mathematically easier to handle than reserves:

1 We start with the expectation of the ultimate payments conditioned on all currently available information.
2 Estimate
3 One year later we do the same, but of course with more available information.
4 The difference is the observed claims development result (CDR) at time $I + 1$.
5 Taking the expectation conditioned on all currently available information we expect to get zero. From the business point of view this means, we assume that the CDR is zero within the planning framework at time $I$. 

Mathematically that could be interpreted like:

$$E\left[\sum_{k=0}^{I} S_{m,k} \bigg| p^I\right] - E\left[\sum_{k=0}^{I} S_{m,k} \bigg| p^{I+1}\right] = 0.$$ 

estimated claims development result 
estimated at time $I$
Uncertainty of the Best Estimate

- The Holy Grail of loss reserving is to estimate the ($D^I$-conditional) distribution of the reserves. Unfortunately, to do so we need very restrictive model assumptions.
- At least we would like to estimate beside the Best Estimate the corresponding uncertainty. Often this is done via the mean squared error of prediction (mse):

**Definition 1.14 (mse)**

The $B$-conditional mean square error of prediction of the estimate $\hat{Y}$ of a square integrable random variable $Y$ is defined by

$$\text{mse}_B[\hat{Y}] := \mathbb{E}[(Y - \hat{Y})^2 | B].$$

In practice one often fits some distribution to the estimates of the first two centred moments $\hat{Y}$ and $\text{mse}_B[\hat{Y}]$. In loss reserving one often takes a log-normal distribution.

**Lemma 1.15 (Random and parameter error)**

*The mean squared error of prediction can be split into random the parameter error:*

$$\text{mse}_B[\hat{Y}] = \text{Var}[Y | B] + \left( \mathbb{E}[Y - \hat{Y} | B] \right)^2.$$
A proof of the split of the mse will be given in Lecture 3.
Definition 1.16 (Ultimate uncertainty)

The ultimate uncertainty of the estimated ultimate (or reserves) of accident period \( i \) is defined by

\[
\text{mse}_{D^I} \left[ \sum_{k=0}^{J} \hat{S}_{i,k} \right] = E \left[ \left( \sum_{k=0}^{J} (S_{i,k} - \hat{S}_{i,k}) \right)^2 \right] I
\]

and analogously we define the ultimate uncertainty of the whole ultimate (or reserves) by

\[
\text{mse}_{D^I} \left[ \sum_{i=0}^{I} \sum_{k=0}^{J} \hat{S}_{i,k} \right] = E \left[ \left( \sum_{i=0}^{I} \sum_{k=0}^{J} (S_{i,k} - \hat{S}_{i,k}) \right)^2 \right] I
\]
Definition 1.16 (Ultimate uncertainty)

The ultimate uncertainty of the estimated ultimate (or reserves) of accident period $i$ is defined by

$$\text{mse}_{\text{D}_i} = \mathbb{E} \left[ \left( \sum_{k=0}^{\infty} \left( S_{i,k} - \hat{S}_{i,k} \right) \right)^2 \right]$$

and analogously we define the ultimate uncertainty of the whole ultimate (or reserves) by

$$\text{mse}_{\text{D}_i} = \mathbb{E} \left[ \left( \sum_{i=0}^{\infty} \sum_{k=0}^{\infty} \left( S_{i,k} - \hat{S}_{i,k} \right) \right)^2 \right]$$.
Definition 1.17 (CDR)

The true claims development result (true CDR) of accident period $i$ at time $I + 1$ is the difference of the expected ultimates conditioned on all information at time $I$ and $I + 1$, i.e.:

$$\text{CDR}_{i}^{I+1} := \mathbb{E}\left[\sum_{k=0}^{J} S_{i,k} \mid D^I\right] - \mathbb{E}\left[\sum_{k=0}^{J} S_{i,k} \mid D^{I+1}\right].$$

The (observed) claims development result (CDR) of accident period $i$ at time $I + 1$ is the difference of the two corresponding estimates. If necessary we will denote the time of estimation by an additional upper index:

$$\widehat{\text{CDR}}_{i}^{I+1} := \sum_{k=0}^{J} (\widehat{S}_{i,k}^I - \widehat{S}_{i,k}^{I+1}) = \sum_{k=I+1-i}^{J} \widehat{S}_{i,k}^I - \left(S_{i,I+1-i} + \sum_{k=I+2-i}^{J} \widehat{S}_{i,k}^{I+1}\right).$$

The true and the observed CDR of the aggregation of all accident periods are defined by:

$$\text{CDR}^{I+1} := \sum_{i=0}^{I} \text{CDR}^{I+1}_i \quad \text{and} \quad \widehat{\text{CDR}}^{I+1} := \sum_{i=0}^{I} \widehat{\text{CDR}}^{I+1}_i.$$

- A negative CDR corresponds to a loss and a positive CDR corresponds to a profit.
- If we have a Best Estimate then the estimate of the $D^I$-conditional expectation of the observed CDR equals zero.
For the true CDR we have

\[
E\left[ CDRI_{i+1} \mid D^I \right] = E \left[ \sum_{k=I+1-i}^{J} S_{i,k} \mid D^I \right] - \left( S_{i,I+1-i} + E \left[ \sum_{k=I+2-i}^{J} S_{i,k} \mid D_{I+1} \right] \right) = 0.
\]

But for the observed CDR it depends on how do we estimate. Best Estimate is implicitly defined by

\[
E\left[ \widehat{CDRI}_{i+1} \mid D^I \right] = 0.
\]
Uncertainty of the CDR

As we have seen in the example of Converium it is very important (in particular for the CFO, Solvency II or SST) to have some estimate of the uncertainty of the claims development result. Often this is done via some kind of mean squared error of prediction:

### Definition 1.18 (Solvency uncertainty)

The solvency uncertainty of the estimated ultimate (or reserves) of accident period $i$ is defined by

$$
\text{mse}_{0|D^I} \left[ \hat{\text{CDR}}_{i}^{I+1} \right] := E \left[ \left( \hat{\text{CDR}}_{i}^{I+1} - 0 \right)^2 \middle| D^I \right]
$$

and analogously we define the solvency uncertainty of the aggregated ultimate (or reserves) by

$$
\text{mse}_{0|D^I} \left[ \hat{\text{CDR}}^{I+1} \right] := E \left[ \left( \hat{\text{CDR}}^{I+1} - 0 \right)^2 \middle| D^I \right].
$$

### Remark 1.19

Since in practice the deviation of the observed CDR from zero is more important than its deviation from the true CDR, we take the difference between the observed CDR and zero instead of the difference between the observed CDR and the true CDR.
Uncertainty of the CDR

As we have seen in the example of Converium it is very important (in particular for the CFO, Solvency II or SST) to have some estimate of the uncertainty of the claims development result. Often this is done via some kind of mean squared error of prediction:

Definition 1.18 (Solvency uncertainty)

The solvency uncertainty of the estimated ultimate (or reserves) of accident period $i$ is defined by

$$
\text{mse} = \mathbb{E}\left[ (\hat{CDR}_i^{i+1} - 0)^2 \right]
$$

and analogously we define the solvency uncertainty of the aggregated ultimate (or reserves) by

$$
\text{mse} = \mathbb{E}\left[ (\hat{CDR}^{i+1} - 0)^2 \right].
$$

Remark 1.19

Since in practice the deviation of the observed CDR from zero is more important than its deviation from the true CDR, we take the difference between the observed CDR and zero instead of the difference between the observed CDR and the true CDR.

- SST means Swiss Solvency Test
- It is also possible to look at the deviation of the observed CDR from the true CDR. The corresponding uncertainty will always be less or equal to the one we are looking at.
Best Estimate reserves and ultimate and solvency uncertainty will be the main objects of interest for these lectures. When estimating them you should always keep in mind:

- Best Estimate reserves can be compared with the real world. We only have to wait some (maybe very long) time. Moreover, observing the CDR and other statistics we can learn from the past in order to get better estimates in the future.
- But uncertainties cannot be compared with observations from the real world. They will always be a result of a model. Therefore, we cannot learn from the past in order to get better estimates in the future (we even cannot determine if some estimate is better than another).
- Best Estimate reserves and the corresponding uncertainties are like position and impulse in physics:
  
  You cannot (should not) measure both simultaneously!

For instance, in order to get a Best Estimate you may apply some expert judgement, which cannot be reflected in the estimation of uncertainties by the underlying model.
Best Estimate reserves and ultimate and solvency uncertainty will be the main objects of interest for these lectures. When estimating them you should always keep in mind:

- Best Estimate reserves can be compared with the real world. We only have to wait some (maybe very long) time. Moreover, observing the CDR and other statistics we can learn from the past in order to get better estimates in the future.
- But uncertainties cannot be compared with observations from the real world. They will always be a result of a model. Therefore, we cannot learn from the past in order to get better estimates in the future (we even cannot determine if some estimate is better than another).
- Best Estimate reserves and the corresponding uncertainties are like position and impulse in physics: You cannot (should not) measure both simultaneously!

For instance, in order to get a Best Estimate you may apply some expert judgement, which cannot be reflected in the estimation of uncertainties by the underlying model.
Conditional expectations and intuition

Let assume a mother has two children.

a) What (approximately) is the probability that she has two girls?

- $\frac{1}{2}$
- $\frac{1}{3}$
- $\frac{1}{4}$

b) Assume in addition that she has at least one daughter. What (approximately) is the probability that she has two girls?

- $\frac{1}{2}$
- $\frac{1}{3}$
- $\frac{1}{4}$

c) Assume in addition that one daughter was born on a Monday. What (approximately) is the probability that she has two girls?

- $\frac{1}{2}$
- $\frac{1}{3}$
- $\frac{1}{4}$
Conditional expectations and intuition

Let assume a mother has two children.

a) What (approximately) is the probability that she has two girls?

   □ 1/2 □ 1/3 □ 1/4

b) Assume in addition that she has at least one daughter.
   What (approximately) is the probability that she has two girls?

   □ 1/2 □ 1/3 □ 1/4

c) Assume in addition that one daughter was born on a Monday.
   What (approximately) is the probability that she has two girls?

   □ 1/2 □ 1/3 □ 1/4

---

- In general insurance and in particular in reserving conditional probabilities and expectations play an important roll. But they are often not easy to understand.
- In order to illustrate this let have a look at an easy exercise.
- Be careful: The human brain is not build for (conditioned) probabilities and expectations.
Reserving in the real world:

- MCL, PIC, ECLRM
- CLM
- CC
- BF, CLRM

underwriter marked news ...

MCL: Munich-Chain-Ladder-Method
PIC: Paid-Incurred-Chain-Claims-Method
CLM: Chain-Ladder-Method
CC: Cape-Cod-Method
CLRM: Complementary-Loss-Ratio-Method
ECLRM: Extended-Complementary-Loss-Ratio-Method
BFM: Bornhuetter-Ferguson-Method

...
Stochastic Reserving

Basics of claim reserving

Basic terms and definitions

- On the one hand there are information. If actuaries speak of reserving they often thinks in triangles or vectors, containing the usual candidates like payments, reported amounts and number of reported claims, or more exotic things like payments just before closing a claim.
- But often we forget that there are a lot of other very important sources of information, which even may not be numerical.
- On the other hand there are a lot of reserving methods which may help us to get a Best Estimate:
  - Most of them are based on one triangle only, like Chain-Ladder or Cape Code.
  - In recent years some methods, which combine several (in most cases two) triangles, have been propagated. For instance, Munich-Chain-Ladder, Extended-Complementary-Loss-Ratio-Method and Paid-Incurred-Chain-Claims-Method.
  - But at the end the actuary has to include all the other information in order to get his or hers Best Estimate. And to be honest, often this has more to do with fortune telling than with mathematics or statistics.
Literature


*Probability theory. Translated from the German by Robert B. Burckel.*
Berlin: de Gruyter, 1996.

*Measure and integration theory. Transl. from the German by Robert B. Burckel.*

[4] Schmidt, Klaus D.
A Bibliography on Loss Reserving (permanent update).

*Loss reserving: an actuarial perspective.*
Includes bibliographical references and index.

*Stochastic claims reserving methods in insurance.*
Stochastic Reserving

Basics of claim reserving

- Literature and software

Literature


Probability theory. Translated from the German by Robert B. Burckel.

Measure and integration theory. Transl. from the German by Robert B. Burckel.

[4] Schmidt, Klaus D.
A Bibliography on Loss Reserving (permanent update).

Loss reserving: an actuarial perspective.
Includes bibliographical references and index.

Stochastic claims reserving methods in insurance.
• Free software:
  * R (www.cran.r-project.org), in particular the packages actuar and ChainLadder.
  * LSRM Tools (http://sourceforge.net/projects/lsrmtools/)
  * …

• Commercial software:
  * IBNRS by Addactis
  * CROS by Deloitte (not for sale any more)
  * ResQ by Towers Watson (almost no further development)
  * …
Stochastic Reserving

- Basics of claim reserving
- Literature and software

Free software:
- R (www.cran.r-project.org), in particular the packages actuar and ChainLadder
- LSRM Tools (http://sourceforge.net/projects/lsrmtools/)
- ...

Commercial software:
- IBNRIS by Addactis
- CRDS by Deloitte (not for sale any more)
- ResQ by Towers Watson (almost no further development)
- ...
Stochastic Reserving
Lecture 2
**Chain-Ladder method**

René Dahms

ETH Zurich, Spring 2019

27 February 2019
(Last update: 18 February 2019)
2 Chain-Ladder-Method (CLM)

2.1 How does the Chain-Ladder method work
2.1.1 Chain-Ladder method without stochastic
2.1.2 Stochastic behind the Chain-Ladder method

2.2 Future development
2.2.1 Projection of the future development

2.3 Validation and examples (part 1 of 3)
2.3.1 Chain-Ladder method on Payments and on Incurred
2.3.2 How to validate the Chain-Ladder assumptions

2.4 Ultimate uncertainty
2.4.1 Ultimate uncertainty of accident period $i$
2.4.2 Ultimate uncertainty of the aggregation of all accident periods

2.5 Validation and examples (part 2 of 3)
2.5.1 Ultimate uncertainty

2.6 Solvency uncertainty
2.6.1 Solvency uncertainty of a single accident period
2.6.2 Solvency uncertainty of all accident periods
2.6.3 Uncertainties of further CDR’s

2.7 Validation and examples (part 3 of 3)
2.7.1 Solvency uncertainty

2.8 Literature
2 Chain-Ladder-Method (CLM)

2.1 How does the Chain-Ladder method work
2.1.1 Chain-Ladder method without stochastic
2.1.2 Stochastic behind the Chain-Ladder method

2.2 Future development
2.2.1 Projection of the future development

2.3 Validation and examples (part 1 of 3)
2.3.1 Chain-Ladder method on Payments and on Incurred
2.3.2 How to validate the Chain-Ladder assumptions

2.4 Ultimate uncertainty
2.4.1 Ultimate uncertainty of accident period
2.4.2 Ultimate uncertainty of the aggregation of all accident periods

2.5 Validation and examples (part 2 of 3)
2.5.1 Ultimate uncertainty

2.6 Solvency uncertainty
2.6.1 Solvency uncertainty of a single accident period
2.6.2 Solvency uncertainty of all accident periods
2.6.3 Uncertainty of further CDR’s

2.7 Validation and examples (part 3 of 3)
2.7.1 Solvency uncertainty

2.8 Literature
Basic idea behind the Chain-Ladder method

The Chain-Ladder method is based on a single triangle. Originally it was formulated in terms of the cumulative payments

\[ C_{i,k} := \sum_{j=0}^{k} S_{i,j} \]

instead of the payments \( S_{i,k} \) during the development period \( k \).

The Chain-Ladder method is based on the idea that:

- cumulative payments of the next development period are approximately proportional to the cumulative payments of the current period, i.e.
  \[ C_{i,k+1} \approx f_k C_{i,k}; \text{ and} \]

- accident period are independent.

In particular that means that all accident periods are comparable with respect to their development.
Basic idea behind the Chain-Ladder method

The Chain-Ladder method is based on a single triangle. Originally it was formulated in terms of the cumulative payments $C_{i,k}$ instead of the payments $S_{i,k}$ during the development period $k$.

The Chain-Ladder method is based on the idea that:

• cumulative payments of the next development period are approximately proportional to the cumulative payments of the current period, i.e.
  
  $C_{i,k+1} \approx f(k) C_{i,k}$;

• accident period are independent.

In particular that means that all accident periods are comparable with respect to their development.

2019-02-18

Stochastic Reserving

Chain-Ladder-Method (CLM)

How does the Chain-Ladder method work?
Simple example

<table>
<thead>
<tr>
<th>i \ k</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>ultimate</th>
<th>reserves</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>100</td>
<td>190</td>
<td>304</td>
<td>380</td>
<td>380</td>
<td>380</td>
<td>0 = 380 - 380</td>
</tr>
<tr>
<td>1</td>
<td>120</td>
<td>265</td>
<td>424</td>
<td>530</td>
<td>530</td>
<td>530</td>
<td>0 = 530 - 530</td>
</tr>
<tr>
<td>2</td>
<td>200</td>
<td>405</td>
<td>648</td>
<td>810</td>
<td>810</td>
<td>810</td>
<td>162 = 810 - 648</td>
</tr>
<tr>
<td>3</td>
<td>150</td>
<td>280</td>
<td>448</td>
<td>560</td>
<td>560</td>
<td>560</td>
<td>280 = 560 - 280</td>
</tr>
<tr>
<td>4</td>
<td>200</td>
<td>400</td>
<td>640</td>
<td>800</td>
<td>800</td>
<td>800</td>
<td>600 = 800 - 200</td>
</tr>
</tbody>
</table>

\[ \hat{f}_{0} = \frac{190 + 265 + 405 + 280}{100 + 120 + 200 + 150} = 2.0 \]
\[ \hat{f}_{1} = \frac{304 + 424 + 648}{190 + 265 + 405} = 1.6 \]
\[ \hat{f}_{2} = \frac{380 + 530}{304 + 424} = 1.2 \]
\[ \hat{f}_{3} = \frac{380}{380} = 1.0 \]
How does the Chain-Ladder method work
Definition 2.1 (σ-algebras)

- $\mathcal{B}_{i,k}$ is the σ-algebra of all information of accident period $i$ up to development period $k$:
  $$\mathcal{B}_{i,k} := \sigma (S_{i,j} : 0 \leq j \leq k) = \sigma (C_{i,j} : 0 \leq j \leq k)$$

- $\mathcal{D}_{i,k}$ is the σ-algebra containing all information up to accident period $i$ and development period $k$:
  $$\mathcal{D}_{i,k} := \sigma (S_{i,j} : 0 \leq h \leq i, 0 \leq j \leq k) = \sigma (B_{h,k} : 0 \leq h \leq i)$$

- $\mathcal{D}^{n}$ is the σ-algebra of all information up to calendar period $n$:
  $$\mathcal{D}^{n} := \sigma (S_{i,k} : 0 \leq i \leq I, 0 \leq k \leq J \land (n - i)) = \sigma (C_{i,k} : 0 \leq i \leq I, 0 \leq k \leq J \land (n - i))$$
  $$= \sigma \left( \bigcup_{i=0}^{I} \bigcup_{k=0}^{J \land (n-i)} \mathcal{B}_{i,k} \right)$$

- $\mathcal{D}_{k}$ is the σ-algebra of all information up to development period $k$:
  $$\mathcal{D}_{k} := \sigma (S_{i,j} : 0 \leq i \leq I, 0 \leq j \leq k)$$
  $$= \sigma (C_{i,j} : 0 \leq i \leq I, 0 \leq j \leq k)$$
  $$= \sigma \left( \bigcup_{i=0}^{I} \mathcal{B}_{i,k} \right)$$

- $\mathcal{D}_{k}^{n} := \sigma (\mathcal{D}^{n} \cup \mathcal{D}_{k})$
The σ-algebra $\mathcal{D}_n^k$ is used in order to enable us to separate two arbitrary payments $S_{i_1,k_1}$ and $S_{i_2,k_2}$ with $(i_1, k_1) \neq (i_2, k_2)$. That means, for all $(i_1, k_1) \neq (i_2, k_2)$ there exists $n$ and $k$ such that

$$S_{i_1,k_1} \in \mathcal{D}_n^k \quad \text{and} \quad S_{i_2,k_2} \notin \mathcal{D}_n^k.$$
Assumption 2.A (Mack’s Chain-Ladder method)

There exist development factors $f_k$ and variance parameters $\sigma_k^2$ such that the cumulative payments

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

satisfy

\[ \begin{align*}
\text{i)}^{\text{CLM}} & \quad \mathbb{E} \left[ C_{i,k+1} \big| B_{i,k} \right] = f_k C_{i,k}, \\
\text{ii)}^{\text{CLM}} & \quad \text{Var} \left[ C_{i,k+1} \big| B_{i,k} \right] = \sigma_k^2 C_{i,k} \quad \text{and} \\
\text{iii)}^{\text{CLM}} & \quad \text{accident periods are independent.}
\end{align*} \]
Assumption 2.A (Mack’s Chain-Ladder method)

There exist development factors $f_k$ and variance parameters $\sigma_k^2$ such that the cumulative payments $C_{i,k}$ satisfy

1. $\mathbb{E}[C_{i,k+1} | D_{i+k}] = f_k C_{i,k}$
2. $\text{Var}[C_{i,k+1} | D_{i+k}] = \sigma_k^2 C_{i,k}$
3. Accident periods are independent.

How does the Chain-Ladder method work

If $B_{i,k}$ are replaced by $D_{i+k}$, then the last assumption about independence is not necessary, i.e. it is enough to assume

1. $\mathbb{E}[C_{i,k+1} | D_{i+k}] = f_k C_{i,k}$
2. $\text{Var}[C_{i,k+1} | D_{i+k}] = \sigma_k^2 C_{i,k}$

We will see later that we can replace the exposure $C_{i,k}$ on the right side by more arbitrary exposures, which will leads to a wide class of reserving methods, called Linear Stochastic Reserving methods (LSRMs), see section 4.
Remark 2.2

- Since accident periods are independent, $B_{i,k}$ could be replaced by $D_k$, $D_{i,k}$, or $D_{i+k}^k$.
- Published by Thomas Mack in 1991, see [21]. But other actuaries have used at least parts of the stochastic model before. The reserving method itself is much older.
- From a statistical point of view the estimation of development factors and variance parameters is critical, because we have to estimate $2J$ parameters by only $J(I - \frac{J-1}{2})$ observed development factors. Therefore, in practise the reserving actuary has to include other information in order to overcome the lack of observed data (over parametrised model).
- The method cannot deal with incomplete triangles, where payments for early calendar periods are missing and therefore the cumulative payments for early accident periods are not complete (usually too small).
- There are other stochastic models that lead to the same estimates of the reserves. For instance, the over-dispersed Poisson model, see [10].
Stochastic Reserving

Remark 2.2

• Since accident periods are independent, $B_{i,k}$ could be replaced by $D_k$, $D_{i,k}$ or $D_{i+k}$.

• Published by Thomas Mack in 1991, see [21]. But other actuaries have used at least parts of the stochastic model before. The reserving method itself is much older.

• From a statistical point of view the estimation of development factors and variance parameters is critical, because we have to estimate $J^2$ parameters by only $J(J+1)/2$ observed development factors. Therefore, in practice the reserving actuary has to include other information in order to overcome the lack of observed data (over-parametrised model).

• The method cannot deal with incomplete triangles, where payments for early calendar periods are missing and therefore the cumulative payments for early accident periods are not complete (usually too small).

• There are other stochastic models that lead to the same estimates of the reserves. For instance, the over-dispersed Poisson model, see [10].
Corollary 2.3

- The parts \( i^{\text{CLM}} \) and \( ii^{\text{CLM}} \) of Assumption 2.A can be rewritten in terms of the incremental payments \( S_{i,k} \):
  
  \( i')^{\text{CL}} \quad \mathbb{E}[S_{i,k+1} | B_{i,k}] = (f_k - 1) C_{i,k} \) and
  
  \( ii')^{\text{CL}} \quad \text{Var}[S_{i,k+1} | B_{i,k}] = \sigma_k^2 C_{i,k}. \)

Therefore, Assumption 2.A means that under the knowledge of \( B_{i,k} \) the cumulative payments \( C_{i,k} \) are a good exposure for next periods payments \( S_{i,k+1} \).

- Iterating part \( i^{\text{CLM}} \) of Assumption 2.A we get

\[
\mathbb{E}[C_{i,k+n} | B_{i,k}] = \mathbb{E}[\mathbb{E}[C_{i,k+n} | B_{i,k+n-1}] | B_{i,k}] \\
= f_{k+n-1} \mathbb{E}[C_{i,k+n-1} | B_{i,k}] \\
= \ldots \\
= f_{k+n-1} \cdot \ldots \cdot f_k C_{i,k}.
\]
Corollary 2.3
• The parts i)\textsubscript{CLM} and ii)\textsubscript{CLM} of Assumption 2.A can be rewritten in terms of the incremental payments $S_{i,k}$:
  \[ E[S_{i,k+1}|B_{i,k}] = (f_{k}-1)C_{i,k} \]
  \[ \text{Var}[S_{i,k+1}|B_{i,k}] = \sigma^{2}kC_{i,k} \]

Therefore, Assumption 2.A means that under the knowledge of $B_{i,k}$ the cumulative payments $C_{i,k}$ are a good exposure for next periods payments $S_{i,k+1}$.

• Iterating part i)\textsubscript{CLM} of Assumption 2.A we get

Proof of i')\textsubscript{CLM}:

\[
E[S_{i,k+1}|B_{i,k}] = E[C_{i,k+1} - C_{i,k}|B_{i,k}]
\]
\[
= E[C_{i,k+1}|B_{i,k}] - C_{i,k}
\]
\[
= f_{k}C_{i,k} - C_{i,k}
\]

Proof of ii')\textsubscript{CLM}:

\[
\text{Var}[S_{i,k+1}|B_{i,k}] = \text{Var}[C_{i,k+1} - C_{i,k}|B_{i,k}]
\]
\[
= \text{Var}[C_{i,k+1}|B_{i,k}]
\]
\[
C_{i,k} \text{ is } B_{i,k} \text{ measurable}
\]
Lemma 2.4 (Chain-Ladder development factors)

Let Assumption 2.A be fulfilled and take arbitrary $D^I \cap D_k$-measurable weights $0 \leq w_{i,k} \leq 1$ with

• $w_{i,k} = 0$ if $C_{i,k} = 0$ and
• $\sum_{i=0}^{I-1-k} w_{i,k} = 1$ if $C_{i,k} \neq 0$ for at least one $0 \leq i \leq I - 1 - k$.

Then:

1. The weighted means

\[
\hat{f}_k := \sum_{i=0}^{I-1-k} w_{i,k} \frac{C_{i,k+1}}{C_{i,k}}
\]  

(2.1)

of the observed development factors $\frac{C_{i,k+1}}{C_{i,k}}$ are $D_k$-conditional unbiased estimators of the development factors $f_k$. In order to shorten notations we use here and in the following the definition $0 \div 0 := 0$.

Moreover, the weights

\[
w_{i,k} := \frac{C_{i,k}}{\sum_{h=0}^{I-1-k} C_{h,k}}
\]  

(2.2)

result in estimators $\hat{f}_k$ with the smallest ($D_k$-conditional) variance of all estimators of the form (2.1).

2. For all $k$ and all $k_n > k_{n-1} > \ldots > k_0 \geq 0$ we have

\[
\text{Var} \left[ \hat{f}_k \bigg| D_k \right] = \sum_{i=0}^{I-1-k} \frac{\sigma_k^2 w_{i,k}^2}{C_{i,k}} \quad \text{and} \quad E \left[ \hat{f}_{k_n} \hat{f}_{k_{n-1}} \ldots \hat{f}_{k_0} \bigg| D_{k_0} \right] = f_{k_n} f_{k_{n-1}} \ldots f_{k_0}.
\]
For all $\pi(k)$

The weighted means $\bar{x}$ unbiased:

$$\bar{x} = \sum_i w_i x_i$$

• minimal variance: $\text{Var} = \sum_i w_i (x_i - \bar{x})^2$

• uncorrelated: $\text{E} = \sum_i w_i (x_i - \bar{x})(x_j - \bar{x})$

### Stochastic Reserving

#### Chain-Ladder-Method (CLM)

**Future development**

- unbiased:
  
  $$\text{E} [\hat{f}_k | D_k] = \text{E} \left[ \sum_{i=0}^{I-1-k} w_{i,k} \frac{C_{i,k+1}}{C_{i,k}} | D_k \right] = \sum_{i=0}^{I-1-k} w_{i,k} \frac{\text{E} [C_{i,k+1} | D_k]}{C_{i,k}} = \sum_{i=0}^{I-1-k} w_{i,k} \frac{\hat{f}_k C_{i,k}}{C_{i,k}} = f_k$$

  measurable with respect to $D_k$

- minimal variance:
  
  $$\text{Var} [\hat{f}_k | D_k] = \text{Var} [\hat{f}_k | D_k] + \text{Var} [\text{E} [\hat{f}_k | D_k]] = \text{E} [\text{Var} [\hat{f}_k | D_k]] + 0$$

  $$\text{Var} [\hat{f}_k | D_k] = \text{Var} \left[ \sum_{i=0}^{I-1-k} w_{i,k} \frac{C_{i,k+1}}{C_{i,k}} | D_k \right] = \sum_{i=0}^{I-1-k} w_{i,k}^2 \frac{\text{Var} [C_{i,k+1} | D_k]}{C_{i,k}^2} = \sigma_k^2 \sum_{i=0}^{I-1-k} w_{i,k}^2 \frac{1}{C_{i,k}}$$

  measurable with respect to $D_k$ and iii) CLM

  $$\text{CLM}$$

**Lagrange:** minimize

$$\sum_{i=0}^{I-1-k} w_{i,k}^2 \frac{1}{C_{i,k}} - \lambda (1 - \sum_{i=0}^{I-1-k} w_{i,k})$$

$$w_{i,k} = \frac{\lambda}{2} C_{i,k}$$

and

$$\lambda = \frac{2}{\sum_{i=0}^{I-1-k} C_{i,k}^2}$$

$$w_{i,k} = \frac{C_{i,k}}{\sum_{i=0}^{I-1-k} C_{i,k}}$$

$$\sum_{i=0}^{I-1-k} w_{i,k} = 1$$

- uncorrelated:
  
  $$\text{E} [\hat{f}_{k_n} \hat{f}_{k_{n-1}} \ldots \hat{f}_0 | D_{k_0}] = \text{E} [\text{E} [\hat{f}_{k_n} | D_{k_n}] \hat{f}_{k_{n-1}} \ldots \hat{f}_0 | D_{k_0}]$$

  $$= f_{k_n} \text{E} [\hat{f}_{k_{n-1}} \ldots \hat{f}_0 | D_k] = \ldots = f_{k_n} f_{k_{n-1}} \ldots f_0$$

**Lemma 2.4 (Chain-Ladder development factors)**

Let Assumption 2.1 be fulfilled and take arbitrary $D_k$, measurable weights $0 \leq w_{i,k} \leq 1$ with

- $w_{i,k} = 0$ if $i < k$
- $w_{i,k} = 1$ if $i \geq k$

Then:

1. The weighted means $\hat{f}_k = \sum_{i=0}^{I-1-k} w_{i,k} f_i$ of the observed development factors $f_i$ are $D_k$-measurable $\hat{f}_k$, \(\hat{f}_kn\)-conditional unbiased estimators of the development factors $f_i$. In order to shorten notation we use here and in the following the definition $\hat{f}_k := \hat{f}_kn$.

2. Moreover, the weights $w_{i,k}$ are measurable with respect to $D_k$.

3. For all $k$ and all $k_n \geq k_{n-1} \geq k_0 \geq 0$ we have

$$\text{Var} [\hat{f}_{k_n} | D_{k_n}] = \sum_{i=0}^{I-1-k_n} w_{i,k_n}^2 \frac{1}{C_{i,k_n}}$$

$$\text{Var} [\hat{f}_{k_n} | D_{k_n}] = \sum_{i=0}^{I-1-k_n} w_{i,k_n}^2 \frac{1}{C_{i,k_n}}$$

and $\text{E} [\hat{f}_{k_n}, \hat{f}_{k_{n-1}}, \ldots, \hat{f}_0] = \hat{f}_{k_n}, \hat{f}_{k_{n-1}}, \ldots, \hat{f}_0$.
Estimator 2.5 (Chain-Ladder Ultimate)

Let Assumption 2.A be fulfilled. Then the estimates

\[ \hat{C}_{i,k} := \hat{f}_{k-1} \cdot \cdots \cdot \hat{f}_{I-i} C_{i,I-k} \]

are \( D_{I-i} \)-conditional unbiased estimators of \( C_{i,k} \), for \( I - i < k \leq J \).

In order to shorten notation we define

\[ \hat{C}_{i,k} := C_{i,k}, \]

for \( 0 \leq k \leq I - i \).

Theorem 2.6 (Chain-Ladder Best Estimate)

The Estimator 2.5 with the variance minimizing weights (2.2) satisfies the condition of a Best Estimate, i.e.

\[ \hat{E} \left[ \hat{C}_{i,J+1}^I - \hat{C}_{i,J}^I \bigg| D^I \right] = 0, \]

where the additional upper index specifies the time of estimation.
• Proof of unbiasedness:

\[
E[\hat{C}_{i,k} | D_{I-i}] = E[\hat{f}_{k-1} \cdots \hat{f}_{I-i} C_{i,I-i} | D_{I-i}] = E[E[\hat{f}_{k-1} | D_{k-1}] \hat{f}_{k-2} \cdots \hat{f}_{I-i} C_{i,I-i} | D_{I-i}]
\]

\[
= E[f_{k-1} \hat{f}_{k-2} \cdots \hat{f}_{I-i} C_{i,I-i} | D_{I-i}] = \cdots = f_{k-1} \cdots f_{I-i} C_{i,I-i}
\]

\[
= E[C_{i,k} | D_{I-i}]
\]

• Best Estimate: \[
\hat{f}_{k+1}^{I+1} := \sum_{i=0}^{I-k} C_{i,k} \frac{C_{i,k+1}}{\sum_{h=0}^{I-k} C_{h,k}} = \left(1 - \frac{C_{I-k,k}}{\sum_{h=0}^{I-k} C_{h,k}}\right) \hat{f}_I + \frac{C_{I-k,k}}{\sum_{h=0}^{I-k} C_{h,k}} f_k := \bar{f}_I
\]

\[
\Rightarrow E[\hat{f}_{k+1}^{I+1} | D_k] = \left(1 - \frac{C_{I-k,k}}{\sum_{h=0}^{I-k} C_{h,k}}\right) \hat{f}_I + \frac{C_{I-k,k}}{\sum_{h=0}^{I-k} C_{h,k}} f_k =: \bar{f}_I
\]

\[
\Rightarrow \hat{E}[\hat{f}_{k+1}^{I+1} | D_k] = \bar{f}_I
\]

\[
E[\hat{C}_{I,J}+1 | D^I] = E[\hat{f}_{J-1}^{I+1} \cdots \hat{f}_{I+1-I}^{I+1} C_{i,I+1-I} | D^I]
\]

\[
= E[E[\hat{f}_{J-1}^{I+1} | D_{J-1}] \hat{f}_{I+1-I}^{I+1} \cdots \hat{f}_{I+1-I}^{I+1} C_{i,I+1-I} | D^I]
\]

\[
= E[\hat{f}_{J-1} \hat{f}_{J-2} \cdots \hat{f}_{I+1-I} C_{i,I+1-I} | D^I]
\]

\[
= \cdots = \hat{f}_{J-1} \cdots \hat{f}_{I+1-I} E[C_{i,I+1-I} | D^I] = \hat{f}_{J-1} \cdots \hat{f}_{I+1-I} f_{I-I} C_{i,I-i}
\]

\[
\Rightarrow \hat{E}[\hat{C}_{i,J}+1 | D^I] = \hat{f}_{J-1} \cdots \hat{f}_{I-i} C_{i,I-i} = \hat{C}_{i,J} = \hat{E}[\hat{C}_{i,J} | D^I].
\]
Chain-Ladder method in practice

- The Chain-Ladder method is probably the most popular reserving method in general insurance and usually works fine for most of the standard business, provided we take care of:
  - The size of the portfolio (has to be large enough to get the law of large numbers working).
  - The homogeneity of the portfolio (for example exclude extraordinary large or late claims).

- But it has problems with:
  - Inflation or other diagonal effects, because such effects contradict the assumption of independent accident periods.
  - Too large or too small values at the last (known) diagonal. Because the values of the last diagonal are realisations of random variables, this may even happen if the portfolio satisfies Assumption 2.A perfectly.
The Chain-Ladder method is probably the most popular reserving method in general insurance and usually works fine for most of the standard business, provided we take care of:

* The size of the portfolio (has to be large enough to get the law of large numbers working).
* The homogeneity of the portfolio (for example exclude extraordinary large or late claims).

But it has problems with:

* Inflation or other diagonal effects, because such effects contradict the assumption of independent accident periods.
* Too large or too small values at the last (known) diagonal. Because the values of the last diagonal are realisations of random variables, this may even happen if the portfolio satisfies Assumption 2A perfectly.
Example 2.7 (Chain-Ladder method on payments)

- We took the variance minimizing weights (2.2).
- For the calculation of the IBNR we used the corresponding incurred from Example 2.8.

### Payments

<table>
<thead>
<tr>
<th>AP/DP</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>Current</th>
<th>Ultimate</th>
<th>Reserves</th>
<th>IBN(e/y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1'216'632</td>
<td>1'347'072</td>
<td>1'786'877</td>
<td>2'281'606</td>
<td>2'656'224</td>
<td>2'909'307</td>
<td>3'283'388</td>
<td>3'587'549</td>
<td>3'754'403</td>
<td>3'921'258</td>
<td>3'921'258</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>798'924</td>
<td>1'051'912</td>
<td>1'215'785</td>
<td>1'349'939</td>
<td>1'655'312</td>
<td>1'926'210</td>
<td>2'132'833</td>
<td>2'287'311</td>
<td>2'567'056</td>
<td>2'681'142</td>
<td>2'681'142</td>
<td>114'086</td>
<td>-238'813</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1'115'636</td>
<td>1'387'387</td>
<td>1'930'867</td>
<td>2'177'002</td>
<td>2'513'171</td>
<td>2'931'930</td>
<td>3'047'368</td>
<td>3'182'511</td>
<td>3'424'441</td>
<td>3'576'632</td>
<td>3'576'632</td>
<td>318'050</td>
<td>394'121</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1'052'161</td>
<td>1'321'206</td>
<td>1'700'132</td>
<td>1'971'303</td>
<td>2'298'349</td>
<td>2'645'113</td>
<td>3'003'425</td>
<td>3'214'137</td>
<td>3'458'471</td>
<td>3'612'174</td>
<td>3'612'174</td>
<td>303'425</td>
<td>608'749</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>808'864</td>
<td>1'029'523</td>
<td>1'229'626</td>
<td>1'590'338</td>
<td>1'842'662</td>
<td>2'150'351</td>
<td>2'368'112</td>
<td>2'534'252</td>
<td>2'726'902</td>
<td>2'848'093</td>
<td>2'848'093</td>
<td>215'035</td>
<td>284'893</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1'016'862</td>
<td>1'251'420</td>
<td>1'608'052</td>
<td>2'105'143</td>
<td>2'385'339</td>
<td>2'732'771</td>
<td>3'009'512</td>
<td>3'220'652</td>
<td>3'465'481</td>
<td>3'619'496</td>
<td>3'619'496</td>
<td>238'539</td>
<td>361'949</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>948'312</td>
<td>1'108'791</td>
<td>1'315'524</td>
<td>1'487'577</td>
<td>1'730'732</td>
<td>1'982'819</td>
<td>2'183'614</td>
<td>2'336'811</td>
<td>2'514'452</td>
<td>2'626'200</td>
<td>2'626'200</td>
<td>148'7577</td>
<td>262'6200</td>
<td>1'138'623</td>
</tr>
<tr>
<td>7</td>
<td>917'530</td>
<td>1'082'426</td>
<td>1'484'405</td>
<td>1'769'095</td>
<td>2'058'267</td>
<td>2'358'060</td>
<td>2'596'855</td>
<td>2'779'043</td>
<td>2'990'302</td>
<td>3'123'198</td>
<td>3'123'198</td>
<td>148'405</td>
<td>3'123'198</td>
<td>163'8793</td>
</tr>
<tr>
<td>8</td>
<td>1'001'238</td>
<td>1'376'124</td>
<td>1'775'689</td>
<td>2'116'244</td>
<td>2'462'160</td>
<td>2'820'781</td>
<td>3'106'435</td>
<td>3'324'374</td>
<td>3'577'088</td>
<td>3'736'063</td>
<td>3'736'063</td>
<td>137'6124</td>
<td>3'736'063</td>
<td>2'359'939</td>
</tr>
<tr>
<td>9</td>
<td>841'930</td>
<td>1'039'196</td>
<td>1'340'932</td>
<td>1'598'106</td>
<td>1'859'328</td>
<td>2'130'146</td>
<td>2'345'860</td>
<td>2'510'439</td>
<td>2'701'279</td>
<td>2'821'331</td>
<td>2'821'331</td>
<td>841'930</td>
<td>2'821'331</td>
<td>1'979'401</td>
</tr>
</tbody>
</table>

| Total   | 223'999'76 | 32'565'588 | 10'165'612 | -3'239'141 |

### Observed development factors (ratios)

<table>
<thead>
<tr>
<th>AP/DP</th>
<th>0-&gt;1</th>
<th>1-&gt;2</th>
<th>2-&gt;3</th>
<th>3-&gt;4</th>
<th>4-&gt;5</th>
<th>5-&gt;6</th>
<th>6-&gt;7</th>
<th>7-&gt;8</th>
<th>8-&gt;9</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.10721</td>
<td>1.32649</td>
<td>1.27687</td>
<td>1.16419</td>
<td>1.09528</td>
<td>1.12858</td>
<td>1.09264</td>
<td>1.04651</td>
<td>1.04444</td>
</tr>
<tr>
<td>1</td>
<td>1.31666</td>
<td>1.15579</td>
<td>1.11034</td>
<td>1.22621</td>
<td>1.16365</td>
<td>1.10727</td>
<td>1.07243</td>
<td>1.12230</td>
<td>1.13456</td>
</tr>
<tr>
<td>2</td>
<td>1.24358</td>
<td>1.39173</td>
<td>1.13774</td>
<td>1.15442</td>
<td>1.16663</td>
<td>1.03937</td>
<td>1.04435</td>
<td>1.04435</td>
<td>1.04435</td>
</tr>
<tr>
<td>3</td>
<td>1.25571</td>
<td>1.28680</td>
<td>1.15950</td>
<td>1.16590</td>
<td>1.15088</td>
<td>1.13546</td>
<td>1.13456</td>
<td>1.13456</td>
<td>1.13456</td>
</tr>
<tr>
<td>4</td>
<td>1.27280</td>
<td>1.19436</td>
<td>1.29335</td>
<td>1.15866</td>
<td>1.16686</td>
<td>1.13310</td>
<td>1.13310</td>
<td>1.13310</td>
<td>1.13310</td>
</tr>
<tr>
<td>5</td>
<td>1.23067</td>
<td>1.35690</td>
<td>1.23974</td>
<td>1.13310</td>
<td>1.13079</td>
<td>1.13079</td>
<td>1.13079</td>
<td>1.13079</td>
<td>1.13079</td>
</tr>
<tr>
<td>6</td>
<td>1.16923</td>
<td>1.18645</td>
<td>1.3079</td>
<td>1.13079</td>
<td>1.13079</td>
<td>1.13079</td>
<td>1.13079</td>
<td>1.13079</td>
<td>1.13079</td>
</tr>
<tr>
<td>7</td>
<td>1.17972</td>
<td>1.37137</td>
<td>1.13742</td>
<td>1.13742</td>
<td>1.13742</td>
<td>1.13742</td>
<td>1.13742</td>
<td>1.13742</td>
<td>1.13742</td>
</tr>
<tr>
<td>8</td>
<td>1.37442</td>
<td>1.37442</td>
<td>1.37442</td>
<td>1.37442</td>
<td>1.37442</td>
<td>1.37442</td>
<td>1.37442</td>
<td>1.37442</td>
<td>1.37442</td>
</tr>
</tbody>
</table>

### Estimated development factors

<table>
<thead>
<tr>
<th>cum.</th>
<th>3.351028</th>
<th>2.714917</th>
<th>2.104007</th>
<th>1.765421</th>
<th>1.517393</th>
<th>1.324478</th>
<th>1.202685</th>
<th>1.12384</th>
<th>1.04444</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-&gt;1</td>
<td>1.23430</td>
<td>1.29036</td>
<td>1.19179</td>
<td>1.16346</td>
<td>1.14565</td>
<td>1.10127</td>
<td>1.07016</td>
<td>1.07602</td>
<td>1.04444</td>
</tr>
</tbody>
</table>
Example 2.7 (Chain-Ladder method on payments)

- We took the variance minimizing weights (2.3).
- For the calculation of the IBNR we used the corresponding incurred from Example 2.8.
Example 2.8 (Chain-Ladder method on incurred losses)

- We took the variance minimizing weights (2.2).
- For the calculation of the reserves we used the corresponding payments from Example 2.7.

<table>
<thead>
<tr>
<th>Incurred</th>
<th>AP\DP</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>Current</th>
<th>Ultimate</th>
<th>Reserves</th>
<th>IBN(e/y)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>3'362'115</td>
<td>5'217'243</td>
<td>4'754'900</td>
<td>4'381'677</td>
<td>4'136'883</td>
<td>4'094'140</td>
<td>4'018'736</td>
<td>3'971'591</td>
<td>3'941'391</td>
<td>3'921'258</td>
<td>3921258</td>
<td>3'921'258</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2'640'443</td>
<td>4'643'860</td>
<td>3'869'954</td>
<td>3'248'558</td>
<td>3'102'002</td>
<td>3'019'980</td>
<td>2'976'064</td>
<td>2'946'941</td>
<td>2'919'955</td>
<td>2'905'040</td>
<td>2919955</td>
<td>2'905'040</td>
<td>337'984</td>
<td>-14'915</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2'879'697</td>
<td>4'785'531</td>
<td>4'045'448</td>
<td>3'467'322</td>
<td>3'377'340</td>
<td>3'341'934</td>
<td>3'238'928</td>
<td>3'257'827</td>
<td>3'230'899</td>
<td>3'214'395</td>
<td>3257827</td>
<td>3'214'395</td>
<td>31'884</td>
<td>-43'432</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2'933'345</td>
<td>5'299'146</td>
<td>4'451'963</td>
<td>3'700'809</td>
<td>3'553'391</td>
<td>3'469'505</td>
<td>3'413'921</td>
<td>3'379'921</td>
<td>3'351'984</td>
<td>3'334'861</td>
<td>3413921</td>
<td>3'334'861</td>
<td>331'436</td>
<td>-79'060</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2'768'181</td>
<td>4'658'833</td>
<td>3'936'455</td>
<td>3'512'735</td>
<td>3'385'129</td>
<td>3'288'998</td>
<td>3'243'821</td>
<td>3'211'515</td>
<td>3'184'970</td>
<td>3'168'701</td>
<td>3298998</td>
<td>3'168'701</td>
<td>1'018'350</td>
<td>-130'297</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3'228'439</td>
<td>5'271'304</td>
<td>4'484'946</td>
<td>3'798'384</td>
<td>3'702'427</td>
<td>3'632'746</td>
<td>3'571'987</td>
<td>3'536'413</td>
<td>3'507'182</td>
<td>3'489'267</td>
<td>3702427</td>
<td>3'489'267</td>
<td>1'103'928</td>
<td>-213'160</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2'927'033</td>
<td>5'067'768</td>
<td>4'066'526</td>
<td>3'704'113</td>
<td>3'561'274</td>
<td>3'494'250</td>
<td>3'435'807</td>
<td>3'401'589</td>
<td>3'373'473</td>
<td>3'356'241</td>
<td>3704113</td>
<td>3'356'241</td>
<td>1'868'664</td>
<td>-347'872</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3'038'329</td>
<td>4'790'944</td>
<td>4'058'097</td>
<td>3'842'969</td>
<td>3'694'775</td>
<td>3'625'239</td>
<td>3'564'605</td>
<td>3'529'104</td>
<td>3'499'934</td>
<td>3'482'056</td>
<td>4408097</td>
<td>3'482'056</td>
<td>1'997'651</td>
<td>-926'041</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2'761'163</td>
<td>4'132'757</td>
<td>3'538'198</td>
<td>3'084'593</td>
<td>2'965'643</td>
<td>2'909'829</td>
<td>2'861'611</td>
<td>2'832'666</td>
<td>2'809'252</td>
<td>2'794'903</td>
<td>4132757</td>
<td>2'794'903</td>
<td>1'418'779</td>
<td>-1'337'854</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3'045'376</td>
<td>5'025'345</td>
<td>4'302'373</td>
<td>3'750'799</td>
<td>3'606'160</td>
<td>3'538'291</td>
<td>3'479'112</td>
<td>3'444'462</td>
<td>3'415'991</td>
<td>3'398'542</td>
<td>3045376</td>
<td>3'398'542</td>
<td>2'556'612</td>
<td>353'166</td>
</tr>
</tbody>
</table>

Total 35804729 33'065'263 10'665'287 -2'739'466

<table>
<thead>
<tr>
<th>Observed development factors (ratios)</th>
<th>AP\DP</th>
<th>0-1</th>
<th>1-2&gt;</th>
<th>2-3&gt;</th>
<th>3-4&gt;</th>
<th>4-5&gt;</th>
<th>5-6&gt;</th>
<th>6-7&gt;</th>
<th>7-8&gt;</th>
<th>8-9&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.55177</td>
<td>0.91138</td>
<td>0.92151</td>
<td>0.94413</td>
<td>0.98967</td>
<td>0.98158</td>
<td>0.98827</td>
<td>0.99240</td>
<td>0.99489</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1.75874</td>
<td>0.83335</td>
<td>0.83943</td>
<td>0.95489</td>
<td>0.97356</td>
<td>0.98546</td>
<td>0.99021</td>
<td>0.99084</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1.66182</td>
<td>0.84535</td>
<td>0.85722</td>
<td>0.97397</td>
<td>0.98946</td>
<td>0.98264</td>
<td>0.99205</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1.80652</td>
<td>0.84013</td>
<td>0.83128</td>
<td>0.96017</td>
<td>0.97639</td>
<td>0.98398</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1.68303</td>
<td>0.84493</td>
<td>0.89236</td>
<td>0.96367</td>
<td>0.97456</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1.63277</td>
<td>0.85082</td>
<td>0.84692</td>
<td>0.97474</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1.73137</td>
<td>0.80243</td>
<td>0.91088</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>1.55377</td>
<td>0.92009</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>1.49675</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Estimated development factors</th>
<th>0-1</th>
<th>1-2&gt;</th>
<th>2-3&gt;</th>
<th>3-4&gt;</th>
<th>4-5&gt;</th>
<th>5-6&gt;</th>
<th>6-7&gt;</th>
<th>7-8&gt;</th>
<th>8-9&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.16016</td>
<td>0.85613</td>
<td>0.87180</td>
<td>0.96144</td>
<td>0.98118</td>
<td>0.98327</td>
<td>0.99004</td>
<td>0.99173</td>
<td>0.99489</td>
<td></td>
</tr>
<tr>
<td>1.115968</td>
<td>0.67628</td>
<td>0.789923</td>
<td>0.906085</td>
<td>0.94247</td>
<td>0.960504</td>
<td>0.976842</td>
<td>0.986669</td>
<td>0.99489</td>
<td></td>
</tr>
</tbody>
</table>
Example 2.8 (Chain-Ladder method on incurred losses)

- We took the variance minimizing weights (2.3).
- For the calculation of the reserves we used the corresponding payments from Example 2.7.
Comparison of the two results

- Both, payments and incurred losses, will eventually result in the same ultimate. But the estimates are not the same! This gap is a systematic problem of projecting payments and incurred losses independently of each other. For more information see [7].

- Although in total the difference is only 5% we have much larger differences per accident period, which almost cancel each other.
Comparison of the two results

• Both, payments and incurred losses, will eventually result in the same ultimate. But the estimates are not the same! This gap is a systematic problem of projecting payments and incurred losses independently of each other. For more information see [7].

• Although in total the difference is only 5% we have much larger differences per accident period, which almost cancel each other.
2.3.1 Chain-Ladder method on Payments and on Incurred

Payments and Incurred inclusive corresponding projection

- 0 Payments
- 1 Payments
- 2 Payments
- 3 Payments
- 4 Payments
- 5 Payments
- 6 Payments
- 7 Payments
- 8 Payments
- 9 Payments
- 0 Incurred
- 1 Incurred
- 2 Incurred
- 3 Incurred
- 4 Incurred
- 5 Incurred
- 6 Incurred
- 7 Incurred
- 8 Incurred
- 9 Incurred

Development Period
Stochastic Reserving
- Chain-Ladder-Method (CLM)
- Validation and examples (part 1 of 3)
Validation of Chain-Ladder Assumption 2.A

- Since we only have very few data, any statistical validation of Assumption 2.A will usually fail.
- There are some helpful statistics and graphical presentations that can be used to get a feeling about which estimate we should trust more. In the following slides we will show some of them.
- The most important information is the knowledge about the composition of the underlying portfolio and the corresponding risks. We usually face the problem of splitting up the portfolio in subportfolios, which are as homogeneous as possible, but are not too small in order to get the law of large numbers working. Typical criteria for separation are:
  * Type of the risk insured.
  * Type of claims, like property damage or bodily injury.
  * Type of payments, like lump sums, annuities, salvage and subrogation or deductibles.
  * Type of case reserves, like automatically generated, set individually by a normal claims manager or set individually by an expert.
  * Complexity of the claims, often the size of the claim may be a criteria for its complexity.
  * ... 
- Finally, actuaries have to use other information, too, in order to determine their ultimates.
Since we only have very few data, any statistical validation of Assumption 2.A will usually fail.

There are some helpful statistics and graphical presentations that can be used to get a feeling about which estimate we should trust more. In the following slides we will show some of them.

The most important information is the knowledge about the composition of the underlying portfolio and the corresponding risks. We usually face the problem of splitting up the portfolio in subportfolios, which are as homogeneous as possible, but are not too small in order to get the law of large numbers working. Typical criteria for separation are:

- Type of the risk insured.
- Type of claims, like property damage or bodily injury.
- Type of payments, like lump sums, annuities, salvage and subrogation or deductibles.
- Type of case reserves, like automatically generated, set individually by a normal claims manager or set individually by an expert.
- Complexity of the claims, often the size of the claim may be a criteria for its complexity.

Finally, actuaries have to use other information, too, in order to determine their ultimates.
The projection of incurred is more stable and closer to the estimated ultimate than the projection of payments. This may be an indication to trust it more.
The projection of incurred is more stable and closer to the estimated ultimate than the projection of payments. This may be an indication to trust it more.
Plot of residuals

The residuals are defined by

\[ \frac{C_{i,k+1} - \hat{f}_k}{\sqrt{\text{Var} \left[ \frac{C_{i,k+1}}{C_{i,k}} | D_k \right]}} = \frac{C_{i,k+1} - \hat{f}_k}{\sqrt{\frac{\hat{\sigma}_k^2}{C_{i,k}}}}. \]

Payments

Incurred

The residual plots are very similar, except that the incurred residuals look a bit more symmetric.
Plot of residuals
The residuals are defined by
$C_{i,k+1} - C_{i,k} - \hat{f}_k \sqrt{\hat{\sigma}^2_k} = C_{i,k}$.

The residual plots are very similar, except that the incurred residuals look a bit more symmetric.
Backtesting step by step

Here we compare the observed values with the one step backwards projected estimate, i.e.

\[ C_{i,k} \quad \text{with} \quad \frac{C_{i,k+1}^{\hat{f}_k}}{f_k}. \]

### Payments

<table>
<thead>
<tr>
<th>Back testing (step by step)</th>
<th>AP/DD</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual</td>
<td>1'216'632</td>
<td>1'347'072</td>
<td>1'786'877</td>
<td>2'281'606</td>
<td>2'656'224</td>
<td>2'909'307</td>
<td>3'283'388</td>
<td>3'587'549</td>
<td>3'754'403</td>
<td>3'921'258</td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>1'370'166</td>
<td>1'448'338</td>
<td>1'857'710</td>
<td>2'322'146</td>
<td>2'784'208</td>
<td>3'060'500</td>
<td>3'260'420</td>
<td>3'489'362</td>
<td>3'754'403</td>
<td>3'921'258</td>
<td></td>
</tr>
<tr>
<td>1 Actual</td>
<td>789'924</td>
<td>1'051'912</td>
<td>1'215'785</td>
<td>1'349'339</td>
<td>1'655'112</td>
<td>1'926'210</td>
<td>2'132'833</td>
<td>2'287'311</td>
<td>2'567'056</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>800'095</td>
<td>867'559</td>
<td>1'274'033</td>
<td>1'518'488</td>
<td>1'766'940</td>
<td>2'024'301</td>
<td>2'229'297</td>
<td>2'385'698</td>
<td>2'567'056</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Actual</td>
<td>1'115'636</td>
<td>1'387'387</td>
<td>1'930'867</td>
<td>2'177'002</td>
<td>2'513'171</td>
<td>2'931'930</td>
<td>3'047'368</td>
<td>3'182'511</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>1'067'324</td>
<td>1'317'400</td>
<td>1'593'314</td>
<td>2'205'937</td>
<td>2'337'091</td>
<td>2'700'405</td>
<td>2'973'972</td>
<td>3'182'511</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Actual</td>
<td>1'052'161</td>
<td>1'322'206</td>
<td>1'700'132</td>
<td>1'971'303</td>
<td>2'298'149</td>
<td>2'645'133</td>
<td>3'003'415</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>1'077'980</td>
<td>1'330'492</td>
<td>1'716'607</td>
<td>2'246'960</td>
<td>2'380'514</td>
<td>2'727'244</td>
<td>3'003'425</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 Actual</td>
<td>808'864</td>
<td>1'029'523</td>
<td>1'229'626</td>
<td>1'590'338</td>
<td>1'842'662</td>
<td>2'150'351</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>849'916</td>
<td>1'049'059</td>
<td>1'333'584</td>
<td>1'913'265</td>
<td>2'176'963</td>
<td>2'519'351</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Actual</td>
<td>1'016'862</td>
<td>1'251'420</td>
<td>1'698'052</td>
<td>2'105'143</td>
<td>2'385'339</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>1'001'118</td>
<td>1'318'168</td>
<td>1'720'287</td>
<td>2'050'216</td>
<td>2'385'339</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 Actual</td>
<td>948'312</td>
<td>1'108'791</td>
<td>1'315'524</td>
<td>1'487'577</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>783'700</td>
<td>967'322</td>
<td>1'248'190</td>
<td>1'487'577</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 Actual</td>
<td>917'530</td>
<td>1'082'426</td>
<td>1'484'405</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>932'012</td>
<td>1'150'384</td>
<td>1'484'405</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 Actual</td>
<td>1'001'138</td>
<td>1'376'124</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>1'124'901</td>
<td>1'376'124</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9 Actual</td>
<td>841'930</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>841'930</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Incurred

<table>
<thead>
<tr>
<th>Back testing (step by step)</th>
<th>AP/DD</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual</td>
<td>3'962'119</td>
<td>4'217'243</td>
<td>4'754'900</td>
<td>4'381'677</td>
<td>4'136'883</td>
<td>4'094'140</td>
<td>4'018'736</td>
<td>3'971'591</td>
<td>3'941'391</td>
<td>3'921'258</td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>3'510'773</td>
<td>3'797'232</td>
<td>4'164'013</td>
<td>3'727'994</td>
<td>3'416'808</td>
<td>3'402'501</td>
<td>3'401'219</td>
<td>3'374'241</td>
<td>3'541'391</td>
<td>3'921'258</td>
<td></td>
</tr>
<tr>
<td>1 Actual</td>
<td>2'640'443</td>
<td>4'643'860</td>
<td>3'869'984</td>
<td>3'248'558</td>
<td>3'102'002</td>
<td>3'019'980</td>
<td>2'976'064</td>
<td>2'964'941</td>
<td>2'919'555</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>2'603'157</td>
<td>4'295'914</td>
<td>3'677'265</td>
<td>3'206'145</td>
<td>3'082'509</td>
<td>3'042'495</td>
<td>2'973'909</td>
<td>2'944'291</td>
<td>2'919'555</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Actual</td>
<td>2'879'667</td>
<td>4'785'531</td>
<td>4'045'448</td>
<td>3'467'822</td>
<td>3'377'540</td>
<td>3'341'934</td>
<td>3'285'928</td>
<td>3'257'827</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>2'880'766</td>
<td>4'753'051</td>
<td>4'069'233</td>
<td>3'547'865</td>
<td>3'410'763</td>
<td>3'346'572</td>
<td>3'290'599</td>
<td>3'257'827</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Actual</td>
<td>2'933'345</td>
<td>5'288'146</td>
<td>4'451'983</td>
<td>3'700'399</td>
<td>3'553'391</td>
<td>3'469'505</td>
<td>3'413'921</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>2'988'313</td>
<td>4'931'811</td>
<td>4'321'797</td>
<td>3'680'518</td>
<td>3'538'586</td>
<td>3'471'991</td>
<td>3'419'291</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 Actual</td>
<td>2'768'181</td>
<td>4'658'933</td>
<td>3'936'455</td>
<td>3'512'735</td>
<td>3'385'129</td>
<td>3'298'998</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>2'839'419</td>
<td>4'685'483</td>
<td>4'021'406</td>
<td>3'497'235</td>
<td>3'362'277</td>
<td>3'308'996</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Actual</td>
<td>3'228'439</td>
<td>5'271'304</td>
<td>4'484'946</td>
<td>3'798'384</td>
<td>3'702'427</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>3'129'579</td>
<td>5'169'486</td>
<td>4'417'226</td>
<td>3'950'928</td>
<td>3'702'427</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 Actual</td>
<td>2'927'033</td>
<td>5'067'768</td>
<td>4'066'526</td>
<td>3'704'113</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>3'007'470</td>
<td>4'962'794</td>
<td>4'248'822</td>
<td>3'704'113</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 Actual</td>
<td>3'083'429</td>
<td>4'790'944</td>
<td>4'048'097</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>3'120'211</td>
<td>5'148'834</td>
<td>4'408'097</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 Actual</td>
<td>2'761'169</td>
<td>4'132'757</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>2'504'465</td>
<td>4'132'757</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9 Actual</td>
<td>3'045'376</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>3'045'376</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Incurred seems to be a bit more stable, in particular for later development periods.
Stochastic Reserving

Chain-Ladder-Method (CLM)

Validation and examples (part 1 of 3)
2 Chain-Ladder-Method (CLM)

2.3 Validation and examples (part 1 of 3)

2.3.2 How to validate the Chain-Ladder assumptions (5/6)

Backtesting the ultimate

Here we compare the projected ultimate starting at development period $k$ with the one starting at development period $I - i$ (the estimated ultimate), i.e.

$$C_{i,k} \prod_{j=k}^{J-1} \hat{f}_j \quad \text{with} \quad C_{i,I-i} \prod_{j=I-i}^{J-1} \hat{f}_j = \hat{C}_{i,J}.$$ 

Payments

Incurred

Again, incurred seems to be a bit more stable, in particular for later development periods.
Backtesting the ultimate

Here we compare the projected ultimate starting at development period $i$ with the one starting at development period $i - 1$ (the estimated ultimate), i.e.

$$c_{i,k} \prod_{j=k}^{i} \hat{\delta}_j \quad \text{with} \quad c_{i,I-k} \prod_{j=I-k}^{I-1} \hat{\delta}_j = \hat{c}_{i,J}.$$

Payments

Incurred

Again, incurred seems to be a bit more stable, in particular for later development periods.
Sensitivity to exclusion or inclusion of individual observed development factors

Here we compare the projected ultimate based on the selected development factors with the projected ultimate if we exclude (or include) a observed development factor within the estimation of $\hat{f}_k$.

Payments

<table>
<thead>
<tr>
<th>AP/DP</th>
<th>0 Ratio</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 Ratio</td>
<td>1.10721</td>
<td>1.32649</td>
<td>1.27647</td>
<td>1.16419</td>
<td>1.09528</td>
<td>1.22858</td>
<td>1.09264</td>
<td>1.04651</td>
<td>1.04444</td>
</tr>
<tr>
<td>Change of reserves</td>
<td>0.45%</td>
<td>-0.03%</td>
<td>-1.34%</td>
<td>-0.22%</td>
<td>2.20%</td>
<td>-1.78%</td>
<td>-2.93%</td>
<td>10.99%</td>
<td>-11.99%</td>
</tr>
<tr>
<td>1 Ratio</td>
<td>1.31666</td>
<td>1.15579</td>
<td>1.11034</td>
<td>1.22621</td>
<td>1.16365</td>
<td>1.10727</td>
<td>1.07243</td>
<td>1.12230</td>
<td>1.04435</td>
</tr>
<tr>
<td>Change of reserves</td>
<td>-0.18%</td>
<td>0.83%</td>
<td>0.82%</td>
<td>-0.87%</td>
<td>-0.44%</td>
<td>-0.25%</td>
<td>-0.16%</td>
<td>-7.00%</td>
<td></td>
</tr>
<tr>
<td>2 Ratio</td>
<td>1.24358</td>
<td>1.39173</td>
<td>1.12747</td>
<td>1.15442</td>
<td>1.16663</td>
<td>1.09937</td>
<td>1.04435</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change of reserves</td>
<td>-0.03%</td>
<td>-0.86%</td>
<td>1.11%</td>
<td>0.22%</td>
<td>-0.85%</td>
<td>4.07%</td>
<td>2.99%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Ratio</td>
<td>1.25571</td>
<td>1.29680</td>
<td>1.15950</td>
<td>1.16590</td>
<td>1.15088</td>
<td>1.13546</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change of reserves</td>
<td>-0.06%</td>
<td>0.03%</td>
<td>0.48%</td>
<td>-0.05%</td>
<td>-0.19%</td>
<td>-1.95%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 Ratio</td>
<td>1.27280</td>
<td>1.19436</td>
<td>1.29335</td>
<td>1.15866</td>
<td>1.16698</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change of reserves</td>
<td>-0.09%</td>
<td>0.58%</td>
<td>-1.03%</td>
<td>0.06%</td>
<td>-0.39%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Ratio</td>
<td>1.23067</td>
<td>1.35690</td>
<td>1.23974</td>
<td>1.13310</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change of reserves</td>
<td>0.01%</td>
<td>-0.50%</td>
<td>-0.71%</td>
<td>0.71%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 Ratio</td>
<td>1.14923</td>
<td>1.18645</td>
<td>1.13079</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change of reserves</td>
<td>0.18%</td>
<td>0.68%</td>
<td>0.67%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 Ratio</td>
<td>1.17972</td>
<td>1.37137</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change of reserves</td>
<td>0.14%</td>
<td>-0.52%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 Ratio</td>
<td>1.37442</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change of reserves</td>
<td>-0.40%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Incurred

<table>
<thead>
<tr>
<th>AP/DP</th>
<th>0 Ratio</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 Ratio</td>
<td>1.55177</td>
<td>0.91138</td>
<td>0.92151</td>
<td>0.94413</td>
<td>0.98967</td>
<td>0.98158</td>
<td>0.98827</td>
<td>0.99240</td>
<td>0.99489</td>
<td></td>
</tr>
<tr>
<td>Change of reserves</td>
<td>-1.07%</td>
<td>-2.21%</td>
<td>3.85%</td>
<td>-2.12%</td>
<td>1.61%</td>
<td>-0.52%</td>
<td>-0.97%</td>
<td>0.86%</td>
<td>-5.46%</td>
<td></td>
</tr>
<tr>
<td>1 Ratio</td>
<td>1.75874</td>
<td>0.83335</td>
<td>0.83943</td>
<td>0.95489</td>
<td>0.97356</td>
<td>0.98546</td>
<td>0.99021</td>
<td>0.99084</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change of reserves</td>
<td>0.90%</td>
<td>-0.86%</td>
<td>-1.97%</td>
<td>-0.56%</td>
<td>-1.01%</td>
<td>0.44%</td>
<td>0.06%</td>
<td>-0.64%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Ratio</td>
<td>1.66182</td>
<td>0.84535</td>
<td>0.85722</td>
<td>0.97397</td>
<td>0.98946</td>
<td>0.98264</td>
<td>0.99025</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change of reserves</td>
<td>0.11%</td>
<td>-0.33%</td>
<td>-0.33%</td>
<td>1.15%</td>
<td>1.21%</td>
<td>-0.15%</td>
<td>0.80%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Ratio</td>
<td>1.80562</td>
<td>0.84013</td>
<td>0.83128</td>
<td>0.96017</td>
<td>0.97639</td>
<td>0.98398</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change of reserves</td>
<td>1.46%</td>
<td>-0.65%</td>
<td>-2.91%</td>
<td>-0.13%</td>
<td>-0.75%</td>
<td>0.17%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 Ratio</td>
<td>1.68303</td>
<td>0.84493</td>
<td>0.89236</td>
<td>0.96367</td>
<td>0.97456</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change of reserves</td>
<td>0.28%</td>
<td>-0.39%</td>
<td>1.28%</td>
<td>0.21%</td>
<td>-0.97%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Ratio</td>
<td>1.63277</td>
<td>0.85082</td>
<td>0.84692</td>
<td>0.97474</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change of reserves</td>
<td>-0.18%</td>
<td>-0.21%</td>
<td>-1.80%</td>
<td>1.36%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 Ratio</td>
<td>1.73137</td>
<td>0.80243</td>
<td>0.91088</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change of reserves</td>
<td>0.76%</td>
<td>-2.07%</td>
<td>2.52%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 Ratio</td>
<td>1.55377</td>
<td>0.92090</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change of reserves</td>
<td>-0.95%</td>
<td>2.32%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 Ratio</td>
<td>1.49675</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change of reserves</td>
<td>-1.34%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Again, incurred seems to be a bit more stable, in particular for later development periods.
Here we compare the projected ultimate based on the selected development factors with the projected ultimate if we exclude (or include) a observed development factor within the estimation of $\hat{f}_k$.
Stochastic Reserving
Lecture 3 (Continuation of Lecture 2)

Chain-Ladder method

René Dahms

ETH Zurich, Spring 2019

6 March 2019

(Last update: 18 February 2019)
2 Chain-Ladder-Method (CLM)

2.1 How does the Chain-Ladder method work
2.1.1 Chain-Ladder method without stochastic
2.1.2 Stochastic behind the Chain-Ladder method

2.2 Future development
2.2.1 Projection of the future development

2.3 Validation and examples (part 1 of 3)
2.3.1 Chain-Ladder method on Payments and on Incurred
2.3.2 How to validate the Chain-Ladder assumptions

2.4 Ultimate uncertainty
2.4.1 Ultimate uncertainty of accident period $i$
2.4.2 Ultimate uncertainty of the aggregation of all accident periods

2.5 Validation and examples (part 2 of 3)
2.5.1 Ultimate uncertainty

2.6 Solvency uncertainty
2.6.1 Solvency uncertainty of a single accident period
2.6.2 Solvency uncertainty of all accident periods
2.6.3 Uncertainties of further CDR’s

2.7 Validation and examples (part 3 of 3)
2.7.1 Solvency uncertainty

2.8 Literature
2 Chain-Ladder-Method (CLM)

2.1 How does the Chain-Ladder method work
2.1.1 Chain-Ladder method without stochastic
2.1.2 Stochastic behind the Chain-Ladder method

2.2 Future development
2.2.1 Projection of the future development
2.3 Validation and examples (part 1 of 3)
2.3.1 Chain-Ladder method on Payments and on Incurred
2.3.2 How to validate the Chain-Ladder assumptions

2.4 Ultimate uncertainty
2.4.1 Ultimate uncertainty of accident period i
2.4.2 Ultimate uncertainty of the aggregation of all accident periods

2.5 Validation and examples (part 2 of 3)
2.5.1 Ultimate uncertainty

2.6 Solvency uncertainty
2.6.1 Solvency uncertainty of a single accident period
2.6.2 Solvency uncertainty of all accident periods
2.6.3 Uncertainty of further CDR’s

2.7 Validation and examples (part 3 of 3)

2.8 Literature
Ultimate uncertainty of a single accident period (repetition)

The ultimate uncertainty of the estimated ultimate (or reserves) of accident period $i$ is defined by

$$
mse_{D^I} \left[ \hat{C}_{i,J} \right] = E \left[ \left( C_{i,J} - \hat{C}_{i,J} \right)^2 \mid D^I \right].$$

The mse can be split into random and parameter error

$$
mse_{D^I} \left[ \hat{C}_{i,J} \right] = \text{Var} \left[ C_{i,J} \mid D^I \right] + E \left[ C_{i,J} - \hat{C}_{i,J} \mid D^I \right]^2
$$

random error \hspace{1cm} parameter error

and analogously for the ultimate uncertainty of the whole reserves.
The ultimate uncertainty of the estimated ultimate (or reserves) of accident period $i$ is defined by

$$\text{mse}_{DI} \left[ \hat{C}_{i,J} \right] = \mathbb{E} \left[ (C_{i,J} - \hat{C}_{i,J})^2 \mid DI \right].$$

The mse can be split into random and parameter error

$$\text{mse}_{DI} \left[ \hat{C}_{i,J} \right] = \text{Var} \left[ C_{i,J} \mid DI \right] + \mathbb{E} \left[ (C_{i,J} - \hat{C}_{i,J})^2 \mid DI \right]$$

and analogously for the ultimate uncertainty of the whole reserves.

$$\text{Var} \left[ C_{i,J} \mid DI \right] = \text{Var} \left[ C_{i,J} - \hat{C}_{i,J} \mid DI \right]$$

$$= \mathbb{E} \left[ (C_{i,J} - \hat{C}_{i,J})^2 \mid DI \right] - \mathbb{E} \left[ C_{i,J} - \hat{C}_{i,J} \mid DI \right]^2$$

$$= \text{mse}_{DI} \left[ \hat{C}_{i,J} \right]$$
Taylor approximation of the mse (introduced by Ancus Röhr in [30])

Let's look at the (multi-linear) functional:

\[ U_i(g) \, x := g_{J-1} \cdots g_{I-i} \, x. \]

Then we get:

\[
\begin{align*}
\frac{\partial}{\partial g_j} U_i(g) \, x &= g_{J-1} \cdots g_{j+1} g_{j-1} \cdots g_{I-i} \, x = \frac{U_i(g) \, x}{g_j}, \\
U_i(\hat{f}) \, C_{i,I-i} &= \hat{f}_{J-1} \cdots \hat{f}_{I-i} C_{i,I-i} = \hat{C}_{i,J}, \\
U_i(F_i) \, C_{i,I-i} &= F_{i,J-1} \cdots F_{i,I-i} C_{i,I-i} = C_{i,J} \text{ and} \\
C_{i,J} - \hat{C}_{i,J} &\approx \sum_{k=I-i}^{J-1} \frac{\partial}{\partial F_{i,k}} U_i(F_i) \bigg|_{\hat{f}} C_{i,I-i} \left( F_{i,k} - \hat{f}_k \right) \\
&= \sum_{k=I-i}^{J-1} \frac{\hat{C}_{i,J}}{\hat{f}_k} \left( F_{i,k} - \hat{f}_k \right),
\end{align*}
\]

where we used a first order Taylor approximation and \( F_i \) and \( \hat{f} \) denote the vector of all link ratios \( F_{i,k} := C_{i,k+1}/C_{i,k} \) of accident period \( i \) and the vector of all estimated development factors \( \hat{f}_k \), respectively.

Note, for \( i + k \geq I \) we have:

\[
\begin{align*}
\mathbb{E}[F_{i,k}|D^I] &= \hat{f}_k, & \text{Var}[F_{i,k}|D^I] &\approx \frac{\hat{\sigma}^2_k}{C_{i,k}} \text{ and} & \text{Cov}[F_{i,k}, F_{h,j}|D^I] &= 0 \text{ for } (i, k) \neq (h, j).
\end{align*}
\]
Sine \( F_{i,k} = C_{i,k+1}/C_{i,k} \), we get for \( i + k \geq I \)

\[
\begin{align*}
\mathbb{E}\left[F_{i,k} \mid \mathcal{D}^I\right] &= \mathbb{E}\left[\mathbb{E}\left[F_{i,k} \mid D_{k}^{i+k}\right] \mid \mathcal{D}^I\right] = \mathbb{E}\left[\mathbb{E}\left[\frac{C_{i,k+1}}{C_{i,k}} \mid D_{k}^{i+k}\right] \mid \mathcal{D}^I\right] = \mathbb{E}\left[f_k \mid \mathcal{D}^I\right] = f_k \\
\text{Var}\left[F_{i,k} \mid \mathcal{D}^I\right] &= \mathbb{E}\left[\text{Var}\left[F_{i,k} \mid D_{k}^{i+k}\right] \mid \mathcal{D}^I\right] + \text{Var}\left[\mathbb{E}\left[F_{i,k} \mid D_{k}^{i+k}\right] \mid \mathcal{D}^I\right] \\
&= \mathbb{E}\left[\text{Var}\left[\frac{C_{i,k+1}}{C_{i,k}} \mid D_{k}^{i+k}\right] \mid \mathcal{D}^I\right] + 0 = \mathbb{E}\left[\frac{\sigma_k^2 C_{i,k}}{C_{i,k}^2} \mid \mathcal{D}^I\right] \approx \frac{\hat{\sigma}_k^2}{C_{i,k}}.
\end{align*}
\]

For the covariance statement we get: If \( i + k < I \) or \( h + j < I \) then \( F_{i,k} \in \mathcal{D}^I \) or \( F_{h,j} \in \mathcal{D}^I \) and we are done. Otherwise, since \((i, k) \neq (h, j)\), either \( F_{i,k} \in \mathcal{D}_{j}^{h+j} \) or \( F_{h,j} \in \mathcal{D}_{k}^{i+k} \). Let's assume the first:

\[
\begin{align*}
\text{Cov}\left[F_{i,k}, F_{h,j} \mid \mathcal{D}^I\right] &= \mathbb{E}\left[\text{Cov}\left[F_{i,k}, F_{h,j} \mid \mathcal{D}_{j}^{h+j}\right] \mid \mathcal{D}^I\right] + \text{Cov}\left[\mathbb{E}\left[F_{i,k} \mid \mathcal{D}_{j}^{h+j}\right], \mathbb{E}\left[F_{h,j} \mid \mathcal{D}_{j}^{h+j}\right] \mid \mathcal{D}^I\right] \\
&= 0 + \text{Cov}\left[F_{i,k}, f_j \mid \mathcal{D}^I\right] = 0
\end{align*}
\]
Estimator 2.9 (Linear approximation of the ultimate uncertainty of accident period $i$)

$$\text{mse}_{D^I} \left[ \hat{C}_{i,J} \right] = E \left[ \left( C_{i,J} - \hat{C}_{i,J} \right)^2 \mid D^I \right]$$

$$\approx E \left[ \left( \sum_{k=I-i}^{J-1} \frac{\hat{C}_{i,J}}{\hat{f}_k} \left( F_{i,k} - \hat{f}_k \right) \right)^2 \mid D^I \right]$$

(Taylor approximation)

$$= \sum_{k_1,k_2=I-i}^{J-1} \frac{\hat{C}_{i,J}}{\hat{f}_{k_1}} \frac{\hat{C}_{i,J}}{\hat{f}_{k_2}} E \left[ \left( F_{i,k_1} - \hat{f}_{k_1} \right) \left( F_{i,k_2} - \hat{f}_{k_2} \right) \mid D^I \right]$$

$$= \sum_{k_1,k_2=I-i}^{J-1} \frac{\hat{C}_{i,J}}{\hat{f}_{k_1}} \frac{\hat{C}_{i,J}}{\hat{f}_{k_2}} \left( \text{Cov} \left[ F_{i,k_1}, F_{i,k_2} \mid D^I \right] + \left( \hat{f}_{k_1} - f_{k_1} \right) \left( \hat{f}_{k_2} - f_{k_2} \right) \right)$$

$$\approx \sum_{k_1,k_2=I-i}^{J-1} \frac{\hat{C}_{i,J}}{\hat{f}_{k_1}} \frac{\hat{C}_{i,J}}{\hat{f}_{k_2}} \left( \left\{ \text{Cov} \left[ F_{i,k_1}, F_{i,k_2} \mid D^I \right] \right\} + \left\{ \text{Cov} \left[ \hat{f}_{k_1}, \hat{f}_{k_2} \mid D_{k_1 \land k_2} \right] \right\} \right)$$

random error

parameter error

$$\approx \sum_{k=I-i}^{J-1} \frac{\hat{C}_{i,J}^2}{\hat{f}_k^2} \frac{1}{\hat{C}_{i,k}} + \sum_{k=I-i}^{J-1} \frac{\hat{C}_{i,J}^2}{\hat{f}_k^2} \sum_{h=0}^{I-k-1} \frac{w_{h,k}^2}{C_{h,k}} = \hat{C}_{i,J}^2 \sum_{k=I-i}^{J-1} \frac{\hat{C}_{i,k}^2}{\hat{f}_k^2} \left( \frac{1}{\hat{C}_{i,k}} + \sum_{h=0}^{I-k-1} \frac{w_{h,k}^2}{C_{h,k}} \right)$$

random error

parameter error
Stochastic Reserving

---

Chain-Ladder-Method (CLM)

Ultimate uncertainty

\[ E[(F_{i,k_1} - \hat{f}_{k_1})(F_{i,k_2} - \hat{f}_{k_2})|D^I] = E[((F_{i,k_1} - f_{k_1}) - (\hat{f}_{k_1} - f_{k_1}))((F_{i,k_2} - f_{k_2}) - (\hat{f}_{k_2} - f_{k_2}))|D^I] \]

\[ = E[(F_{i,k_1} - f_{k_1})(F_{i,k_2} - f_{k_2})|D^I] - E[(F_{i,k_1} - f_{k_1})(\hat{f}_{k_2} - f_{k_2})|D^I] \]

\[ - E[(\hat{f}_{k_1} - f_{k_1})(F_{i,k_2} - f_{k_2})|D^I] + E[(\hat{f}_{k_1} - f_{k_1})(\hat{f}_{k_2} - f_{k_2})|D^I] \]

\[ = \text{Cov}[F_{i,k_1}, F_{i,k_2}|D^I] - \text{E}[F_{i,k_1} - f_{k_1}|D^I](\hat{f}_{k_2} - f_{k_2}) \]

\[ - \text{E}[F_{i,k_2} - f_{k_2}|D^I](\hat{f}_{k_1} - f_{k_1}) + (\hat{f}_{k_1} - f_{k_1})(\hat{f}_{k_2} - f_{k_2}) \]

\[ = \text{Cov}[F_{i,k_1}, F_{i,k_2}|D^I] - 0 - 0 + (\hat{f}_{k_1} - f_{k_1})(\hat{f}_{k_2} - f_{k_2}) \]

For \( k_1 < k_2 \) we have \( \hat{f}_{k_1} \in D_{k_2} \) and \( F_{i,k_1} \in D^I_{k_2} \). This leads to

\[ \text{Cov}[\hat{f}_{k_1}, \hat{f}_{k_2}|D_{k_1 \land k_2}] = \text{E}[\text{Cov}[\hat{f}_{k_1}, \hat{f}_{k_2}|D_{k_2}]|D_{k_1 \land k_2}] + \text{Cov}[\text{E}[\hat{f}_{k_1}|D_{k_2}], \text{E}[\hat{f}_{k_2}|D_{k_2}]|D_{k_1 \land k_2}] \]

\[ = 0 + \text{Cov}[\hat{f}_{k_1}, f_{k_2}|D_{k_1 \land k_2}] = 0 \]

\[ \text{Cov}[F_{i,k_1}, F_{i,k_2}|D^I] = \text{E}[\text{Cov}[F_{i,k_1}, F_{i,k_2}|D^I_{k_2}]|D^I] + \text{Cov}[\text{E}[F_{i,k_1}|D^I_{k_2}], \text{E}[F_{i,k_2}|D^I_{k_2}]|D^I] \]

\[ = 0 + \text{Cov}[F_{i,k_1}, f_{k_2}|D^I] = 0 \]
Corollary 2.10

- If we use the variance minimizing weights

\[
    w_{i,k} = \frac{C_{i,k}}{I-k-1} \sum_{h=0}^{I-k-1} C_{h,k}
\]

we get

\[
    \text{mse}_{D|} \left[ \hat{C}_{i,J} \right] \approx \hat{C}_{i,J}^2 \sum_{k=i-I}^{J-1} \frac{\hat{\sigma}_k^2}{\hat{f}_k^2} \left( \frac{1}{\hat{C}_{i,k}} + \frac{1}{\sum_{h=0}^{I-k-1} C_{h,k}} \right).
\]

- For the estimated coefficient of variation (and variance minimizing weights) we get

\[
    \sqrt{\text{Var} \left[ C_{i,J} \right]} \approx \sqrt{\text{mse}_{C_{i,J}|D} \left[ \hat{C}_{i,J} \right]}
\]

\[
    \sqrt{\sum_{k=i-I}^{J-1} \frac{\hat{\sigma}_k^2}{\hat{f}_k^2}} \left( \frac{1}{\hat{C}_{i,k}} + \frac{1}{\sum_{h=0}^{I-k-1} C_{h,k}} \right)
\]

which means the coefficient of variation of the ultimate uncertainty (or at least of the parameter error) vanishes with increasing volume. Usually, this is not valid in practice. Therefore, you should always consider in addition some model error.
If we always use only the \( n \) diagonals in order to estimate the development factors the parameter error term in the coefficient of variation will not converge to zero for \( I \to \infty \).

In practice, this is often a reasonable approach, because the comparability of the development of very old (calender) periods in respect to the expected future is very questionable.

Nevertheless, you should always consider some **model error**.
Corollary 2.11

- Instead of using the Taylor approximation you can directly estimate the random and the parameter error like Mack has done in the original approach, see [21]. The results are the same.

- For the process error we have made three approximations:
  * Taylor approximation,
  * \( \text{Var}\left[\frac{1}{C_{i,k}}|D^I\right] \approx \frac{1}{\hat{C}_{i,k}} \) and
  * \( \sigma^2_k \approx \hat{\sigma}^2_k \).

  Following the original calculation of Mack one can see that the first two approximation cancel each other.
Corollary 2.11

Instead of using the Taylor approximation you can directly estimate the random and the parameter error like Mack has done in the original approach, see [21]. The results are the same.

For the process error we have made three approximations:

1. Taylor approximation
2. \( \frac{1}{C_{i,k}|D^I} \approx \frac{1}{\hat{C}_{i,k}} \)
3. \( \sigma^2_k \approx \hat{\sigma}^2_k \)

Following the original calculation of Mack one can see that the first two approximation cancel each other.

Original estimation of the random error:

\[
\text{Var} \left[ C_{i,J} \mid D^I \right] = \underbrace{\text{Var} \left[ \hat{C}_{i,J} \mid B_{i,I-i} \right]}_{\text{iii) CLM}}
\]

\[
= \text{Var} \left[ E \left[ C_{i,J} \mid B_{i,J-1} \right] \right] + E \left[ \text{Var} \left[ C_{i,J} \mid B_{i,J-1} \right] \right] B_{i,I-i}
\]

\[
= \text{Var} \left[ f_{J-1} C_{i,J-1} \mid B_{i,I-i} \right] + \underbrace{E \left[ \sigma^2_{J-1} C_{i,J-1} \mid B_{i,I-i} \right]}_{\text{ii) CLM}}
\]

\[
= f_{J-1}^2 \text{Var} \left[ C_{i,J-1} \mid B_{i,I-i} \right] + \sigma^2_{J-1} \prod_{j=I-i}^{J-2} f_j C_{i,I-i}
\]

\[
= \ldots = \sum_{k=I-i}^{J-1} \prod_{j=k+1}^{J-1} f_j^2 \sigma_k^2 \prod_{j=I-i}^{k-1} f_j C_{i,I-i}
\]

\[
= \sum_{k=I-i}^{J-1} f_k^2 \prod_{j=I-i}^{k-1} f_j C_{i,I-i} \left( \prod_{j=I-i}^{J-1} f_j C_{i,I-i} \right)^2
\]

\[
\approx \sum_{k=I-i}^{J-1} \widehat{\sigma}_k^2 \prod_{j=I-i}^{k-1} \widehat{f}_j C_{i,I-i}
\]

\[
= \sum_{k=I-i}^{J-1} \frac{\widehat{\sigma}_k^2}{f_k^2 \widehat{C}_{i,k}}
\]
Estimator 2.12 (Variance parameter)

Let Assumption 2.A be fulfilled. Then

$$
\hat{\sigma}^2_k := \frac{1}{Z_k} \sum_{i=0}^{I-1-k} C_{i,k} \left( \frac{C_{i,k+1}}{C_{i,k}} - \hat{f}_k \right)^2,
$$

with

$$
Z_k := I - 2 - k + \sum_{i=0}^{I-1-k} w_{i,k}^2 \frac{1}{C_{i,k}} \sum_{h=0}^{I-1-k} C_{h,k},
$$

are $D_k$-conditional unbiased estimators for the variance parameters $\sigma^2_k$, provided that $Z_k > 0$.

If $Z_k \leq 0$ one could take

$$
\tilde{\sigma}^2_k := \min \left( \frac{(\hat{\sigma}^2_k - 1)^2}{\hat{\sigma}^2_{k-2}}, \hat{\sigma}^2_{k-2}, \hat{\sigma}^2_{k-1} \right).
$$

Variance minimizing weights of (2.2) lead to $Z_k = I - k - 1$. 
Taking the variance minimizing weights we get

Unbiasedness:

\[
E \left[ C_{i,k} \left( \frac{C_{i,k+1}}{C_{i,k}} - \hat{f}_k \right)^2 \Bigg| \mathcal{D}_k \right] = C_{i,k} \text{Var} \left[ \frac{C_{i,k+1}}{C_{i,k}} - \hat{f}_k \Bigg| \mathcal{D}_k \right]
\]

\[
= C_{i,k} \text{Var} \left[ \frac{C_{i,k+1}}{C_{i,k}} - \sum_{h=0}^{I-k-1} w_{h,k} \frac{C_{h,k+1}}{C_{h,k}} \Bigg| \mathcal{D}_k \right]
\]

\[
= C_{i,k} \text{Var} \left[ \sum_{h=0}^{I-k-1} \left( \frac{C_{i,k+1}}{(I-k)C_{i,k}} - w_{h,k} \frac{C_{h,k+1}}{C_{h,k}} \right) \Bigg| \mathcal{D}_k \right]
\]

\[
= C_{i,k} \sum_{h_1=0}^{I-1-k} \sum_{h_2=0}^{I-1-k} \text{Cov} \left[ \left( \frac{C_{i,k+1}}{(I-k)C_{i,k}} - w_{h_1,k} \frac{C_{h_1,k+1}}{C_{h_1,k}} \right), \left( \frac{C_{i,k+1}}{(I-k)C_{i,k}} - w_{h_2,k} \frac{C_{h_2,k+1}}{C_{h_2,k}} \right) \Bigg| \mathcal{D}_k \right]
\]

\[
= C_{i,k} \sum_{h_1=0}^{I-1-k} \sum_{h_2=0}^{I-1-k} \left( \frac{\sigma_k^2}{(I-k)^2 C_{i,k}} - \frac{\sigma_k^2 w_{i,k}}{(I-k)C_{i,k}} \right) - \frac{\sigma_k^2 w_{i,k}}{(I-k)C_{i,k}} \mathbf{1}_{h_1=i} \mathbf{1}_{h_2=i} + \frac{\sigma_k^2 w_{i,k} w_{h_1,k} w_{h_2,k}}{C_{h_1,k}} \mathbf{1}_{h_1=h_2}
\]

\[
= \sigma_k^2 \left( 1 - 2 w_{i,k} + C_{i,k} \sum_{h=0}^{I-1-k} \frac{w_{h,k}}{C_{h,k}} \right)
\]

\[
\Rightarrow \sum_{i=0}^{I-1-k} E \left[ C_{i,k} \left( \frac{C_{i,k+1}}{C_{i,k}} - \hat{f}_k \right)^2 \Bigg| \mathcal{D}_k \right] = Z_k
\]

Taking the variance minimizing weights we get

\[
Z_k = I - 2 - k + \sum_{i=0}^{I-1-k} \frac{C_{i,k}^2}{\left( \sum_{h=0}^{I-1-k} C_{h,k} \right)^2} \sum_{h=0}^{I-1-k} C_{h,k} = I - 2 - k + \sum_{i=0}^{I-1-k} \frac{C_{i,k}}{\sum_{h=0}^{I-1-k} C_{h,k}} = I - 1 - k.
\]
Ultimate uncertainty of all accident periods

Analogue to the procedure we used for a single accident period, we split the ultimate uncertainty of the aggregation of all accident periods into:

\[
\text{mse}_{D^I} \left[ \sum_{i=0}^{I} \hat{C}_{i,J} \right] = \text{Var} \left[ \sum_{i=I-J+1}^{I} C_{i,J} \bigg| D^I \right] + \mathbb{E} \left[ \sum_{i=I-J+1}^{I} \left( C_{i,J} - \hat{C}_{i,J} \right) \bigg| D^I \right]^2
\]

Since accident periods are independent, the random error of the sum of all accident periods is simply the sum of all single periods random errors. But for the parameter error this is not the case, because the accident periods are coupled via the same estimated development factors.
Analogue to the procedure we used for a single accident period, we split the ultimate uncertainty of the aggregation of all accident periods into:

\[ \text{mse} = \text{Var} \left( \sum_{i=J}^{I(J+1)} \hat{C}_{i,J} \right) + \text{E} \left[ \sum_{i=J}^{I(J+1)} (C_{i,J} - \hat{C}_{i,J})^2 \right] \]

Since accident periods are independent, the random error of the sum of all accident periods is simply the sum of all single periods random errors. But for the parameter error this is not the case, because the accident periods are coupled via the same estimated development factors.
Estimator 2.13 (of the ultimate uncertainty of all accident periods)

$$\text{mse}_{D^I} \left[ \hat{C}_{i,J} \right] = \mathbb{E} \left[ (C_{i,J} - \hat{C}_{i,J})^2 \bigg| D^I \right]$$

$$\approx \mathbb{E} \left[ \left( \sum_{i=0}^{I} \sum_{k=I-i}^{J-1} \frac{\hat{C}_{i,J}}{\hat{f}_k} (F_{i,k} - \hat{f}_k) \right)^2 \bigg| D^I \right]$$

$$(\text{Taylor approximation})$$

$$= \sum_{k_1,k_2=0}^{J-1} \sum_{i_1=I-k_1}^{I} \sum_{i_2=I-k_2}^{I} \frac{\hat{C}_{i_1,J}}{\hat{f}_{k_1}} \frac{\hat{C}_{i_2,J}}{\hat{f}_{k_2}} \mathbb{E} \left[ (F_{i_1,k_1} - \hat{f}_{k_1}) (F_{i_2,k_2} - \hat{f}_{k_2}) \bigg| D^I \right]$$

$$+ \sum_{k_1,k_2=0}^{J-1} \sum_{i_1=I-k_1}^{I} \sum_{i_2=I-k_2}^{I} \frac{\hat{C}_{i_1,J}}{\hat{f}_{k_1}} \frac{\hat{C}_{i_2,J}}{\hat{f}_{k_2}} \left( \text{Cov} [F_{i_1,k_1}, F_{i_2,k_2} | D^I] + (\hat{f}_{k_1} - f_{k_1}) (\hat{f}_{k_2} - f_{k_2}) \right)$$

$$\approx \sum_{k=0}^{J-1} \sum_{i=I-k}^{I} \frac{\sigma_k^2}{\hat{f}_k^2} \frac{\hat{C}_{i,J}^2}{\hat{C}_{i,k}} + \sum_{k=0}^{J-1} \frac{\sigma_k^2}{\hat{f}_k^2} \left( \sum_{i=I-k}^{I} \frac{\hat{C}_{i,J}}{\hat{f}_k} \right)^2 \sum_{h=0}^{I-k-1} \frac{w_{h,k}^2}{C_{h,k}}$$

random error

parameter error
For $i_1 + k_1 \geq I$ and $i_2 + k_2 \geq I$ we get

$$\text{Cov} \left[ F_{i_1, k_1}, F_{i_2, k_2} \mid D^I \right] = \text{Cov} \left[ \mathbf{E} \left[ F_{i_1, k_1} \mid D^I_{k_1 \lor k_2} \right], \mathbf{E} \left[ F_{i_2, k_2} \mid D^I_{k_1 \lor k_2} \right] \right] \mid D^I + \mathbf{E} \left[ \text{Cov} \left[ F_{i_1, k_1}, F_{i_2, k_2} \mid D^I_{k_1 \lor k_2} \right] \right] \mid D^I$$

$$= 0 + \mathbf{1}_{i_1=i_2} 1_{k_1}=k_2 \sigma^2_{k_1} \mathbf{E} \left[ \frac{1}{C_{i_1, k_1}} \right] \mid D^I \approx \mathbf{1}_{i_1=i_2} 1_{k_1}=k_2 \sigma^2_{k_1} \frac{1}{C_{i_1, k_1}}$$

$$\text{Cov} \left[ \hat{f}_{k_1}, \hat{f}_{k_2} \mid D_{k_1 \land k_2} \right] = \mathbf{1}_{k_1}=k_2 \sigma^2_{k_1} \sum_{h=0}^{I - k_1 - 1} \frac{w^2_{h, k_1}}{C_{h, k}}.$$

Therefore,

$$\text{mse} \sum_{i=0}^{I} C_{i, J} \mid D^I \left[ \hat{C}_{i, J} \right] \approx \sum_{k_1, k_2=0}^{J-1} \sum_{i_1=I-k_1}^{I} \sum_{i_2=I-k_2}^{I} \frac{\hat{C}_{i_1, J}}{\hat{f}_{k_1}} \frac{\hat{C}_{i_2, J}}{\hat{f}_{k_2}} \sigma^2_{k_1} \mathbf{1}_{k_1}=k_2 \left( \mathbf{1}_{i_1=i_2} \frac{1}{C_{i_1, k_2}} + \sum_{h=0}^{I - k_1 - 1} \frac{w^2_{h, k_1}}{C_{h, k}} \right)$$

$$= \sum_{k=0}^{J-1} \frac{\sigma^2_{k}}{\hat{f}_{k}^2} \left( \sum_{i=I-k}^{I} \hat{C}_{i, J} \frac{1}{\hat{C}_{i, k}} + \left( \sum_{i=I-k}^{I} \hat{C}_{i, J} \right) \sum_{h=0}^{I-k-1} \frac{w^2_{h, k}}{C_{h, k}} \right).$$
Corollary 2.14

If we use the variance minimizing weights

\[ w_{i,k} = \frac{C_{i,k}}{\sum_{h=0}^{I-k-1} C_{h,k}} \]

and the notation \( \hat{C}_{i,k} := C_{i,k} \), for \( i + k \leq I \), we get

\[
\text{mse}_{D_I} \left[ \hat{C}_{i,J} \right] \\
\approx J^{-1} \sum_{k=0}^{J-1} \sum_{i=I-k}^{I} \left( \sum_{h=0}^{I} \frac{\hat{C}_{i,J}^2}{\hat{C}_{i,k}} \right) \left( \sum_{h=0}^{I} \hat{C}_{h,J} \right)^2 \left( \sum_{i=I-k}^{I} \hat{C}_{i,J} \right)^2 \left( \sum_{h=0}^{I} \hat{C}_{h,J} \right)^2 + \left( \sum_{i=I-k}^{I} \hat{C}_{i,J} \right)^2 \left( \sum_{h=0}^{I} \hat{C}_{h,J} \right)^2 \left( \sum_{h=0}^{I-k-1} C_{h,k} \right)^2 \\
= J^{-1} \sum_{k=0}^{J-1} \sum_{i=I-k}^{I} \left( \sum_{h=0}^{I} \hat{C}_{i,k} \right)^2 \left( \sum_{h=0}^{I} \hat{C}_{h,J} \right)^2 \left( \sum_{i=I-k}^{I} \hat{C}_{i,J} \right)^2 \left( \sum_{h=0}^{I} \hat{C}_{h,J} \right)^2 + \left( \sum_{i=I-k}^{I} \hat{C}_{i,J} \right)^2 \left( \sum_{h=0}^{I} \hat{C}_{h,J} \right)^2 \left( \sum_{h=0}^{I-k-1} C_{h,k} \right)^2 \\
= \left( \sum_{i=0}^{I} \hat{C}_{i,J} \right)^2 \sum_{k=0}^{J-1} \sum_{i=0}^{I-k-1} \frac{\hat{C}_{i,J}^2}{\hat{C}_{i,k}} \left( \frac{1}{\sum_{i=0}^{I-k-1} \hat{C}_{i,k}} - \frac{1}{\sum_{i=0}^{I} \hat{C}_{i,k}} \right)
\]
For each $k < J$ we have

$$\sum_{i=0}^{I} \hat{C}_{i,k+1} = \sum_{i=0}^{I-k-1} C_{i,k+1} + \sum_{i=I-k}^{I} C_{i,k+1} = \frac{\sum_{i=0}^{I-k-1} C_{i,k+1}}{\sum_{i=0}^{I-k-1} C_{i,k}} \sum_{i=0}^{I-k-1} C_{i,k} + \sum_{i=I-k}^{I} \hat{f}_k C_{i,k} = \hat{f}_k \sum_{i=0}^{I} \hat{C}_{i,k}.$$ 

Therefore, we get for each $k \geq I - i$

$$\frac{\hat{C}_{i,J}}{\sum_{i=0}^{I} C_{i,J}} = \frac{\hat{f}_{J-1} \cdots \hat{f}_k \hat{C}_{i,k}}{\hat{f}_{J-1} \cdots \hat{f}_k \sum_{i=0}^{I} C_{i,k}} = \frac{\hat{C}_{i,k}}{\sum_{i=0}^{I} C_{i,k}}.$$ 

Finally,

$$\sum_{i=I-k}^{I} \frac{\hat{C}_{i,k}^2}{\left(\sum_{h=0}^{I} \hat{C}_{h,k}\right)^2} \frac{1}{\hat{C}_{i,k}} + \frac{\left(\sum_{i=I-k}^{I} \hat{C}_{i,k}\right)^2}{\left(\sum_{h=0}^{I} \hat{C}_{h,k}\right)^2} \frac{1}{\sum_{h=0}^{I-k-1} C_{h,k}}$$

$$= \frac{\sum_{i=I-k}^{I} \hat{C}_{i,k}}{\left(\sum_{i=0}^{I} \hat{C}_{i,k}\right)^2} \sum_{i=0}^{I-k-1} C_{i,k} + \left(\sum_{i=I-k}^{I} \hat{C}_{i,k}\right)^2 \frac{1}{\sum_{h=0}^{I-k-1} C_{h,k}}$$

$$= \frac{\sum_{i=I-k}^{I} \hat{C}_{i,k} \left(\sum_{i=0}^{I-k-1} \hat{C}_{i,k} + \sum_{i=I-k}^{I} \hat{C}_{i,k}\right)}{\left(\sum_{i=0}^{I} \hat{C}_{i,k}\right)^2} \sum_{i=0}^{I-k-1} C_{i,k}$$

$$= \frac{\sum_{i=I-k}^{I} \hat{C}_{i,k}}{\sum_{h=0}^{I} \hat{C}_{h,k} \sum_{h=0}^{I-k-1} C_{h,k}} \sum_{i=0}^{I-k-1} C_{i,k}$$

$$= \frac{1}{\sum_{i=0}^{I-k-1} \hat{C}_{i,k} - \sum_{i=0}^{I} \hat{C}_{i,k}}.$$
Credibility like weighting of ultimates

One way of combining (two or more) estimates for the same ultimate is to use a credibility like weighting. This means, for an estimated ultimate we take the lesser credibility the further away it is from the last known value. In formula:

**Estimator 2.15 (Credibility like weighted ultimate)**

Let $\hat{C}_{i,J}^m$, $0 \leq m \leq M$, be estimates of the same (unknown) ultimate. Then

\[
\sum_{m=0}^{M} \min \left( \frac{\hat{C}_{i,J}^m}{C_{i,I-i}^m}, \frac{C_{i,I-i}^m}{\hat{C}_{i,J}^m} \right) \left( \sum_{l=0}^{M} \min \left( \frac{\hat{C}_{i,J}^l}{C_{i,I-i}^l}, \frac{C_{i,I-i}^l}{\hat{C}_{i,J}^l} \right) \right)^{-1} \hat{C}_{i,J}^m
\]

is a credibility like weighted mean of these estimates.

**Remark 2.16 (Credibility like weighted ultimate uncertainty)**

We will see later, see Section 4, that it is possible to transfer the weighting of ultimates to the corresponding ultimate uncertainties.
Credibility like weighting of ultimates

One way of combining (two or more) estimates for the same ultimate is to use a credibility like weighting. This means, for an estimated ultimate we take the lesser credibility the further away it is from the last known value. In formula:

\[
\text{Estimator 2.15 (Credibility like weighted ultimate)}
\]

Let \( \hat{C}_{m}^{i,J} \), \( 0 \leq m \leq M \), be estimates of the same (unknown) ultimate. Then

\[
\sum_{m=0}^{M} \min \left( \frac{\hat{C}_{m}^{i,J}}{C_{m}^{i,I}} - i, \frac{C_{m}^{i,I} - \hat{C}_{m}^{i,J}}{i} \right) \left( \sum_{l=0}^{M} \min \left( \frac{\hat{C}_{l}^{i,J}}{C_{l}^{i,I}} - i, \frac{C_{l}^{i,I} - \hat{C}_{l}^{i,J}}{i} \right) \right) \hat{C}_{m}^{i,J}
\]

is a credibility like weighted mean of these estimates.

Remark 2.16 (Credibility like weighted ultimate uncertainty)

We will see later, see Section 4, that it is possible to transfer the weighting of ultimates to the corresponding ultimate uncertainties.
Credibility like weighting of ultimates from Examples 2.7 and 2.8

- We used the standard estimators for the variance parameters, see Estimator 2.12.
- Since the incurred values are much faster near the ultimate, the corresponding projection gets more weight.

<table>
<thead>
<tr>
<th>AP</th>
<th>Payments</th>
<th>Incurred</th>
<th>Credibility like weighting</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ultimate</td>
<td>Reserves</td>
<td>Ultimate</td>
</tr>
<tr>
<td>0</td>
<td>3'921'258</td>
<td>---</td>
<td>3'921'258</td>
</tr>
<tr>
<td>1</td>
<td>2'681'142</td>
<td>114'086</td>
<td>2'905'040</td>
</tr>
<tr>
<td>2</td>
<td>3'576'632</td>
<td>394'121</td>
<td>3'214'395</td>
</tr>
<tr>
<td>3</td>
<td>3'612'174</td>
<td>608'749</td>
<td>3'334'861</td>
</tr>
<tr>
<td>4</td>
<td>2'848'093</td>
<td>697'742</td>
<td>3'168'701</td>
</tr>
<tr>
<td>5</td>
<td>3'619'496</td>
<td>1'234'157</td>
<td>3'489'267</td>
</tr>
<tr>
<td>6</td>
<td>2'626'200</td>
<td>1'138'623</td>
<td>3'356'241</td>
</tr>
<tr>
<td>7</td>
<td>3'123'198</td>
<td>1'638'793</td>
<td>3'482'056</td>
</tr>
<tr>
<td>8</td>
<td>3'736'063</td>
<td>2'359'939</td>
<td>2'794'903</td>
</tr>
<tr>
<td>9</td>
<td>2'821'331</td>
<td>1'979'401</td>
<td>3'398'542</td>
</tr>
<tr>
<td>Total</td>
<td>32'565'588</td>
<td>10'165'612</td>
<td>33'065'263</td>
</tr>
</tbody>
</table>

© R. Dahms (ETH Zurich, Spring 2019)
Credibility like weighting of ultimates from Examples 2.7 and 2.8

- We used the standard estimators for the variance parameters, see Estimator 2.12.
- Since the incurred values are much faster near the ultimate, the corresponding projection gets more weight.

### Validation and examples (part 2 of 3)

<table>
<thead>
<tr>
<th>AP</th>
<th>Payments</th>
<th>Incurred</th>
<th>Credibility like weighting</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ultimate</td>
<td>Reserves</td>
<td>Ultimate</td>
</tr>
<tr>
<td>0</td>
<td>3921258</td>
<td>—</td>
<td>3921258</td>
</tr>
<tr>
<td>1</td>
<td>2881142</td>
<td>114086</td>
<td>2905040</td>
</tr>
<tr>
<td>2</td>
<td>3575632</td>
<td>354121</td>
<td>3524785</td>
</tr>
<tr>
<td>3</td>
<td>3812774</td>
<td>608746</td>
<td>3334961</td>
</tr>
<tr>
<td>4</td>
<td>2848093</td>
<td>697742</td>
<td>3168701</td>
</tr>
<tr>
<td>5</td>
<td>3819486</td>
<td>1234157</td>
<td>3469267</td>
</tr>
<tr>
<td>6</td>
<td>2968200</td>
<td>1338323</td>
<td>3556241</td>
</tr>
<tr>
<td>7</td>
<td>3123196</td>
<td>1838793</td>
<td>3482056</td>
</tr>
<tr>
<td>8</td>
<td>3738663</td>
<td>2558936</td>
<td>2794903</td>
</tr>
<tr>
<td>9</td>
<td>2821331</td>
<td>1979401</td>
<td>3398542</td>
</tr>
<tr>
<td>Total</td>
<td>32565358</td>
<td>10165612</td>
<td>33065263</td>
</tr>
</tbody>
</table>
The projection of incurred is much faster very close and stable to the estimated ultimate than the projection of payments. This may be an indication to trust it more.
The projection of incurred is much faster very close and stable to the estimated ultimate than the projection of payments. This may be an indication to trust it more.
Ultimate uncertainties for Examples 2.7 and 2.8

- We used the standard estimators for the variance parameters, see Estimator 2.12.
- Since the incurred values are a bit more stable, in particular for later development periods, the corresponding uncertainties are lower.
- The linear approximation for the (parameter estimation) uncertainty results in almost the same values like without approximation.

<table>
<thead>
<tr>
<th>AP</th>
<th>Ultimate uncertainty for payments</th>
<th>Ultimate uncertainty for incurred</th>
<th>Credibility like weighting</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Proc Var</td>
<td>Para Err</td>
<td>Total</td>
</tr>
<tr>
<td>0</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>1</td>
<td>68'914</td>
<td>56'985</td>
<td>89'423</td>
</tr>
<tr>
<td>2</td>
<td>184'912</td>
<td>144'485</td>
<td>324'666</td>
</tr>
<tr>
<td>3</td>
<td>203'838</td>
<td>154'232</td>
<td>358'062</td>
</tr>
<tr>
<td>4</td>
<td>223'462</td>
<td>135'431</td>
<td>358'893</td>
</tr>
<tr>
<td>5</td>
<td>270'501</td>
<td>178'156</td>
<td>448'657</td>
</tr>
<tr>
<td>6</td>
<td>241'283</td>
<td>131'817</td>
<td>373'090</td>
</tr>
<tr>
<td>7</td>
<td>330'933</td>
<td>173'453</td>
<td>504'386</td>
</tr>
<tr>
<td>8</td>
<td>437'284</td>
<td>227'437</td>
<td>664'721</td>
</tr>
<tr>
<td>9</td>
<td>430'953</td>
<td>182'846</td>
<td>613'799</td>
</tr>
<tr>
<td>Total</td>
<td>865'025</td>
<td>1'247'250</td>
<td>2112'275</td>
</tr>
</tbody>
</table>

Linear approximation

<table>
<thead>
<tr>
<th></th>
<th>Ultimate uncertainty for payments</th>
<th>Ultimate uncertainty for incurred</th>
<th>Credibility like weighting</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Proc Var</td>
<td>Para Err</td>
<td>Total</td>
</tr>
<tr>
<td></td>
<td>865'025</td>
<td>1'246'787</td>
<td>2111'812</td>
</tr>
</tbody>
</table>

We always show the square root of uncertainties.
The uncertainty of the weighing has been calculated using a LSRM coupling of both CLM via the exposure $R_{i,k}^{0,1} = R_{i,k}^{1,0} := R_{i,k}^{0,0} + R_{i,k}^{1,1}$, see Section 4.

One can derive estimators for the ultimate uncertainty without a first order Taylor approximation, see [20]. In practice, the resulting figures do only deviate marginally.
Density plot of the distribution of the estimated reserves using Lognormal distributions (dotted lines representing the Best Estimate)
The projection of incurred losses results in a more symmetric and tight distribution than the projection of payments. Therefore, if we believe in the incurred projection and the corresponding estimate of the ultimate uncertainty we would expect that the true future payments will only deviate from the estimated reserves by a small amount. Whereas the projection of payments indicates much larger differences (uncertainty).

The uncertainty of the weighing has been calculated using a LSRM coupling of both CLM via the exposure $R_{i,k}^{0,1} = R_{i,k}^{1,0} := R_{i,k}^{0,0} + R_{i,k}^{1,1}$, see Section 4.
Problem 2.17 (Fitting a distribution to Best Estimate reserves and a mse)

Assume that for a portfolio we have

- A Best Estimate for the reserves,
- an estimate for uncertainties in terms of mse and the corresponding estimate of the reserves $R$. That means the method, which was used for the estimation of the uncertainty gives us a corresponding estimate of the reserves, which will usually differ from the Best Estimate reserves.

How to fit a distribution to those estimates?

Fitting a distribution to Best Estimate reserves and a mse

- **Shifting the distribution**: Means we fit the distribution with
  
  \[
  \text{Expectation} = \text{Best Estimate reserves (or ultimate)} \\
  \text{Variance} = \hat{\text{mse}}
  \]

- **Stretching the distribution**: Means we fit the distribution with
  
  \[
  \text{Expectation} = \text{Best Estimate reserves (or ultimate)} \\
  \text{Variance} = \frac{\hat{\text{mse}} \cdot (\text{Best Estimate reserves})^2}{R^2}
  \]
Problem 2.17 (Fitting a distribution to Best Estimate reserves and a mse)

Assume that for a portfolio we have

- A Best Estimate for the reserves,
- An estimate for uncertainties in terms of mse and the corresponding estimate of the reserves \( R \). That means the method, which was used for the estimation of the uncertainty gives us a corresponding estimate of the reserves, which will usually differ from the Best Estimate reserves.

How to fit a distribution to those estimates?

- **Shifting the distribution:** Means we fit the distribution with
  
  \[
  \text{Expectation} = \text{Best Estimate reserves (or ultimate)}
  
  \text{Variance} = \hat{\text{mse}}
  \]

- **Stretching the distribution:** Means we fit the distribution with
  
  \[
  \text{Expectation} = \text{Best Estimate reserves (or ultimate)}
  
  \text{Variance} = \hat{\text{mse}} \cdot \frac{(\text{Best Estimate reserves})^2}{R^2}
  \]

I prefer the stretching, as long as it leads to plausible results, which in particular is not the case if \( R \approx 0 \).
Densitiy plot of the Lognormal distributions

Best Estimate reserves (BE) = 10
\( \hat{\text{mse}} = 0.5 \quad \text{and} \quad R = 8 \)

\( \sigma^2 = 0.5 \)

- Shifting \( \sigma^2 = 0.5 \)

- Stretching \( \sigma^2 = 0.5 \frac{100^2}{64^2} \approx 1.22 \)
Stretching means to keep the coefficient of variation \( \frac{\sqrt{\text{Variance}}}{\text{Expectation}} \) constant.
Stochastic Reserving

Chain-Ladder-Method (CLM)

Validation and examples (part 2 of 3)
2 Chain-Ladder-Method (CLM)

2.1 How does the Chain-Ladder method work
2.1.1 Chain-Ladder method without stochastic
2.1.2 Stochastic behind the Chain-Ladder method

2.2 Future development
2.2.1 Projection of the future development

2.3 Validation and examples (part 1 of 3)
2.3.1 Chain-Ladder method on Payments and on Incurred
2.3.2 How to validate the Chain-Ladder assumptions

2.4 Ultimate uncertainty
2.4.1 Ultimate uncertainty of accident period \( i \)
2.4.2 Ultimate uncertainty of the aggregation of all accident periods

2.5 Validation and examples (part 2 of 3)
2.5.1 Ultimate uncertainty

2.6 Solvency uncertainty
2.6.1 Solvency uncertainty of a single accident period
2.6.2 Solvency uncertainty of all accident periods
2.6.3 Uncertainties of further CDR’s

2.7 Validation and examples (part 3 of 3)
2.7.1 Solvency uncertainty

2.8 Literature
2 Chain-Ladder-Method (CLM)

2.1 How does the Chain-Ladder method work
2.1.1 Chain-Ladder method without stochastic
2.1.2 Stochastic behind the Chain-Ladder method

2.2 Future development
2.2.1 Projection of the future development

2.3 Validation and examples (part 1 of 3)
2.3.1 Chain-Ladder method on Payments and on Incurred
2.3.2 How to validate the Chain-Ladder assumptions

2.4 Ultimate uncertainty
2.4.1 Ultimate uncertainty of accident period
2.4.2 Ultimate uncertainty of the aggregation of all accident periods

2.5 Validation and examples (part 2 of 3)
2.5.1 Ultimate uncertainty

2.6 Solvency uncertainty
2.6.1 Solvency uncertainty of a single accident period
2.6.2 Solvency uncertainty of all accident periods
2.6.3 Uncertainty of further CDR’s

2.7 Validation and examples (part 3 of 3)
2.7.1 Solvency uncertainty

2.8 Literature
Claims development result and solvency uncertainty (repetition)

The observed claims development result (CDR) at time $I + 1$ of a single accident period $i$ is the (observed) difference of the estimated ultimates of estimation time $I$ and estimation time $I + 1$:

$$
\hat{\text{CDR}}_{i}^{I+1} := \hat{C}_{i,J}^{I} - \hat{C}_{i,J}^{I+1}.
$$

Here and in the following we denote (if necessary) the time of estimation by an additional upper index.

A negative CDR corresponds to a loss and a positive CDR corresponds to a profit. Moreover, in the Best Estimate case the estimate of the conditionally expected CDR is zero, i.e.

$$
\hat{E} \left[ \hat{\text{CDR}}_{i}^{I+1} \mid \mathcal{D}^I \right] = 0.
$$

The solvency uncertainty of a single accident period $i$ is defined as the mse of the $\hat{\text{CDR}}_{i}^{I+1}$ conditioned under all information at time $I$, i.e.

$$
mse_{0\mid \mathcal{D}^I} \left[ \hat{\text{CDR}}_{i}^{I+1} \right] := E \left[ \left( \hat{\text{CDR}}_{i}^{I+1} - 0 \right)^2 \mid \mathcal{D}^I \right] = Var \left[ \hat{C}_{i,J}^{I+1} \mid \mathcal{D}^I \right] + E \left[ \hat{C}_{i,J}^{I+1} - \hat{C}_{i,J}^{I} \mid \mathcal{D}^I \right]^2.
$$

random error \hspace{1cm} parameter error
Claims development result and solvency uncertainty

The observed claims development result (CDR) at time after a single accident period $i$ is the (observed) difference of the estimated ultimates of estimation time $I$ and estimation time $I+1$:

$$\hat{\text{CDR}}_{I+1}^i := \hat{C}_{I,i,J} - \hat{C}_{I+1,i,J}.$$  

Here and in the following we denote (if necessary) the time of estimation by an additional upper index.

A negative CDR corresponds to a loss and a positive CDR corresponds to a profit. Moreover, in the Best Estimate case the estimate of the conditionally expected CDR is zero, i.e.

$$\mathbb{E}\left[\hat{\text{CDR}}_{I+1}^i | \mathcal{D}_I\right] = 0.$$  

The solvency uncertainty of a single accident period $i$ is defined as the mse of the CDR $\hat{\text{CDR}}_{I+1}^i$ conditioned under all information at time $I$, i.e.

$$\text{mse} \left[\hat{\text{CDR}}_{I+1}^i | \mathcal{D}_I\right] := \mathbb{E} \left[ (\hat{\text{CDR}}_{I+1}^i - 0)^2 | \mathcal{D}_I \right] = \mathbb{E} \left[ (\hat{C}_{I,i,J} - \hat{C}_{I+1,i,J})^2 | \mathcal{D}_I \right] \triangleq \text{random error} + \mathbb{E} \left[ (\hat{C}_{I+1,i,J} - \hat{C}_{I,i,J})^2 | \mathcal{D}_I \right] \triangleq \text{parameter error}.$$
Assumption 2.B (Consistent estimates over time)

In order to have consistent estimates at times \( I \) and \( I + 1 \) we assume that there exist \( D^I \cap D^k \)-measurable weights \( 0 \leq w^I_{I-k,k} \leq 1 \) with

- \( C_{I-k,k} = 0 \) implies \( w^I_{I-k,k} = 0 \),
- \( w^I_{i,k} := (1 - w^I_{I-k,k})w^I_{i,k} \), for \( 0 \leq i \leq I - 1 - k \).

Remark 2.18

The above assumption means that we do not change our (relative) believes into the old development periods and only put some credibility \( w^I_{I-k,k} \) to the new encountered development. The variance minimizing weights, introduced in Lemma 2.4, satisfy Assumption 2.B.

Lemma 2.19 (Consistent estimates over time)

Let Assumptions 2.A and 2.B be fulfilled. Then we have

1. \( \hat{f}^I_{k} + 1 = (1 - w^I_{I-k,k})\hat{f}^I_{k} + w^I_{I-k,k} C_{I-k,k+1} = (1 - w^I_{I-k,k})\hat{f}^I_{k} + w^I_{I-k,k} F_{I-k,k} \),
2. \( \bar{f}_k := E[\hat{f}^I_{k} + 1 | D^I] = E[\hat{f}^I_{k} + 1 | D^k] = (1 - w^I_{I-k,k})\hat{f}^I_{k} + w^I_{I-k,k} f_k \approx \hat{f}^I_{k} \),
3. \( \bar{C}_{i,J} := E[\hat{C}^I_{i,J} | D^I] = \prod_{k=I+1-i}^{J-1} \bar{f}_k f_{I-i} C_{i,I-i} \),
4. \( \hat{E}[\hat{\text{CDR}}_i^I + 1 | D^I] = 0 \), which means we have a Best Estimate.
1. \( \hat{f}_{k+1} := \sum_{i=0}^{I-k} w_{i,k}^+ \frac{C_{i,k+1}}{C_{i,k}} = (1 - w_{I-k,k}^+) \sum_{i=0}^{I-k-1} w_{i,k}^+ \frac{C_{i,k+1}}{C_{i,k}} + w_{I-k,k}^+ \frac{C_{I-k,k+1}}{C_{I-k,k}} \)

\( = (1 - w_{I-k,k}^+) \hat{f}_k^+ + w_{I-k,k}^+ \frac{C_{I-k,k+1}}{C_{I-k,k}} = (1 - w_{I-k,k}^+) \hat{f}_k^+ + w_{I-k,k}^+ F_{I-k,k} \)

2. \( \Rightarrow \mathbb{E}[\hat{f}_{k+1}|D^I] = \mathbb{E}[\hat{f}_k^+|D^I] = (1 - w_{I-k,k}^+) \hat{f}_k^+ + w_{I-k,k}^+ F_{I-k,k} \)

\( \Rightarrow \hat{f}_k = \hat{f}_k^+ \)

3. \( \mathbb{E}[\hat{C}_{i,j}^I|D^I] = \mathbb{E} \left[ \prod_{k=I+1-i}^{J-1} \hat{f}_{k,k+1} C_{i,k+1-i} \Big| D^I \right] = \mathbb{E} \left[ \mathbb{E}[\hat{f}_{k+1}|D^I] \prod_{k=I+1-i}^{J-2} \hat{f}_{k,k+1} C_{i,k+1-i} \Big| D^I \right] \)

\( = \hat{f}_{J-1} \mathbb{E} \left[ \prod_{k=I+1-i}^{J-2} \hat{f}_{k,k+1} C_{i,k+1-i} \Big| D^I \right] = \ldots \)

\( = \prod_{k=I+1-i}^{J-1} \hat{f}_k \mathbb{E}[C_{i,k+1-i}|B_i,k-I-i] = \prod_{k=I+1-i}^{J-1} \hat{f}_k F_{I-I-k} C_{i,k+1-i} \)

4. \( \mathbb{E}[\hat{C}_{i,j}^I|D^I] = \mathbb{E}[\hat{C}_{i,j}^I|D^I] - \mathbb{E}[\hat{C}_{i,j}^I|D^I] = \prod_{k=I-i}^{J-1} \hat{f}_k C_{i,i-k} - \prod_{k=I+1-i}^{J-1} \hat{f}_k F_{I-I-k} C_{i,i-k} \)

\( \approx \prod_{k=I-i}^{J-1} \hat{f}_k C_{i,i-k} - \prod_{k=I+1-i}^{J-1} \hat{f}_k F_{I-I-k} C_{i,i-k} = 0 \)
Taylor approximation of next years estimates

Recall the (multi-linear) functional:

\[ U_i(g) x := g_{J-1} \cdots g_{I-i} x. \]

Then we get:

\[ \frac{\partial}{\partial g_j} U_i(g) x = g_{J-1} \cdots g_{j+1} g_{j-1} \cdots g_{I-i} x = \frac{U_i(g) x}{g_j}, \]

\[ U_i\left(\hat{f}^I\right) C_{i,I-i} = \hat{f}_{J-1}^I \cdots \hat{f}_{I-i}^I C_{i,I-i} = \hat{C}_{i,J}^I, \]

\[ U_i\left(F_{i+1}^I\right) C_{i,I-i} = \hat{f}_{J-1}^{I+1} \cdots \hat{f}_{I-i+1}^{I+1} F_{i,I-i} C_{i,I-i} = \hat{C}_{i,J}^{I+1} \text{ and} \]

\[ \hat{C}_{i,J}^{I+1} - \hat{C}_{i,J}^I \approx \sum_{k=I-i}^{J-1} \left( \frac{\partial}{\partial F_{i+1,k}^I} U_i\left(F_{i+1}^I\right) \right) \bigg|_{\hat{f}^I} C_{i,I-i}^I \left( F_{i+1,k}^I - \hat{f}_k^I \right), \]

\[ = \frac{C_{i,J}^I}{\hat{f}_{I-i}^I} \left( F_{i,I-i} - \hat{f}_{I-i}^I \right) + \sum_{k=I-i+1}^{J-1} \frac{C_{i,J}^I}{\hat{f}_k^I} w_{I-k,k}^{I+1} \left( F_{I-k,k}^I - \hat{f}_k^I \right) \]

where we used a first order Taylor approximation and \( \hat{f}^I \) denotes the vector of the at time \( I \) estimated development factors and \( F_{i+1}^I \) is a vector with components

\[ F_{i,k}^{I+1} := \begin{cases} \hat{f}_k^{I+1}, & \text{for } i + k > I, \\ F_{i,k}, & \text{for } i + k = I. \end{cases} \]

The red parts are the difference to the ultimate uncertainty case.
Taylor approximation of next year’s estimates

Recall the (multi-linear) functional:

\[ U_i(g) = g_{J-1} \cdots g_{I-i} x. \]

Then we get:

\[ \frac{\partial}{\partial g_j} U_i(g) = g_{J-1} \cdots g_{j+1} g_j - 1 \cdots g_{I-i} x = U_i(g) x g_j, \]

\[ U_i(\hat{f}_I) = C_i,i - \hat{f}_I \]

and

\[ U_i(F_{I+1}) = C_i,i - \hat{f}_I + \sum_{k=I-i}^{I} \hat{f}_k \]

where we used a first order Taylor approximation and \( \hat{f} \) denotes the vector of the at time \( I \) estimated development factors and \( F_I^{(1)} \) is a vector with components

\[ \hat{f}_I = \begin{cases} \hat{f}_I & \text{for } i+k > I, \\ F_{i,k} & \text{for } i+k = I. \end{cases} \]

The red parts are the difference to the ultimate uncertainty case.

For \( k = I - i \) we get

\[ F_{i,I-i}^{I+1} - \hat{f}_{I-i} = F_{i,I-i} - \hat{f}_{I-i} \]

and for \( k > I - i \) it is

\[ F_{i,k}^{I+1} - \hat{f}_k = \hat{f}_k^{I+1} - \hat{f}_k = (1 - w_{I-k,k}^{I+1}) \hat{f}_k + w_{I-k,k}^{I+1} F_{I-k,k} - \hat{f}_k = w_{I-k,k}^{I+1} \left( F_{I-k,k} - \hat{f}_k \right) \]
2.6 Solvency uncertainty

2.6.1 Solvency uncertainty of a single accident period

Linear approximation of the CDR

If we replace in the linear approximation of the ultimate, i.e. in

\[ C_{i,J} - \hat{C}_{i,J} \approx \sum_{k=I-i}^{J-1} \frac{\hat{C}_{i,J}}{\hat{f}_k^I} \left( F_{i,k} - \hat{f}_k^I \right), \]

the term \( \left( F_{i,k} - \hat{f}_k^I \right) \) by

\[
\tilde{F}_{i,k}^I - \tilde{f}_i^I := \begin{cases} 
F_{I-k,k} - \hat{f}_k^I, & \text{for } k = I - i, \\
\omega_{I-k,k}^{I+1} \left( F_{I-k,k} - \hat{f}_k^I \right), & \text{for } k > I - i, 
\end{cases}
\]

we get the linear approximation of the CDR, i.e.

\[
\hat{C}_{i+1,J} - \hat{C}_{i,J} \approx \frac{\hat{C}_{i,J}}{\hat{f}_{I-i}^I} \left( F_{i,I-i} - \hat{f}_{I-i}^I \right) + \sum_{k=I-i+1}^{J-1} \frac{\hat{C}_{i,J}}{\hat{f}_k^I} \omega_{I-k,k}^{I+1} \left( F_{I-k,k} - \hat{f}_k^I \right)
\]

\[
= \sum_{k=I-i}^{J-1} \frac{\hat{C}_{i,J}}{\hat{f}_k^I} \left( \tilde{F}_{i,k}^I - \tilde{f}_i^I \right).
\]
Linear approximation of the CDR

If we replace in the linear approximation of the ultimate, i.e., in
\[ C_{i,J}^I \approx \sum_{k=I-i}^{J-1} \hat{C}_{i,J}^I \hat{f}_{I,k} (F_{i,k} - \hat{f}_{I,k}), \]
the term \((F_{i,k} - \hat{f}_{I,k})\) by
\[ \tilde{F}_{I,i,k} - \tilde{\hat{f}}_{I,i,k} = \begin{cases} F_{i,k} - \hat{f}_{I,k} & \text{for } k = I-i, \\ w_{I+1}^{I-i} (F_{i,k} - \hat{f}_{I,k}) & \text{for } k > I-i, \end{cases} \]
we get the linear approximation of the CDR, i.e.,
\[ C_{i,J}^{I+1} - C_{i,J}^I \approx \sum_{k=I-i}^{J-1} \hat{C}_{i,J}^I \hat{f}_{I,k} (\tilde{F}_{I,i,k} - \tilde{\hat{f}}_{I,i,k}). \]

The term \( \tilde{F}_{i,k} - \tilde{\hat{f}}_{i,k} \) depends on the accident period \( i \) only via the indicator functions \( 1_{k=I-i} \) and \( 1_{k> I-i} \).
Estimator 2.20 (Solvency uncertainty of accident period $i$)

$$\text{mse}_{0|D} \left[ \hat{\text{CDR}}_i \right] = E \left[ \left( \hat{C}_{i,J}^{I+1} - \hat{C}_{i,J}^I \right)^2 \right]$$

$$\approx E \left[ \left( \sum_{k=I-i}^{J-1} \frac{\hat{C}_{i,J}^I}{\hat{f}_k^I} \left( F_{k,k}^{I+1} - \hat{f}_k^I \right) \right)^2 \right] (\text{Taylor approximation})$$

$$= \sum_{k_1,k_2=I-i}^{J-1} \frac{\hat{C}_{i,J}^I}{\hat{f}_{k_1}^I} \frac{\hat{C}_{i,J}^I}{\hat{f}_{k_2}^I} \left( 1_{k_1=I-i} + 1_{k_1>I-i} w_{I-k_1}^{I+1} \right) \left( 1_{k_2=I-i} + 1_{k_2>I-i} w_{I-k_2}^{I+1} \right)$$

$$E \left[ \left( F_{I-k_1,k_1} - \hat{f}_{k_1}^I \right) \left( F_{I-k_2,k_2} - \hat{f}_{k_2}^I \right) \right]$$

$$\approx \sum_{k=I-i}^{J-1} \frac{\hat{\sigma}_k^2}{\left( \hat{f}_k^I \right)^2} \left( 1_{k=I-i} + 1_{k>I-i} w_{I-k}^{I+1} \right) \left( \hat{C}_{i,J}^I \right)^2 \frac{1}{C_{I-k,k}}$$

random error

$$+ \sum_{k=I-i}^{J-1} \frac{\hat{\sigma}_k^2}{\left( \hat{f}_k^I \right)^2} \left( 1_{k=I-i} + 1_{k>I-i} w_{I-k}^{I+1} \right) \left( \hat{C}_{i,J}^I \right)^2 \sum_{h=0}^{I-k-1} \frac{(w_{h,k})^2}{C_{h,k}}$$

parameter error
From the derivation of the ultimate uncertainty we already know

\[
\mathbb{E}\left[ (F_{I-k_f},k_f - \hat{f}_{k_f}) (F_{I-k_f},k_f - \hat{f}_{k_f}) \mid D^I \right] = \text{Cov}\left[ F_{I-k_f},k_f , F_{I-k_f},k_f \mid D^I \right] + (\hat{f}_{k_f} - f_{k_f}) (\hat{f}_{k_f} - f_{k_f}) \\
\approx \text{Cov}\left[ F_{I-k_f},k_f , F_{I-k_f},k_f \mid D^I \right] + \text{Cov}\left[ \hat{f}_{k_f} , \hat{f}_{k_f} \mid D_{k_f} \cap k_f \right] \\
\approx 1_{k_f=k_f} \left( \frac{\hat{\sigma}^2_{k_f}}{C_{I-k_f},k_f} + \sum_{h=0}^{I-k_f-1} \frac{\hat{\sigma}^2_{k_f} (w_{I,k_f})^2}{C_{h,k_f}} \right).
\]

Therefore (the red terms are the differences to the ultimate uncertainty case),

\[
\text{mse}_{\mid D^I} \left[ \text{CDR}_i \right] \\
\approx \sum_{k=I-i}^{J-1} \left( \frac{\hat{C}_{I,J}}{\hat{f}_{k_f}} \hat{C}_{I,J} \right)\left( 1_{k_f=I-k_f} + 1_{k_f>I-k_f} w_{I-k_f,k_f} \right)^2 \left( 1_{k_f=I-k_f} + 1_{k_f>I-k_f} w_{I-k_f,k_f} \right) \\
\quad \times 1_{k_f=k_f} \left( \frac{\hat{\sigma}^2_{k_f}}{C_{I-k_f},k_f} + \sum_{h=0}^{I-k_f-1} \frac{\hat{\sigma}^2_{k_f} (w_{I,h,k_f})^2}{C_{h,k_f}} \right) \\
= \sum_{k=I-i}^{J-1} \left( \frac{\hat{C}_{I,J}}{\hat{f}_{k_f}} \right)^2 \left( 1_{k_f=I-k_f} + 1_{k_f>I-k_f} w_{I-k_f,k_f} \right)^2 \left( \frac{\hat{\sigma}^2_{k_f}}{C_{I-k_f,k_f}} + \sum_{h=0}^{I-k_f-1} \frac{\hat{\sigma}^2_{k_f} (w_{I,h,k_f})^2}{C_{h,k_f}} \right).
\]
Ultimate uncertainty for accident period $i$

$$
\text{mse}_{D_i} \left[ \hat{C}_{i,J} \right] \approx \sum_{k=I-i}^{J-1} \frac{\hat{\sigma}_k^2}{(\hat{f}_k^I)^2} \left( \hat{C}_{i,J} \right)^2 \frac{1}{\hat{C}_{i,k}} + \sum_{k=I-i}^{J-1} \frac{\hat{\sigma}_k^2}{(\hat{f}_k^I)^2} \left( \hat{C}_{i,J} \right)^2 \sum_{h=0}^{I-k-1} \frac{\left( w_{h,k}^I \right)^2}{C_{h,k}}
$$

Random error

Parameter error

Solvency uncertainty for accident period $i$

$$
\text{mse}_{0|D_i} \left[ \text{CDR}_{i} \right] \approx \sum_{k=I-i}^{J-1} \frac{\hat{\sigma}_k^2}{(\hat{f}_k^I)^2} \left( 1_{k=I-i} + 1_{k>I-i} w_{I-k,k}^I \right) \left( \hat{C}_{i,J} \right)^2 \frac{1}{C_{I-k,k}} + \sum_{k=I-i}^{J-1} \frac{\hat{\sigma}_k^2}{(\hat{f}_k^I)^2} \left( 1_{k=I-i} + 1_{k>I-i} w_{I-k,k}^I \right) \left( \hat{C}_{i,J} \right)^2 \sum_{h=0}^{I-k-1} \frac{\left( w_{h,k}^I \right)^2}{C_{h,k}}
$$

Random error

Parameter error
Stochastic Reserving

Chain-Ladder-Method (CLM)

Solvency uncertainty

It almost looks like a simple multiplication by the factor

\[
(1_{k=I-i} + 1_{k>0} w_{I-k,k}^{I+1})
\]

except for the index replacement \(i\) by \(I-k\) in the random error part.
Corollary 2.21

If we use the variance minimizing weights

\[ w_{i,k}^I = \frac{C_{i,k}}{\sum_{h=0}^{I-k-1} C_{h,k}} \quad \text{and} \quad w_{i,k}^{I+1} = \frac{C_{i,k}}{\sum_{h=0}^{I-k} C_{h,k}} \]

we get for the solvency uncertainty of accident period \( i \)

\[
\text{mse}_{0|D_I} \left[ \hat{\text{CDR}}_i \right] \\
\approx \hat{C}_{i,J}^2 \sum_{k=I-i}^{J-1} \frac{\hat{\sigma}_{k}^2}{\left( \hat{f}_k^I \right)^2} \left( 1_{k=I-i} + 1_{k>I-i} \frac{C_{I-k,k}^2}{\left( \sum_{h=0}^{I-k} C_{h,I-k} \right)^2} \right) \left( \frac{1}{C_{I-k,k}} + \sum_{h=0}^{I-k-1} \frac{C_{h,k}^2}{C_{h,k} \left( \sum_{v=0}^{I-k-1} C_{v,k} \right)^2} \right)
\]

\[
= \hat{C}_{i,J}^2 \sum_{k=I-i}^{J-1} \frac{\hat{\sigma}_{k}^2}{\left( \hat{f}_k^I \right)^2} \left( 1_{k=I-i} + 1_{k>I-i} \frac{C_{I-k,k}^2}{\left( \sum_{h=0}^{I-k} C_{h,I-k} \right)^2} \right) \left( \frac{1}{C_{I-k,k}} + \frac{1}{\sum_{h=0}^{I-k-1} C_{h,k}} \right),
\]

where the red terms indicate the differences to the ultimate uncertainty case.
Corollary 2.21
If we use the variance minimizing weights

\[ w_{i,k} = \frac{C_{i,k}}{\sum_c C_{i,k}} \quad \text{and} \quad w_{i,k+1} = \frac{C_{i,k}}{\sum_c C_{i,k}} \]

we get for the solvency uncertainty of accident period

\[ \text{mse}_0 | D = \hat{C}_{CDR,i} \approx \hat{C}^2_{i,J} - \hat{\sigma}^2_k (\hat{f}_{I,k}) \]

where the red terms indicate the differences to the ultimate uncertainty case.
Dependent accident periods

Since $\tilde{F}_{i,k}^I$ and $\tilde{f}_k^I$ depend on $F_{I-k,k}^I = C_{I-k,k+1}/C_{I-k,k}$, for all $i$, accident periods are not independent. Therefore, we cannot simply take the sum over all accident periods in order to derive the solvency uncertainty of the aggregation of all accident periods.

But the Taylor approximation still works:
Since \( \tilde{F}_{i,k} \) and \( \tilde{f}_{i,k} \) depend on \( \tilde{F}_{i-1,k} = C_{i-1,k+1}/C_{i-1,k} \) for all \( i \), accident periods are not independent. Therefore, we cannot simply take the sum over all accident periods in order to derive the solvency uncertainty of the aggregation of all accident periods. But the Taylor approximation still works.
Estimator 2.22 (Solvency uncertainty of all accident periods)

\[
\text{mse}_{0|D^I} \left[ \sum_{i=0}^{I} \widehat{\text{CDR}}_i \right] = \mathbb{E} \left[ \left( \sum_{i=0}^{I} \left( \widehat{C}_{i,J}^{I+1} - \widehat{C}_{i,J}^I \right) \right)^2 \right| D^I \\
\approx \mathbb{E} \left[ \left( \sum_{i=0}^{I} \sum_{k=I-i}^{J-1} \frac{\widehat{C}_{i_1,J}^I}{\widehat{f}^I_{k_1}} \left( \widehat{F}_{i,k}^I - \bar{f}_{i,k}^I \right) \right)^2 \right| D^I \\
= \sum_{i_1,i_2=0}^{I} \sum_{k_1=I-i_1}^{J-1} \sum_{k_2=I-i_2}^{J-1} \frac{\widehat{C}_{i_1,J}^I \widehat{C}_{i_2,J}^I}{\widehat{f}^I_{k_1} \widehat{f}^I_{k_2}} \mathbb{E} \left[ \left( F_{I-k_1,k_1} - \bar{f}^I_{k_1} \right) \left( F_{I-k_2,k_2} - \bar{f}^I_{k_2} \right) \right| D^I \\
\approx \sum_{k=0}^{J-1} \frac{\hat{\sigma}_k^2}{\widehat{f}^I_k} \left( \sum_{i=I-k}^{I} \left( 1_{k_1=I-i_1} + 1_{k_1>I-i_1} w_{I-k_1,k_1}^{I+1} \right) \widehat{C}_{i,J}^I \right)^2 \frac{1}{C_{I-k,k}} \\
\quad + \sum_{k=0}^{J-1} \frac{\hat{\sigma}_k^2}{\widehat{f}^I_k} \left( \sum_{i=I-k}^{I} \left( 1_{k_2=I-i_2} + 1_{k_2>I-i_2} w_{I-k_2,k_2}^{I+1} \right) \widehat{C}_{i,J}^I \right)^2 I-k-1 \sum_{h=0}^{I-k-1} \left( w_{h,k}^I \right)^2 \frac{1}{C_{h,k}}
\]

random error

\[\text{parameter error}\]
Therefore, we get

\[ E \left[ (F_{I-k_1,k_1} - \hat{f}_{k_1}^I)(F_{I-k_2,k_2} - \hat{f}_{k_2}^I) | \mathcal{D}^I \right] \approx 1_{k_1=k_2} \left( \frac{\hat{\sigma}^2_{k_1}}{C_{I-k_1,k_1}} + \sum_{h=0}^{I-k_1-1} \frac{\hat{\sigma}^2_{k_1}}{C_{h,k_1}} \left( w_{h,k_1}^I \right)^2 \right) . \]

From the single accident period case we know

\[ \text{mse}_{0|\mathcal{D}^I} \left[ \sum_{i=0}^{I} \hat{CDR}_i \right] \]

\[ \approx \sum_{i_1,i_2=0}^{I} \sum_{k_1=I-i_1}^{J-1} \sum_{k_2=I-i_2}^{J-1} \frac{\hat{C}_{i_1,J}^I}{\hat{f}_{k_1}^I} \frac{\hat{C}_{i_2,J}^I}{\hat{f}_{k_2}^I} 1_{k_1=k_2} \left( 1_{k_1=I-i_1} + 1_{k_1>I-i_1} w_{I-k_1,k_1}^{I+1} \right) \left( 1_{k_2=I-i_2} + 1_{k_2>I-i_2} w_{I-k_2,k_2}^{I+1} \right) \]

\[ = \sum_{k_1,k_2=0}^{J-1} \sum_{i_1=I-k_1}^{I} \sum_{i_2=I-k_2}^{I} \frac{\hat{C}_{i_1,J}^I}{\hat{f}_{k_1}^I} \frac{\hat{C}_{i_2,J}^I}{\hat{f}_{k_2}^I} 1_{k_1=k_2} \left( 1_{k_1=I-i_1} + 1_{k_1>I-i_1} w_{I-k_1,k_1}^{I+1} \right) \left( 1_{k_2=I-i_2} + 1_{k_2>I-i_2} w_{I-k_2,k_2}^{I+1} \right) \]

\[ = \sum_{k=0}^{J-1} \frac{\hat{\sigma}^2_{k}}{\left( \hat{f}_{k}^I \right)^2} \left( \sum_{i=I}^{I} \left( 1_{k_1=I-i} + 1_{k_1>I-i} w_{I-k_1,k_1}^{I+1} \right) \hat{C}_{i,J}^I \right)^2 \left( \frac{1}{C_{I-k_1}} + \sum_{h=0}^{I-k_1-1} \frac{\hat{\sigma}^2_{k}}{C_{h,k_1}} \left( w_{h,k_1}^I \right)^2 \right) . \]
2.6 Solvency uncertainty

2.6.2 Solvency uncertainty of all accident periods

Ultimate uncertainty of all accident periods

\[
\text{mse}_D \left[ \hat{C}_{i,J} \right] \approx \sum_{k=0}^{J-1} \frac{\hat{\sigma}_k^2}{(\hat{f}_k^I)^2} \sum_{i=I-k}^{I} \left( \hat{C}_{i,J} \right)^2 \frac{1}{\hat{C}_{i,k}}
\]

random error

\[
+ \sum_{k=0}^{J-1} \frac{\hat{\sigma}_k^2}{(\hat{f}_k^I)^2} \left( \sum_{i=I-k}^{I} \hat{C}_{i,J} \right)^2 \sum_{h=0}^{I-k-1} \frac{k \left( w_{h,k}^I \right)^2}{C_{h,k}}
\]

parameter error

Estimator 2.23 (Solvency uncertainty of all accident periods)

\[
\text{mse}_{0,D} \left[ \sum_{i=0}^{I} \overline{\text{CDR}}_i \right] \approx \sum_{k=0}^{J-1} \frac{\hat{\sigma}_k^2}{(\hat{f}_k^I)^2} \left( \sum_{i=I-k}^{I} \left( 1_{k=I-i} + 1_{k>I-i} w_{I-k,k}^{I+1} \right) \hat{C}_{i,J} \right)^2 \frac{1}{C_{I-k,k}}
\]

random error

\[
+ \sum_{k=0}^{J-1} \frac{\hat{\sigma}_k^2}{(\hat{f}_k^I)^2} \left( \sum_{i=I-k}^{I} \left( 1_{k=I-i} + 1_{k>I-i} w_{I-k,k}^{I+1} \right) \hat{C}_{i,J} \right)^2 \sum_{h=0}^{I-k-1} \frac{\left( w_{h,k}^I \right)^2}{C_{h,k}}
\]

parameter error
Stochastic Reserving

- Chain-Ladder-Method (CLM)
- Solvency uncertainty
Corollary 2.24

If we use the variance minimizing weights

\[ w^I_{i,k} = \frac{C_{i,k}}{I-k-1} \sum_{h=0}^{I-k-1} C_{h,k} \]

and

\[ w^{I+1}_{i,k} = \frac{C_{i,k}}{I-k} \sum_{h=0}^{I-k} C_{h,k} \]

we get for the solvency uncertainty of all accident periods

\[
\text{mse}_{0|D^I \left[ \hat{\text{CDR}}_i \right]} \approx \left( \sum_{i=0}^{I} \hat{C}_{i,J}^I \right)^2 \sum_{k=0}^{J-1} \left( \frac{\hat{\sigma}_k^2}{\left( \hat{f}_k^I \right)^2} \right) \left( \frac{1}{\sum_{h=0}^{I-k-1} C_{h,k}} - \frac{1}{\sum_{h=0}^{I-k} C_{h,k}} \right),
\]

where the red term indicate the difference to the ultimate uncertainty case.
Corollary 2.24

If we use the variance minimizing weights

\[ w_{i,k} = \frac{C_{i,k}}{\sum_{h=0}^{I-1} C_{h,k}} \]

and

\[ w_{i+1,k} = \frac{C_{i,k}}{\sum_{h=0}^{I-1} C_{h,k}} \]

we get for the solvency uncertainty of all accident periods

\[
\text{mse}_{0|D[I]}[\hat{\text{CDR}}_i] \approx \sum_{k=0}^{I-1} \frac{\sigma_k^2}{(\bar{f}_k)^2} \left( \sum_{i=0}^{I} \hat{C}_{i,J} \right)^2 \left( \frac{C_{I-k,k}}{\sum_{h=0}^{I-k} C_{h,k}} \right)^2 \left( \frac{1}{C_{I-k,k}} + \sum_{h=0}^{I-k-1} \frac{1}{C_{h,k}} \right)
\]

Therefore, we get for the solvency uncertainty of all accident periods

\[
\text{mse}_{0|D[I]}[\hat{\text{CDR}}_i] \approx \sum_{k=0}^{I-1} \frac{\sigma_k^2}{(\bar{f}_k)^2} \left( \sum_{i=0}^{I} \hat{C}_{i,J} \right)^2 \left( \frac{C_{I-k,k}}{\sum_{h=0}^{I-k} C_{h,k}} \right)^2 \left( \frac{1}{C_{I-k,k}} + \sum_{h=0}^{I-k-1} \frac{1}{C_{h,k}} \right)
\]
Estimation at time $n \geq I$

Analogously to the next years estimation we can look at the estimation of the ultimate at any time $n \geq I$

$$\hat{C}_{i,J}^n := C_{i,n-i} \prod_{k=n-i}^{J-1} \hat{f}_k^n = C_{i,I-i} \prod_{k=I-i}^{n-i-1} F_{i,k} \prod_{k=n-i}^{J-1} \hat{f}_k^n.$$ 

The development factors are estimated by

$$\hat{f}_k^n := \sum_{h=0}^{n-k-1} w_{h,k}^n F_{h,k}$$

with consistent future weights $w_{i,k}^n$. That means for $I - k \leq i \leq n - k - 1$, there exists $\mathcal{D}_k^n$-measurable weights $0 \leq w_{i,k}^n \leq 1$ with

- $C_{i,k} = 0$ implies $w_{i,k}^n = 0$,
- $w_{i,k}^n = (1 - w_{n-k,k}^n)w_{i,k}^{n-1}$, for $i + k < n$. 

Estimation at time $n \geq I$

Analogously to the next years estimation we can look at the estimation of the ultimate at any time $n \geq I$

$$\hat{C}_{n,i,J} := C_{i,n} - i \prod_{k=n-i}^{n-I} \hat{f}_{k}$$

The development factors are estimated by

$$\hat{f}_{n,k} := \sum_{h=0}^{n-k-1} w_{n,i,k} F_{h,k}$$

with consistent future weights $w_{n,i,k}$. That means for $I - k \leq i \leq n - k - 1$, there exists $\Gamma_{n}$-measurable weights $0 \leq w_{n,i,k} \leq 1$ with

- $C_{i,k} = 0$ implies $w_{n,i,k} = 0$.
- $w_{n,i,k} = (1 - w_{n,i,k}) w_{n,i,k}$, for $i + k < n$. 

Stochastic Reserving

---

Chain-Ladder-Method (CLM)

Solvency uncertainty
Claims development result between two estimation time $I \leq n_1 < n_2$

Since formulas will get very tedious (see for instance [30]), if one analyses the CDR with respect to two time periods $I \leq n_1 < n_2$ analogously to the next year claim development result, we will only consider the special case of variance minimizing weights

$$w_{i,k}^n := \frac{C_{i,k}}{\sum_{h=0}^{n-k-1} C_{h,k}},$$

which leads to the following estimates (at time $n$) of the development factors

$$\hat{f}_k^n := \sum_{i=0}^{n-k-1} w_{i,k}^n \frac{C_{i,k+1}}{C_{i,k}} = \frac{\sum_{i=0}^{n-k-1} C_{i,k+1}}{\sum_{i=0}^{n-k-1} C_{i,k}}.$$

In this case we have

$$\sum_{i=0}^{I} \hat{C}_{i,J}^n = \sum_{i=0}^{I} C_{i,0} \prod_{k=0}^{J-1} \hat{f}_k^n.$$
For each $k > 0$ we have

$$
\sum_{i=0}^{I} \widehat{C}_{i,k+1}^{n} = \sum_{i=0}^{n-k-1} C_{i,k+1} + \sum_{i=n-k}^{I} f_{k}^{n} \widehat{C}_{i,k}^{n}
$$

$$
= \frac{\sum_{i=0}^{n-k-1} C_{i,k+1}}{\sum_{i=0}^{n-k-1} C_{i,k}} \sum_{i=0}^{n-k-1} C_{i,k} + f_{k}^{n} \sum_{i=n-k}^{I} \widehat{C}_{i,k}^{n}
$$

$$
= \widehat{f}_{k}^{n} \sum_{i=0}^{n-k-1} C_{i,k} + \widehat{f}_{k}^{n} \sum_{i=n-k}^{I} \widehat{C}_{i,k}^{n}
$$

$$
= \widehat{f}_{k}^{n} \sum_{i=0}^{I} \widehat{C}_{i,k}^{n},
$$

which by induction proves (2.4).
Estimator 2.25 (Uncertainty of the CDR\(^{n_1,n_2}\) with variance minimizing weights)

In the case of variance minimizing weights (2.3) the uncertainty of the claims development result \(\sum_{i=0}^{I}(\hat{C}_{i,J}^{n_2} - \hat{C}_{i,J}^{n_1})\) between two time periods \(I \leq n_1 < n_2\) can be estimated by

\[
\text{mse}_{0|D_I} \left[ \text{CDR}_{n_1,n_2} \right] = E \left[ \left( \sum_{i=0}^{I} \left( \hat{C}_{i,J}^{n_2} - \hat{C}_{i,J}^{n_1} \right) \right)^2 \right| D_I]
\]

\[
= \left( \sum_{i=0}^{I} C_{i,0} \right) \left( \sum_{i=0}^{I} C_{i,J} \right) \left( \prod_{k=0}^{J-1} \hat{f}_{k}^{n_2} - \prod_{k=0}^{J-1} \hat{f}_{k}^{n_1} \right)^2 \left| D_I \right]
\]

\[
\approx \left( \sum_{i=0}^{I} C_{i,J} \right) \left( \prod_{k=0}^{J-1} \hat{f}_{k}^{n_2} - \prod_{k=0}^{J-1} \hat{f}_{k}^{n_1} \right)^2 \left( 1 + \frac{\hat{\sigma}_k^2}{(\hat{f}_{k}^{I})^2} \left( \frac{1}{\sum_{i=0}^{n_1-k-1} \hat{C}_{i,k}^{I}} - \frac{1}{\sum_{i=0}^{n_2-k-1} \hat{C}_{i,k}^{I}} \right) \right) - 1
\]

\[
\approx \left( \sum_{i=0}^{I} \hat{C}_{i,J}^{I} \right) \sum_{k=0}^{J-1} \frac{\hat{\sigma}_k^2}{\hat{f}_{k}^{I}} \left( \frac{1}{\sum_{i=0}^{n_1-k-1} \hat{C}_{i,k}^{I}} - \frac{1}{\sum_{i=0}^{n_2-k-1} \hat{C}_{i,k}^{I}} \right).
\]
Stochastic Reserving

In order to estimate the remaining expectations of the square of \( f_k^n \), we will look at the corresponding variance and expectation of \( \hat{f}_k^n \) and always replace all future weights \( w_{i,k}^2 \) by their estimates at time \( I \), i.e. by

\[
\hat{w}_{i,k}^2 := \frac{\hat{C}_{i,k}}{\sum_{h=0}^{n_2-k-1} \hat{C}_{h,k}}.
\]

We get

\[
E \left[ \hat{f}_k^n \middle| D^{n_1} \right] = E \left[ \sum_{i=0}^{n_2-k-1} \hat{w}_{i,k}^2 \frac{C_{i,k+1}}{C_{i,k}} \middle| D^{n_1} \right] \approx E \left[ \sum_{i=0}^{n_2-k-1} \hat{w}_{i,k}^2 \frac{C_{i,k+1}}{C_{i,k}} \middle| D^{n_1} \right] = \sum_{i=0}^{n_1-k-1} \hat{w}_{i,k}^2 \frac{C_{i,k+1}}{C_{i,k}} + \sum_{i=n_1-k}^{n_2-k-1} \hat{w}_{i,k}^2 f_k \\
= \sum_{i=0}^{n_1-k-1} \frac{\hat{C}_{i,k}}{\sum_{h=0}^{n_2-k-1} \hat{C}_{h,k}} \frac{C_{i,k+1}}{C_{i,k}} \hat{f}_k + \sum_{h=n_1-k}^{n_2-k-1} \frac{\hat{C}_{h,k}}{\sum_{h=0}^{n_2-k-1} \hat{C}_{h,k}} \frac{\hat{f}_k^n}{\hat{f}_k} \approx \hat{f}_k + \Omega_{k,n_1,n_2} (f_k - \hat{f}_k^n)
\]
This leads to

\[
\left( \mathbb{E}[\hat{f}_{k}^{n_2} | D^{n_1}] \right)^2 = \left( \hat{f}_{k}^{n_1} \right)^2 + 2 \Omega_{k}^{n_1,n_2} \hat{f}_{k}^{n_1} (f_k - \hat{f}_{k}^{n_1}) + \left( \Omega_{k}^{n_1,n_2} \right)^2 (f_k - \hat{f}_{k}^{n_1})^2
\]

\[\approx \left( \hat{f}_{k}^{n_1} \right)^2 + 0 + \left( \Omega_{k}^{n_1,n_2} \right)^2 \text{Var}[\hat{f}_{k}^{n_1} | D_k] \approx \left( \hat{f}_{k}^{n_1} \right)^2 + \bar{\sigma}_k^2 \frac{\left( \sum_{i=n_1-k}^{n_2-k-1} \hat{C}_{i,k} I \right)^2}{\sum_{i=0}^{n_1-k-1} \hat{C}_{i,k} \left( \sum_{i=0}^{n_2-k-1} \hat{C}_{i,k} I \right)^2}, \]

For the variance we get

\[
\text{Var}[\hat{f}_{k}^{n_2} | D^{n_1}] = \text{Var} \left[ \sum_{i=0}^{n_2-k-1} w_{i,k}^{n_2} \frac{C_{i,k+1}}{C_{i,k}} | D^{n_1} \right] \approx \text{Var} \left[ \sum_{i=0}^{n_2-k-1} \hat{w}_{i,k}^{n_2} \frac{C_{i,k+1}}{C_{i,k}} | D^{n_1} \right]
\]

\[= \sum_{i=n_1-k}^{n_2-k-1} \left( \hat{w}_{i,k}^{n_2} \right)^2 \text{Var} \left[ \frac{C_{i,k+1}}{C_{i,k}} | D^{n_1} \right]
\]

\[= \sum_{i=n_1-k}^{n_2-k-1} \left( \hat{w}_{i,k}^{n_2} \right)^2 \left( \text{Var} \left[ \mathbb{E} \left[ \frac{C_{i,k+1}}{C_{i,k}} | D_k \right] | D^{n_1} \right] + \mathbb{E} \left[ \text{Var} \left[ \frac{C_{i,k+1}}{C_{i,k}} | D_k \right] | D^{n_1} \right] \right)
\]

\[= \sum_{i=n_1-k}^{n_2-k-1} \left( \hat{w}_{i,k}^{n_2} \right)^2 \left( 0 + \mathbb{E} \left[ \frac{\sigma_k^2}{C_{i,k}} | D^{n_1} \right] \right) \approx \sum_{i=n_1-k}^{n_2-k-1} \left( \hat{w}_{i,k}^{n_2} \right)^2 \frac{\bar{\sigma}_k^2}{\hat{C}_{i,k} I} = \bar{\sigma}_k^2 \frac{\sum_{i=n_1-k}^{n_2-k-1} \hat{C}_{i,k} I}{\left( \sum_{i=0}^{n_2-k-1} \hat{C}_{i,k} I \right)^2}.
\]
Stochastic Reserving

- Chain-Ladder-Method (CLM)
- Solvency uncertainty

Both estimates together lead to

$$
\text{E}
\left[
\left( \hat{f}_k n^2 \right)^2 \middle| D^{n_1} \right] \approx \left( \hat{f}_k \right)^2 + \sigma_k^2 \left( \frac{\sum_{i=n_1-k}^{n_2-k-1} \hat{C}_{i,k}}{\sum_{i=0}^{n_1-k-1} \hat{C}_{i,k}} \right)^2 + \frac{\sum_{i=n_1-k}^{n_2-k-1} \hat{C}_{i,k}}{\left( \sum_{i=0}^{n_2-k-1} \hat{C}_{i,k} \right)^2}
$$

Combining all we get

$$
\text{mse}_0|_{DI} \left[ \hat{\text{CDR}}^{n_1,n_2} \right] \approx \left( \sum_{i=0}^{I} C_{i,0} \right)^2 \left( \prod_{k=0}^{J-1} \left( \hat{f}_k \right)^2 + \sigma_k^2 \left( \frac{1}{\sum_{i=0}^{n_1-k-1} \hat{C}_{i,k}} - \frac{1}{\sum_{i=0}^{n_2-k-1} \hat{C}_{i,k}} \right) \right) - \prod_{k=0}^{J-1} \left( \hat{f}_k \right)^2
$$

where we used in the last step a Taylor approximation in $\sigma_k^2$ at zero.
Remark 2.26

- All summation over accident periods stop at $I$, but we skipped $\land I$ in order to keep the formulas a bit simpler.
- The red parts are the differences to our estimators for the solvency and ultimate uncertainty, i.e.
  * If we take $n_2 = I + 1$ and $n_1 = I$ Estimator 2.25 leads to the same formulas as in the solvency uncertainty case, see Corollary 2.24.
  * If we take $n_2 = \infty$ and $n_1 = I$ Estimator 2.25 leads to the same formulas as in the ultimate uncertainty case, see Corollary 2.14.
- The derivation of Estimator 2.25 is based on the article [30] by Ancus Röhr and discussion with Alois Gisler.
- In practise the differences between the last two lines of Estimator 2.25 are usually very very small.
Remark 2.26

- All summation over accident periods stop at $I$, but we skipped $\wedge I$ in order to keep the formulas a bit simpler.
- The red parts are the differences to our estimators for the solvency and ultimate uncertainty, i.e.
  - If we take $n_2 = I + 1$ and $n_1 = I$ Estimator 2.25 leads to the same formulas as in the solvency uncertainty case, see Corollary 2.24.
  - If we take $n_2 = \infty$ and $n_1 = I$ Estimator 2.25 leads to the same formulas as in the ultimate uncertainty case, see Corollary 2.14.
- The derivation of Estimator 2.25 is based on the article [30] by Ancus Röhr and discussion with Alois Gisler.
- In practice the differences between the last two lines of Estimator 2.25 are usually very very small.
Solvency uncertainties for Examples 2.7 and 2.8

- We used the standard estimators for the variance parameters, see Estimator 2.12.
- Since the incurred values are a bit more stable, in particular for later development periods, the corresponding uncertainties are lower.
- The linear approximation for the (parameter estimation) uncertainty results in almost the same values like without approximation.

<table>
<thead>
<tr>
<th>AP</th>
<th>Solvency uncertainty for payments</th>
<th>Solvency uncertainty for incurred</th>
<th>Credibility like weighting</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Proc Var</td>
<td>Para Err</td>
<td>Total</td>
</tr>
<tr>
<td>0</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>1</td>
<td>68'914</td>
<td>56'985</td>
<td>89'423</td>
</tr>
<tr>
<td>2</td>
<td>171'037</td>
<td>126'690</td>
<td>212'847</td>
</tr>
<tr>
<td>3</td>
<td>109'318</td>
<td>73'276</td>
<td>182'605</td>
</tr>
<tr>
<td>4</td>
<td>143'337</td>
<td>73'807</td>
<td>216'223</td>
</tr>
<tr>
<td>5</td>
<td>126'341</td>
<td>73'120</td>
<td>199'475</td>
</tr>
<tr>
<td>6</td>
<td>92'633</td>
<td>49'013</td>
<td>141'947</td>
</tr>
<tr>
<td>7</td>
<td>212'791</td>
<td>89'328</td>
<td>302'120</td>
</tr>
<tr>
<td>8</td>
<td>261'148</td>
<td>111'014</td>
<td>372'162</td>
</tr>
<tr>
<td>9</td>
<td>215'464</td>
<td>78'066</td>
<td>293'530</td>
</tr>
<tr>
<td>Total</td>
<td>847'287</td>
<td>539'524</td>
<td>1'004'481</td>
</tr>
</tbody>
</table>

We always show the square root of uncertainties.
The uncertainty of the weighing has been calculated using a LSRM coupling of both CLM via the exposure $R_{i,k}^{0,1} = R_{i,k}^{1,0} := R_{i,k}^{0,0} + R_{i,k}^{1,1}$, see Section 4.
Density plot of the distribution of the CDR using Lognormal distributions (dotted lines representing the Best Estimate)

- Projection of Incurred
- Credibility like weighting
- Projection of Payments

in million
The incurred projection results in a very symmetric and tight distribution of the CDR. Therefore, if we believe in it we would expect only very small amounts for the CDR.

The uncertainty of the weighing has been calculated using a LSRM coupling of both CLM via the exposure $R_{i,k}^{0,1} = R_{i,k}^{1,0} := R_{i,k}^{0,0} + R_{i,k}^{1,1}$, see Section 4.
Ultimate vs. solvency uncertainties for Examples 2.7 and 2.8

- We used the standard estimators for the variance parameters, see Estimator 2.12.
- In total the square root of the solvency uncertainty is about 70% of the square root of the ultimate uncertainty, whereas it is higher in older and lesser in recent accident periods. That means during one business period we gain information that is worth about 30% of the uncertainty.
- For standard business one usually expects that the square root of the solvency uncertainty lies between 50% and 90% of the square root of the ultimate uncertainty.

<table>
<thead>
<tr>
<th>AP</th>
<th>Uncertainty for payments</th>
<th>Credibility like weighting</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ultimate</td>
<td>Solvency</td>
</tr>
<tr>
<td>0</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>1</td>
<td>89'423</td>
<td>89'423</td>
</tr>
<tr>
<td>2</td>
<td>234'666</td>
<td>212'847</td>
</tr>
<tr>
<td>3</td>
<td>255'612</td>
<td>131'605</td>
</tr>
<tr>
<td>4</td>
<td>261'298</td>
<td>161'223</td>
</tr>
<tr>
<td>5</td>
<td>323'899</td>
<td>145'975</td>
</tr>
<tr>
<td>6</td>
<td>274'942</td>
<td>104'800</td>
</tr>
<tr>
<td>7</td>
<td>373'634</td>
<td>230'780</td>
</tr>
<tr>
<td>8</td>
<td>492'894</td>
<td>283'765</td>
</tr>
<tr>
<td>9</td>
<td>468'137</td>
<td>229'170</td>
</tr>
<tr>
<td>Total</td>
<td>1'517'861</td>
<td>1'004'481</td>
</tr>
</tbody>
</table>

We always show the square root of uncertainties.
The uncertainty of the weighing has been calculated using a LSRM coupling of both CLM via the exposure $R_{i,k}^{0,1} = R_{i,k}^{1,0} := R_{i,k}^{0,0} + R_{i,k}^{1,1}$, see Section 4.
Density plot of the distribution of the CDR (solid curves) and estimated reserves (dotted curves) using Lognormal distributions (dotted lines representing the Best Estimate).
Note, distributions of the estimated reserves have been obtained by fitting the Lognormal distribution to the estimated reserves as mean and the corresponding uncertainty as variance. Like expected, the densities of the solvency uncertainty are much tighter than the one of the ultimate uncertainty.

The uncertainty of the weighing has been calculated using a LSRM coupling of both CLM via the exposure $R_{i,k}^{0,1} = R_{i,k}^{1,0} := R_{i,k}^{0,0} + R_{i,k}^{1,1}$, see Section 4.
Literature

The prediction error of the chain ladder method applied to correlated run-off triangles.

Linear stochastic reserving methods.

Distribution-free calculation of the standard error of chain ladder reserving estimates.

[10] Markus Buchwalder; Hans Bühlmann; Michael Merz and Mario V. Wütrich.
The mean square error of prediction in the chain ladder reserving method (Mack and Murphy revisited).

A stochastic model underlying the chain-ladder technique.

Chain Ladder and Error Propagation.

*Stochastic claims reserving methods in insurance.*
Literature


Stochastic Reserving
Lecture 5
Other Reserving Methods

René Dahms

ETH Zurich, Spring 2019

20 March 2019
(Last update: 18 February 2019)
3 Other classical reserving methods

3.1 Complementary-Loss-Ration method (CLRM)
3.1.1 CLRM without stochastic
3.1.2 Stochastic behind CLRM

3.2 Bornhuetter-Ferguson method (BFM)
3.2.1 BFM without stocastics
3.2.2 Stochastic behind BFM

3.3 Benktander-Hovinen method (BHM)

3.4 Cape-Cod method

3.5 Extended-Complementary-Loss-Ration method (ECLRM)
3.5.1 ECLRM without stochastic
3.5.2 Stochastic behind ECLRM

3.6 Other methods

3.7 Literature
3 Other classical reserving methods
3.1 Complementary-Loss-Ration method (CLRM)
3.1.1 CLRM without stochastic
3.1.2 Stochastic linked CLRM
3.2 Bornhuetter-Ferguson method (BFM)
3.2.1 BFM without stochastic
3.2.2 Stochastic linked BFM
3.3 Benktander-Hovinen method (BHM)
3.4 Cape-Cod method
3.5 Extended-Complementary-Loss-Ration method (ECLRM)
3.5.1 ECLRM without stochastic
3.5.2 Stochastic linked ECLRM
3.6 Other methods
3.7 Literature
Basic idea behind the Complementary-Loss-Ration method

The Complementary-Loss-Ration method is based on a single triangle and an exposure $P_i$ depending on accident periods $i$. Often pricing information like the risk premium is taken as exposure.

The Complementary-Loss-Ration method is based on the idea that:

- The payments of the next development period are proportional to the given exposure, i.e. 
  \[ S_{i,k+1} \approx f_k P_i. \]
- Accident periods are independent.

In particular, that means that all accident periods are comparable with respect to their development.
Basic idea behind the Complementary-Loss-Ration method

The Complementary-Loss-Ration method is based on a single triangle and an exposure $P_i$ depending on accident periods $i$. Often pricing information like the risk premium is taken as exposure.

The Complementary-Loss-Ration method is based on the idea that:
- The payments of the next development period are proportional to the given exposure, i.e. $S_{i+1} \approx f_k P_i$.
- Accident periods are independent.

In particular, that means that all accident periods are comparable with respect to their development.
### Simple example

<table>
<thead>
<tr>
<th>$i \backslash k$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>exposure</th>
<th>ultimate</th>
<th>reserves</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>100</td>
<td>3.8</td>
<td>380</td>
<td>2.8</td>
<td>280</td>
<td>1.0</td>
<td>100</td>
<td>860</td>
</tr>
<tr>
<td>1</td>
<td>120</td>
<td>3.6</td>
<td>360</td>
<td>2.6</td>
<td>260</td>
<td>1.2</td>
<td>120</td>
<td>860</td>
</tr>
<tr>
<td>2</td>
<td>200</td>
<td>3.9</td>
<td>780</td>
<td>2.3</td>
<td>460</td>
<td>1.1</td>
<td>200</td>
<td>1660</td>
</tr>
<tr>
<td>3</td>
<td>140</td>
<td>3.8</td>
<td>570</td>
<td>2.5</td>
<td>375</td>
<td>1.1</td>
<td>150</td>
<td>1250</td>
</tr>
<tr>
<td>4</td>
<td>200</td>
<td>3.8</td>
<td>836</td>
<td>2.5</td>
<td>550</td>
<td>1.1</td>
<td>220</td>
<td>1828</td>
</tr>
</tbody>
</table>

| $\hat{f}_k$   | 3.8 | 2.5 | 1.1 | 0.0 | 770 | 6458 | 2388 |

\[
\hat{f}_0 = \frac{380+360+780+570}{100+100+200+150} = 3.8 = \sum_{i=0}^{I-1} \frac{P_i}{\sum_{h=0}^{I-1} P_h} \left( S_{i,1} \right) \\
\hat{f}_1 = \frac{280+260+460}{100+100+200} = 2.5 \\
\hat{f}_2 = \frac{100+120}{100+100} = 1.1 \\
\hat{f}_3 = \frac{0}{100} = 0.0
\]
<table>
<thead>
<tr>
<th>i/k</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>exposure</th>
<th>ultimate</th>
<th>reserves</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>200</td>
<td>280</td>
<td>360</td>
<td>440</td>
<td>520</td>
<td>860</td>
<td>860</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>120</td>
<td>120</td>
<td>240</td>
<td>120</td>
<td>0</td>
<td>860</td>
<td>860</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>200</td>
<td>240</td>
<td>280</td>
<td>220</td>
<td>0</td>
<td>100</td>
<td>1660</td>
<td>220</td>
</tr>
<tr>
<td>3</td>
<td>140</td>
<td>375</td>
<td>165</td>
<td>0</td>
<td>0</td>
<td>1250</td>
<td>540</td>
<td>860</td>
</tr>
<tr>
<td>4</td>
<td>200</td>
<td>342</td>
<td>242</td>
<td>0</td>
<td>0</td>
<td>1828</td>
<td>1660</td>
<td>1660</td>
</tr>
</tbody>
</table>

Simple example

\[
\hat{f}_k = \frac{\text{exposure}_k + \text{ultimate}_k}{\text{exposure}_k + \text{ultimate}_k} - 0
\]

- **Complementary-Loss-Ration method (CLRM)**

Other classical reserving methods
Assumption 3.A (CLRM)

There exist exposures $P_i$, development factors $f_k$ and variance parameters $\sigma^2_k$ such that

i)\text{CLRM} \quad E \left[ S_{i,k+1} \bigg| B_{i,k} \right] = f_k P_i,

ii)\text{CLRM} \quad \text{Var} \left[ S_{i,k+1} \bigg| B_{i,k} \right] = \sigma^2_k P_i \quad \text{and}

iii)\text{CLRM} \quad \text{accident periods are independent.}

Remark 3.1

- Since accident periods are independent, $B_{i,k}$ could be replaced by $D_k$ or by $D_{i+k}^k$.
- Often the assumptions are formulated without conditioning. The difference between both ways are:
  * In taking unconditional expectations we take the average over all possible triangles and therefore ignore the observed past $B_{i,k}$ completely.
  * In taking conditional expectations we explicitly assume that the observed past $B_{i,k}$ has no influence on the expected future development.
Assumption 3.A (CLRM)
There exist exposures $P$, development factors $f_k$, and variance parameters $\sigma^2_k$ such that

1) $\text{CLRM } E[S_{i,k+1}|B_{i,k}] = f_k P_i$,
2) $\text{CLRM } \text{Var}[S_{i,k+1}|B_{i,k}] = \sigma^2_k P_i$,
3) accident periods are independent.

Remark 3.1
- Since accident periods are independent, $B_{i,k}$ could be replaced by $D_{i,k}^0$ or by $D_{i,k}^0 + D_{i,k}$.
- Often the assumptions are formulated without conditioning. The difference between both ways are:
  - In taking unconditional expectations we take the average over all possible triangles and therefore ignore the observed past $B_{i,k}$ completely.
  - In taking conditional expectations we explicitly assume that the observed past $B_{i,k}$ has no influence on the expected future development.
Estimator 3.2 (Future development for CLRM)

Let Assumption 3.A be fulfilled. Then for every set of $D_k$-conditionally unbiased estimators $\hat{f}_k$ of $f_k$ the estimator

$$\hat{C}_{i,J}^{\text{CLRM}} := C_{i,(I-i)\land J} + \sum_{k=I-i}^{J-1} \hat{f}_k P_i$$

is a $D_{I-i}$-conditionally unbiased estimator for the ultimate outcome $C_{i,J}$.

Remark 3.3

- Usually one takes

$$\hat{f}_k := \sum_{i=0}^{I-k-1} \frac{P_i}{\sum_{h=0}^{I-k-1} P_h} \frac{S_{i,k+1}}{P_i}.$$

- Because of the additive structure of Estimator 3.2 the Complementary-Loss-Ratio method is often called additive method.
Estimator 3.2 (Future development for CLRM)

Let Assumption 3.A be fulfilled. Then for every set of $D_k$-conditionally unbiased estimators $\hat{f}_k$ of $f_k$ the estimator

$$\hat{C}_{CLRM}^{i,J} = C_{i,(I-i)^\wedge J} + \sum_{k=I-i}^{J-1} E\left[ E\left[ \hat{f}_k | D_k \right] | D_{I-i} \right] P_i$$

is a $D_{I-i}$-conditionally unbiased estimator for the ultimate outcome $C_{i,J}$.

Remark 3.3

• Usually one takes

$$\hat{f}_k = \sum_{h=0}^{I-I-k-1} \frac{P}{P_k} S_{i,k+1} P_h$$

• Because of the additive structure of Estimator 3.2 the Complementary-Loss-Ratio method is often called additive method.

Other classical reserving methods

Complementary-Loss-Ratio method (CLRM)

$$E\left[ \hat{C}_{CLRM}^{i,J} | D_{I-i} \right] = C_{i,(I-i)^\wedge J} + \sum_{k=I-i}^{J-1} E\left[ \hat{f}_k | D_{I-i} \right] P_i$$

$$= C_{i,(I-i)^\wedge J} + \sum_{k=I-i}^{J-1} E\left[ E\left[ \hat{f}_k | D_k \right] | D_{I-i} \right] P_i$$

$$= C_{i,(I-i)^\wedge J} + \sum_{k=I-i}^{J-1} \hat{f}_k P_i$$

$$= C_{i,(I-i)^\wedge J} + \sum_{k=I-i}^{J-1} E\left[ E\left[ S_{i,k+1} | D_k \right] | D_{I-i} \right] P_i^{i)CLRM}$$

$$= E[C_{i,J} | D_{I-i}]$$
Remark 3.4

- The method itself is well known and often used. But, because of its simplicity, corresponding stochastic models haven’t been studied so much as like for the Chain-Ladder method.
- From a statistical point of view the estimation of the development factors and the variance parameters is critical since we have to estimate $2J$ parameters based on $J(I - \frac{J-1}{2})$ observed development factors. Therefore, in practise the reserving actuary has to include other information in order to overcome the lack of observed data (over parametrised model).
- The method can deal with some kind of incomplete triangle, where some upper left sub-triangle is missing.
- Since the exposures $P_i$ are given and fixed over (development) time, the method cannot really react on observed changes in the data. For instance, assume we take the risk premium as exposure and observe at time $k = 1$, that the frequency of claims has doubled. Therefore, we would expect twice the payments compared to those that have been projected with CLRM.
- Often the CLRM is used for the early development periods, where we do not have so much information within the observed data. And for later development periods other methods like CLM are used in order to take the information contained in $B_{i,k}$ into account.
- Because of part iii) of Assumption 3.A, CLRM cannot deal with diagonal effects like inflation.
- Analogously to what we have done for the Chain-Ladder method, see Section 2, we could derive formulas for the ultimate uncertainty as well as for the solvency uncertainty.
Remark 3.4

- The method itself is well known and often used. But, because of its simplicity, corresponding stochastic models haven’t been studied so much as like for the Chain-Ladder method.

- From a statistical point of view the estimation of the development factors and the variance parameters is critical since we have to estimate $2J$ parameters based on $(J^2 - J + 1)$ observed development factors. Therefore, in practice the reserving actuary has to include other information in order to overcome the lack of observed data (over parametrized model).

- The method can deal with some kind of incomplete triangle, where some upper left sub-triangle is missing.

- Since the exposures $P_i$ are given and fixed over (development) time, the method cannot really react on observed changes in the data. For instance, assume we take the risk premium as exposure and observe at time $k = 1$, that the frequency of claims has doubled. Therefore, we would expect twice the payments compared to those that have been projected with CLRM.

- Often the CLRM is used for the early development periods, where we do not have so much information within the observed data. And for later development periods other methods like CLM are used in order to take the information contained in $B_{i,k}$ into account.

- Because of part iii) of Assumption 3.A, CLRM cannot deal with diagonal effects like inflation.

- Analogously to what we have done for the Chain-Ladder method, see Section 2, we could derive formulas for the ultimate uncertainty as well as for the solvency uncertainty.
Problem 3.5 (How to include an experts opinion about the ultimate?)

We have often repeated that an actuary has to use all available information in order to determine a Best Estimate. But how to combine an experts opinion $U_{i}^{pri}$ about the ultimate $C_{i,J}$ with the observed data.

Bornhuetter-Ferguson method

One solution is to used the Bornhuetter-Ferguson method, introduced by Bornhuetter and Ferguson in [14]. The basic idea is that we take the last observed data $C_{i,I-i}$ and add a fraction $1 - l_{i}$ of the external given a priori ultimate $U_{i}^{pri}$, i.e.

$$\hat{C}_{i,J}^{BFM} := C_{i,I-i} + (1 - \hat{l}_{i})U_{i}^{pri}$$

(3.1)

where the factors $l_{i}$ are called link ratios and should represent the proportion of the ultimate that has already developed.

Problem 3.6 (Where to get the link ratios?)

Possible answers:

- Experts opinion.
- Use a reserving method and take $\hat{l}_{i} := \frac{C_{i,I-i}}{C_{i,J}}$. In the case of CLM we would get

$$\hat{l}_{i} = \prod_{k=I-i}^{J-1}(f_{k}^{CLM})^{-1},$$

which was the original idea behind BFM.
- Use a stochastic model that leads to estimators which have the same shape like (3.1).
Problem 3.5 (How to include an expert's opinion about the ultimate?)

We have often repeated that an actuary has to use all available information in order to determine a Best Estimate. But how to combine an expert's opinion $U_{pri}$ about the ultimate $C_i$ with the observed data.

**Bornhuetter-Ferguson method**

One solution is to use the Bornhuetter-Ferguson method, introduced by Bornhuetter and Ferguson in [14]. The basic idea is that we take the last observed data $C_{i,I}$ and add a fraction $l_i$ of the external given a priori ultimate $U_{pri}$, i.e.

$$\hat{C}_{BFM,i,J} := C_{i,I} + (1 - \hat{l}_i) U_{pri},$$

where the factors $l_i$ are called link ratios and should represent the proportion of the ultimate that has already developed.

Problem 3.6 (Where to get the link ratios?)

Possible answers:

- Experts opinion.
- Use a reserving method and take $\hat{l}_i := \frac{C_{i,I}}{\hat{C}_{i,J}}$. In the case of CLM we would get $\hat{l}_i := \prod_{k=1}^{J-I} (C_{i,k}/\hat{C}_{i,k,J})^{-1}$, which was the original idea behind BFM.
- Use a stochastic model that leads to estimators which have the same shape like (3.1).
Remark 3.7

- Since the link ratios $l_i$ should represent the proportion of the ultimate that has already developed we expect that $l_{I-J} = 1$, provided we have no tail development.
- As actuaries we have to be very careful in using experts opinions, in particular, if we take the a priori ultimate and the link ratios from the same expert. The reason is that those experts often have own interests in a profitable (or sometimes non profitable) outcome of the portfolio.

BFM as credibility weighted average

If we take a reserving method in order to determine the link ratios $l_i := \frac{C_{i,I-i}}{C_{i,J}}$ and if all link ratios $0 \leq l_i \leq 1$ then $C_{i,J}^{BFM}$ could be looked at as credibility like weighted average of the a priori ultimate $U_{i}^{pri}$ and the estimated ultimate $\hat{C}_{i,J}$ with credibility weights $(1-l_i)$ and $l_i$, respectively:

$$C_{i,J}^{BFM} = C_{i,I-i} + (1-l_i)U_{i}^{pri} = \frac{C_{i,I-i}}{\hat{C}_{i,J}} \hat{C}_{i,J} + (1-l_i)U_{i}^{pri} = l_i \hat{C}_{i,J} + (1-l_i)U_{i}^{pri}.$$  

Note, this formula is similar to the credibility like weighting of ultimates proposed in Estimator 2.15.
Remark 3.7

• Since the link ratios $l_i$ should represent the proportion of the ultimate that has already developed we expect that $l_i - l_j = 1$, provided we have no tail development.

• As actuaries we have to be very careful in using experts opinions, in particular, if we take the a priori ultimate and the link ratios from the same expert. The reason is that those experts often have own interests in a profitable (or sometimes non-profitable) outcome of the portfolio.

BFM as credibility weighted average

If we take a reserving method in order to determine the link ratios $l_i := \frac{C_{i,I} - C_{i,J}}{\hat{C}_{i,J}}$ and if all link ratios $0 \leq l_i \leq 1$ then $C_{BFM}^{i,j}$ could be looked at as credibility like weighted average of the a priori ultimate $U_{Pri}^{i}$ and the estimated ultimate $\hat{C}_{i,J}$ with credibility weights $(1 - l_i)$ and $l_i$, respectively:

$$C_{BFM}^{i,j} = C_{i,I} - C_{i,J} + (1 - l_i) U_{Pri}^{i} = \hat{C}_{i,J} + l_i U_{Pri}^{i}.$$

Note, this formula is similar to the credibility like weighting of ultimates proposed in Estimator 2.15.
Remark 3.8 (BFM as Complementary-Loss-Ratio method)

If we take the Complementary-Loss-Ratio method with exposure $P_i := U_i^{pri}$ we get the estimate (see 3.2)

$$\hat{C}_{i,J}^{CLRM} = C_{i,(I-i)\wedge J} + \sum_{k=I-i}^{J-1} \hat{f}_k P_i.$$  

Defining the link ratios via

$$\hat{l}_i := 1 - \sum_{k=I-i}^{J-1} \hat{f}_k$$

we get the same form as in (3.1). Therefore, the Bornhuetter-Ferguson method can be looked at as Complementary-Loss-Ratio method with exposures $U_i^{pri}$.

Remark 3.9

There are other stochastic models that lead to estimators of the form (3.1), see for instance [17, Section 6.6].
Remark 3.8 (BFM as Complementary-Loss-Ratio method)

If we take the Complementary-Loss-Ratio method with exposure \( P_i = U_{pri} \) we get the estimate (see 3.2)

\[
\hat{C}_{CLRM}^{i,j} = C_i, (I - i) \land J + J - 1 \sum_{k=I-i}^{J} \hat{f}_k P_i.
\]

Defining the link ratios via

\[
\hat{l}_i = 1 - J - 1 \sum_{k=I-i}^{J} \hat{f}_k
\]

we get the same form as in (3.1). Therefore, the Bornhuetter-Ferguson method can be looked at as Complementary-Loss-Ratio method with exposures \( U_{pri} \).

Remark 3.9

There are other stochastic models that lead to estimators of the form (3.1), see for instance [17, Section 6.6].
Basic idea behind the Benktander-Hovinen method

The basic idea of BHM is to apply the Bornhuetter-Ferguson method on the Chain-Ladder method estimation with the weighted a priori ultimate

\[ U_{i}^{BHM pri} := \hat{l}_{i} \hat{C}_{i,J}^{CLM} + (1 - \hat{l}_{i})U_{i}^{pri} = C_{i,I-i} + (1 - \hat{l}_{i})U_{i}^{pri} = \hat{C}_{i,J}^{BFM}, \]

and the link ratios \( \hat{l}_{i} \) of the Chain-Ladder method. Therefore, we assume that \( 0 < \hat{l}_{i} \leq 1 \).

Then we get the estimate

\[ \hat{C}_{i,J}^{BHM} := C_{i,I-i} + (1 - \hat{l}_{i})\hat{C}_{i,J}^{BFM}. \]

Remark 3.10

Connection between BHM, BFM and CLM

- BHM was independently developed by Benktander, see [13], and Hovinen, see [15].
- The BHM is a twice iterated BFM with Chain-Ladder link ratios.
- Iterating BFM further will finally lead to the CLM Best Estimate, see [16].
Basic idea behind the Benktander-Hovinen method

The basic idea of BHM is to apply the Bornhuetter-Ferguson method on the Chain-Ladder method estimation with the weighted a priori ultimate

$$U_{pri}^i := \hat{C}_{i,J}^{BFM} + (1 - \hat{l}_i)U_{pri}^i = C_{i,I}^i - \hat{l}_i. \hat{C}_{J,J}^{BFM}.$$  

and the link ratios $\hat{l}_i$ of the Chain-Ladder method. Therefore, we assume that $0 < \hat{l}_i \leq 1$.

Then we get the estimate

$$\hat{C}_{i,J}^{BHM} = C_{i,J}^i + (1 - \hat{l}_i)\hat{C}_{J,J}^{BFM}.$$  

Remark 3.10

Connection between BHM, BFM and CLM

- BHM was independently developed by Benktander, see [13], and Hovinen, see [15].
- The BHM is a twice iterated BFM with Chain-Ladder link ratios.
- Iterating BFM further will finally lead to the CLM Best Estimate, see [16].
Basic idea behind the Cape-Cod method (CCM)

We have seen that the Best Estimate reserves of the Chain-Ladder method depend heavily on the last known diagonal, which makes this method vulnerable to outliers of $C_{i,I-i}$. The Cape-Cod method uses an external given exposure $P_i$ to smooth the last diagonal. Therefore,

1. We assume that there exists a $\kappa$ with

$$C_{i,I-i} \approx \kappa \hat{l}_i P_i,$$

where $\hat{l}_i := \prod_{k=I-i}^{J-1} (\hat{f}_k^{CLM})^{-1}$ are the link ratios of the CLM.

2. Then we estimate $\kappa$ by

$$\hat{\kappa} := \frac{\sum_{i=I-J}^{I} C_{i,I-i}}{\sum_{i=I-J}^{I} \hat{l}_i P_i}.$$

3. Finally, we replace the value $C_{i,I-i}$ in the Chain-Ladder estimate for the reserves by

$$\hat{C}_{i,I-i}^{CCM} := \hat{\kappa} \hat{l}_i P_i.$$

Then we get

$$\hat{C}_{i,J}^{CCM} := C_{i,I-i} - \hat{C}_{i,I-i}^{CCM} + \prod_{k=I-i}^{J-1} \hat{f}_k^{CLM} \hat{C}_{i,I-i}^{CCM} = C_{i,I-i} + (1 - \hat{l}_i) \hat{\kappa} P_i. \quad (3.2)$$
Basic idea behind the Cape-Cod method (CCM)

We have seen that the Best Estimate reserves of the Chain-Ladd er method depend heavily on the last known diagonal, which makes this method vulnerable to outliers of $C_{i,I}$. The Cape-Cod method uses an external given exposure $P_i$ to smooth the last diagonal. Therefore,

1. We assume that there exists a $\kappa$ with 
   $C_{i,I} \approx \kappa \hat{l}_i P_i$,
   where $\hat{l}_i = \prod_{J=I}^{J-1} (\hat{f}_{CLM}^k)$ are the link ratios of the CLM.
2. Then we estimate $\kappa$ by 
   $\hat{\kappa} = \frac{\sum_{I=I}^{I-1} C_{i,I} \hat{l}_i P_i}{\sum_{I=I}^{I-1} \hat{l}_i P_i}$.
3. Finally, we replace the value $C_{i,I}$ in the Chain-Ladder estimate for the reserves by 
   $\hat{C}_{i,J}^{CCM} = \hat{\kappa} \hat{l}_i P_i$.

Then we get 
$\hat{C}_{i,J}^{CCM} = C_{i,J} - \hat{C}_{i,J}^{CCM} + \prod_{J=I}^{J-1} \hat{C}_{i,J}^{CCM} = C_{i,J} + (1 - \hat{\kappa}) \hat{l}_i P_i$. (3.2)
Remark 3.11

- The name Cape-Cod refers to the place where this method has been introduced for the first time.
- Because of (3.2), CCM can also be seen as a BFM with (by $\hat{\kappa}$) modified a priory ultimate $\hat{\kappa}P_i$. 
Remark 3.11

- The name Cape-Cod refers to the place where this method has been introduced for the first time.
- Because of (3.2), CCM can also be seen as a BFM with (by $\hat{\kappa}$) modified a priori ultimate $\hat{\kappa}_i$. 
Basic idea behind the Extended-Complementary-Loss-Ration method

The Extended-Complementary-Loss-Ration method is based on a triangle of payments $S_{i,k}^1$ and a triangle of the corresponding (changes of the) incurred losses $S_{i,k}^0$. The Extended-Complementary-Loss-Ration method is based on the idea that:

- The payments of the next development period are proportional to the case reserves at the end of the current development period, i.e.

$$S_{i,k+1}^1 \approx f_k^1 \sum_{j=0}^{k} (S_{i,j}^0 - S_{i,j}^1).$$

- The changes of the incurred losses during the next development period $k \geq 1$ are proportional to the case reserves at the end of the current development period, i.e.

$$S_{i,k+1}^0 \approx f_k^0 \sum_{j=0}^{k} (S_{i,j}^0 - S_{i,j}^1).$$

- Accident period are independent.

In particular, that means that all accident periods are comparable with respect to their development.
Basic idea behind the Extended-Complementary-Loss-Ration method

The Extended-Complementary-Loss-Ration method is based on a triangle of payments $S_{1}$ and a triangle of the corresponding (changes of the) incurred losses $S_{0}$.

The Extended-Complementary-Loss-Ration method is based on the idea that:

• The payments of the next development period are proportional to the case reserves at the end of the current development period, i.e.

$$S_{1,k+1} = f_{k} \sum_{j=0}^{k} (S_{0,i,j} - S_{1,i,j})$$

• The changes of the incurred losses during the next development period $k \geq 1$ are proportional to the case reserves at the end of the current development period, i.e.

$$S_{0,k+1} = f_{k} \sum_{j=0}^{k} (S_{0,i,j} - S_{1,i,j})$$

• Accident period are independent. In particular, that means that all accident periods are comparable with respect to their development.
### Simple example

**Changes of incurred losses** $S_{i,k}^0$

<table>
<thead>
<tr>
<th>i\k</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>500</td>
<td>0.5</td>
<td>200</td>
<td>-0.4</td>
</tr>
<tr>
<td>1</td>
<td>700</td>
<td>0.4</td>
<td>160</td>
<td>-0.4</td>
</tr>
<tr>
<td>2</td>
<td>900</td>
<td>0.3</td>
<td>120</td>
<td>-0.4</td>
</tr>
<tr>
<td>3</td>
<td>550</td>
<td>0.4</td>
<td>120</td>
<td>-0.4</td>
</tr>
</tbody>
</table>

| $\hat{f}_0$ | 0.4   | -0.4  | 0.0   |

\[
\hat{f}_0 = \frac{200 + 160 + 120}{400 + 400 + 400} = 0.4
\]

\[
\hat{f}_1 = \frac{-160 - 160}{400 + 400} = -0.4
\]

\[
\hat{f}_2 = \frac{0}{40} = 0.0
\]

**Payments** $S_{i,k}^1$

<table>
<thead>
<tr>
<th>i\k</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>100</td>
<td>0.5</td>
<td>200</td>
<td>0.5</td>
</tr>
<tr>
<td>1</td>
<td>300</td>
<td>0.4</td>
<td>160</td>
<td>0.5</td>
</tr>
<tr>
<td>2</td>
<td>500</td>
<td>0.6</td>
<td>240</td>
<td>0.5</td>
</tr>
<tr>
<td>3</td>
<td>250</td>
<td>0.5</td>
<td>150</td>
<td>0.5</td>
</tr>
</tbody>
</table>

| $\hat{f}_0$ | 0.5   | 0.5   | 1.0   |

\[
\hat{f}_0 = \frac{200 + 160 + 240}{400 + 400 + 400} = 0.5
\]

\[
\hat{f}_1 = \frac{200 + 200}{400 + 400} = 0.5
\]

\[
\hat{f}_2 = \frac{40}{40} = 1.0
\]

**Case reserves** $\sum_{j=0}^{k}(S_{i,j}^1 - S_{i,j}^1)$

<table>
<thead>
<tr>
<th>i\k</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>400</td>
<td>1.0</td>
<td>400</td>
<td>0.1</td>
</tr>
<tr>
<td>1</td>
<td>400</td>
<td>1.0</td>
<td>400</td>
<td>0.1</td>
</tr>
<tr>
<td>2</td>
<td>400</td>
<td>0.7</td>
<td>280</td>
<td>0.1</td>
</tr>
<tr>
<td>3</td>
<td>300</td>
<td>0.9</td>
<td>270</td>
<td>0.1</td>
</tr>
</tbody>
</table>

| $\hat{f}_0$ | 0.9   | 0.1   | 0.0   |

\[
\hat{f}_0 = 1 + 0.4 - 0.5 = 0.9
\]

\[
\hat{f}_1 = 1 - 0.4 - 0.5 = 0.1
\]

\[
\hat{f}_2 = 1 + 0.0 - 1.0 = 0.0
\]

- The case reserves develop according to the Chain-Ladder method with $\hat{f}_k = 1 + \hat{f}_0^0 - \hat{f}_1^1$.
- If we use CLM we would get

<table>
<thead>
<tr>
<th>i</th>
<th>Ultimate Reserves</th>
<th>IBNR</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>540</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>700</td>
<td>40</td>
</tr>
<tr>
<td>2</td>
<td>908</td>
<td>168</td>
</tr>
<tr>
<td>3</td>
<td>562</td>
<td>212</td>
</tr>
</tbody>
</table>

| $\sum$ | 2710 | 520 | -100 |

<table>
<thead>
<tr>
<th>CLM on Payments</th>
<th>CLM on Incurred</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reserves</td>
<td>969</td>
</tr>
</tbody>
</table>

Simple example

Changes of incurred losses

\[ S_{0}^{i,k} = \begin{array}{c|c|c|c|c|c|c|c|c|c|c} i & k & 0 & 1 & 2 & 3 \\ \hline 0 & 500 & 0.5 & 200 & -0.4 & 160 & 0.0 & 0 \\ 1 & 700 & 0.4 & 160 & -0.4 & 160 & 0.0 & 0 \\ 2 & 900 & 0.3 & 120 & -0.4 & 112 & 0.0 & 0 \\ 3 & 550 & 0.4 & 120 & -0.4 & 108 & 0.0 & 0 \end{array} \]

\[ \hat{f}_{0}^{k} = \begin{array}{c} 0.4 \quad -0.4 \quad 0.0 \end{array} \]

\[ \hat{f}_{0}^{0} = 200 + 160 + 120 = 400 + 400 + 400 = 0. \]

\[ \hat{f}_{0}^{1} = -160 - 160 = -320 = \frac{-320}{400 + 400 + 400} = -0.8. \]

\[ \hat{f}_{0}^{2} = 0. \]

Payments

\[ S_{1}^{i,k} = \begin{array}{c|c|c|c|c|c|c|c|c|c|c} i & k & 0 & 1 & 2 & 3 \\ \hline 0 & 100 & 0.5 & 200 & 0.5 & 200 & 1.0 & 40 \\ 1 & 300 & 0.4 & 160 & 0.5 & 200 & 1.0 & 40 \\ 2 & 500 & 0.6 & 240 & 0.5 & 140 & 1.0 & 28 \\ 3 & 250 & 0.5 & 150 & 0.5 & 135 & 1.0 & 27 \end{array} \]

\[ \hat{f}_{1}^{k} = \begin{array}{c} 0.5 \quad 0.5 \quad 1.0 \end{array} \]

\[ \hat{f}_{1}^{0} = 200 + 160 + 240 = 400 + 400 + 400 = 0. \]

\[ \hat{f}_{1}^{1} = 200 + 200 = 0.5 \]

\[ \hat{f}_{1}^{2} = 40 = 1. \]

Case reserves

\[ \sum_{k=0}^{j} (S_{1}^{i,j} - S_{1}^{i,j}) \]

\[ \begin{array}{c|c|c|c|c|c} i & k & 0 & 1 & 2 & 3 \\ \hline 0 & 400 & 1 & 400 & 0 & 400 \end{array} \]

\[ \begin{array}{c|c|c|c|c|c} i & k & 1 & 2 & 3 & 4 \\ \hline 0 & 400 & 1 & 400 & 0 & 400 \end{array} \]

\[ \begin{array}{c|c|c|c|c|c} i & k & 1 & 2 & 3 & 4 \\ \hline 0 & 400 & 1 & 400 & 0 & 400 \end{array} \]

\[ \hat{f}_{k}^{0} = 1 + 0.4 - 0.5 = 0.9 \]

\[ \hat{f}_{k}^{1} = 1 - 0.4 - 0.5 = 0.1 \]

\[ \hat{f}_{k}^{2} = 1 + 0.0 - 1.0 = 0.0 \]

Ultimate reserves

IBNR

\[ \begin{array}{c|c|c|c|c|c} i & k & 0 & 1 & 2 & 3 \\ \hline 0 & 540 & 0 & 0 & 0 \end{array} \]

\[ \begin{array}{c|c|c|c|c|c} i & k & 1 & 2 & 3 & 4 \\ \hline 0 & 700 & 40 & 0 & 0 \end{array} \]

\[ \begin{array}{c|c|c|c|c|c} i & k & 2 & 3 & 4 & 5 \\ \hline 0 & 908 & 168 & -112 & 0 \end{array} \]

\[ \begin{array}{c|c|c|c|c|c} i & k & 3 & 4 & 5 & 6 \\ \hline 0 & 562 & 212 & 12 & 0 \end{array} \]

\[ \sum = 2710 \]

\[ \begin{array}{c|c|c|c|c|c} i & k & 0 & 1 & 2 & 3 \\ \hline 0 & 540 & 1 & 0 & 0 \end{array} \]

Stochastic Reserving

Other classical reserving methods

Extended-Complementary-Loss-Ration method (ECLRM)
Assumption 3.B (ECLRM)

There exist development factors \( f_k^m \), \( m \in \{0, 1\} \), and covariance parameters \( \sigma_k^{m_1,m_2} \), \( m_1, m_2 \in \{0, 1\} \), such that

\[
\begin{align*}
\text{i)} & \quad \text{ECLRM} \quad E \left[ S_{i,k+1}^m \mid B_{i,k} \right] = f_k^m \sum_{j=0}^k \left( S_{i,j}^0 - S_{i,j}^1 \right), \\
\text{ii)} & \quad \text{ECLRM} \quad \text{Cov} \left[ S_{i,k+1}^{m_1}, S_{i,k+1}^{m_2} \mid B_{i,k} \right] = \sigma_k^{m_1,m_2} \sum_{j=0}^k \left( S_{i,j}^0 - S_{i,j}^1 \right) \quad \text{and} \\
\text{iii)} & \quad \text{ECLRM} \quad \text{accident periods are independent.}
\end{align*}
\]

Remark 3.12

- Since accident periods are independent, \( B_{i,k} \) could be replaced by \( D_k \) or by \( D_{i+k} \).
- Usually, \( S_{i,k}^0 \) and \( S_{i,k}^1 \) representing changes of incurred losses and payments during development period \( k \) for claims of accident period \( i \), respectively. Then \( \sum_{j=0}^k \left( S_{i,j}^0 - S_{i,j}^1 \right) \) are the case reserves at the end of development period \( k \) for claims of accident period \( i \).
Assumption 3B (ECLRM)

There exist development factors $f_{m}^{k}$, $m \in \{0, 1\}$, and covariance parameters $\sigma_{m_1, m_2}^{k}$, $m_1, m_2 \in \{0, 1\}$, such that

$\text{i)} ECLRM \quad E[S_{m_1, m_2}^{k+1} | B_{k}] = f_{m}^{k} \sum_{j=0}^{k} (S_{0}^{i,j} - S_{1}^{i,j})$,

$\text{ii)} ECLRM \quad Cov[S_{m_1, m_2}^{k+1}, S_{m_2, m_2}^{k+1} | B_{k}] = \sigma_{m_1, m_2}^{k} \sum_{j=0}^{k} (S_{0}^{i,j} - S_{1}^{i,j})$ and

$\text{iii)} ECLRM \quad \text{accident periods are independent.}$

Remark 3.12

- Since accident periods are independent, $B_{k}$ could be replaced by $D_{k}$ or by $D_{k+1}$.
- Usually, $S_{0}^{i,j}$ and $S_{1}^{i,j}$ representing changes of incurred losses and payments during development period $k$ for claims of accident period $i$, respectively. Then $\sum_{j=0}^{k} (S_{0}^{i,j} - S_{1}^{i,j})$ are the case reserves at the end of development period $k$ for claims of accident period $i$. 

---

Stochastic Reserving

Other classical reserving methods

Extended-Complementary-Loss-Ration method (ECLRM)
Assume Assumption 3.B is fulfilled. Then for every set of $D_k$-conditionally unbiased estimators $\hat{f}^m_k$ of $f^m_k$ the estimator

$$\hat{C}_{i,J}^{m,ECLRM} := C_{i,(I-i)}^m \land J + \sum_{k=I-i}^{J-1} \hat{f}^m_k \prod_{j=I-i}^{k-1} (1 + \hat{f}^0_j - \hat{f}^1_j) (C^0_{i,I-i} - C^1_{i,I-i})$$

is a $D_{I-i}$-conditionally unbiased estimator for the ultimate outcome $C_{i,J}^m$.

**Remark 3.14**

Usually one takes

$$\hat{f}^m_k := \sum_{i=0}^{I-k-1} \frac{R_{i,k}}{\sum_{h=0}^{I-k-1} R_{i,k}} \frac{S^m_{i,k+1}}{R_{i,k}},$$

where

$$R_{i,k} := (C^0_{i,k} - C^1_{i,k}) = \sum_{j=0}^{k} (S^0_{i,j} - S^1_{i,j}).$$

denote the case reserves.
From Assumption 3.B.i) \(^{ECLRM}\), it follows that \(E[R_{i,k+1} | D_i] = (1 + f_{k}^{0} - f_{k}^{1}) R_{i,k}\). Therefore, we get

\[
\begin{align*}
E[\hat{C}_{i,J}^{ECLRM} | D_{I-i}] &= C_{i,(I-i)\wedge J}^m + \sum_{k=I-i}^{J-1} E \left[ \hat{f}_{k}^{m} \prod_{j=I-i}^{k-1} (1 + \hat{f}_{j}^{0} - \hat{f}_{j}^{1}) | D_{I-i} \right] R_{i,I-i} \\
&= C_{i,(I-i)\wedge J}^m + \sum_{k=I-i}^{J-1} \left( \prod_{j=I-i}^{k-1} (1 + \hat{f}_{j}^{0} - \hat{f}_{j}^{1}) \right) R_{i,I-i} \\
&= \ldots = C_{i,(I-i)\wedge J}^m + \sum_{k=I-i}^{J-1} f_{k}^{m} \left( \prod_{j=I-i}^{k-1} (1 + f_{j}^{0} - f_{j}^{1}) \right) R_{i,I-i} \\
&= C_{i,(I-i)\wedge J}^m + \sum_{k=I-i}^{J-1} f_{k}^{m} \left( \prod_{j=I-i+1}^{k-1} (1 + f_{j}^{0} - f_{j}^{1}) \right) E[R_{i,I-i+1} | D_{I-i}] \\
&= \ldots = C_{i,(I-i)\wedge J}^m + \sum_{k=I-i}^{J-1} f_{k}^{m} E[R_{i,k} | D_{I-i}] \\
&= C_{i,(I-i)\wedge J}^m + \sum_{k=I-i}^{J-1} E[S_{i,k+1}^m | D_{I-i}] = E[C_{i,J}^{m} | D_{I-i}]
\end{align*}
\]
Remark 3.15

- ECLRM couples payments and incurred losses in a natural way via the case reserves such that the projections of both triangles lead to the same ultimate, provided we don’t have any tail development. But we will still get two estimates for the ultimate uncertainty as well as for the solvency uncertainty.
- The method can deal with incomplete triangles, where some upper left sub-triangles are missing, as long as case reserves are available for all recent calendar periods.
- It depends heavily on the case reserves. In particular, it may have problems dealing with portfolios with a high reopening rate, because in such situation the case reserves may be very small or even equal to zero.
- The method itself is not so well known, in particular under the name ECLRM.
- From a statistical point of view the estimation of the development factors and the variance parameters is critical since we have to estimate $5J$ parameters based on $2J(I - \frac{J-1}{2})$ observed development factors. Therefore, in practise the reserving actuary has to include other information in order to overcome the lack of observed data (over parametrised model).
- Because of part iii)\textsuperscript{ECLRM} of Assumption 3.B, ECLRM cannot deal with diagonal effects like inflation.
- Analogously to what we have done for the Chain-Ladder method, see Section 2, we could derive formulas for the ultimate uncertainty as well as for the solvency uncertainty.
Remark 3.15

- ECLRM couples payments and incurred losses in a natural way via the case reserves such that the projections of both triangles lead to the same ultimate, provided we don't have any tail development. But we will still get two estimates for the ultimate uncertainty as well as for the solvency uncertainty.

- The method can deal with incomplete triangles, where some upper left sub-triangles are missing, as long as case reserves are available for all recent calendar periods.

- It depends heavily on the case reserve. In particular, it may have problems dealing with portfolios with a high reopening rate, because in such situations the case reserves may be very small or even equal to zero.

- The method itself is not so well known, in particular under the name ECLRM.

- From a statistical point of view the estimation of the development factors and the variance parameters is critical since we have to estimate \( \frac{5}{2} \) parameters based on \( 2(2-J-I-J) \) observed development factors. Therefore, in practice the reserving actuary has to include other information in order to overcome the lack of observed data (over-parametrised model).

- Analogously to what we have done for the Chain-Ladder method, see Section 2, we could derive formulae for the ultimate uncertainty as well as for the solvency uncertainty.
Other methods

There are many more methods used for reserving. Some of them are based on a stochastic model and some not. For instance:

- Frequency severity models, which model the claim frequency and the severity separately.
- Generalised linear models (GLMs) are sometimes used for reserving.
- Munich-Chain-Ladder method, which tries to project payments and incurred losses simultaneously.
- Bayesian models, which model development factors as random variables.
- Distribution based models, which assume some kind of distribution and fit the corresponding parameters based on the observed data.
- The over-dispersed Poisson model, which leads to the same estimates for the reserves like the Chain-Ladder method we have discussed. But the estimates for the corresponding ultimate (or solvency) uncertainties are different.
- ...
There are many more methods used for reserving. Some of them are based on a stochastic model and some not. For instance:

- Frequency severity models, which model the claim frequency and the severity separately.
- Generalised linear models (GLMs) are sometimes used for reserving.
- Munich-Chain-Ladder method, which tries to project payments and incurred losses simultaneously.
- Bayesian models, which model development factors as random variables.
- Distribution based models, which assume some kind of distribution and fit the corresponding parameters based on the observed data.
- The over-dispersed Poisson model, which leads to the same estimates for the reserves like the Chain-Ladder method we have discussed. But the estimates for the corresponding ultimate (or solvency) uncertainties are different.
Literature

An approach to credibility in calculating IBNR for casualty excess reinsurance.  

The actuary and IBNR.  

Additive and continuous IBNR.  

Credible claims reserves: the Benktander method.  

[18] Mario V. Wüthrich and Michael Merz.  
*Stochastic claims reserving methods in insurance*.  
Stochastic Reserving

- Other classical reserving methods

Literature

An approach to credibility in calculating IBNR for casualty excess reinsurance.  

The actuary and IBNR.  

Additive and continuous IBNR.  

Credible claims reserves: the Benktander method.  
CAS Forum (Fall), pages 141–157, 2000.

[18] Mario V. Wüthrich and Michael Merz.  
Stochastic claims reserving methods in insurance.  
Stochastic Reserving
Lecture 6
Linear-Stochastic-Reserving methods

René Dahms

ETH Zurich, Spring 2019

27 March 2019
(Last update: 18 February 2019)
4 Linear-Stochastic-Reserving methods

4.1 How do Linear-Stochastic-Reserving methods (LSRM) work
4.1.1 LSRM without stochastic
4.1.2 Stochastic behind LSRMs

4.2 Future development
4.2.1 Projection of the future development
4.2.2 Examples

4.3 Ultimate uncertainty
4.3.1 Mixing of claim properties
4.3.2 Ultimate uncertainty
4.3.3 Estimation of the covariance parameters
4.3.4 Examples

4.4 Solvency uncertainty
4.4.1 Estimation at time $I + 1$
4.4.2 Solvency uncertainty
4.4.3 Uncertainties of further CDR’s

4.5 Examples

4.6 Estimation of correlation of reserving Risks
4.6.1 Avoiding correlation matrices for the reserving risks
4.6.2 Using LSRMs to estimate a correlation matrix

4.7 Literature
4 Linear-Stochastic-Reserving methods

4.1 How do Linear-Stochastic-Reserving methods (LSRM) work
4.1.1 LSRM without stochastic
4.1.2 Stochastic behind LSRMs
4.2 Future development
4.2.1 Projection of the future development
4.2.2 Examples

4.3 Ultimate uncertainty
4.3.1 Merging of claim properties
4.3.2 Ultimate uncertainty
4.3.3 Estimation of the covariance parameters
4.3.4 Examples

4.4 Solvency uncertainty
4.4.1 Estimation at time t = 1
4.4.2 Solvency uncertainty
4.4.3 Uncertainties of further CDR’s

4.5 Examples

4.6 Estimation of correlation of reserving Risks
4.6.1 Avoiding correlation matrices for the reserving risks
4.6.2 Using LSRMs to estimate a correlation matrix

4.7 Literature
Motivation for LSRMs

All the methods we have seen up to now can only handle one or at most two triangles. In order to estimate Best Estimate reserves we could simply add the estimates of all portfolios, but how to deal with the uncertainties? Depending of the portfolios we would expect some diversification effects, caused by the law of large numbers, and some dependencies, caused for instance by:

- same underlying risk (hail storms for property and motor hull)
- monetary and superimposed inflation
- changes in insurance contracts (deductibles)
- ...

In practice one often takes a covariance matrix to couple the uncertainties of portfolios, but how to estimate such covariance matrices?

Moreover, there are simple dependencies, which cannot be modelled even for the ultimate outcome. For instance, it is intuitive that future subrogation (regress) may be approximately proportional to the sum of all payments up to know.
Stochastic Reserving

- Linear-Stochastic-Reserving methods
- How do Linear-Stochastic-Reserving methods (LSRM) work

Motivation for LSRMs

All the methods we have seen up to now can only handle one or at most two triangles. In order to estimate Best Estimate reserves we could simply add the estimates of all portfolios, but how to deal with the uncertainties? Depending of the portfolios we would expect some diversification effects, caused by the law of large numbers, and some dependencies, caused for instance by:

- same underlying risk (hail storms for property and motor hull)
- monetary and superimposed inflation
- changes in insurance contracts (deductibles)
- ...

In practice one often takes a covariance matrix to couple the uncertainties of portfolios, but how to estimate such covariance matrices?

Moreover, there are simple dependencies, which cannot be modelled even for the ultimate outcome. For instance, it is intuitive that future subrogation (recess) may be approximately proportional to the sum of all payments up to know.
Basic idea behind Linear-Stochastic-Reserving methods

Linear-Stochastic-Reserving methods are reserving methods for a whole collection of claim properties $S_{i,k}^m$ (triangles), which may be

- payments
- incurred losses
- number of reported claims
- small or large claims
- ...

of the same or different portfolios.

The basic assumption behind LSRMs is that the changes of each claim property $S_{i,k}^m$ are approximately proportional to an exposure $R_{i,k}^m$, which is a linear combination of claim properties of the past.

For instance, denote subrogation by $S_{i,k}^0$ and other payments by $S_{i,k}^1$. Then we could take

$$S_{i,k+1}^1 \approx f_k^1 \sum_{j=0}^{k} S_{i,j}^1$$

and

$$S_{i,k+1}^0 \approx f_k^0 \sum_{j=0}^{k} (S_{i,j}^0 + S_{i,j}^1).$$
Basic idea behind Linear-Stochastic-Reserving methods

Linear-Stochastic-Reserving methods are reserving methods for a whole collection of claim properties $S_{m}^{i,k}$ (triangles), which may be:
- payments
- incurred losses
- number of reported claims
- small or large claims
- ...

of the same or different portfolios.

The basic assumption behind LSRMs is that the changes of each claim property $S_{m}^{i,k}$ are approximately proportional to an exposure $R_{m}^{i,k}$, which is a linear combination of claim properties of the past.

For instance, denote subrogation by $S_{m}^{0}$ and other payments by $S_{m}^{1}$. Then we could take

$$S_{m}^{1} (i,k) + 1 \approx f_{1} \sum_{j=0}^{k} S_{m}^{1} (i,j)$$
$$S_{m}^{0} (i,k) + 1 \approx f_{0} \sum_{j=0}^{k} \left( S_{m}^{0} (i,j) + S_{m}^{1} (i,j) \right)$$.
4 Linear-Stochastic-Reserving methods

4.1 How do Linear-Stochastic-Reserving methods (LSRM) work

4.1.2 Stochastic behind LSRMs (1/8)

- $\mathcal{B}_{i,k}$ is the $\sigma$-algebra of all information of accident period $i$ up to development period $k$:

$$\mathcal{B}_{i,k} := \sigma (S_{i,j} : 0 \leq j \leq k) = \sigma (C_{i,j} : 0 \leq j \leq k)$$

- $\mathcal{D}_{i,k}$ is the $\sigma$-algebra containing all information up to accident period $i$ and development period $k$:

$$\mathcal{D}_{i,k} := \sigma (S_{i,j} : 0 \leq h \leq i, \ 0 \leq j \leq k) = \sigma (B_{h,k} : 0 \leq h \leq i)$$

- $\mathcal{D}^n$ is the $\sigma$-algebra of all information up to calendar period $n$:

$$\mathcal{D}^n := \sigma (S_{i,k} : 0 \leq i \leq I, \ 0 \leq k \leq J \land (n - i))$$

$$= \sigma (C_{i,k} : 0 \leq i \leq I, \ 0 \leq k \leq J \land (n - i))$$

$$= \sigma \left( \bigcup_{i=0}^{I} \bigcup_{k=0}^{J \land (n-i)} \mathcal{B}_{i,k} \right)$$

- $\mathcal{D}_k$ is the $\sigma$-algebra of all information up to development period $k$:

$$\mathcal{D}_k := \sigma (S_{i,j} : 0 \leq i \leq I, \ 0 \leq j \leq k)$$

$$= \sigma (C_{i,j} : 0 \leq i \leq I, \ 0 \leq j \leq k)$$

$$= \sigma \left( \bigcup_{i=0}^{I} \mathcal{B}_{i,k} \right)$$

- $\mathcal{D}_k^n := \sigma (\mathcal{D}^n \cup \mathcal{D}_k)$
The $\sigma$-algebra $D^n_k$ is used in order to enable us to separate two arbitrary payments $S_{i_1,k_1}^{m_1}$ and $S_{i_2,k_2}^{m_2}$ with $(i_1,k_1) \neq (i_2,k_2)$. That means, for all $(i_1,k_1) \neq (i_2,k_2)$ there exists $n$ and $k$ such that

$$S_{i_1,k_1}^{m_1} \in D^n_k \quad \text{and} \quad S_{i_2,k_2}^{m_2} \notin D^n_k.$$
Assumption 4.A (Linear-Stochastic-Reserving method)

We call the stochastic model of the increments $S_{m}^{i,k}$ a Linear-Stochastic-Reserving method (LSRM) with

- development exposures $R_{m}^{i,k} \in D_{i,k}$, which depend linearly on the claim properties, and
- covariance exposures $R_{m1,m2}^{i,k} \in D_{i,k},$

if there exist constants $f_{k}^{m}$ and $\sigma_{k}^{m1,m2}$ such that

i)\textsuperscript{LSRM} for all $m$, $i$ and $k$, the expectation of the claim property $S_{m}^{i,k+1}$ under the condition of all information of its past $D_{i,k}^{i+k}$ is proportional to $R_{m}^{i,k}$, i.e.

$E\left[S_{m}^{i,k+1}|D_{i,k}^{i+k}\right] = f_{k}^{m} R_{i,k}^{m}.$

ii)\textsuperscript{LSRM} for all $m_1$, $m_2$, $i$ and $k$, the covariance of the claim properties $S_{m1}^{i,k+1}$ and $S_{m2}^{i,k+1}$ under the condition of all information of their past $D_{i,k}^{i+k}$ is proportional to $R_{i,k}^{m1,m2}$, i.e.

$\text{Cov}\left[S_{m1}^{i,k+1}, S_{m2}^{i,k+1}|D_{i,k}^{i+k}\right] = \sigma_{k}^{m1,m2} R_{i,k}^{m1,m2}.$
Stochastic Reserving

Linear-Stochastic-Reserving methods

How do Linear-Stochastic-Reserving methods (LSRM) work

- We will call the parameters \( f^m_k \) and \( \sigma^m_{1,2} \) development factors and covariance parameters, respectively.

- The stochastic model of LSRMs was introduced in [20]. Unfortunately, this article contains some typing errors, which make the implementation very hard. Therefore, a corrected version can be obtained by the lecturer. But, in the next lectures we will use a different approach to derive estimators of the uncertainties.

- A GPL-licensed implementation of LSRMs (ActiveX component and a corresponding Excel interface) can be obtained from [http://sourceforge.net/projects/lsrmtools/](http://sourceforge.net/projects/lsrmtools/).

- The choice of the exposures \( R^m_i,k \) and \( R^{m1,2}_{i,k} \) is of great importance. Unfortunately, we neither can provide a statistical nor a general heuristic concept for this choice. In some cases there is portfolio based information that may help with the choice of exposures, for instance for subrogation. Another useful technique is back-testing, that means to look for exposures for which we see now that the corresponding projections would have been reliable in the past.
4 Linear-Stochastic-Reserving methods

4.1 How do Linear-Stochastic-Reserving methods (LSRM) work

4.1.2 Stochastic behind LSRMs

LSRM step by step
$$\Gamma_{i,k}^m$$ denotes the linear operator that generates $$R_{i,k}^m$$. 

Stochastic Reserving

- Linear-Stochastic-Reserving methods
- How do Linear-Stochastic-Reserving methods (LSRM) work
Remark 4.1 (Dependencies of accident periods)

- There is no additional assumption about independent accident periods necessary! 😊

- Roughly spoken, part ii)\(_{\text{LSRM}}\) means something like: ‘accident periods are uncorrelated up to the first column’. This means LSRMs are affected by (changes in) inflation, too! 😞

- But known diagonal effects can be easily compensated by changing the exposures. 😊

- The choice of the exposures \(R_{i,k}^{m_1,m_2}\) is not completely free. They have to fulfil the covariance assumption ii)\(_{\text{LSRM}}\), which means that all resulting corresponding covariance matrices have to be positive semi-definite. 😞
Remark 4.1 (Dependencies of accident periods)

• There is no additional assumption about independent accident periods necessary!

• Roughly spoken, part ii)\textsubscript{LSRM} means something like: ‘accident periods are uncorrelated up to the first column’.
   This means LSRM are affected by (changes in) inflation, too.

• But known diagonal effects can be easily compensated by changing the exposures.

• The choice of the exposures $R_{mk}^{\text{LSRM}}$ is not completely free. They have to fulfill the covariance assumption ii)\textsubscript{LSRM}, which means that all resulting corresponding covariance matrices have to be positive semi-definite.
Lemma 4.2

Assume $S_{i,k}^m$ satisfy Assumption 4.A. Then

a) $E\left[S_{i,k+1}^m | D^{i+k}\right] = E\left[S_{i,k+1}^m | D_k\right] = E\left[S_{i,k+1}^m | D_{i,k}\right] = E\left[S_{i,k+1}^m | D^{i+k} \cap D_k\right] = f^m_k R_{i,k}^m$.

b) $\text{Cov}\left[S_{i,k+1}^{m_1}, S_{i,k+1}^{m_2} | D^{i+k}\right] = \text{Cov}\left[S_{i,k+1}^{m_1}, S_{i,k+1}^{m_2} | D_k\right] = \text{Cov}\left[S_{i,k+1}^{m_1}, S_{i,k+1}^{m_2} | D_{i,k}\right] = \text{Cov}\left[S_{i,k+1}^{m_1}, S_{i,k+1}^{m_2} | D^{i+k} \cap D_k\right] = \sigma_{k}^{m_1,m_2} R_{i,k}^{m_1,m_2}$.

c) $\text{Cov}\left[S_{n+1-j_1,j_1}^{m_1}, S_{n+1-j_2,j_2}^{m_2} | D^n\right] = 0$, for $j_1 \neq j_2$.

d) provided that all exposures $R_{i,k}^m$ depend only on the $i$-th accident period, all accident periods will be uncorrelated under the knowledge of some past, i.e. for all $\sigma$-algebras $D^n_k$, all $i_1 \neq i_2$ and arbitrary $k_1, k_2, m_1$ and $m_2$ we have $\text{Cov}\left[S_{i_1,k_1}^{m_1}, S_{i_2,k_2}^{m_2} | D^n_k\right] = 0$.

e) If we have independent accident periods the conditioning on $D^{i+k}_k$ could be replaced by conditioning on $B_{i,k}$.
a), b) Follows from the measurability of $R_{i,k}^m$ and $R_{i,k}^{m_1,m_2}$ with respect to $D_{i,k}$.

c) Assume that $j_1 > j_2$. Then $S_{n+1-j_1,j_1}^{m1}$ is $D_{j_1-1}^n$-measurable and we get

$$ \text{Cov}[S_{n+1-j_1,j_1}^{m1}, S_{n+1-j_2,j_2}^{m2} | D^n] = \text{Cov}[E[S_{n+1-j_1,j_1}^{m1} | D_{j_1-1}^n], E[S_{n+1-j_2,j_2}^{m2} | D_{j_1-1}^n] | D^n] + E[\text{Cov}[S_{n+1-j_1,j_1}^{m1}, S_{n+1-j_2,j_2}^{m2} | D_{j_1-1}^n] | D^n] $$

is $D^n$ measurable.}

$$ S_{n+1-j_1,j_1}^{m1} \text{ is } D_{j_1-1}^n \text{-measurable}$$

$$ = 0 $$

d) If $S_{i_1,k_1}^{m1}$ or $S_{i_2,k_2}^{m2}$ is measurable with respect to $D_k^n$ we are done. Otherwise, $D_k^n$ is a subset of $D_{i_1+k_1-1}$ and $D_{i_2+k_2-1}$, and $S_{i_1,k_1}^{m1}$ is measurable with respect to the past of $S_{i_2,k_2}^{m2}$, or vice versa.

Without loss of generality assume that $S_{i_1,k_1}^{m1}$ is $D_{k_2-1}^{i_2+k_2-1}$-measurable. Then we get

$$ \text{Cov}[S_{i_1,k_1}^{m1}, S_{i_2,k_2}^{m2} | D_k^n] = E[\text{Cov}[S_{i_1,k_1}^{m1}, S_{i_2,k_2}^{m2} | D_{k_2-1}^{i_2+k_2-1}] | D_k^n] $$

$$ + \text{Cov}[E[S_{i_1,k_1}^{m1} | D_{k_2-1}^{i_2+k_2-1}], E[S_{i_2,k_2}^{m2} | D_{k_2-1}^{i_2+k_2-1}] | D_k^n] $$

$$ = 0 + \text{Cov}[S_{i_1,k_1}^{m1}, f_{k_2-1}^{m2} R_{i_2,k_2-1}^{m2} | D_k^n]. $$

Since $R_{i_2,k_2-1}^{m2} \in B_{i_2,k_2-1}$ and depends linearly on $S$ it is enough to show that $S_{i_1,k_1}^{m1}$ and $S_{i_2,k_2-1}^{m2}$ are $D_k^n$-conditional uncorrelated. Iteration until $S_{i_1,k_1}^{m1} - j_1$ or $S_{i_2,k_2-1}^{m2} - j_2$ is $D_k^n$-measurable proves part d).

e) Because of independent accident periods.
Remark 4.3 (CLM as LSRM)

Because of Corollary 2.3, i.e.

\[ E\left[ S_{i,k+1}^0 | D_{k}^{i+k} \right] = (f_k - 1) \sum_{j=0}^{k} S_{i,j}^0 = (f_k - 1)C_{i,k}, \]

\[ \text{Cov}\left[ S_{i,k+1}^0, S_{i,k+1}^0 | D_{k}^{i+k} \right] = \sigma_k^2 \sum_{j=0}^{k} S_{i,j}^0 = \sigma_k^2C_{i,k}, \]

the Chain-Ladder method is a LSRM with exposures

\[ R_{i,k}^0 = R_{i,k}^{0,0} = C_{i,k} \]

and parameters

\[ f_k^0 = f_k - 1, \]

\[ \sigma_k^{0,0} = \sigma_k^2. \]
Stochastic Reserving

---

Linear-Stochastic-Reserving methods

How do Linear-Stochastic-Reserving methods (LSRM) work

Remark 4.3 (CLM as LSRM)
Because of Corollary 2.3, i.e.

\[
E[S_{0,i,k+1} | D_{i,k}] = (f_k - 1) \sum_{j=0}^{f_k} S_{j,i} - (f_k - 1) C_{i,k},
\]

\[
Cov[S_{0,i,k+1}, S_{0,i,k} | D_{i,k}] = \sigma_k^2 \sum_{j=0}^{f_k} S_{j,i} = \sigma_k^2 C_{i,k},
\]

the Chain-Ladder method is a LSRM with exposures

\[
R_{0,i,k} = R_{0,i,k}^0 - C_{i,k},
\]

and parameters

\[
f_k^0 = f_k - 1,
\]

\[
\sigma_k^0 = \sigma_k^2.
\]
Remark 4.4 (CLRM as LSRM)

If we set

\[ S_{i,k}^1 := \begin{cases} P_i, & \text{for } k = 0, \\ 0, & \text{otherwise,} \end{cases} \]

then the Complementary-Loss-Ratio method can be rewritten as

\[
E \left[ S_{i,k+1}^m \mid D_k^{i+k} \right] = f_k^m P_i \\
E \left[ S_{i,k+1}^{m_1}, S_{i,k+1}^{m_2} \mid D_k^{i+k} \right] = \sigma_k^{m_1,m_2} P_i
\]

with parameters

\[
f_k^0 = f_k \quad \text{and} \quad f_k^1 = 0, \\
\sigma_k^{0,0} = \sigma_k^2 \quad \text{and} \quad \sigma_k^{0,1} = \sigma_k^{1,0} = \sigma_k^{1,1} = 0.
\]

Therefore, it is a LSRM with exposure parameters

\[
R_{i,k}^0 = R_{i,k}^1 = R_{i,k}^{0,0} = R_{i,k}^{0,1} = R_{i,k}^{1,0} = R_{i,k}^{1,1} = P_i.
\]
Remark 4.4 (CLRM as LSRM)

If we set
\[ S_{i,k} := \begin{cases} P_i, & \text{for } k = 0, \\ 0, & \text{otherwise,} \end{cases} \]
then the Complementary-Loss-Ratio method can be rewritten as
\[
E \left[ S_{m,i,k} + 1 \mid D_i^k \right] = f_{m,k} P_i
\]
\[
E \left[ S_{m,i,k} + 1, S_{m,i,k}^{11} \mid D_i^k \right] = \sigma_{m,k} P_i
\]
with parameters
\[
f_0^k = f_k \quad \text{and} \quad f_1^k = 0,
\]
\[
\sigma_{0,0}^k = \sigma_k \quad \text{and} \quad \sigma_{0,1}^k = \sigma_{1,0}^k = \sigma_{1,1}^k = 0.
\]
Therefore, it is a LSRM with exposure parameters
\[
R_{1,k} = R_{1,k}^0 = R_{1,k}^{01} = R_{1,k}^{10} = R_{1,k}^{11} = P_i.
\]
Remark 4.5 (BFM as LSRM)

Since we can look at BFM as a Complementary-Loss-Ratio method (see Remark 3.8), it can also be interpreted as LSRM.

Remark 4.6 (ECLRM as LSRM)

By definition is the Extended-Complementary-Loss-Ratio method a LSRM with exposure parameters

\[ R_{i,k}^{0,0} = R_{i,k}^{0,1} = R_{i,k}^{1,0} = R_{i,k}^{1,1} = \sum_{j=0}^{k} (S_{i,k}^1 - S_{i,k}^0) \]

and parameters \( f_k^m \) and \( \sigma_k^{m_1,m_2} \).
Stochastic Reserving

Linear-Stochastic-Reserving methods

How do Linear-Stochastic-Reserving methods (LSRM) work

Remark 4.5 (BFM as LSRM)
Since we can look at BFM as a Complementary-Loss-Ratio method (see Remark 3.8), it can also be interpreted as LSRM.

Remark 4.6 (ECLRM as LSRM)
By definition is the Extended-Complementary-Loss-Ratio method a LSRM with exposure parameters

\[ R_{i,k}^{0} - R_{i,k}^{1} = R_{i,k}^{2} - R_{i,k}^{3} - \sum_{j=0}^{k} (S_{i,k}^{1} - S_{i,k}^{0}) \]

and parameters \( f^{m}_{k} \) and \( \sigma^{m}_{j,k} \).
Estimator 4.7 (of the development parameter $f^m_k$)

Let $S^m_{i,k}$ satisfy Assumption 4.A. Then for each set of $\mathcal{D}^I \cap \mathcal{D}_k$-measurable weights $w^m_{i,k}$ with

- $w^m_{i,k} \geq 0$ and $R^m_{i,k} = 0$ implies $w^m_{i,k} = 0$,
- $\sum_{i=0}^{I-1-k} w^m_{i,k} = 1$ if at least one $R^m_{i,k} \neq 0$

we get that

$$\hat{f}^m_k := \sum_{i=0}^{I-1-k} w^m_{i,k} \frac{S^m_{i,k+1}}{R^m_{i,k}}$$

(4.1)

is a $\mathcal{D}_k$-conditionally unbiased estimator of the development factor $f^m_k$ and the weights

$$w^m_{i,k} := \left( \frac{R^m_{i,k}}{R^m_{i,k}} \right)^2 \left( \sum_{h=0}^{I-1-k} \left( \frac{R^m_{h,k}}{R^m_{h,k}} \right)^2 \right)^{-1}$$

(4.2)

result in estimators $\hat{f}^m_k$ with minimal ($\mathcal{D}_k$-conditional) variance of all estimators of the form (4.1).

- For every tuple $\hat{f}^m_{k_1}, \ldots, \hat{f}^m_{k_r}$ with $k_1 < k_2 < \cdots < k_r$ we get

$$\mathbb{E} \left[ \hat{f}^m_{k_1} \cdots \hat{f}^m_{k_r} \mid \mathcal{D}_{k_1} \right] = f^m_{k_1} \cdots f^m_{k_r} = \mathbb{E} \left[ \hat{f}^m_{k_1} \mid \mathcal{D}_{k_1} \right] \cdots \mathbb{E} \left[ \hat{f}^m_{k_r} \mid \mathcal{D}_{k_r} \right],$$

which implies that the estimators are pairwise $\mathcal{D}_{k_1}$-conditionally uncorrelated.
For every tuple $(i,k)$ the unbiased: $E\left[ \hat{f}^m_k \mid D_k \right] = \sum_{i=0}^{I-1-k} w_{i,k}^m \frac{E\left[ S_{i,k+1}^m \mid D_k \right]}{R_{i,k}^m m} = \sum_{i=0}^{I-1-k} w_{i,k}^m \frac{f_k^m R_{i,k}^m}{R_{i,k}^m m} = f_k^m$

minimal variance: $\text{Var}[\hat{f}_k^m] = E\left[ \text{Var}[\hat{f}_k^m \mid D_k] \right] + \text{Var}[E[\hat{f}_k^m \mid D_k]] = E\left[ \text{Var}[\hat{f}_k^m \mid D_k] \right] + 0$

$\text{Var}[\hat{f}_k^m \mid D_k] = \text{Var} \left[ \sum_{i=0}^{I-1-k} w_{i,k}^m \frac{S_{i,k+1}^m}{R_{i,k}^m} \mid D_k \right] = \sum_{i=0}^{I-1-k} (w_{i,k}^m)^2 \frac{\text{Var}[S_{i,k+1}^m \mid D_k]}{(R_{i,k}^m)^2} = \sum_{i=0}^{I-1-k} (w_{i,k}^m)^2 \frac{\text{Var}[S_{i,k+1}^m \mid D_k]}{(R_{i,k}^m)^2} = \sigma_k^m \sum_{i=0}^{I-1-k} (w_{i,k}^m)^2 \frac{R_{i,k}^m}{(R_{i,k}^m)^2} = \sigma_k^m \sum_{i=0}^{I-1-k} (w_{i,k}^m)^2 \frac{R_{i,k}^m}{(R_{i,k}^m)^2}$

measurable with respect to $D_k$ and Remark 4.1

Lagrange: minimize

$\sum_{i=0}^{I-1-k} (w_{i,k}^m)^2 R_{i,k}^m \frac{R_{i,k}^m}{(R_{i,k}^m)^2} + \lambda \left(1 - \sum_{i=0}^{I-1-k} w_{i,k}^m\right)$

$\frac{\partial}{\partial w_{i,k}^m} = 2w_{i,k}^m \frac{R_{i,k}^m}{(R_{i,k}^m)^2} - \lambda$ \quad $\implies$ \quad $w_{i,k}^m = \frac{\lambda (R_{i,k}^m)^2}{2 R_{i,k}^m m} \frac{R_{i,k}^m}{R_{i,k}^m m}$ \quad $\text{and}$ \quad $\lambda = 2 \left( \sum_{i=0}^{I-1-k} \frac{(R_{i,k}^m)^2}{R_{i,k}^m m} \right)^{-1}$

\[ \sum_{i=0}^{I-1-k} w_{i,k}^m = 1 \]

uncorrelated: $E\left[ \hat{f}^{m_1}_{k_1} \cdots \hat{f}^{m_r}_{k_r} \mid D_{k_1} \right] = E\left[ E\left[ \hat{f}^{m_1}_{k_1} \cdots \hat{f}^{m_r}_{k_r} \mid D_{k_r} \right] \mid D_{k_1} \right] = E\left[ \hat{f}^{m_1}_{k_1} \cdots \hat{f}^{m_{r-1}}_{k_{r-1}} E\left[ \hat{f}^{m_r}_{k_r} \mid D_{k_r} \right] \mid D_{k_1} \right] = E\left[ \hat{f}^{m_1}_{k_1} \cdots \hat{f}^{m_{r-1}}_{k_{r-1}} \mid D_{k_1} \right] f^{m_r}_{k_r} = \cdots = f^{m_1}_{k_1} \cdots f^{m_r}_{k_r}$
**Definition 4.8 (Diagonal by diagonal projection)**

Since the exposures $R_{i,k}^m$ depend linearly on claim properties, there exist linear operators $\Gamma_{i,k}^m$, which generate these exposures. We now want to formalise the diagonal by diagonal projection. Therefore, we denote by

$$\#^n := \#\{(m, i, k): 0 \leq m \leq M, 0 \leq i \leq I, 0 \leq k \leq J - 1, 0 \leq i + k \leq n\}$$

the number of claim properties below or on the diagonal $n$ and define

$$F_{i,k}^m(g): \mathbb{R}^{\#^{i+k}} \rightarrow \mathbb{R}: \quad F_{i,k}^m(g) := g_{i,k}^m \Gamma_{i,k}^m,$$

$$F^n(g): \mathbb{R}^{\#^n} \rightarrow \mathbb{R}^{\#^{n+1}}: \quad (F^n(g) x)_{i,k}^m := \begin{cases} x_{i,k}^m, & \text{if } i + k \leq n, \\ F_{i,k-1}^m(g) x, & \text{otherwise,} \end{cases}$$

$$F_{n2\leftarrow n1}^n(g): \mathbb{R}^{\#^{n1}} \rightarrow \mathbb{R}^{\#^{n2+1}}: \quad F_{n2\leftarrow n1}^n(g) := \begin{cases} F_{n2}^n(g) \circ \cdots \circ F_{n1}^n(g), & \text{if } n_2 \geq n_1, \\ \Pi^{\#_{n2+1}}, & \text{otherwise,} \end{cases}$$

$$F_{i,k}^{m,n}(g): \mathbb{R}^{\#^n} \rightarrow \mathbb{R}: \quad F_{i,k}^{m,n}(g) x := \left(F_{i+k\leftarrow n}^1(g) x\right)_{i,k+1}^m,$$

where $\Pi^{\#^n}$ denotes the projection onto $\mathbb{R}^{\#^n}$ and $g$ is any large enough vector with co-ordinates $g_{i,k}^m$. 
Definition 4.8 (Diagonal by diagonal projection)

Since the exposures $R_{m,i,k}$ depend linearly on claim properties, there exist linear operators $\Gamma_{m,i,k}$ which generate these exposures. We now want to formalise the diagonal by diagonal projection. Therefore, we denote by $\# n := \{(m, i, k): 0 \leq m \leq M, 0 \leq i \leq I, 0 \leq k \leq J - 1, 0 \leq i + k \leq n\}$ the number of claim properties below or on the diagonal $n$ and define $F_{m,i,k}(g): \mathbb{R}^{\# i + k} \rightarrow \mathbb{R}^{\# m + 1}$:

$$F_{m,i,k}(g):= g_{m,i,k} \Gamma_{m,i,k},$$

$$F_{n}(g): \mathbb{R}^{\# n} \rightarrow \mathbb{R}^{\# n + 1}: (F_{n}(g)x)_{m,i,k} := \begin{cases} x_{m,i,k}, & \text{if } i + k \leq n, \\ F_{m,i,k - 1}(g)x, & \text{otherwise}, \end{cases}$$

$$F_{m,n}(g): \mathbb{R}^{\# n} \rightarrow \mathbb{R}^{\# m + 1}: (F_{m,n}(g)x)_{i} := \begin{cases} F_{n}(g) \circ \cdots \circ F_{n}(g), & \text{if } n \geq n_1, \\ \Pi^{\# n}, & \text{otherwise}, \end{cases}$$

where $\Pi^{\# n}$ denotes the projection onto $\mathbb{R}^{\# n}$ and $g$ is any large enough vector with coordinates $g_{m,i,k}$.

- Since the operators $\Gamma_{m,i,k}$ and $F_{m,i,k}(g)$ only depend on coordinates $(l, h, j)$ with $0 \leq m \leq M, 0 \leq h \leq i, 0 \leq j \leq k$ they could be defined on a smaller domain, but than concatenation would not be possible.

- We added the parameter $g$ in order to denote which development factors we are using, for instance the real, but unknown, development vectors $f_{m,k}$ or their estimates $\hat{f}_{m,k}$.

- In most cases we will use parameters $g$, which do not depend on the accident period $i$ and will skip the index $i$ in $g_{i,k}$. 

Stochastic Reserving

- Linear-Stochastic-Reserving methods

  Future development

2019-02-18
Lemma 4.9 (multi-linear structure of $\mathcal{F}$)

For all $i + k \geq n$ and for all $Y \in \mathcal{D}^n$ there exist random variables $X_{i,k,h_1,...,h_r,j_1,...,j_r} \in \mathcal{D}_{i,k} \cap \mathcal{D}^n$, which depend linearly on the coordinates of $Y$, such that for all $g$

$$\mathcal{F}_{i,k}^{m,n}(g) Y = \sum_{r=1}^{k+i+1-n} \sum_{0 \leq l_1,...,l_r \leq M} \sum_{0 \leq h_1,...,h_r \leq i} \sum_{n-i \leq j_1 < \cdots < j_r \leq k} g_{l_1} h_{l_1} \cdots g_{l_r} h_{l_r} X_{i,k,l_1,...,l_r}.$$

Remark 4.10

That means we have a multi-linear structure in the development factors as well as in the claims properties, like in the Chain-Ladder case. Most of the following derivation could be done using this representation of the operator $\mathcal{F}_{i,k}^{m,n}(g)$, but since it is not really handy we will use conditional expectations instead.
Lemma 4.9 (multi-linear structure of $\mathcal{F}$)

For all $i + k \geq n$ and for all $Y \in D^n$, there exist random variables

$$X_{n,m,l}^{i,k,h_1},\ldots,l_r^{i,k,h_r,j_1},\ldots,j_r \in D_{i,k} \cap D_{n},$$

which depend linearly on the coordinates of $Y$, such that for all $g$

$$\mathcal{F}_{i,k}^{m,n}(g) Y = g_{i,k}^{m} \Gamma_{i,k}^{m} Y \quad \in D_{i,k} \cap D_{i+k}.$$

Remark 4.10

That means we have a multi-linear structure in the development factors as well as in the claims properties, like in the Chain-Ladder case. Most of the following derivation could be done using this representation of the operator $\mathcal{F}_{i,k}^{m,n}(g)$, but since it is not really handy we will use conditional expectations instead.

If $i + k = n$ we get

$$\mathcal{F}_{i,k}^{m,n}(g) Y = g_{i,k}^{m} \Gamma_{i,k}^{m} Y \quad \in D_{i,k} \cap D_{i+k}.$$

Now assume that the statement is fulfilled for all $n \leq h + j < i + k$. Then we get

$$\mathcal{F}_{i,k}^{m,n}(g) Y = \mathcal{F}_{i,k}^{m}(g) \circ \mathcal{F}_{i+k-1}^{i+k-1,n}(g) Y = g_{i,k}^{m} \Gamma_{i,k}^{m} \circ \mathcal{F}_{i+k-1}^{i+k-1,n}(g) Y.$$

By assumption the statement is fulfilled for each coordinate of $\mathcal{F}_{i+k-1}^{i+k-1,n}(g) Y$ and since $\Gamma_{i,k}^{m}$ depends only on coordinates $h \leq i$ and $j \leq k$, only development factors $g_{h,j}^{l}$ with $n - i \leq j < k$ are involved, which by induction proves our statement.
Remark 4.11

- The mapping $\mathcal{F}^n(g)$ fills the $(n+1)$-th diagonal of all claim property triangles based on all diagonals up to the $n$-th diagonal.
- The functional $\mathcal{F}^m_{i,k}(g)$ does depend on coordinates up to accident period $i$ and development period $k$, only.
- $\mathcal{F}^m_{i,k}(g) \mathbf{x} = (\mathcal{F}^{i+k}(g) \mathbf{x})^m_{i,k+1}$,
- $R^m_{i,k} = \Gamma^m_{i,k} \mathbf{S}^{i+k}$,
- $E\left[ S^m_{i,k+1} \middle| \mathcal{D}^{i+k}_k \right] = \mathcal{F}^m_{i,k}(f) \mathbf{S}^{i+k}$,
- $E\left[ S^{n_1+n_2+1}_{n_1} \middle| \mathcal{D}^{n_1}_n \right] = \mathcal{F}^{n_2 \leftarrow n_1}(f) \mathbf{S}^{n_1}$,
- $E\left[ S^m_{i,k+n+1} \middle| \mathcal{D}^{i+k}_k \cap \mathcal{D}_k \right] = E\left[ S^m_{i,k+n+1} \middle| \mathcal{D}^{i+k}_k \right] = \mathcal{F}^m_{i,k+n}(f) \mathbf{S}^{i+k}$,

where $f := (f^m_k)_{0 \leq m \leq M}^{0 \leq k < J}$ denotes the vector of the real (but unknown) development factors and

$$\mathbf{S}^n := (S^m_{i,k})_{0 \leq i \leq I, 0 \leq k < J, 0 \leq i+k \leq n}^{0 \leq m \leq M}$$

is the vector of all claim properties below or on the diagonal $n$. 
Remark 4.11

- The mapping \( F_n(g) \) fills the \((n+1)\)-th diagonal of all claim property triangles based on all diagonals up to the \( n \)-th diagonal.
- The functional \( F_m,i,k(g) \) does depend on coordinates up to accident period \( i \) and development period \( k \), only.
- \( F_m,i,k(g) \times = \left( F_i + k(g) \right) m \) \( i,k+n+1 \),
- \( R_m,i,k = \Gamma_m,i,k S_i + k \),
- \( E\left[ S_{m,i,k+n+1} | D_{i+k} \right] = F_m,i,k+1(f) S_{i,k+1} \),
- \( E\left[ S_{m,i,k+n+1} | D_{i+k} \right] = F_m,i,k+1(f) S_{i,k+1} \),
- \( S^n := \left( S_{m,i,k+n+1} \right) \),

where \( f := \left( f_{m,i,k+n} \right)_{0 \leq m \leq M, 0 \leq i \leq I, 0 \leq k < J} \) denotes the vector of the real (but unknown) development factors and \( S^n := \left( S_{m,i,k+n} \right) \) is the vector of all claim properties below or on the diagonal \( n \).
Estimator 4.12 (of the future development)

Let \( S_{i,k}^m \) satisfy Assumption 4.A. Then

\[
\hat{S}_{i,k+1}^m := \mathcal{F}^m_{i,k}(\hat{f}) S^I, \quad I - i \leq k < J,
\]

are \( \mathcal{D}_{I-i} \)-conditional unbiased estimators of \( \mathbb{E}[S_{i,k+1}^m | \mathcal{D}^I] \).

Moreover, we define \( \hat{S}_{i,k}^m := S_{i,k}^m \), for \( i + k \leq I \), and

\[
\hat{R}_{i,k}^m := \Gamma_{i,k}^m \hat{S}^{i+k} \quad \text{and} \quad \hat{R}_{i,k}^{m_1,m_2} := \Gamma_{i,k}^{m_1,m_2} \hat{S}^{i+k},
\]

where \( \Gamma_{i,k}^{m_1,m_2} \) denotes the operator that generates \( R_{i,k}^{m_1,m_2} \) based on \( S^{i+k} \).
We will even prove that \( \hat{S}_{i,k+1}^m \) is an \( D^I_{-h} \)-conditionally unbiased estimator of \( E[S_{i,k+1}^m D^I] \) for all \( h \geq i \). We will do that by induction. If \( i + k = I \) we get for all \( h \geq i \)

\[
E[\hat{S}_{i,k+1}^m | D^I_{-h}] = E[E[\hat{S}_{i,k+1}^m | D_k] | D^I_{-h}] = E[E[\hat{f}_k^m \Gamma_{i,k}^m S^I | D_k] | D^I_{-h}] = E[E[\hat{f}_k^m | D_k] \Gamma_{i,k}^m S^I | D^I_{-h}]
\]

**Estimator 4.7**

Now assume that the statement is fulfilled for all \( i + k < n \). Then we get for \( i + k = n \) and all \( h \geq i \)

\[
E[\hat{S}_{i,k+1}^m | D^I_{-h}] = E[F^m_{i,k}(\hat{f}) \hat{S}^{i+k} | D^I_{-h}] = E[E[F^m_{i,k}(\hat{f}) | D_k] \hat{S}^{i+k} | D^I_{-h}]
\]

**Remark 4.11**

\[
= E[F^m_{i,k}(\hat{f}) \hat{S}^{i+k} | D^I_{-h}] = F^m_{i,k}(\hat{f}) E[\hat{S}^{i+k} | D^I_{-h}].
\]

Since \( F^m_{i,k}(\hat{f}) \) depends only on accident periods \( h_1 \leq i \), all coordinates \( E[\hat{S}_{h_1,j}^l | D^I_{-h}] \) of \( E[\hat{S}^{i+k} | D^I_{-h}] \) with \( h_1 > i \) will not be taken into account. For all others we can apply the induction hypotheses and proceed with

\[
= F^m_{i,k}(\hat{f}) E[S^{i+k} | D^I_{-h}] = E[E[S_{i,k+1}^m D^I | D^I_{-h}] = E[E[S_{i,k+1}^m D^I] | D^I_{-h}].
\]

**induction hypothesis**

\[
\]

**Remark 4.11**

Note, since \( R_{1,i,k}^m,m_2 \) is \( D^{i+k} \) measurable, there always exists an operator \( \Gamma_{1,i,k}^{m_1,m_2} \) such that

\[
R_{1,i,k}^m,m_2 = \Gamma_{1,i,k}^{m_1,m_2} S^{i+k}.
\]
Example 4.13 (Swiss mandatory accident portfolio: part 1 of 3, see LSRM_Accident_ActiveX.xlsx)

We have the following three main types of (non annuity) payments:

- **Medical expenses (ME)** will be estimated by CLM, because it worked fine in the past.
- **Payments for incapacitation for work (IW)** are by law proportional to the insured salary $P_i$, which is limited to a maximum amount. Moreover, during accident period 7 the maximum insured salary has been increased by about 20%, valid for all claims happening afterwards. Therefore, we would like to take CLRM with the insured salary as external exposure.
  
  On the other side, we know from the past that the claim frequency is influenced by the economic situation, which is better reflected by CLM than by CLRM. Combining both we take a mixture of the exposures of both methods, whereas the weight of the insured salary is $\kappa^{k+1}$.

- **Subrogation (Sub)** possibilities are huge, because many claims are caused by car accidents and by law the accident insurer of the insured persons has to pay first and may take subrogation against the motor liability insurer afterwards.
  
  Therefore, we assume that the amount of possible subrogation is proportional to the total amount that already has been paid, i.e. to ME+IW+Sub.
We have the following three main types of (non annuity) payments:

- **Medical expenses (ME)** will be estimated by CLM, because it worked fine in the past.
- **Payments for incapacitation for work (IW)** are by law proportional to the insured salary $P_i$, which is limited to a maximum amount. Moreover, during accident period 7 the maximum insured salary has been increased by about 20%, valid for all claims happening afterwards. Therefore, we would like to take CLRM with the insured salary as external exposure. On the other side, we know from the past that the claim frequency is influenced by the economic situation, which is better reflected by CLM than by CLRM. Combining both we take a mixture of the exposures of both methods, whereas the weight of the insured salary is $\kappa + 1$.
- **Subrogation (Sub)** possibilities are huge, because many claims are caused by car accidents and by law the accident insurer of the insured person has to pay first and may take subrogation against the motor liability insurer afterwards. Therefore, we assume that the amount of possible subrogation is proportional to the total amount that already has been paid, i.e. to $ME + IW + Sub$. 

---

Example 4.13 (Swiss mandatory accident portfolio: part 1 of 3, see LSNF_Accident_ActiveX.xlsx)
Mathematical that means:

We have four claim properties with exposures

- **ME:** $R_{i,k}^0 = R_{i,k}^{0,0} = \sum_{j=0}^{k} S_{i,j}^0$
- **IW:** $R_{i,k}^1 = R_{i,k}^{1,1} = \sum_{l=0}^{k} \left( \kappa^{l+1} S_{i,l}^3 + (1 - \kappa^{l+1}) S_{i,l}^1 \right)$
- **Sub:** $R_{i,k}^2 = R_{i,k}^{2,2} = \sum_{l=0}^{k} \left( S_{i,l}^0 + S_{i,l}^1 + S_{i,l}^2 \right)$
- **Salary:** $S_{i,0}^3 = P_i$, $S_{i,l}^3 = 0$, for $l > 0$, and $R_{i,k}^3 = R_{i,k}^{3,0} = R_{i,k}^{0,3} = R_{i,k}^{3,1} = R_{i,k}^{1,3} = R_{i,k}^{3,2} = R_{i,k}^{2,3} = R_{i,k}^{3,3} = 0$

For the not yet defined exposures we take the total payments up to now, i.e.

$R_{i,k}^{0,1} = R_{i,k}^{1,0} = R_{i,k}^{0,2} = R_{i,k}^{2,0} = R_{i,k}^{1,2} = R_{i,k}^{2,1} = \sum_{l=0}^{k} \left( S_{i,l}^0 + S_{i,l}^1 + S_{i,l}^2 \right)$.

Resulting Best Estimate reserves

- Depend almost linear on $\kappa$, because it practically influences only the first development period, that means the most recent accident period $i = 8$.
- Are much higher than the CLM on total payments, if $\kappa = 1$. The main difference is in the most recent accident period $i = 8$.
- Are slightly smaller than CLM, if $\kappa = 0$. This may be a consequence of the more detailed modelling of subrogation.
We have four claim properties with exposures:

**Mathematical that means:**

- **ME:**
  \[ R_{0}^{i,k} = R_{0}, \quad S_{0}^{i,j} = \sum_{k=0}^{j} S_{0}^{i,k} \]

- **IW:**
  \[ R_{1}^{i,k} = R_{1}, \quad \sum_{l=0}^{k} (\kappa_{l} + 1) S_{0}^{i,l} + (1 - \kappa_{l} + 1) S_{1}^{i,l} \]

- **Sub:**
  \[ R_{2}^{i,k} = R_{2}, \quad \sum_{l=0}^{k} (S_{0}^{i,l} + S_{1}^{i,l} + S_{2}^{i,l}) \]

- **Salary:**
  \[ S_{3}^{i,0} = P_{i}, \quad S_{3}^{i,l} = 0, \text{ for } l > 0 \]

For the not yet defined exposures we take the total payments up to now, i.e.

**Resulting Best Estimate reserves**

- Depend almost linear on \( \kappa \), because it practically influences only the first development period, that means the most recent accident period \( i = 8 \).
- Are much higher than the CLM on total payments, if \( \kappa = 1 \). The main difference is in the most recent accident period \( i = 8 \).
- Are slightly smaller than CLM, if \( \kappa = 0 \). This may be a consequence of the more detailed modeling of subrogation.
Example 4.13: Best Estimate reserves in dependence of $\kappa$

- CLM

claim reserves

$\kappa$
• The estimated covariance parameters $\hat{\sigma}_{k}^{m_1,m_2}$ together with the estimated exposures $\hat{R}_{i,k}^{m_1,m_2}$ lead to covariance matrices which are slightly non-positive definite for development periods $k \in \{5, 6, 7\}$. Since the corresponding negative eigenvalues are almost zero we believe that it is not a model but an estimation problem. We could change the estimated covariance parameters slightly in order to get non-negative covariance matrices without changing uncertainties a lot.
Example 4.14 (ECLRM vs. CLM, see Examples 2.7 and 2.8: part 1 of 3, see LSRM_Examples_ActiveX.xlsx)

We have seen that the Chain-Ladder method leaves a gap between the Best Estimate reserves based on payments and the one based on incurred losses. Moreover, we have closed this gap by a credibility like weighting.

Now we want to look at the corresponding results, if we take the case reserves as exposure (ECLRM):

<table>
<thead>
<tr>
<th>AP</th>
<th>CLM paid</th>
<th>CLM incurred</th>
<th>CLM weighting</th>
<th>ECLRM</th>
<th>Case Reserve</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>1</td>
<td>114 086</td>
<td>337 984</td>
<td>228 182</td>
<td>314 902</td>
<td>352 899</td>
</tr>
<tr>
<td>2</td>
<td>394 121</td>
<td>31 884</td>
<td>203 653</td>
<td>66 994</td>
<td>75 316</td>
</tr>
<tr>
<td>3</td>
<td>608 749</td>
<td>331 436</td>
<td>458 946</td>
<td>359 384</td>
<td>410 496</td>
</tr>
<tr>
<td>4</td>
<td>697 742</td>
<td>1 018 350</td>
<td>877 247</td>
<td>981 883</td>
<td>1 148 647</td>
</tr>
<tr>
<td>5</td>
<td>1 234 157</td>
<td>1 103 928</td>
<td>1 157 520</td>
<td>1 115 768</td>
<td>1 317 088</td>
</tr>
<tr>
<td>6</td>
<td>1 138 623</td>
<td>1 868 664</td>
<td>1 587 838</td>
<td>1 786 947</td>
<td>2 216 536</td>
</tr>
<tr>
<td>7</td>
<td>1 638 793</td>
<td>1 997 651</td>
<td>1 862 844</td>
<td>1 942 518</td>
<td>2 923 692</td>
</tr>
<tr>
<td>8</td>
<td>2 359 939</td>
<td>1 418 779</td>
<td>1 750 635</td>
<td>1 569 657</td>
<td>2 756 633</td>
</tr>
<tr>
<td>9</td>
<td>1 979 401</td>
<td>2 556 612</td>
<td>2 412 410</td>
<td>2 590 718</td>
<td>2 203 446</td>
</tr>
<tr>
<td>Total</td>
<td>10 165 612</td>
<td>10 665 287</td>
<td>10 539 276</td>
<td>10 728 771</td>
<td>13 404 753</td>
</tr>
</tbody>
</table>
We have seen that the Chain-Ladder method leaves a gap between the Best Estimate reserves based on payments and the one based on incurred losses. Moreover, we have closed this gap by a credibility like weighting.

Now we want to look at the corresponding results, if we take the case reserves as exposure (ECLRM):

<table>
<thead>
<tr>
<th>Year</th>
<th>Best Estimate reserves</th>
<th>ECLRM reserves</th>
<th>ECLRM weighting</th>
<th>Case Reserve</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>1</td>
<td>114,086</td>
<td>337,984</td>
<td>228,182</td>
<td>314,902</td>
</tr>
<tr>
<td>2</td>
<td>394,121</td>
<td>31,884</td>
<td>203,653</td>
<td>66,994</td>
</tr>
<tr>
<td>3</td>
<td>608,749</td>
<td>331,436</td>
<td>458,946</td>
<td>359,384</td>
</tr>
<tr>
<td>4</td>
<td>697,742</td>
<td>1,018,350</td>
<td>877,247</td>
<td>981,883</td>
</tr>
<tr>
<td>5</td>
<td>1,234,157</td>
<td>1,103,928</td>
<td>1,157,520</td>
<td>1,115,768</td>
</tr>
<tr>
<td>6</td>
<td>1,138,623</td>
<td>1,868,664</td>
<td>1,587,838</td>
<td>1,786,947</td>
</tr>
<tr>
<td>7</td>
<td>1,638,793</td>
<td>1,997,651</td>
<td>1,862,944</td>
<td>1,942,518</td>
</tr>
<tr>
<td>8</td>
<td>2,359,939</td>
<td>2,556,612</td>
<td>2,412,463</td>
<td>2,590,446</td>
</tr>
<tr>
<td>9</td>
<td>1,979,401</td>
<td>2,356,612</td>
<td>2,264,652</td>
<td>2,308,146</td>
</tr>
<tr>
<td>Total</td>
<td>10,165,612</td>
<td>10,665,287</td>
<td>10,539,276</td>
<td>10,728,771</td>
</tr>
</tbody>
</table>

2019-02-18

- CLM on incurred, CLM weighting and ECLRM lead to similar results, whereas the later reflects the information contained in the case reserves at best (see third accident period $i = 2$).

- In total the results of CLM on payments are in the same range like the others, but the estimated reserves for individual accident periods are quite different.
Stochastic Reserving
Lecture 7 (Continuation of Lecture 4)
Linear-Stochastic-Reserving methods

René Dahms

ETH Zurich, Spring 2019

3 April 2019

(Last update: 18 February 2019)
4 Linear-Stochastic-Reserving methods

4.1 How do Linear-Stochastic-Reserving methods (LSRM) work
4.1.1 LSRM without stochastic
4.1.2 Stochastic behind LSRMs

4.2 Future development
4.2.1 Projection of the future development
4.2.2 Examples

4.3 Ultimate uncertainty
4.3.1 Mixing of claim properties
4.3.2 Ultimate uncertainty
4.3.3 Estimation of the covariance parameters
4.3.4 Examples

4.4 Solvency uncertainty
4.4.1 Estimation at time $I + 1$
4.4.2 Solvency uncertainty
4.4.3 Uncertainties of further CDR’s

4.5 Examples

4.6 Estimation of correlation of reserving Risks
4.6.1 Avoiding correlation matrices for the reserving risks
4.6.2 Using LSRMs to estimate a correlation matrix

4.7 Literature
Linear-Stochastic-Reserving methods

4.1 How do Linear-Stochastic Reserving methods (LSRM) work
4.1.1 LSRM without stochastic
4.1.2 Stochastic behind LSRMs
4.2 Future development
4.2.1 Projection of the future development
4.2.2 Examples
4.3 Ultimate uncertainty
4.3.1 Mixing of claim properties
4.3.2 Ultimate uncertainty
4.3.3 Estimation of the covariance parameters
4.3.4 Examples
4.4 Solvency uncertainty
4.4.1 Estimate at time \( I + 1 \)
4.4.2 Solvency uncertainty
4.4.3 Uncertainties of further CDR’s
4.5 Examples
4.6 Estimation of correlation of reserving Risks
4.6.1 Avoiding correlation matrices for the reserving risks
4.6.2 Using LSRMs to estimate a correlation matrix
4.7 Literature
Mixing weights

In the last lecture we derived unbiased estimators for the future development of Linear-Stochastic-Reserving methods. Now we want to look at the corresponding ultimate uncertainty.

We have seen in Estimator 2.15 and Examples 4.13 that we are often interested in a linear combination of claim properties. Since claim reserves are expectations such mixing can be transferred to the corresponding Best Estimate reserves. But, because of diversification and dependencies, the mixing of claim properties has an influence on the estimated uncertainties. Therefore, we will look at the ultimate uncertainty of

\[
\sum_{m=0}^{M} \alpha_{i}^{m} \sum_{k=I-i}^{J-1} \hat{S}_{i,k+1}^{m} \quad \text{and} \quad \sum_{m=0}^{M} \sum_{i=0}^{I} \alpha_{i}^{m} \sum_{k=I-i}^{J-1} \hat{S}_{i,k+1}^{m},
\]

where \( \alpha_{i}^{m} \) are \( D^{I} \)-measurable real numbers.

That means we want to estimate

\[
\text{mse}_{D^{I}} \left[ \sum_{m=0}^{M} \alpha_{i}^{m} \sum_{k=I-i}^{J-1} \hat{S}_{i,k+1}^{m} \right] \quad \text{and} \quad \text{mse}_{D^{I}} \left[ \sum_{m=0}^{M} \sum_{i=0}^{I} \alpha_{i}^{m} \sum_{k=I-i}^{J-1} \hat{S}_{i,k+1}^{m} \right].
\]
In the last lecture we derived unbiased estimators for the future development of Linear-Stochastic-Reserving methods. Now we want to look at the corresponding ultimate uncertainty.

We have seen in Estimator 2.15 and Examples 4.13 that we are often interested in a linear combination of claim properties. Since claim reserves are expectations such mixing can be transferred to the corresponding Best Estimate reserves. But, because of diversification and dependencies, the mixing of claim properties has an influence on the estimated uncertainties. Therefore, we will look at the ultimate uncertainty of

\[\sum_{m=0}^{M} \sum_{i=0}^{J-1} \alpha_{m}^{i} \hat{S}_{m,i,k+1} \] and \[\sum_{m=0}^{M} \sum_{i=0}^{J-1} \alpha_{m}^{i} \sum_{k=0}^{I-1} \hat{S}_{m,i,k+1} \]

where \(\alpha_{m}^{i}\) are \(D^{n}\)-measurable real numbers.

That means we want to estimate

\[\text{mse}_{D^{n}} \left( \sum_{m=0}^{M} \sum_{i=0}^{J-1} \alpha_{m}^{i} \hat{S}_{m,i,k+1} \right) \] and \[\text{mse}_{D^{n}} \left( \sum_{m=0}^{M} \sum_{i=0}^{J-1} \alpha_{m}^{i} \sum_{k=0}^{I-1} \hat{S}_{m,i,k+1} \right)\]

Examples for mixing:

- Combination of two Chain-Ladder projections, one for payments and one for incurred losses.
- Adding dependent payments, for instance subrogation and normal payments, which are projected separately.
Decomposition of the ultimate uncertainty

\[ \text{mse}_{D^I} \left[ \sum_{m=0}^{M} \sum_{i=0}^{I} \alpha_i^m \sum_{k=I-i}^{J-1} \hat{S}_{i,k+1}^m \right] = \text{Var} \left[ \sum_{m=0}^{M} \sum_{i=0}^{I} \alpha_i^m \sum_{k=I-i}^{J-1} S_{i,k+1}^m \right]_{D^I} + \left( \sum_{m=0}^{M} \sum_{i=0}^{I} \alpha_i^m \sum_{k=I-i}^{J-1} \mathbb{E} \left[ S_{i,k+1}^m - \hat{S}_{i,k+1}^m \bigg| D^I \right] \right)^2 \]

random error

parameter error

**Remark 4.15**

The ultimate uncertainty of a single accident period or a single claim property can easily be obtained from the general formula by setting some of the \( \alpha_i^m \) to zero.
Decomposition of the ultimate uncertainty

\[
\text{mse} = \text{Var} \left[ \sum_{m=0}^{M} \sum_{i=0}^{I} \alpha_{m,i} \sum_{k=I}^{J-1} \left( S_{m,i,k}^{(n)} + 1 \right) \right] + \left( \sum_{m=0}^{M} \sum_{i=0}^{I} \sum_{k=I}^{J-1} \mathbb{E} \left[ \left( S_{m,i,k+1}^{(n)} - \hat{S}_{m,i,k+1}^{(n)} \right)^2 \right] \right)^2
\]

Remark 4.15
The ultimate uncertainty of a single accident period or a single claim property can easily be obtained from the general formula by setting some of the \( \alpha_{m,i} \) to zero.
Taylor approximation

Like in the Chain-Ladder case we will look at the functional

\[ U(g)x := \sum_{m=0}^{M} \sum_{i=0}^{I} \alpha_i^m \left( \sum_{k=0}^{I-i} x_{i,k}^m + \sum_{k=I-i}^{J-1} F_{i,k}^m(g)x \right). \]

Then we get:

\[ \partial_{h,j}^l U(\bar{g})x := \left. \frac{\partial U(g)x}{\partial g_{h,j}^l} \right|_{g=\bar{g}} = \frac{U(\bar{g})x - U(\bar{g}_{h,j}|0)x}{\bar{g}_{h,j}^l} = U(\bar{g}_{h,j}|1)x - U(\bar{g}_{h,j}|0)x, \]

where \( \bar{g}_{h,j}|a \) denotes the vector \( \bar{g} \) with exchanged coordinate \( \bar{g}_{h,j} = a \).

Moreover, we have

\[ U(f)S^I = \sum_{m=0}^{M} \sum_{i=0}^{I} \sum_{k=0}^{J} \alpha_i^m \sum_{k=0}^{J} \mathbb{E}[S_{i,k}^m|D^I], \]

\[ U(\hat{f})S^I = \sum_{m=0}^{M} \sum_{i=0}^{I} \sum_{k=0}^{J} \hat{S}_{i,k}^m, \]

\[ U(F)S^I = \sum_{m=0}^{M} \sum_{i=0}^{I} \sum_{k=0}^{J} S_{i,k}^m, \]

\[ \sum_{m=0}^{M} \sum_{i=0}^{I} \sum_{k=0}^{J} \alpha_i^m \sum_{k=0}^{J} \left( \hat{S}_{i,k}^m - S_{i,k}^m \right) \approx \sum_{l=0}^{M} \sum_{h=0}^{I} \sum_{j=I-h}^{J-1} \partial_{h,j}^l U(\hat{f})S^I \left( F_{h,j}^l - \hat{f}_{j}^l \right). \]

where we used a first order Taylor approximation and \( F \) and \( \hat{f} \) denote the vector of all link ratios \( F_{i,k}^m := S_{i,k+1}^m/R_{i,k}^m \) and the vector of all estimated development factors \( \hat{f}_{k}^m \), respectively.
Because of Lemma 4.9 \( U(g) \) is an affine operator in each coordinate \( g^{T_n}_{i,k} \) of \( g \). This implies the formula for its partial derivative.

Moreover, the representations of the expected, estimated and real ultimate are a direct consequence of the definitions of \( U \) and \( \mathcal{F} \).
Comparison with Chain-Ladder

Except for some additional summations (and the mixing parameters $\alpha^m_i$) we have the same form like in the Chain-Ladder case:

\[ \sum_{i=0}^{I} \sum_{k=0}^{J} \left( \hat{S}_{i,k}^m - S_{i,k}^m \right) = \sum_{i=0}^{I} \left( \hat{C}_{i,J} - C_{i,J} \right) \]

\[ \approx \sum_{h=0}^{I} \sum_{j=I-i}^{J-1} \frac{\hat{C}_{h,J}}{f_j} \left( F_{h,j} - \hat{f}_j \right) \cdot \partial_{h,j} U(\hat{f}) S^I \]

LSRM case:

\[ \sum_{m=0}^{M} \sum_{i=0}^{I} \alpha^m_i \sum_{k=0}^{J} \left( \hat{S}_{i,k}^m - S_{i,k}^m \right) \approx \sum_{l=0}^{M} \sum_{h=0}^{I} \sum_{j=I-h}^{J-1} \partial_{h,j} U(\hat{f}) S^I \left( F_{h,j}^l - \hat{f}_j^l \right) . \]
The partial derivative of the ultimate is a bit simpler in the Chain-Ladder case, because if we set some development factor $f_j$ to zero we get $U(\hat{f}_j|0) = 0$. 

<table>
<thead>
<tr>
<th>Comparison with Chain-Ladder</th>
</tr>
</thead>
<tbody>
<tr>
<td>Except for some additional summations (and the mixing parameters $\alpha_{mj}$) we have the same form like in the Chain-Ladder case:</td>
</tr>
<tr>
<td>$\sum_{i=0}^{I} \sum_{k=0}^{J} (\hat{S}<em>{mfi,k} - S</em>{mfi,k}) = \sum_{i=0}^{I} (\hat{C}<em>{i,J} - C</em>{i,J})$</td>
</tr>
<tr>
<td>$\approx \sum_{i=0}^{I} \sum_{j=i}^{I} \sum_{k=0}^{J} \alpha_{mj}(\hat{r}<em>{j}) \hat{r}</em>{j}^k \cdot (\hat{r}<em>{i} - \hat{r}</em>{j})$.</td>
</tr>
<tr>
<td>LSRM case:</td>
</tr>
<tr>
<td>$\sum_{i=0}^{I} \sum_{k=0}^{J} (\hat{S}<em>{mfi,k} - S</em>{mfi,k}) \approx \sum_{i=0}^{I} \sum_{j=i}^{I} \sum_{k=0}^{J} \alpha_{mj}(\hat{r}<em>{j}) \hat{r}</em>{j}^k \cdot (\hat{r}<em>{i} - \hat{r}</em>{j})$.</td>
</tr>
</tbody>
</table>
Preparation for the derivation of the ultimate uncertainty

Like in the Chain-Ladder case we need some expectations and covariances of \( \hat{f}_k \) and \( F_{i,k}^m \):

\[
E[F_{i,k}^m | D^I] = E[\hat{f}_k | D_k] = f_k^m \quad i + k \geq I
\]

\[
\text{Cov}[F_{i,k}^m, F_{i,k}^m | D^I] = E\left[\sigma_k^{m_1,m_2} R_{i,k}^{m_1,m_2} \left| D^I \right. \right] 
\approx \frac{\hat{\sigma}_k^{m_1,m_2} \hat{R}_{i,k}^{m_1,m_2}}{\hat{R}_{i,k}^{m_1} \hat{R}_{i,k}^{m_2}} \quad i + k \geq I
\]

\[
\text{Cov}[\hat{f}_k^m, \hat{f}_k^m | D_k] = \sigma_k^{m_1,m_2} \sum_{i=0}^{I-1-k} \omega_{i,k}^{m_1} \omega_{i,k}^{m_2} \frac{R_{i,k}^{m_1,m_2}}{R_{i,k}^{m_1} R_{i,k}^{m_2}}
\]

\[
\text{Cov}[F_{i_1,k_1}^{m_1}, F_{i_2,k_2}^{m_2} | D^I] = 0 \quad (i_1, k_1) \neq (i_2, k_2)
\]

\[
\text{Cov}[\hat{f}_{k_1}^{m_1}, \hat{f}_{k_2}^{m_2} | D_{k_1}] = 0 \quad k_1 < k_2
\]

\[
E\left[\left(F_{i_1,k_1}^{m_1} - \hat{f}_{k_1}^{m_1}\right)\left(F_{i_2,k_2}^{m_2} - \hat{f}_{k_2}^{m_2}\right) \right| D^I] 
= \text{Cov}[F_{i_1,k_1}^{m_1}, F_{i_2,k_2}^{m_2} | D^I] + (\hat{f}_{k_1}^{m_1} - f_{k_1}^{m_1}) (\hat{f}_{k_2}^{m_2} - f_{k_2}^{m_2})
\]
Stochastic Reserving

Linear-Stochastic-Reserving methods

Ultimate uncertainty

$$E\left[ F^m_{i,k} \mid D^I \right] = E\left[ E\left[ F^m_{i,k} \mid D^i_{k} \right] \mid D^I \right] = E\left[ f^m_k \mid D^I \right] = f^m_k = E\left[ \hat{f}^m_k \mid D_k \right]$$

because $\hat{f}^m_k$ is $D_k$-unbiased

$$\text{Cov}\left[ F^m_{i,k} \mid D^I \right] = \text{Cov}\left[ E\left[ F^m_{i,k} \mid D^i_{k} \right] , E\left[ F^m_{i,k} \mid D^i_{k} \right] \right] + \text{E}\left[ \text{Cov}\left[ F^m_{i,k} , F^m_{i,k} \mid D^i_{k} \right] \right] \mid D^I$$

$$= \text{Cov}\left[ f^m_k , f^m_k \mid D^I \right] + \text{E}\left[ \frac{\sigma^{m_1,m_2 \cdot R^m_{i,k}}}{R^m_{i,k} R^m_{i,k}} \mid D^I \right] = 0 + \text{E}\left[ \frac{\sigma^{m_1,m_2 \cdot R^m_{i,k}}}{R^m_{i,k} R^m_{i,k}} \mid D^I \right] \approx \frac{\sigma^{m_1,m_2 \cdot \hat{R}^m_{i,k}}}{\hat{R}^m_{i,k} \hat{R}^m_{i,k}}$$

$$\text{Cov}\left[ \hat{f}^m_{k_1} , \hat{f}^m_{k_2} \mid D_{k_1} \right] = \text{Cov}\left[ E\left[ \hat{f}^m_{k_1} \mid D_{k_2} \right] , E\left[ \hat{f}^m_{k_2} \mid D_{k_2} \right] \mid D_{k_1} \right] + \text{E}\left[ \text{Cov}\left[ \hat{f}^m_{k_1} , \hat{f}^m_{k_2} \mid D_{k_2} \right] \right] \mid D_{k_1} = 0$$

$$\text{E}\left[ \left( F^m_{i,k_1} - \hat{f}^m_{k_1} \right) \left( F^m_{i,k_2} - \hat{f}^m_{k_2} \right) \mid D^I \right]$$

$$= \text{E}\left[ \left( F^m_{i,k_1} - f^m_k \right) \left( F^m_{i,k_2} - f^m_k \right) \left( F^m_{i,k_2} - f^m_k \right) \right] \mid D^I$$

$$= \text{E}\left[ \left( F^m_{i,k_1} - f^m_k \right) \left( f^m_k \right) \left( f^m_k \right) \right] \mid D^I - \text{E}\left[ \left( F^m_{i,k_1} - f^m_k \right) \left( f^m_k \right) \left( f^m_k \right) \right] \mid D^I$$

$$= \text{Cov}\left[ F^m_{i,k_1} , F^m_{i,k_2} \mid D^I \right] - 0 + \left( f^m_k - f^m_k \right) \left( f^m_k - f^m_k \right)$$

If $i_1 + k_1 < I$ or $i_2 + k_2 < I$ then $F^m_{i_1,k_1} \in D^I$ or $F^m_{i_2,k_2} \in D^I$ and we are done. Otherwise, since $(i_1, k_1) \neq (i_2, k_2)$, either $F^m_{i_1,k_1} \in D^i_{k_2} + k_2$ or $F^m_{i_2,k_2} \in D^i_{k_1} + k_1$. Let's assume the first:

$$\text{Cov}\left[ F^m_{i,k_1} , F^m_{i,k_2} \mid D^I \right] = \text{E}\left[ \text{Cov}\left[ F^m_{i,k_1} , F^m_{i,k_2} \mid D^i_{k_2} \right] \mid D^I \right] + \text{E}\left[ \text{E}\left[ F^m_{i,k_1} \mid D^i_{k_2} \right] , \text{E}\left[ F^m_{i,k_2} \mid D^i_{k_2} \right] \right] \mid D^I$$

$$= 0 + \text{Cov}\left[ F^m_{i,k_1} , f^m_{k_2} \mid D^I \right] = 0$$
Estimator 4.16 (Linear approximation of the ultimate uncertainty)

\[
\text{mse}_{U(f)S^I|D^I} \left[ U(f) S^I \right] = E \left[ \left( \sum_{m=0}^{M} \sum_{i=0}^{I} \sum_{k=0}^{J} \alpha_i^m \left( \hat{S}_{i,k}^m - S_{i,k}^m \right) \right)^2 \right] \quad \text{(Taylor approximation)}
\]

\[
\approx E \left[ \left( \sum_{l=0}^{M} \sum_{h=0}^{I} \sum_{k=I-i}^{J-1} \partial_{h,j}^l U(\hat{f}) S^I \left( F_{h,j}^l - \hat{f}_j^l \right) \right)^2 \right]
\]

\[
\approx \sum_{l_1,l_2=0}^{M} \sum_{j_1,j_2=0}^{J-1} \sum_{h_1,I-j}^{I} \sum_{h_2,I-j}^{I} \partial_{h_1,j_1}^{l_1} U(\hat{f}) S^I \partial_{h_2,j_2}^{l_2} U(\hat{f}) S^I 
\]

\[
\left( \text{Cov} \left[ F_{h_1,j_1}^{l_1}, F_{h_2,j_2}^{l_2} \right] | D^I \right) + \text{Cov} \left[ \hat{f}_{j_1}^{l_1}, \hat{f}_{j_2}^{l_2} \right] | D_{j_1 \wedge j_2} \right)
\]

\[
\approx \sum_{l_1,l_2=0}^{M} \sum_{j=0}^{J-1} \sum_{h=I-j}^{I} \partial_{h,j}^{l_1} U(\hat{f}) S^I \partial_{h,j}^{l_2} U(\hat{f}) S^I \hat{\sigma}_{j}^{l_1,l_2} \frac{\hat{R}_{h,j}^{l_1,l_2}}{\hat{R}_{h,j}^{l_1} \hat{R}_{h,j}^{l_2}}
\]

\[
\text{random error} \quad \text{parameter error}
\]

\[
+ \sum_{l_1,l_2=0}^{M} \sum_{j=0}^{J-1} \sum_{h_1,h_2=I-j}^{I} \partial_{h_1,j}^{l_1} U(\hat{f}) S^I \partial_{h_2,j}^{l_2} U(\hat{f}) S^I \hat{\sigma}_{j}^{l_1,l_2} \sum_{h=0}^{I-j-1} w_{h,j}^{l_1} w_{h_2,j}^{l_2} \frac{\hat{R}_{h,j}^{l_1,l_2}}{\hat{R}_{h,j}^{l_1} \hat{R}_{h,j}^{l_2}}
\]

\[
\text{random error} \quad \text{parameter error}
\]
In the second approximation we used

\[
E \left[ (F_{h_1,j_1}^{l_1} - \hat{f}_{j_1}^{l_1}) (F_{h_2,j_2}^{l_2} - \hat{f}_{j_2}^{l_2}) \right| D^I] = \text{Cov} \left[ F_{h_1,j_1}^{l_1}, F_{h_2,j_2}^{l_2} \right| D^I] + (\hat{f}_{j_1}^{l_1} - f_{j_1}^{l_1})(\hat{f}_{j_2}^{l_2} - f_{j_2}^{l_2})
\]

\[
\approx \text{Cov} \left[ F_{h_1,j_1}^{l_1}, F_{h_2,j_2}^{l_2} \right| D^I] + \text{Cov} \left[ \hat{f}_{j_1}^{l_1}, \hat{f}_{j_2}^{l_2} \right| D_{j_1 \wedge j_2}
\]

and from the preparations above we know

\[
\text{Cov} \left[ F_{h_1,j_1}^{l_1}, F_{h_2,j_2}^{l_2} \right| D^I] \approx 1_{j_1 = j_2} 1_{h_1 = h_2} \hat{\sigma}_{j_1}^{l_1,l_2} \frac{\hat{R}_{h_1,j_1}^{l_1,l_2}}{\hat{R}_{h_1,j_1}^{l_1} \hat{R}_{h_2,j_2}^{l_2}}
\]

\[
\text{Cov} \left[ \hat{f}_{j_1}^{l_1}, \hat{f}_{j_2}^{l_2} \right| D_{j_1 \wedge j_2}] \approx 1_{j_1 = j_2} \hat{\sigma}_{j_1}^{l_1,l_2} \sum_{h=0}^{I-j_1-1} w_{h,j_1}^{l_1} w_{h,j_2}^{l_2} \frac{\hat{R}_{h,j_1}^{l_1,l_2}}{\hat{R}_{h,j_1}^{l_1} \hat{R}_{h,j_2}^{l_2}}.
\]

This leads directly to the stated estimator.
Chain-Ladder estimator for the ultimate uncertainty

$$\text{mse}_{\sum_{i=0}^{I} C_{i,J} | D^I} \left[ \hat{C}_{i,J} \right] \approx \sum_{k=0}^{I} \frac{\hat{\sigma}_k^2}{f_k^2} \sum_{i=I-k}^{I} \hat{C}_{i,J}^2 \frac{1}{\hat{C}_{i,k}} + \sum_{k=0}^{I-1} \frac{\hat{\sigma}_k^2}{f_k^2} \left( \sum_{i=I-k}^{I} \hat{C}_{i,J} \right)^2 \sum_{h=0}^{I-k-1} \frac{w_{h,k}^2}{C_{h,k}}$$

LSRM estimator for the ultimate uncertainty

$$\text{mse}_{U(f)S^I | D^I} \left[ U(\hat{f}) S^I \right]$$

$$\approx \sum_{l_1,l_2=0}^{M} \sum_{j=0}^{J-1} \sum_{h=I-j}^{I} \partial_{h,j}^{l_1} U(\hat{f}) S^I \partial_{h,j}^{l_2} U(\hat{f}) S^I \hat{\sigma}_j^{l_1,l_2} \frac{\hat{R}_{h,j}^{l_1,l_2}}{\hat{R}_{h,j}^{l_1} \hat{R}_{h,j}^{l_2}} + \sum_{l_1,l_2=0}^{M} \sum_{j=0}^{J-1} \sum_{h_1,h_2=I-j}^{I} \partial_{h_1,j}^{l_1} U(\hat{f}) S^I \partial_{h_2,j}^{l_2} U(\hat{f}) S^I \hat{\sigma}_j^{l_1,l_2} \sum_{h=0}^{I-j-1} \frac{w_{h_1,j} w_{h_2,j}^2}{\hat{R}_{h,j}^{l_1} \hat{R}_{h,j}^{l_2}} \frac{\hat{R}_{h,j}^{l_1,l_2}}{\hat{R}_{h,j}^{l_1} \hat{R}_{h,j}^{l_2}}$$
Because we have several claim properties, squared terms for Chain-Ladder are replaced by products of claim properties and the corresponding double sum.
Change of the variance exposures in Chain-Ladder

The Chain-Ladder method assumes variances to be proportional to the cumulative payments, i.e.

\[ \text{Var} \left[ C_{i,k+1} \mid D_{k}^{i+k} \right] = \sigma_k^2 C_{i,k}, \]

which leads to vanishing coefficient of variation of (ultimate) uncertainties with increasing volume, see Corollary 2.10. This is one of many arguments against Chain-Ladder.

One way to solve this is to change the variance exposure, for instance to \( C_{i,k}^2 \). Then we get

\[
\left( \text{VaC} \left( \sum_{i=0}^{I} C_{i,J} \right) \right)^2 \approx \frac{\sum_{i=0}^{I} \hat{C}_{i,J}^2}{\left( \sum_{i=0}^{I} \hat{C}_{i,J} \right)^2} \sum_{k=I-i}^{J-1} \frac{\hat{\sigma}_k^2}{\hat{f}_k^2} \\
+ \sum_{k=0}^{J-1} \frac{\hat{\sigma}_k^2}{\hat{f}_k^2} \left( \sum_{i=I-k}^{I} \hat{C}_{i,J} \right)^2 \sum_{h=0}^{I-k-1} w_{i,k}^2 \left( \sum_{i=0}^{I} \hat{C}_{i,J} \right)^2
\]

which does not decrease with increasing volume. Nevertheless, you should always add some model error.
In practice, the choice of the variance exposure does not matter so much, because the estimation of the variance parameters $\sigma_k^2$ will change, too, which compensates some effects.
Estimator 4.17 (of covariance parameter $\sigma_{k,1,2}^{m_1,m_2}$)

If the normalizing constant

$$Z_{k}^{m_1,m_2} := \sum_{i=0}^{I-1-k} \frac{w_{i,k}^m w_{i,k}^m}{R_{i,k}^m R_{i,k}^m} \left( 1 - w_{i,k}^m - w_{i,k}^m + R_{i,k}^m \sum_{h=0}^{I-1-k} \frac{w_{h,k}^m w_{h,k}^m}{R_{h,k}^m R_{h,k}^m} \right) > 0$$

then the covariance parameter $\sigma_{k,1,2}^{m_1,m_2}$ can be estimated by the following $D_k$-unbiased estimator

$$\hat{\sigma}_{k}^{m_1,m_2} := \frac{1}{Z_{k}^{m_1,m_2}} \sum_{i=0}^{I-1-k} \frac{w_{i,k}^m w_{i,k}^m}{R_{i,k}^m R_{i,k}^m} \left( \frac{S_{i,k+1}^m}{R_{i,k}^m} - \hat{f}_{k}^m \right) \left( \frac{S_{i,k+1}^m}{R_{i,k}^m} - \hat{f}_{k}^m \right)$$

For $Z_{k}^{m_1,m_2} = 0$ and in particular for $k = I - 1$ one could take the following extrapolations,

$$\hat{\sigma}_{k}^{m,m} := \min \left( \frac{(\hat{\sigma}_{k-1}^{m,m})^2}{\hat{\sigma}_{k-1}^{m,m}}, \hat{\sigma}_{k-2}^{m,m}, \hat{\sigma}_{k-1}^{m,m} \right),$$

$$\hat{\sigma}_{k}^{m_1,m_2} := \hat{\sigma}_{k-1}^{m_1,m_2} \left( \frac{\hat{\sigma}_{k}^{m_1,m_1} \hat{\sigma}_{k}^{m_2,m_2}}{\hat{\sigma}_{k-1}^{m_1,m_1} \hat{\sigma}_{k-1}^{m_2,m_2}} \right)^{1/2}, \text{ for } m_1 \neq m_2.$$
Stochastic Reserving

Linear-Stochastic-Reserving methods

Ultimate uncertainty

\[
E \left[ \left( \frac{S_{i,k+1}^{m_1}}{R_{i,k}^{m_1}} - \hat{f}^{m_1}_k \right) \left( \frac{S_{i,k+1}^{m_2}}{R_{i,k}^{m_2}} - \hat{f}^{m_2}_k \right) | D_k \right] = \text{Cov} \left[ \left( \frac{S_{i,k+1}^{m_1}}{R_{i,k}^{m_1}} - \hat{f}^{m_1}_k \right), \left( \frac{S_{i,k+1}^{m_2}}{R_{i,k}^{m_2}} - \hat{f}^{m_2}_k \right) | D_k \right]
\]

\[
= \text{Cov} \left[ \frac{S_{i,k+1}^{m_1}}{R_{i,k}^{m_1}}, \frac{S_{i,k+1}^{m_2}}{R_{i,k}^{m_2}} | D_k \right] - \sum_{h=0}^{I-k-1} \text{Cov} \left[ \frac{S_{i,k+1}^{m_1}}{R_{i,k}^{m_1}}, \frac{S_{h,k+1}^{m_2}}{R_{h,k}^{m_2}} | D_k \right] - \sum_{h=0}^{I-k-1} \text{Cov} \left[ \frac{S_{i,k+1}^{m_1}}{R_{i,k}^{m_1}}, w_{h,k}^{m_2} \frac{S_{h,k+1}^{m_2}}{R_{h,k}^{m_2}} | D_k \right] + \sum_{h_1,h_2=0}^{I-k-1} \text{Cov} \left[ w_{h_1,k}^{m_1} \frac{S_{h_1,k+1}^{m_1}}{R_{h_1,k}^{m_1}}, w_{h_2,k}^{m_2} \frac{S_{h_2,k+1}^{m_2}}{R_{h_2,k}^{m_2}} | D_k \right]
\]

\[
= \sigma_{k}^{m_1,m_2} \frac{R_{i,k}^{m_1,m_2}}{R_{i,k}^{m_1} R_{i,k}^{m_2}} \left( 1 - w_{i,k}^{m_1} - w_{i,k}^{m_2} + \frac{R_{i,k}^{m_1} R_{i,k}^{m_2}}{R_{i,k}^{m_1,m_2}} \sum_{h=0}^{I-k-1} \frac{w_{h,k}^{m_1} w_{h,k}^{m_2} R_{h,k}^{m_1,m_2}}{R_{h,k}^{m_1} R_{h,k}^{m_2}} \right)
\]

\[
\Rightarrow E \left[ \sigma_k^{m_1,m_2} | D_k \right] = \frac{\sigma_k^{m_1,m_2}}{Z_k^{m_1,m_2}} \sum_{i=0}^{I-1-k} \frac{w_{i,k}^{m_1} w_{i,k}^{m_2}}{R_{i,k}^{m_1} R_{i,k}^{m_2}} \left( 1 - w_{i,k}^{m_1} - w_{i,k}^{m_2} + \frac{R_{i,k}^{m_1} R_{i,k}^{m_2}}{R_{i,k}^{m_1,m_2}} \sum_{h=0}^{I-k-1} \frac{w_{h,k}^{m_1} w_{h,k}^{m_2} R_{h,k}^{m_1,m_2}}{R_{h,k}^{m_1} R_{h,k}^{m_2}} \right)
\]

\[
= \sigma_k^{m_1,m_2}
\]
Remark 4.18 (estimation of the covariance parameter $\sigma_{k}^{m_1,m_2}$)

- Even if the real covariance parameter $\sigma_{k}^{m_1,m_2}$ lead to positive semi-defined covariance matrices

\[
\begin{pmatrix}
\sigma_{k}^{m_1,m_2} & R_{i,k}^{m_1,m_2}
\end{pmatrix}_{0 \leq m_1,m_2 \leq M}
\]

the estimated values may not. In particular this may be the case if one eigenvalue of the real covariance matrix is (almost) equal to zero. Therefore, we always have to check the positive semi-definiteness of the estimated covariance matrices.

- The first part of the extrapolation goes back to Mack [21]. Roughly spoken it assumes that the variance parameter decay exponentially for later development periods.

- Depending on the data we may get better estimators if we introduce weights or use other extrapolations.
Remark 4.18 (estimation of the covariance parameter $\sigma_{m_1,m_2,k}$)

- Even if the real covariance parameter $\sigma_{m_1,m_2,k}$ lead to positive semi-defined covariance matrices

$$
\left(\sigma_{m_1,m_2,k}^{i,j}\right)_{i,j=1,\ldots,M}
$$

the estimated values may not. In particular this may be the case if one eigenvalue of the real covariance matrix is (almost) equal to zero. Therefore, we always have to check the positive semi-definiteness of the estimated covariance matrices.

- The first part of the extrapolation goes back to Mack [21]. Roughly spoken it assumes that the variance parameter decay exponentially for later development periods.

- Depending on the data we may get better estimators if we introduce weights or use other extrapolations.

Stochastic Reserving

Linear-Stochastic-Reserving methods

Ultimate uncertainty
Swiss mandatory accident portfolio: part 2 of 3, see Example 4.13

We have four claim properties with exposures

**ME:** \( R_{i,k}^0 = R_{i,k}^{0,0} = \sum_{j=0}^{k} S_{i,j}^0 \)

**IW:** \( R_{i,k}^1 = R_{i,k}^{1,1} = \sum_{l=0}^{k} \left( \kappa^{l+1} S_{i,l}^3 + (1 - \kappa^{l+1}) S_{i,l}^1 \right) \)

**Sub:** \( R_{i,k}^2 = R_{i,k}^{2,2} = \sum_{l=0}^{k} \left( S_{i,l}^0 + S_{i,l}^1 + S_{i,l}^2 \right) \)

**Salary:** \( S_{i,0}^3 = P_i, \quad S_{i,l}^3 = 0, \text{ for } l > 0, \text{ and} \)
\[
R_{i,k}^3 = R_{i,k}^{3,0} = R_{i,k}^{0,3} = R_{i,k}^{3,1} = R_{i,k}^{1,3} = R_{i,k}^{3,2} = R_{i,k}^{2,3} = R_{i,k}^{3,3} = 0
\]

For the not yet defined exposures we take the total payments up to now, i.e.
\[
R_{i,k}^{0,1} = R_{i,k}^{1,0} = R_{i,k}^{0,2} = R_{i,k}^{2,0} = R_{i,k}^{1,2} = R_{i,k}^{2,1} = \sum_{l=0}^{k} \left( S_{i,l}^0 + S_{i,l}^1 + S_{i,l}^2 \right).
\]

**Resulting ultimate uncertainty**

- The estimated ultimate uncertainty varies much less than the Best Estimate reserves (5% vs. 11%).
- Although the estimated ultimate uncertainty is minimal for \( \kappa \approx 0.3 \) you should never use this as criteria to chose the reserving method. For this portfolio, I would go for \( \kappa = 1 \) (at least for the first development periods).
- For \( \kappa = 0 \) the ultimate uncertainty is slightly smaller than CLM on total payments.
We have four claim properties with exposures

ME:

\[ R^0_{i,k} = \sum_{j=0}^{k} S^0_{i,j} \]

IW:

\[ R^1_{i,k} = \sum_{l=0}^{k} (\kappa_l + 1) S^3_{i,l} + (1 - \kappa_l + 1) S^1_{i,l} \]

Sub:

\[ R^2_{i,k} = \sum_{l=0}^{k} (S^0_{i,l} + S^1_{i,l} + S^2_{i,l}) \]

Salary:

\[ S^3_{i,0} = P_i, S^3_{i,l} = 0, \text{ for } l > 0, \text{ and } R^3_{i,k} = R^0_{i,k} = R^1_{i,k} = R^2_{i,k} = 0 \]

For the not yet defined exposures we take the total payments up to now, i.e.

\[ R^0_{i,k} = R^1_{i,k} = R^2_{i,k} = R^3_{i,k} = 0 \]

Resulting ultimate uncertainty

- The estimated ultimate uncertainty varies much less then the Best Estimate reserves (5% vs. 11%).
- Although the estimated ultimate uncertainty is minimal for \( \kappa \approx 0.3 \) you should never use this as criteria to chose the reserving method. For this portfolio, I would go for \( \kappa = 1 \) (at least for the first development periods).
- For \( \kappa = 0 \) the ultimate uncertainty is slightly smaller than CLM on total payments.
Example 4.13: Ultimate uncertainty in dependence of $\kappa$

We always show the square root of uncertainties.
Example 4.13: Ultimate uncertainty in dependence of $\kappa$

Even if it looks tempting you must not use the estimates of the ultimate uncertainty to evaluate which model is the best!
Example 4.19 (ECLRM vs. CLM: part 2 of 3, see Example 4.14)

In the first part we have compared the Best Estimate reserves. Now we want to look at the ultimate uncertainty.

For the weighing of uncertainties we have to define $R_{i,k}^{0,1} = R_{i,k}^{1,0}$. We use the arithmetic mean of payments and incurred losses for CLM and the case reserves for the ECLRM:

Square root of the ultimate uncertainty

<table>
<thead>
<tr>
<th>AP</th>
<th>CLM payments</th>
<th>CLM incurred</th>
<th>CLM weighting</th>
<th>ECLRM payments</th>
<th>ECLRM incurred</th>
<th>ECLRM weighting</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>1</td>
<td>89 423</td>
<td>2 553</td>
<td>43 873</td>
<td>194</td>
<td>14 639</td>
<td>7 695</td>
</tr>
<tr>
<td>2</td>
<td>234 666</td>
<td>5 186</td>
<td>109 257</td>
<td>4 557</td>
<td>5 538</td>
<td>4 825</td>
</tr>
<tr>
<td>3</td>
<td>255 612</td>
<td>9 264</td>
<td>114 052</td>
<td>10 569</td>
<td>12 619</td>
<td>11 170</td>
</tr>
<tr>
<td>4</td>
<td>261 298</td>
<td>10 874</td>
<td>112 326</td>
<td>36 825</td>
<td>38 319</td>
<td>37 335</td>
</tr>
<tr>
<td>5</td>
<td>323 899</td>
<td>33 243</td>
<td>128 299</td>
<td>43 971</td>
<td>44 889</td>
<td>44 284</td>
</tr>
<tr>
<td>6</td>
<td>274 942</td>
<td>55 884</td>
<td>100 542</td>
<td>65 091</td>
<td>65 971</td>
<td>65 477</td>
</tr>
<tr>
<td>7</td>
<td>373 634</td>
<td>165 086</td>
<td>187 882</td>
<td>176 720</td>
<td>176 999</td>
<td>176 835</td>
</tr>
<tr>
<td>8</td>
<td>492 894</td>
<td>209 163</td>
<td>249 463</td>
<td>197 790</td>
<td>197 930</td>
<td>197 841</td>
</tr>
<tr>
<td>9</td>
<td>468 137</td>
<td>321 566</td>
<td>292 510</td>
<td>322 922</td>
<td>323 076</td>
<td>323 017</td>
</tr>
<tr>
<td>Total</td>
<td>1 517 861</td>
<td>455 802</td>
<td>676 047</td>
<td>467 964</td>
<td>472 131</td>
<td>469 518</td>
</tr>
</tbody>
</table>
Example 4.19 (ECLRM vs. CLM: part 2 of 3, see Example 4.14)

In the first part we have compared the Best Estimate reserves. Now we want to look at the ultimate uncertainty.

For the weighing of uncertainties we have to define $R_{0i,k}, 1i,k, 0i,k = R_{1i,k}, 0i,k$. We use the arithmetic mean of payments and incurred losses for CLM and the case reserves for the ECLRM:

\[ \text{Square root of the ultimate uncertainty} \]

<table>
<thead>
<tr>
<th></th>
<th>CLM</th>
<th>ECLRM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>88.68</td>
<td>23.63</td>
</tr>
<tr>
<td>2</td>
<td>294.65</td>
<td>52.16</td>
</tr>
<tr>
<td>3</td>
<td>254.52</td>
<td>154.08</td>
</tr>
<tr>
<td>4</td>
<td>212.33</td>
<td>104.23</td>
</tr>
<tr>
<td>5</td>
<td>343.89</td>
<td>37.23</td>
</tr>
<tr>
<td>6</td>
<td>234.66</td>
<td>51.16</td>
</tr>
<tr>
<td>7</td>
<td>255.94</td>
<td>97.23</td>
</tr>
<tr>
<td>8</td>
<td>261.29</td>
<td>108.74</td>
</tr>
<tr>
<td>9</td>
<td>266.35</td>
<td>114.08</td>
</tr>
</tbody>
</table>

Total: 1,517,861

T aking the arithmetic mean

\[
R_{i,k}^{m1,m2} := \frac{1}{2} \left( R_{i,k}^{m1,m1} + R_{i,k}^{m2,m2} \right)
\]

for the coupling exposures works fine if $R_{i,k}^{m1,m1}$ and $R_{i,k}^{m2,m2}$ are similar. In general the geometric mean

\[
R_{i,k}^{m1,m2} := \sqrt{R_{i,k}^{m1,m1} R_{i,k}^{m2,m2}}
\]

usually works better.

- Although the Best Estimate reserves are similar, the ultimate uncertainties are not, in particular CLM on payments leads to a much higher ultimate uncertainty than the others.
- Again, you must not use estimates of the ultimate uncertainty to evaluate which model is the best.
Stochastic Reserving
Lecture 8 (Continuation of Lecture 4)
Linear-Stochastic-Reserving methods

René Dahms

ETH Zurich, Spring 2019

10 April 2019
(Last update: 18 February 2019)
Stochastic Reserving

- Linear-Stochastic-Reserving methods
- Ultimate uncertainty
4 Linear-Stochastic-Reserving methods

4.1 How do Linear-Stochastic-Reserving methods (LSRM) work

4.1.1 LSRM without stochastic
4.1.2 Stochastic behind LSRMs

4.2 Future development
4.2.1 Projection of the future development
4.2.2 Examples

4.3 Ultimate uncertainty
4.3.1 Mixing of claim properties
4.3.2 Ultimate uncertainty
4.3.3 Estimation of the covariance parameters
4.3.4 Examples

4.4 Solvency uncertainty
4.4.1 Estimation at time $I + 1$
4.4.2 Solvency uncertainty
4.4.3 Uncertainties of further CDR’s

4.5 Examples

4.6 Estimation of correlation of reserving Risks
4.6.1 Avoiding correlation matrices for the reserving risks
4.6.2 Using LSRMs to estimate a correlation matrix

4.7 Literature
4 Linear-Stochastic-Reserving methods

4.1 How do Linear-Stochastic-Reserving methods (LSRM) work
4.1.1 LSRM without stochastic
4.1.2 Stochastic behind LSRMs
4.2 Future development
4.2.1 Projection of the future development
4.2.2 Examples

4.3 Ultimate uncertainty
4.3.1 Mixing of claim properties
4.3.2 Ultimate uncertainty
4.3.3 Estimation of the covariance parameters
4.3.4 Examples

4.4 Solvency uncertainty
4.4.1 Estimation at time \( t+1 \)
4.4.2 Solvency uncertainty
4.4.3 Uncertainties of further CDR’s

4.5 Examples

4.6 Estimation of correlation of reserving Risks
4.6.1 Avoiding correlation matrices for the reserving risks
4.6.2 Using LSRMs to estimate a correlation matrix

4.7 Literature
Consistent estimation over time

In this section we want to look at the solvency uncertainty, i.e. the uncertainty of the claims development result

$$CDR_{i+1}^I := \sum_{m=0}^{M} \alpha_i^m \sum_{k=I-i}^{J-1} \left( \hat{S}_{i,k+1}^m - \hat{S}_{i,k+1}^{m+1} \right) \quad \text{and} \quad CDR_1^I := \sum_{i=0}^{I} CDR_{i+1}^I,$$

where the additional upper index represents the time of estimation and $\alpha_i^m$ are $D_I$-measurable real numbers.

In order to do so the estimates have to be consistent, that means we do not change our (relative) believes into the old development periods and only put some credibility $w_{I-k,k}^{m,I+1}$ to the at time $I + 1$ newly encountered development:

Assumption 4.B

*There exist $D_I \cap D_k$-measurable weights $0 \leq w_{I-k,k}^{m,I+1} \leq 1$ with*

- $R_{I-k,k}^m = 0$ implies $w_{I-k,k}^{m,I+1} = 0$,
- $w_{i,k}^{m,I+1} = (1 - w_{I-k,k}^{m,I+1}) w_{i,k}^m$ for $0 \leq i \leq I - 1 - k$. 
In this section we want to look at the solvency uncertainty, i.e. the uncertainty of the claims development result (CDR).

\[ CDR_{i+1}^I = \sum_{m=0}^M \alpha_m I_{j-1} \sum_{k=I-i}^I (\hat{S}_{m,I,i,k} - \hat{S}_{m,I,i,k+1}) \]

and

\[ CDR_{i+1}^{I+1} = \sum_{i=0}^I CDR_{i+1}^I, \]

where the additional upper index represents the time of estimation and \( \alpha_m \) are \( D_I \)-measurable real numbers.

In order to do so the estimates have to be consistent, that means we do not change our (relative) believes into the old development periods and only put some credibility \( w_m,I^{i+1} \) to the at time \( I+1 \) newly encountered development.

Assumption 4.B

There exist \( D_I \cap D_k \)-measurable weights \( 0 \leq w_m,I^{i+1} \leq 1 \) with

- \( w_{0,I}^{I-k} = 0 \) implies \( w_{I-k}^{I-k} = 0 \),
- \( w_{i,I}^{I-k} = (1 - w_{I-k}^{I-k})w_{i,k}^{I-k} \) for \( 0 \leq i \leq I - 1 - k \).

- We do not allow an estimation time dependence of the mixing weights.
- The variance minimizing weights, defined in Estimator 4.7, fulfil Assumption 4.B.
Lemma 4.20 (Estimation of development factors at time $I + 1$)

Let Assumptions 4.A and 4.B be fulfilled. Then the at time $I + 1$ estimated development factors

$$
\hat{f}_{k}^{m,I+1} := \sum_{i=0}^{I-k} w_{i,k} \frac{S_{i,k+1}^{m}}{R_{i,k}^{m}} = (1 - w_{I-k,k}^{m,I+1}) \hat{f}_{k}^{m,I} + w_{I-k,k}^{m,I+1} \frac{S_{I-k,k+1}^{m}}{R_{I-k,k}^{m}}
$$

satisfy:

1. $E\left[\hat{f}_{k}^{m,I+1} \mid \mathcal{D}^{I}\right] = E\left[\hat{f}_{k}^{m,I+1} \mid \mathcal{D}_{k}^{I}\right] = (1 - w_{I-k,k}^{m,I+1}) \hat{f}_{k}^{m,I} + w_{I-k,k}^{m,I+1} \hat{f}_{k}^{m} =: \bar{f}_{k}^{m}$

2. For every tuple $\hat{f}_{k_{1}}^{m_{1},I+1}, \ldots, \hat{f}_{k_{r}}^{m_{r},I+1}$ with $k_{1} < k_{2} < \cdots < k_{r}$ we get

$$
E\left[\hat{f}_{k_{1}}^{m_{1},I+1} \ldots \hat{f}_{k_{r}}^{m_{r},I+1} \mid \mathcal{D}^{I}\right] = E\left[\hat{f}_{k_{1}}^{m_{1},I+1} \ldots \hat{f}_{k_{r}}^{m_{r},I+1} \mid \mathcal{D}_{k_{1}}^{I}\right] = \bar{f}_{k_{1}}^{m_{1}} \ldots \bar{f}_{k_{r}}^{m_{r}},
$$

which implies that the estimators are pairwise $\mathcal{D}^{I}$-conditionally uncorrelated.

Remark 4.21

Because of part 1. of Lemma 4.20, we will use the estimates $\hat{f}_{k}^{m} := \hat{f}_{k}^{m,I}$. 
Lemma 4.20 (Estimation of development factors at time $I+1$)

Let Assumptions 4.A and 4.B be fulfilled. Then the at time $I+1$ estimated development factors

$$\hat{f}_{m,I}^{k}:=(1-w_{m,I+1}^{I-k,k})\hat{f}_{m,I}^{k}+w_{m,I+1}^{i,I-k}S_{m,I+1}^{I-k,k}R_{m,I+1}^{I-k,k}$$

satisfy:

1. $E\left[\hat{f}_{m,I+1}^{k}\mid D_I\right]=E\left[\hat{f}_{m,I}^{k}\mid D_I^{I-k,k}\right]=\left(1-w_{m,I+1}^{I-k,k}\right)\hat{f}_{m,I}^{k}+w_{m,I+1}^{i,I-k}S_{m,I+1}^{I-k,k}R_{m,I+1}^{I-k,k}$

2. For every tuple $\hat{f}_{m,1,I+1}^{k_1}, ..., \hat{f}_{m,r,I+1}^{k_r}$ with $k_1 < k_2 < \ldots < k_r$ we get

$$E\left[\hat{f}_{m,1,I+1}^{k_1} \ldots \hat{f}_{m,r,I+1}^{k_r} \mid D_I\right]=E\left[\hat{f}_{m,1,I+1}^{k_1} \ldots \hat{f}_{m,r-1,I+1}^{k_r} \mid D_{I,k_r}\right]D_I$$

$$=E\left[\hat{f}_{m,1,I+1}^{k_1} \ldots \hat{f}_{m,r-1,I+1}^{k_r} \hat{f}_{m,r,I+1}^{k_r} \mid D_I\right]$$

$$=\ldots = \hat{f}_{m,1,I+1}^{k_1} \ldots \hat{f}_{m,r,I+1}^{k_r}$$

and similar for $D_{k_1}$ instead of $D_I$. 

Remark 4.21
Because of part 1. of Lemma 4.20, we will use the estimates $\bar{f}_{m,I+1}^{k}:=(1-w_{m,I+1}^{I-k,k})\hat{f}_{m,I}^{k}+w_{m,I+1}^{i,I-k}S_{m,I+1}^{I-k,k}R_{m,I+1}^{I-k,k}$, which implies that the estimators are pairwise $D_I$-conditionally uncorrelated.
Lemma 4.22 (Best Estimate reserves)

Let Assumptions 4.A and 4.B be fulfilled. Then the at time $I + 1$ estimated claim properties satisfy

$$\bar{S}_{i,k}^{m} := E\left[ \hat{S}_{i,k+1}^{m,I+1} \left| \mathcal{D}I \right. \right] = F_{i,k}^{m,I+1}(\bar{f}) \mathcal{F}^{I}(f) S^{I}.$$ 

Hence, we will use the estimates

$$\hat{E}\left[ \hat{S}_{i,k+1}^{m,I+1} \left| \mathcal{D}I \right. \right] = \hat{S}_{i,k+1}^{m} := F_{i,k}^{m,I+1}(\hat{f}) \mathcal{F}^{I}(\hat{f}) S^{I} = F_{i,k}^{m,I}(\hat{f}) S^{I},$$

which implies $\hat{E}[\text{CDR}^{I+1}|\mathcal{D}I] := 0$. That means, we have Best Estimate reserves.

Notation

As always we will use $\hat{S}_{i,k}^{m,I+1} := S_{i,k}^{m}$ for $i + k \leq I + 1$ and

$$\hat{R}_{i,k}^{m,I+1} := \Gamma_{i,k}^{m} \hat{S}_{i+k,I+1}^{m}$$

and

$$\hat{R}_{i,k}^{m_{1},m_{2},I+1} := \Gamma_{i,k}^{m_{1},m_{2}} \hat{S}_{i+k,I+1}^{m_{1},m_{2}}.$$
Lemma 4.22 (Best Estimate reserves)
Let Assumptions 4.A and 4.B be fulfilled. Then the at time $I + 1$ estimated claim properties satisfy
\[
\bar{S}_{m+1}^I : = E\left[\hat{S}_{m+1}^I \mid D_I\right] = \mathcal{F}_{m+1}^I(\bar{f}) S_I.
\]

Hence, we will use the estimates
\[
\hat{E}\left[\hat{S}_{m+1}^I \mid D_I\right] = \hat{E}\left[\hat{S}_{m+1}^I \mid D_I\right] = \mathcal{F}_{m+1}^I(\hat{f}) S_I = \mathcal{F}_{m+1}^I(\hat{f}) S_I,
\]
which implies $\hat{E}[\text{CDR}^I \mid D_I] = 0$. That means, we have Best Estimate reserves.

Notation
As always we will use $\hat{E}\left[\hat{S}_{m+1}^I \mid D_I\right] = \hat{S}_{m+1}^I$ for $i, k < I + 1$ and $\hat{R}_{m+1}^I (\hat{f}) = \mathcal{F}_{m+1}^I(\hat{f}) S_I$.

At estimation time $I + 1$ we have
\[
\hat{S}_{m+1}^I (\hat{f}) S_I.
\]

Induction: If $i + k \leq I$ then $\hat{S}_{m+1}^I (\hat{f}) S_I$

Now assume that Lemma 4.22 is fulfilled for all $i + k < n$. Then we get for $i + k = n$
\[
\hat{S}_{m+1}^I (\hat{f}) S_I.
\]

Note, a proof without induction can be done by a combination of the tower property, the multilinearity of $\mathcal{F}_n (\hat{f})$, see Lemma 4.9, and the product formula of Lemma 4.20.
Decomposition of the solvency uncertainty

\[
\text{mse}_{0|\mathcal{D}^I} [\text{CDR}^{I+1}] = \text{Var} \left[ \sum_{m=0}^{M} \sum_{i=0}^{I} \sum_{k=I-i}^{J-1} \alpha_i^m \hat{S}_{i,k+1} \left| \mathcal{D}^I \right. \right]
\]

random error

\[
+ \left( \sum_{m=0}^{M} \sum_{i=0}^{I} \sum_{k=I-i}^{J-1} \alpha_i^m \left( \hat{S}_{i,k+1} - \mathbb{E} \left[ \hat{S}_{i,k+1} \left| \mathcal{D}^I \right. \right] \right)^2 \right)
\]

parameter error

The solvency uncertainty of a single accident period or a single claim property can be obtained by choosing corresponding mixing parameters \( \alpha_i^m \).
Decomposition of the solvency uncertainty

\[ \text{mse}_\text{R}(\text{CDR}^{i+1}) = \text{Var} \left[ \sum_{m=0}^{M} \sum_{i=0}^{I-1} \sum_{k=I}^{I-1} \alpha_{m} \hat{S}_{m,i,k}^{+1} \right] \]

\[ + \left( \sum_{m=0}^{M} \sum_{i=0}^{I-1} \sum_{k=I}^{I-1} \alpha_{m} \left( \hat{S}_{m,i,k}^{+1} - \mathbb{E} \left[ \hat{S}_{m,i,k}^{+1} \right] \right) \right)^2 \]

The solvency uncertainty of a single accident period or a single claim property can be obtained by choosing corresponding mixing parameters \( \alpha_{m} \).
Taylor approximation of next years estimates

Recall the (multi-linear) functional:

\[ U(g)x := \sum_{m=0}^{M} \sum_{i=0}^{I} \alpha_i^m \left( \sum_{k=0}^{I-i} x_{i,k} + \sum_{k=I-i}^{J-1} F_{i,k}^m(g)x \right). \]

Then we have

\[ U(f)S^I = \sum_{m=0}^{M} \sum_{i=0}^{I} \alpha_i^m \sum_{k=0}^{J} E[S_{i,k}^m|\mathcal{D}^I], \]

\[ U(\hat{f})S^I = \sum_{m=0}^{M} \sum_{i=0}^{I} \alpha_i^m \sum_{k=0}^{J} \hat{S}_{i,k}^m, \quad U(F^{I+1})S^I = \sum_{m=0}^{M} \sum_{i=0}^{I} \alpha_i^m \sum_{k=0}^{J} \hat{S}_{i,k}^{m,I+1}, \]

\[ \text{CDR}^{I+1} \approx \sum_{l=0}^{M} \sum_{h=0}^{I} \sum_{j=I-h}^{J-1} \partial_{h,j} U(\hat{f})S^I \left( F_{h,j}^{l,I+1} - \hat{f}_{j}^{l,I} \right) \]

\[ = \sum_{l=0}^{M} \sum_{h=0}^{I} \partial_{h,I-h} U(\hat{f}) \left( F_{h,I-h}^l - \hat{f}_{I-h}^l \right) + \sum_{l=0}^{M} \sum_{h=0}^{I} \sum_{j=I-h+1}^{J-1} \partial_{h,j} U(\hat{f}) \omega_{I-j,j}^{l,I+1} \left( F_{I-j,j}^l - \hat{f}_{j}^l \right), \]

where \( F^{I+1} \) is the vector with coordinates

\[ F_{i,k}^{m,I+1} := \begin{cases} F_{i,k}^m, & \text{for } i + k = I \\ \hat{f}_{i,k}^m, & \text{for } i + k > I \end{cases} \]
Taylor approximation of next years estimates

Recall the (multi-linear) functional:

\[ U(g) x := M \sum_{m=0}^{\infty} I \sum_{i=0}^{\infty} \alpha_m i \left( I - i \sum_{k=0}^{\infty} x_{m,i,k} + J - 1 \sum_{k=0}^{I-i-1} F_{m,I} i,k (g) x \right). \]

Then we have

\[ U(f) S_I = M \sum_{m=0}^{\infty} I \sum_{i=0}^{\infty} \alpha_m i J \sum_{k=0}^{\infty} E[S_m i,k | D_I], \]

\[ U(\hat{f}) S_I = M \sum_{m=0}^{\infty} I \sum_{i=0}^{\infty} \alpha_m i J \sum_{k=0}^{\infty} \hat{S}_{m,I} i,k, \]

\[ U(F_{I+1}) S_I = M \sum_{m=0}^{\infty} I \sum_{i=0}^{\infty} \alpha_m i J \sum_{k=0}^{\infty} \hat{S}_{m,I+1} i,k, \]

\[ \hat{CDR}_{I+1} \approx M \sum_{l=0}^{\infty} I \sum_{h=0}^{\infty} J - 1 \sum_{j=0}^{I-h} \partial_{l,h,j} U(\hat{f}) S_I (F_{l,I+1} h,j - \hat{f}_{l,I} j) = M \sum_{l=0}^{\infty} I \sum_{h=0}^{\infty} \partial_{l,h,I} - h U(\hat{f}) (F_{l,I} h,j - \hat{f}_{l,I} j) + M \sum_{l=0}^{\infty} I \sum_{h=0}^{\infty} J - 1 \sum_{j=0}^{I-h+1} \partial_{l,h,j} U(\hat{f}) w_{l,I+1} i,j (F_{l,I} j - \hat{f}_{l,I} j). \]

For \( k = I - i \) we get, see Lemma 4.20,

\[ F_{i,I-i}^{m,I+1} - \hat{f}_{i,I-i}^{m,I} = F_{i,I-i}^{m} - \hat{f}_{i,I-i}^{m,I} \]

and for \( k > I - i \) it is

\[ F_{i,k}^{m,I+1} - \hat{f}_{k}^{m,I} = \hat{f}_{k}^{m,I+1} - \hat{f}_{k}^{m,I} = (1 - w_{I-k,k}^{m,I+1}) \hat{f}_{k}^{m,I} + w_{I-k,k}^{m,I+1} F_{I-k,k}^{m} - \hat{f}_{k}^{m,I} = w_{I-k,k}^{m,I+1} \left( F_{I-k,k}^{m} - \hat{f}_{k}^{m,I} \right). \]

Note, since

\[ F_{i,I-i}^{m,I+1} = F_{i,I-i}^{m} = \frac{S_{i,I-i+1}}{R_{i,I-i}} \]

we get

\[ \mathcal{F}_{i,i-I}^{m,I} \left( \mathbf{F}^{I+1} \right) S^I = F_{i,I-i}^{m,I+1} R_{i,I-i} = S_{i,I-i+1}. \]

That means, the operator \( U(\mathbf{F}^{I+1}) \) (re)constructs in the first step the \( I + 1 \)-th diagonal of the claim property triangles.
Linear approximation of the CDR

If we replace in the linear approximation of the ultimate, i.e. in

\[
\sum_{m=0}^{M} \sum_{i=0}^{I} \sum_{k=0}^{J} \alpha_i^m \left( \hat{S}_{i,k}^m - S_{i,k}^m \right) \approx \sum_{l=0}^{M} \sum_{h=0}^{I} \sum_{j=I-h}^{J-1} \partial_{h,j} U \left( \hat{f}^l \right) S^I \left( F_{h,j}^l - \hat{f}_{j}^l \right),
\]

the term \( F_{h,j}^l - \hat{f}_{j}^l \) by

\[
\tilde{F}_{h,j}^l - \hat{f}_{j}^l := \begin{cases} F_{I-j,j}^l - \hat{f}_{j}^l, & \text{for } j = I-h, \\ w_{I-j,j}^l \left( F_{I-j,j}^l - \hat{f}_{j}^l \right), & \text{for } j > I-h, \end{cases}
\]

we get the linear approximation of the CDR, i.e.

\[
\hat{CDR}^{I+1} \approx \sum_{l=0}^{M} \sum_{h=0}^{I} \partial_{h,I-h} U \left( \hat{f}^l \right) \left( F_{h,I-h}^l - \hat{f}_{I-h}^l \right) + \sum_{l=0}^{M} \sum_{h=0}^{I} \sum_{j=I-h+1}^{J-1} \partial_{h,j} U \left( \hat{f}^l \right) w_{I-j,j}^l \left( F_{I-j,j}^l - \hat{f}_{j}^l \right).
\]
Linear approximation of the CDR

If we replace in the linear approximation of the ultimate, i.e. in

\[
\sum_{m} \sum_{i} \alpha_{m} \sum_{k} \left( \hat{S}_{m,i,k} - S_{m,i,k} \right) \approx \sum_{l} \sum_{h} \sum_{j} \partial_{l,h,j} U(\hat{f}) \left( \hat{f} \right) S_{l,h,j} \quad (F_{l,h,j} - \hat{f}_{l,I,j})
\]

we get the linear approximation of the CDR, i.e.

\[
\hat{\text{CDR}}_{I+1} \approx \sum_{l} \sum_{h} \sum_{j} \partial_{l,h,j} U(\hat{f}) \left( \hat{f} \right) S_{l,h,j} \quad (F_{l,h,j} - \hat{f}_{l,I,j})
\]

The term \( \tilde{F}_{l,h,j} - \hat{f}_{l,h,j} \) depends on the accident period \( h \) only via the indicator functions \( 1_{j=I-h} \) and \( 1_{j>I-h} \).
Estimator 4.23 (Solvency uncertainty of all accident periods)

$$\text{mse}_{0|D^I} \left[ \widehat{\text{CDR}} \right] = E \left[ \sum_{m=0}^{M} \left( \sum_{i=0}^{I} \sum_{k=0}^{J-1} \left( \hat{S}^{m,I+1}_{i,k} - \hat{S}^{m,I}_{i,k} \right) \right)^2 | D^I \right]$$

$$\approx E \left[ \left( \sum_{l=0}^{M} \sum_{h=0}^{I} \sum_{j=I-h}^{J-1} \partial_{h,j} U \left( \hat{f}^I \right) S^I \left( \hat{F}_{h,j} - \hat{f}_{h,j} \right) \right)^2 | D^I \right] \quad \text{(Taylor approximation)}$$

$$\approx \sum_{l_1,l_2=0}^{M} \sum_{j=0}^{J-1} \hat{\sigma}_{l_1,l_2,j} \left( \frac{\hat{R}_{l_1}^{l_1,l_2}}{\hat{R}_{l_1}^{l_1}} \frac{\hat{R}_{l_2}^{l_1,l_2}}{\hat{R}_{l_2}^{l_1}} \right) + \sum_{h=0}^{I-j-1} \sum_{l_1,l_2=0}^{I-j-1} \hat{\sigma}_{l_1,l_2,h,j} \left( \frac{\hat{R}_{l_1}^{l_1,l_2}}{\hat{R}_{l_1}^{l_1}} \frac{\hat{R}_{l_2}^{l_1,l_2}}{\hat{R}_{l_2}^{l_1}} \right)$$

random error

parameter error

$$\sum_{h_1=I-j}^{I} \sum_{h_2=I-j}^{I} \left( \mathbf{1}_{j=I-h_1} + \mathbf{1}_{j>I-h_1} w_{l_1,j}^{l_1,I+1} \right) \partial_{h_1,j} U \left( \hat{f}^I \right) S^I$$

$$\left( \mathbf{1}_{j=I-h_2} + \mathbf{1}_{j>I-h_2} w_{l_2,j}^{l_2,I+1} \right) \partial_{h_2,j} U \left( \hat{f}^I \right) S^I$$

The red terms indicate the differences to our estimator of the ultimate uncertainty.
After the Taylor approximation we can exchange expectation and summation to get

\[
\text{mse}_{0|D} \left[ \hat{\mathbf{C}}_{DR} \right] \\
= \sum_{l_1, l_2 = 0}^{M} \sum_{h_1, h_2 = 0}^{I} \sum_{j_1 = I - h_1}^{I - 1} \sum_{j_2 = I - h_2}^{J - 1} \left( 1_{j_1 = I - h_1} + 1_{j_1 > I - h_1} w_{I - j_1, j_1}^{l_1, I + 1} \right) \left( 1_{j_2 = I - h_2} + 1_{j_2 > I - h_2} w_{I - j_2, j_2}^{l_2, I + 1} \right) \partial_{h_1, j_1}^{l_1} U \left( \hat{f}^{I} \right) \mathbf{S}^{I} \partial_{h_2, j_2}^{l_2} U \left( \hat{f}^{I} \right) \mathbf{S}^{I} \mathbb{E} \left[ \left( F_{I - j_1, j_1}^{l_1} - \hat{f}_{j_1}^{l_1} \right) \left( F_{I - j_2, j_2}^{l_2} - \hat{f}_{j_2}^{l_2} \right) \right] | D^{I} \\
\]

and from the estimation of the ultimate uncertainty we know

\[
\mathbb{E} \left[ \left( F_{I - j_1, j_1}^{l_1} - \hat{f}_{j_1}^{l_1} \right) \left( F_{I - j_2, j_2}^{l_2} - \hat{f}_{j_2}^{l_2} \right) \right] | D^{I} \approx 1_{j_1 = j_2} \hat{\sigma}_{j_1}^{l_1, l_2} \left( \frac{\hat{R}_{I - j_1, j_1}^{l_1, l_2}}{\hat{R}_{I - j_1, j_1}^{l_1}} + \sum_{h = 0}^{I - j - 1} w_{h, j}^{l_1} w_{h, j}^{l_2} \frac{\hat{R}_{h, j}^{l_1, l_2}}{\hat{R}_{h, j}^{l_1}} \right).
\]

Both together lead to the stated estimator.

Note, if it wasn’t for different claim properties (indices \( l_1 \) and \( l_2 \)) the last two lines of the estimator would have been a square of a sum over accident periods.

Moreover, for the random error part we had in the ultimate uncertainty case only one sum over accident periods \( h \) from \( I - j \) to \( I \), i.e. we had \( h_1 = h_2 \).
Chain-Ladder estimator for the solvency uncertainty

\[
\text{mse}_{0|D_I} \left[ \sum_{i=0}^{I} \hat{\text{CDR}}_i \right] \approx \sum_{j=0}^{J-1} \frac{\hat{\sigma}^2_j}{(\hat{f}_j^I)^2} \left( \frac{1}{C_{I-j,j}} + \sum_{h=0}^{I-j-1} \frac{(w_{h,j})^2}{C_{h,j}} \right) \left( \sum_{h=I-j}^{I} \left( 1_{j=I-h} + 1_{j>I-h} w_{I-j,j}^{I+1} \right) \hat{C}_{h,J} \right)^2
\]

LSRM estimator for the solvency uncertainty

\[
\text{mse}_{0|D_I} \left[ \hat{\text{CDR}} \right] \approx \sum_{l_1,l_2=0}^{M} \sum_{j=0}^{J-1} \hat{\sigma}_{l_1,l_2}^j \left( \frac{\hat{R}_{l_1,l_2}^{I-j,j}}{\hat{R}_{l_1}^{I-j,j}} \frac{\hat{R}_{l_2}^{I-j,j}}{\hat{R}_{l_2}^{I-j,j}} \right) + \sum_{h=0}^{I-j-1} w_{h,j}^{l_1} w_{h,j}^{l_2} \left( \frac{\hat{R}_{l_1}^{l_1} l_1}{\hat{R}_{h,j}^{l_1}} \frac{\hat{R}_{l_2}^{l_2} l_2}{\hat{R}_{h,j}^{l_2}} \right) \left( \sum_{h_1=I-j}^{I} \sum_{h_2=I-j}^{I} \left( 1_{j=I-h_1} + 1_{j>I-h_1} w_{I-j,j}^{l_1+1} \right) \partial_{h_1,j}^{l_1} U(\hat{f}^I) S^I \right.
\]

\[
\left. \left( 1_{j=I-h_2} + 1_{j>I-h_2} w_{I-j,j}^{l_2+1} \right) \partial_{h_2,j}^{l_2} U(\hat{f}^I) S^I \right)
\]
Because we have several claim properties, squared terms for Chain-Ladder are replaced by products of claim properties and the double sum over them.
Estimation at time $n$

The development factors are estimated by

$$\hat{f}_{k}^{n,n} := \sum_{h=0}^{n-k-1} w_{h,k}^{m,n} F_{h,k}^m$$

with consistent future weights $w_{i,k}^{m,n}$, which means there exists $D_k^n$-measurable weights $0 \leq w_{i,k}^{m,n} \leq 1$, for $I - k \leq i \leq n - k - 1$, with

- $R_{i,k}^m = 0$ implies $w_{i,k}^{m,n} = 0$,
- $w_{i,k}^{m,n} = (1 - w_{n-k,k}^{m,n}) w_{i,k}^{m,n-1}$, for $i + k < n$.

Then the estimate of the ultimate at time $n$ is

$$\sum_{m=0}^{M} \sum_{i=0}^{I} \alpha_i^m \sum_{k=0}^{J} \hat{S}_{i,k}^{m,n} = U(F^n) S^I$$

with

$$F_{i,k}^m := (F^n)_{i,k}^m := \begin{cases} F_{i,k}^m, & \text{for } i + k < n, \\ \hat{f}_{k}^{m,n}, & \text{for } i + k \geq n. \end{cases}$$
Consistent weights mean, that for each future estimation time $n$ we keep our relative believes in the old weights $w_{i,k}^{m,n-1}$ and only choose some weights $w_{n-k,k}^{m,n}$ for the newly observes development.

Note, although if the weights are no longer $D^I$ measurable, we will consider them as constant in our estimations.
Taylor approximation of the $n$-th CDR

$$CDR^n := \sum_{m=0}^{M} \sum_{i=0}^{I} \alpha_i^m \sum_{k=0}^{J-1} \left( \hat{S}_{i,k}^m - \hat{S}_{i,k}^I \right)$$

$$\approx \sum_{l=0}^{M} \sum_{h=0}^{I} \sum_{j=0}^{J-1} \partial_{h,j} U(\hat{f}^I) \left( F_{i,k}^m - \hat{f}_{k,I}^m \right)$$

$$\approx \sum_{l=0}^{M} \sum_{j=0}^{I-1} \sum_{h=I-j}^{(n-j-1)\wedge I} \left( \partial_{h,j} U(\hat{f}^I) S^I + \hat{w}_{h,j}^I \sum_{i=n-j}^{I} \partial_{i,j} U(\hat{f}^I) S^I \right) \left( F_{h,j}^l - \hat{f}_{j}^l \right).$$
Here we used that

\[ \hat{f}^{m,n}_{k,k} := \sum_{i=0}^{n-k-1} w^{m,n}_{i,k} F^{m}_{i,k} = \sum_{i=I-k}^{n-k-1} w^{m,n}_{i,k} F^{m}_{i,k} + \left( 1 - \sum_{i=I-k}^{n-k-1} w^{m,n}_{i,k} \right) \hat{f}^{m,n}_{k,k}, \]

for the Taylor approximation of the \( n \)-th CDR.
Estimator 4.24 (of the uncertainty between two estimation times $n_1$ and $n_2$)

$$ \text{mse}[CDR^{n_1,n_2}] := \sum_{l_1,l_2=0}^{M} \sum_{j=0}^{J-1} \hat{\sigma}_{l_1,l_2}^{h,j} \left[ \sum_{h=n_1-j}^{(n_2-j-1)\wedge I} \left( \partial_{h,j}^{l_1} U(\hat{f}^I) S^I + \hat{w}_{h,j}^{l_1,n_2} \sum_{i=n_2-j}^{I} \partial_{i,j}^{l_1} U(\hat{f}^I) S^I \right) \right] $$

$$ + \sum_{h=0}^{(n_1-j-1)\wedge I} \hat{w}_{h,j}^{l_1,n_1} \hat{w}_{h,j}^{l_2,n_2} \frac{\hat{R}_{l_1,l_2,I}}{R^I_{h,j} R^I_{h,j}} \left[ \sum_{i=n_1-j}^{I} \partial_{i,j}^{l_1} U(\hat{f}^I) S^I - \frac{\Omega_{l_1,n_1}^{l_1,n_2}}{\Omega_{l_1,n_1}^{l_1,n_1}} \sum_{i=n_2-j}^{I} \partial_{i,j}^{l_1} U(\hat{f}^I) S^I \right] $$

$$ + \sum_{h=0}^{(n_1-j-1)\wedge I} \hat{w}_{h,j}^{l_1,n_1} \hat{w}_{h,j}^{l_2,n_2} \frac{\hat{R}_{l_1,l_2,I}}{R^I_{h,j} R^I_{h,j}} \left[ \sum_{i=n_1-j}^{I} \partial_{i,j}^{l_2} U(\hat{f}^I) S^I - \frac{\Omega_{l_2,n_1}^{l_2,n_2}}{\Omega_{l_2,n_1}^{l_2,n_2}} \sum_{i=n_2-j}^{I} \partial_{i,j}^{l_2} U(\hat{f}^I) S^I \right] $$

where $\Omega_{j}^{l,n} = \sum_{i=0}^{I-j-1} \hat{w}_{i,j}^{l,n}$. 
The derivation can be obtained from the lecturer (unpublished working paper).
Remark 4.25

- If we take $n_1 = I$ and $n_2 = I + 1$ we get the same formula as in Estimator 4.23 (solvency uncertainty).
- If we take $n_1 = I$ and $n_2 = \infty$ we get the same formula as in Estimator 4.16 (ultimate uncertainty).
- In the Chain-Ladder case with variance minimizing weights we get the same formula as in Estimator 2.25.
- If the exposures $R_{i,k}^m$ do not depend on other accident periods $h \neq i$ then a similar approach like in the Chain-Ladder case may work to derive Estimator 4.24.
- Estimators for the uncertainty of the CDR between two estimation times are important for SST and Solvency II to estimate the MVM.
Remark 4.25
- If we take $n_1 = I$ and $n_2 = I + 1$ we get the same formula as in Estimator 4.23 (solvency uncertainty).
- If we take $n_1 = I$ and $n_2 = \infty$ we get the same formula as in Estimator 4.16 (ultimate uncertainty).
- In the Chain-Ladder case with variance minimizing weights we get the same formula as in Estimator 2.25.
- If the exposures $R_{m,i,k}$ do not depend on other accident periods $h \neq i$ then a similar approach like in the Chain-Ladder case may work to derive Estimator 4.24.
- Estimators for the uncertainty of the CDR between two estimation times are important for SST and Solvency II to estimate the MVM.
Swiss mandatory accident portfolio: part 3 of 3, see Example 4.13

We have four claim properties with exposures

**ME:** \( R_{i,k}^0 = R_{i,k}^{0,0} = \sum_{j=0}^{k} S_{i,j}^0 \)

**IW:** \( R_{i,k}^1 = R_{i,k}^{1,1} = \sum_{l=0}^{k} (\kappa^{l+1} S_{i,l}^3 + (1 - \kappa^{l+1}) S_{i,l}^1) \)

**Sub:** \( R_{i,k}^2 = R_{i,k}^{2,2} = \sum_{l=0}^{k} (S_{i,l}^0 + S_{i,l}^1 + S_{i,l}^2) \)

**Salary:** \( S_{i,0}^3 = P_i, \ S_{i,l}^3 = 0, \text{ for } l > 0, \) and

\[
R_{i,k}^3 = R_{i,k}^{3,0} = R_{i,k}^{0,3} = R_{i,k}^{3,1} = R_{i,k}^{1,3} = R_{i,k}^{3,2} = R_{i,k}^{2,3} = R_{i,k}^{3,3} = 0
\]

For the not yet defined exposures we take the total payments up to now, i.e.

\[
R_{i,k}^{0,1} = R_{i,k}^{1,0} = R_{i,k}^{0,2} = R_{i,k}^{2,0} = R_{i,k}^{1,2} = R_{i,k}^{2,1} = \sum_{l=0}^{k} (S_{i,l}^0 + S_{i,l}^1 + S_{i,l}^2).
\]

**Resulting solvency uncertainty**

- The estimated ultimate and solvency uncertainties behave almost alike, but on a different level.
- Although the estimated solvency uncertainty is minimal for \( \kappa \approx 0.35 \) you should never use this as criteria to evaluate which model is the best. For this portfolio I, would go for \( \kappa = 1 \) (at least for the first development periods).
- For \( \kappa = 0 \) the solvency uncertainty is slightly smaller than CLM on total payments.
We have four claim properties with exposures

\[ R_{0,i,k} = R, \quad \sum_{j=0}^{k} S_{0,i,j} = \text{EXL}, \quad R_{1,i,k} = R, \quad \sum_{l=0}^{k} (\kappa_l + 1) S_{3,i,l} + (1 - \kappa_l + 1) S_{1,i,l} \]

Sub:

\[ R_{2,i,k} = R, \quad \sum_{l=0}^{k} (S_{0,i,l} + S_{1,i,l} + S_{2,i,l}) \]

Salary:

\[ S_{3,i,0} = P, \quad S_{3,i,l} = 0 \quad \text{for } l > 0, \quad R_{3,i,k} = R, \quad R_{3,i,k} = R, \quad R_{3,i,k} = R, \quad R_{3,i,k} = R, \quad R_{3,i,k} = R \]

For the not yet defined exposures we take the total payments up to now, i.e.

\[ R_{0,1,i,k} = R_{1,0,i,k} = R_{2,0,i,k} = R_{2,0,i,k} = R_{2,0,i,k} = R_{2,0,i,k} = \sum_{l=0}^{k} (S_{0,i,l} + S_{1,i,l} + S_{2,i,l}) \]

Resulting solvency uncertainty

- The estimated ultimate and solvency uncertainties behave almost alike, but on a different level.
- Although the estimated solvency uncertainty is minimal for \( \kappa \approx 0.35 \), you should never use this as criteria to evaluate which model is the best. For this portfolio I would go for \( \kappa = 1 \) (at least for the first development periods).
- For \( \kappa = 0 \) the solvency uncertainty is slightly smaller than CLM on total payments.
Example 4.13: Solvency uncertainty in dependence of $\kappa$

We always show the square root of uncertainties.
Be aware that each curve has its own scale. So although the curve of the solvency and the ultimate uncertainty cross each other, we always have that the solvency uncertainty is smaller than the ultimate uncertainty.

In our example the solvency uncertainty is about 70% of the ultimate uncertainty. In general this ratio usually lies between 50% (general liability) and 90% (NatCat). One minus this ratio represents the gain of information over one year in comparison to all unknown information about the reserves.
Example 4.26 (ECLRM vs. CLM: part 3 of 3, see Example 4.14)

In the first two parts we have compared the Best Estimate reserves and the ultimate uncertainty. Now we want to look at the solvency uncertainty.

For the weighing of uncertainties we have to define $R_{i,k}^{0,1} = R_{i,k}^{1,0}$. We use the arithmetic mean of payments and incurred losses for CLM and the case reserves for the ECLRM:

Square root of the solvency uncertainty

<table>
<thead>
<tr>
<th></th>
<th>CLM</th>
<th>ECLRM</th>
</tr>
</thead>
<tbody>
<tr>
<td>AP</td>
<td>payments</td>
<td>incurred</td>
</tr>
<tr>
<td>0</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>1</td>
<td>89 423</td>
<td>2 553</td>
</tr>
<tr>
<td>2</td>
<td>212 847</td>
<td>4 561</td>
</tr>
<tr>
<td>3</td>
<td>131 605</td>
<td>7 825</td>
</tr>
<tr>
<td>4</td>
<td>161 223</td>
<td>6 666</td>
</tr>
<tr>
<td>5</td>
<td>145 975</td>
<td>31 325</td>
</tr>
<tr>
<td>6</td>
<td>104 800</td>
<td>45 866</td>
</tr>
<tr>
<td>7</td>
<td>230 780</td>
<td>155 175</td>
</tr>
<tr>
<td>8</td>
<td>283 765</td>
<td>150 879</td>
</tr>
<tr>
<td>9</td>
<td>229 170</td>
<td>223 154</td>
</tr>
<tr>
<td>Total</td>
<td>1 004 481</td>
<td>347 709</td>
</tr>
</tbody>
</table>
Example 4.26 (ECLRM vs. CLM: part 3 of 3, see Example 4.14)

In the first two parts we have compared the Best Estimate reserves and the ultimate uncertainty. Now we want to look at the solvency uncertainty.

For the weighing of uncertainties we have to define \( R_{0,1}^{i,k} \) and \( R_{1,0}^{i,k} \). We use the arithmetic mean of payments and incurred losses for CLM and the case reserves for the ECLRM:

\[
\text{Square root of the solvency uncertainty}
\]

<table>
<thead>
<tr>
<th>Year</th>
<th>CLM</th>
<th>ECLRM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>48,423</td>
<td>2,353</td>
</tr>
<tr>
<td>2</td>
<td>212,447</td>
<td>4,361</td>
</tr>
<tr>
<td>3</td>
<td>130,035</td>
<td>7,627</td>
</tr>
<tr>
<td>4</td>
<td>101,223</td>
<td>6,698</td>
</tr>
<tr>
<td>5</td>
<td>165,955</td>
<td>7,954</td>
</tr>
<tr>
<td>6</td>
<td>145,775</td>
<td>31,325</td>
</tr>
<tr>
<td>7</td>
<td>259,768</td>
<td>151,175</td>
</tr>
<tr>
<td>8</td>
<td>254,624</td>
<td>139,734</td>
</tr>
</tbody>
</table>

Total | 1,004,481 | 347,709 |

2019-02-18

- Taking the arithmetic mean

\[
R_{i,k}^{m_1,m_2} := \frac{1}{2} \left( R_{i,k}^{m_1,m_1} + R_{i,k}^{m_2,m_2} \right)
\]

for the coupling exposures works fine if \( R_{i,k}^{m_1,m_1} \) and \( R_{i,k}^{m_2,m_2} \) are similar. In general the geometric mean

\[
R_{i,k}^{m_1,m_2} := \sqrt{R_{i,k}^{m_1,m_1} R_{i,k}^{m_2,m_2}}
\]

usually works better.

- Although the Best Estimate reserves are similar, the solvency uncertainties are not, in particular CLM on payments leads to a much higher solvency uncertainty than the others.

- Again, you must not use estimates of the ultimate uncertainty to evaluate which model is the best.
Measurement of reserving risks under IFRS 17, SST and Solvency II

- In recent years the reserving risk has got more and more attention, for instance under IFRS 17, SST and Solvency II.
- Probably, the most common method to estimate reserving risk is the following:
  1. Make assumptions about the distribution family for the reserves for each portfolio.
  2. Estimate the corresponding parameters, for instance mean (Best Estimate reserves) and variance (mse + model error).
     ⇒ Calculate the reserving risk for each portfolio, for instance, value at risk or expected shortfall.
  3. Make assumptions on the correlation (or copula) of portfolios.
     ⇒ Calculate the reserving risk of the aggregation of all portfolios.
- In particular step 3 is usually based mostly on actuarial judgement.
- LSRMs can be used to avoid correlation matrices or to get some estimates of them.
In recent years the reserving risk has got more and more attention, for instance under IFRS 17, SST and Solvency II.

 Probably, the most common method to estimate reserving risk is the following:

1. Make assumptions about the distribution family for the reserves for each portfolio.
2. Estimate the corresponding parameters, for instance mean (Best Estimate reserves) and variance (mse + model error).
3. Calculate the reserving risk for each portfolio, for instance, value at risk or expected shortfall.
4. Make assumptions on the correlation (or copula) of portfolios.
5. Calculate the reserving risk of the aggregation of all portfolios.

In particular step 3 is usually based mostly on actuarial judgement.

LSRM can be used to avoid correlation matrices or to get some estimates of them.
### Part of the correlation matrix of the SST-Standardmodel 2014

<table>
<thead>
<tr>
<th></th>
<th>MFH</th>
<th>MFK</th>
<th>Sach</th>
<th>ES-Pool</th>
<th>Haft</th>
<th>UVG</th>
</tr>
</thead>
<tbody>
<tr>
<td>MFH</td>
<td>1.00</td>
<td>0.15</td>
<td>0.15</td>
<td>0.15</td>
<td>0.25</td>
<td>0.50</td>
</tr>
<tr>
<td>MFK</td>
<td>0.15</td>
<td>1.00</td>
<td>0.15</td>
<td>0.15</td>
<td>0.15</td>
<td>0.15</td>
</tr>
<tr>
<td>Sach</td>
<td>0.15</td>
<td>0.15</td>
<td>1.00</td>
<td>0.15</td>
<td>0.15</td>
<td>0.15</td>
</tr>
<tr>
<td>ES-Pool</td>
<td>0.15</td>
<td>0.15</td>
<td>0.15</td>
<td>1.00</td>
<td>0.15</td>
<td>0.15</td>
</tr>
<tr>
<td>Haft</td>
<td>0.25</td>
<td>0.15</td>
<td>0.15</td>
<td>0.15</td>
<td>1.00</td>
<td>0.25</td>
</tr>
<tr>
<td>UVG</td>
<td>0.50</td>
<td>0.15</td>
<td>0.15</td>
<td>0.15</td>
<td>0.25</td>
<td>1.00</td>
</tr>
</tbody>
</table>

- The entries are based on actuarial judgement.
- The correlation matrix under Solvency II contains similar entries.
Stochastic Reserving

- Linear-Stochastic-Reserving methods
- Estimation of correlation of reserving Risks

Part of the correlation matrix of the SST-Standardmodel 2014

<table>
<thead>
<tr>
<th></th>
<th>MFH</th>
<th>MFK</th>
<th>Sach</th>
<th>ES-Pool</th>
<th>Haft</th>
<th>UVG</th>
</tr>
</thead>
<tbody>
<tr>
<td>MFH</td>
<td>1.00</td>
<td>0.15</td>
<td>0.15</td>
<td>0.15</td>
<td>0.25</td>
<td>0.50</td>
</tr>
<tr>
<td>MFK</td>
<td>0.15</td>
<td>1.00</td>
<td>0.15</td>
<td>0.15</td>
<td>0.15</td>
<td>0.15</td>
</tr>
<tr>
<td>Sach</td>
<td>0.15</td>
<td>0.15</td>
<td>1.00</td>
<td>0.15</td>
<td>0.15</td>
<td>0.15</td>
</tr>
<tr>
<td>ES-Pool</td>
<td>0.15</td>
<td>0.15</td>
<td>0.15</td>
<td>1.00</td>
<td>0.15</td>
<td>0.15</td>
</tr>
<tr>
<td>Haft</td>
<td>0.25</td>
<td>0.15</td>
<td>0.15</td>
<td>0.15</td>
<td>1.00</td>
<td>0.25</td>
</tr>
<tr>
<td>UVG</td>
<td>0.50</td>
<td>0.15</td>
<td>0.15</td>
<td>0.15</td>
<td>0.25</td>
<td>1.00</td>
</tr>
</tbody>
</table>

- The entries are based on actuarial judgement.
- The correlation matrix under Solvency II contains similar entries.
If we use LSRMs we can avoid correlation matrices for the reserve risks:

1. Set up a LSRM for all portfolios together. That means we have to specify coupling exposures $R_{i,k}^{m_1,m_2}$ for all $m_1 \neq m_2$, too. Here, heuristic arguments can help to do so. For instance, if you use the same method for claim properties $m_1$ and $m_2$ it may be appropriate to take the geometric mean of $R_{i,k}^{m_1,m_1}$ and $R_{i,k}^{m_2,m_2}$.

2. Choose a distribution family for the total reserve of all portfolios.

3. Estimate the corresponding parameter, for instance

   \[
   \text{mean} = \text{Best Estimate reserves and} \\
   \text{variance} = \text{ultimate or solvency uncertainty + model error.}
   \]

Here you may have to scale the variance in case that the Best Estimate reserves are not equal to the reserves estimated by the LSRM, see slide 56.

$\implies$ Calculate the reserving risk.
If we use LSRMs we can avoid correlation matrices for the reserve risks:

1. Set up a LSRM for all portfolios together. That means we have to specify coupling exposures $R_{m_1,m_2}^{i,k}$ for all $m_1 \neq m_2$ too. Here heuristic arguments can help to do so. For instance, if you use the same method for claim properties $m_1$ and $m_2$ it may be appropriate to take the geometric mean of $R_{m_1,m_1}^{i,k}$ and $R_{m_2,m_2}^{i,k}$.

2. Choose a distribution family for the total reserve of all portfolios.

3. Estimate the corresponding parameter, for instance
   - mean = Best Estimate reserves and
   - variance = ultimate or solvency uncertainty + model error.

Here you may have to scale the variance in case that the Best Estimate reserves are not equal to the reserves estimated by the LSRM, see slide 56.

$\Rightarrow$ Calculate the reserving risk.
The formulas for the ultimate and for the solvency uncertainty have the form:

\[
\sum_{m_1,m_2=0}^{M} \sum_{i_1,i_2=0}^{I} \alpha_{i_1}^{m_1} \alpha_{i_2}^{m_2} \beta_{i_1,i_2}^{m_1,m_2},
\]

whereas \(\alpha_i^m\) are arbitrary \(\mathcal{D}^I\)-measurable real numbers. Moreover, since the uncertainties are defined as expectation of the square of some random variable they are non-negative for all collections \((\alpha_i^m)_{0 \leq m \leq M}\) of \(\mathcal{D}^I\)-measurable real numbers, which means that

\[
\begin{pmatrix}
\sum_{i_1,i_2=0}^{I} \alpha_{i_1}^{m_1} \alpha_{i_2}^{m_2} \beta_{i_1,i_2}^{m_1,m_2}
\end{pmatrix}_{0 \leq m_1,m_2 \leq M}
\]

is a positive semidefinite matrix. We already take the diagonal elements of this matrix as variances of the reserving risk of one claim property. Therefore, it is appropriate to use the whole matrix as covariance matrix.
The formulas for the ultimate and for the solvency uncertainty have the form:

$$M \sum_{m_1,m_2=0}^{M} \alpha_{m_1} \alpha_{m_2} \beta_{m_1,m_2} = 0$$

$$I \sum_{i_1,i_2=0}^{I} \alpha_{i_1} \alpha_{i_2} \beta_{i_1,i_2} = 0$$

whereas $\alpha_{m_i}$ are arbitrary $D^I$-measurable real numbers.

Moreover, since the uncertainties are defined as expectation of the square of some random variable they are non negative for all collections $(\alpha_{m_i}, \beta_{m_1,m_2})_{0 \leq m \leq M}$ of $D^I$-measurable real numbers, which means that

$$\left( \sum_{i_1,i_2=0}^{I} \alpha_{i_1} \alpha_{i_2} \beta_{i_1,i_2} \right)_{0 \leq m_1,m_2 \leq M}$$

is a positive semidefinite matrix. We already take the diagonal elements of this matrix as variances of the reserving risk of one claim property. Therefore, it is appropriate to use the whole matrix as covariance matrix.

positive semidefinite: For any vector $x = (x_m)_{0 \leq m \leq M}$ we get

$$x' \left( \sum_{i_1,i_2=0}^{I} \alpha_{i_1}^{m_1} \alpha_{i_2}^{m_2} \beta_{i_1,i_2}^{m_1,m_2} \right)_{0 \leq m_1,m_2 \leq M} x = \sum_{m_1,m_2=0}^{M} x_{m_1} \sum_{i_1,i_2=0}^{I} \alpha_{i_1}^{m_1} \alpha_{i_2}^{m_2} \beta_{i_1,i_2}^{m_1,m_2} x_{m_2}$$

$$= \sum_{m_1,m_2=0}^{M} \sum_{i_1,i_2=0}^{I} \alpha_{i_1}^{m_1} x_{m_1} \alpha_{i_2}^{m_2} x_{m_2} \beta_{i_1,i_2}^{m_1,m_2} \geq 0.$$
4.6 Estimation of correlation of reserving Risks

4.6.2 Using LSRMs to estimate a correlation matrix (2/2)

Estimating correlation of reserve risk, see LSRM_Cor_ActiveX.xlsx

Based on the example of the article [27] by A. Gisler and M. Wüthrich with

\[ R_{m_1,m_2}^{i,k} := \sqrt{\sum_{j=0}^{k} S_{i,j}^{m_1} \sum_{j=0}^{k} S_{i,j}^{m_2}}. \]

Estimated ultimate uncertainty correlation

<table>
<thead>
<tr>
<th>( m_1/m_2 )</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.00</td>
<td>-0.15</td>
<td>0.01</td>
<td>0.23</td>
<td>-0.17</td>
<td>0.26</td>
</tr>
<tr>
<td>1</td>
<td>-0.15</td>
<td>1.00</td>
<td>0.03</td>
<td>0.13</td>
<td>-0.03</td>
<td>-0.00</td>
</tr>
<tr>
<td>2</td>
<td>0.01</td>
<td>0.03</td>
<td>1.00</td>
<td>0.04</td>
<td>0.06</td>
<td>-0.05</td>
</tr>
<tr>
<td>3</td>
<td>0.23</td>
<td>0.13</td>
<td>0.04</td>
<td>1.00</td>
<td>-0.05</td>
<td>0.09</td>
</tr>
<tr>
<td>4</td>
<td>-0.17</td>
<td>-0.03</td>
<td>0.06</td>
<td>-0.05</td>
<td>1.00</td>
<td>0.03</td>
</tr>
<tr>
<td>5</td>
<td>0.26</td>
<td>-0.00</td>
<td>-0.05</td>
<td>0.09</td>
<td>0.03</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Estimated solvency uncertainty correlation

<table>
<thead>
<tr>
<th>( m_1/m_2 )</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.00</td>
<td>0.04</td>
<td>0.05</td>
<td>0.30</td>
<td>-0.26</td>
<td>0.31</td>
</tr>
<tr>
<td>1</td>
<td>0.04</td>
<td>1.00</td>
<td>0.04</td>
<td>0.30</td>
<td>-0.10</td>
<td>0.00</td>
</tr>
<tr>
<td>2</td>
<td>0.05</td>
<td>0.04</td>
<td>1.00</td>
<td>0.09</td>
<td>0.08</td>
<td>-0.06</td>
</tr>
<tr>
<td>3</td>
<td>0.30</td>
<td>0.30</td>
<td>0.09</td>
<td>1.00</td>
<td>-0.08</td>
<td>0.16</td>
</tr>
<tr>
<td>4</td>
<td>-0.26</td>
<td>-0.10</td>
<td>0.08</td>
<td>-0.08</td>
<td>1.00</td>
<td>0.03</td>
</tr>
<tr>
<td>5</td>
<td>0.31</td>
<td>0.00</td>
<td>-0.06</td>
<td>0.16</td>
<td>0.03</td>
<td>1.00</td>
</tr>
</tbody>
</table>
• The calculations can be found in the file ‘LSRM_Cor_Dll.xlsx’ (or ‘LSRM_Cor_ActiveX.xlsx’).

• Most of the correlations are negligible, except for the dependence related to

$$S^5_{i,k} \text{ vs. } S^0_{i,k} \quad \text{and} \quad S^3_{i,k} \text{ vs. } S^0_{i,k}, S^1_{i,k} \text{ and } S^5_{i,k}$$

and some diversification related to

$$S^0_{i,k} \text{ vs. } S^4_{i,k} \quad \text{and maybe} \quad S^0_{i,k} \text{ vs. } S^1_{i,k} \text{ and } S^4_{i,k}.$$

• Strictly taken, the model is not valid, because of some negative eigenvalues of the covariance matrices

$$(\hat{\sigma}_{k}^{m_1,m_2} R_{i,k}^{m_1,m_2})_{0\leq m_1,m_2\leq M} \text{ for } k \in \{6, 8, 9\}.$$ But the results mainly depend on the development periods $k = 0$ and $k = 1$, only. Moreover, except for $k = 6$ the negative eigenvalues are almost zero, which means that it is more a problem of the estimation than a model problem.

• The estimated correlations are estimated under the assumption that the claim properties fulfil Assumptions 4.A and 4.B, which usually is not the case, for instance because of inflation or other diagonal effects. Therefore, in practice we should always think of adding some model error in terms of a positive correlation.
Literature

Credibility for the Chain Ladder Reserving Method.  
_Astin Bull., 38_(2):565–600, 2008._

Implementation of LSRMs under GPL 3. (http://sourceforge.net/projects/lsrmtools/).

Linear stochastic reserving methods.  

[22] Thomas Mack.  
Distribution-free calculation of the standard error of chain ladder reserving estimates.  
Literature
Credibility for the Chain Ladder Reserving Method.

Implementation of LSRMs under GPL 3. (http://sourceforge.net/projects/lsrmtools/).

Linear stochastic reserving methods.

[22] Thomas Mack.
Distribution-free calculation of the standard error of chain ladder reserving estimates.
Stochastic Reserving
Lecture 9
Separation of Large and Small Claims

René Dahms

ETH Zurich, Spring 2019

17 April 2019

(Last update: 18 February 2019)
5 Separation of small and large claims

5.1 What is the problem with large claims

5.2 How to separate small from large claims
   5.2.1 Small and large by latest information
   5.2.2 Ever and never large by latest information
   5.2.3 Small and large now
   5.2.4 Ever and never large up to now
   5.2.5 Ever large up to now and never large by latest information
   5.2.6 Attritional and excess
   5.2.7 Separation methods summary

5.3 Estimation methods for small and large claims

5.4 Modelling the transition from small to large

5.5 Literature
5 Separation of small and large claims

5.1 What is the problem with large claims

5.2 How to separate small from large claims
  5.2.1 Small and large by latest information
  5.2.2 Ever and never large by latest information
  5.2.3 Small and large now
  5.2.4 Ever and never large up to now
  5.2.5 Ever large up to now and never large by latest information
  5.2.6 Attritional and excess
  5.2.7 Separation methods summary

5.3 Estimation methods for small and large claims

5.4 Modelling the transition from small to large

5.5 Literature
### Increments of incurred losses with individual Chain-Ladder development factors

<table>
<thead>
<tr>
<th>i\k</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>i\k</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>296</td>
<td>7.5%</td>
<td>22</td>
<td>-3.0%</td>
<td>-10</td>
<td>4.4%</td>
<td>-14</td>
<td>-6.1%</td>
<td>-18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>285</td>
<td>1.6%</td>
<td>5</td>
<td>-5.9%</td>
<td>-17</td>
<td>-7.2%</td>
<td>-20</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>259</td>
<td>4.4%</td>
<td>11</td>
<td>-8.0%</td>
<td>-22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>277</td>
<td>5.9%</td>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>268</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>269</td>
<td>4.4%</td>
<td>12</td>
<td>-3.8%</td>
<td>-11</td>
<td>-5.7%</td>
<td>-15</td>
<td>-7.9%</td>
<td>-20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>274</td>
<td>1.2%</td>
<td>3</td>
<td>-6.4%</td>
<td>-18</td>
<td>-7.5%</td>
<td>-19</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>250</td>
<td>4.5%</td>
<td>11</td>
<td>-8.0%</td>
<td>-21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>254</td>
<td>3.3%</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>263</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- We see a huge variability within the individual development factors at the first development period.
  - What are the reasons for this behaviour?
  - Are the first two exceptional extremes?
  - How often may they occur? Once in four years or once in 40 years?
- One possible reason is the behaviour of large claims.
- After eliminating all large claims it seems, that only the second observed development factor of the first development period is still out of line.
- Accident period 4 still contains a claim which will become large in three years. But such claims are excluded for accident periods 0 and 1! Therefore accident periods are not comparable!
- The example is taken from [22], but only the first five calendar periods.
Stochastic Reserving

Separation of small and large claims

What is the problem with large claims

We see a huge variability within the individual development factors at the first development period.

- What are the reasons for this behaviour?
- Are the first two exceptional extremes?
- How often may they occur? Once in four years or once in 40 years?

One possible reason is the behaviour of large claims.

After eliminating all large claims it seems, that only the second observed development factor of the first development period is still out of line.

Accident period 4 still contains a claim which will become large in three years. But such claims are excluded for accident periods 0 and 1! Therefore accident periods are not comparable!

The example is taken from [22], but only the first five calendar periods.

### Increments of incurred losses with individual Chain-Ladder development factors

<table>
<thead>
<tr>
<th>i</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>j</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>258</td>
<td>285</td>
<td>259</td>
<td>277</td>
<td>274</td>
<td>269</td>
<td>274</td>
<td>250</td>
<td>254</td>
<td>268</td>
</tr>
<tr>
<td></td>
<td>-32</td>
<td>-12</td>
<td>-11</td>
<td>-17</td>
<td>-15</td>
<td>-12</td>
<td>-11</td>
<td>-17</td>
<td>-15</td>
<td>-12</td>
</tr>
<tr>
<td></td>
<td>-10</td>
<td>-5</td>
<td>-12</td>
<td>-17</td>
<td>-11</td>
<td>-6</td>
<td>-5</td>
<td>-12</td>
<td>-11</td>
<td>-6</td>
</tr>
<tr>
<td></td>
<td>-14</td>
<td>-8</td>
<td>-11</td>
<td>-17</td>
<td>-11</td>
<td>-9</td>
<td>-8</td>
<td>-11</td>
<td>-11</td>
<td>-9</td>
</tr>
<tr>
<td></td>
<td>-18</td>
<td>-12</td>
<td>-11</td>
<td>-17</td>
<td>-11</td>
<td>-13</td>
<td>-12</td>
<td>-11</td>
<td>-11</td>
<td>-13</td>
</tr>
</tbody>
</table>

We see a huge variability within the individual development factors at the first development period.

- What are the reasons for this behaviour?
- Are the first two exceptional extremes?
- How often may they occur? Once in four years or once in 40 years?

One possible reason is the behaviour of large claims.

After eliminating all large claims it seems, that only the second observed development factor of the first development period is still out of line.

Accident period 4 still contains a claim which will become large in three years. But such claims are excluded for accident periods 0 and 1! Therefore accident periods are not comparable!

The example is taken from [22], but only the first five calendar periods.
Aims of separating small and large claims

1. Get a smooth triangle of small claims.
2. Do not transfer too much reserves to the triangle of large claims.

Both aims contradict each other. Therefore, we have to find a good balance.
Aims of separating small and large claims

1. Get a smooth triangle of small claims.
2. Do not transfer too much reserves to the triangle of large claims.

Both aims contradict each other. Therefore, we have to find a good balance.

Aim 1. could be easily fulfilled by defining all claims as large. And on the other hand aim 2. could be easily fulfilled by defining all claims as small.
General problems for separating large and small claims

- Should we compare payments or incurred losses with the threshold? In most cases we should take incurred losses, because payments usually exceed the threshold much later.
- The relations used, i.e. “≤ and >” or “< and ≥”.
- Completeness, i.e. no leftovers and no double counting.
- Consistency over time, i.e. are the separate developments of small and large claims comparable over all accident periods?
- Systematic over- or underestimation. This often goes along with the consistency over time.
- The choice of the threshold, in particular in cases where the separation method is not continuous with respect to the threshold.
- Does the separation lead to better estimates of the reserves? Usually, we would like to take large claims out in order to get a smooth but not trivial triangle of small claims, which then can be analysed by standard methods. Not trivial means that still a reasonable amount of reserves belong to small claims.
Terms like large and small claims are not consistently used in practice as well as in the literature. For instance, you could find

**large claim:**
- large loss
- catastrophic claim (or loss)
- exceptional claim (or loss)
- ...

**small claims:**
- small losses
- normal claims (or losses)
- attritional claims (or losses)
- ...

- Should we compare payments or incurred losses with the threshold? In most cases we should take incurred losses, because payments usually exceed the threshold much later.
- The relations used, i.e. \( \leq \) and \( > \) or \( < \) and \( \geq \).
- Completeness, i.e. no leftovers and no double counting.
- Consistency over time, i.e. are the separate developments of small and large claims comparable over all accident periods?
- Systematic over- or underestimation. This often goes along with the consistency over time.
- The choice of the threshold, in particular in cases where the separation method is not continuous with respect to the threshold.
- Does the separation lead to better estimates of the reserves? Usually, we would like to take large claims out in order to get a smooth but not trivial triangle of small claims, which then can be analysed by standard methods. Not trivial means that still a reasonable amount of reserves belong to small claims.
Discussion of various separation methods

In this lecture we want to discuss various methods to separate small and large claims. Moreover, we want to highlight their advantages and drawbacks. In order to do so we will keep life simple and focus on the following deterministic portfolio (see Excel file “Large_and_Small.xlsx”):

- We fix a threshold of 400.
- The portfolio consists of three types of claims:
  * 100 claims that never exceed the threshold (small claims).
  * One claim that after some time exceeds the threshold, but will be finally settled below it (large claim 1).
  * One claim that exceeds the threshold (large claim 2).

We will illustrate each separation method at the example of large claim 1 and discuss the advantages and drawbacks of the separation at the example of Chain-Ladder projections of separate incurred triangles containing small and large claims. Therefore, we denote by $X_k$ the incurred loss of large claim 1 at (development) time $k$. 
Using CLM is adequate, because we deal with a non random portfolio which is constant over time.
Deterministic development of the example portfolio

The development of payments and incurred losses are as follows:

<table>
<thead>
<tr>
<th>incurred losses</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>small claim</td>
<td>10</td>
<td>15</td>
<td>18</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>large claim 1</td>
<td>300</td>
<td>700</td>
<td>800</td>
<td>350</td>
<td>350</td>
</tr>
<tr>
<td>large claim 2</td>
<td>500</td>
<td>800</td>
<td>900</td>
<td>950</td>
<td>950</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>paid to date</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>small claim</td>
<td>5</td>
<td>13</td>
<td>18</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>large claim 1</td>
<td>10</td>
<td>100</td>
<td>500</td>
<td>350</td>
<td>350</td>
</tr>
<tr>
<td>large claim 2</td>
<td>0</td>
<td>100</td>
<td>250</td>
<td>950</td>
<td>950</td>
</tr>
</tbody>
</table>

Therefore, we expect the following outcome:

<table>
<thead>
<tr>
<th>AP</th>
<th>paid</th>
<th>incurred</th>
<th>ultimate</th>
<th>reserves</th>
<th>IBN(e/y)R</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>3100</td>
<td>3100</td>
<td>3100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>3100</td>
<td>3100</td>
<td>3100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>2550</td>
<td>3500</td>
<td>3100</td>
<td>550</td>
<td>-400</td>
</tr>
<tr>
<td>3</td>
<td>1500</td>
<td>3000</td>
<td>3100</td>
<td>1600</td>
<td>100</td>
</tr>
<tr>
<td>4</td>
<td>510</td>
<td>1800</td>
<td>3100</td>
<td>2590</td>
<td>1300</td>
</tr>
<tr>
<td>total</td>
<td>10760</td>
<td>14500</td>
<td>15500</td>
<td>4740</td>
<td>1000</td>
</tr>
</tbody>
</table>

### Deterministic development of the example portfolio

The development of payments and incurred losses are as follows:

#### Incurred losses

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small</td>
<td>100</td>
<td>150</td>
<td>180</td>
<td>180</td>
<td>180</td>
</tr>
<tr>
<td>Large 1</td>
<td>300</td>
<td>700</td>
<td>800</td>
<td>350</td>
<td>350</td>
</tr>
<tr>
<td>Large 2</td>
<td>500</td>
<td>800</td>
<td>900</td>
<td>950</td>
<td>950</td>
</tr>
</tbody>
</table>

#### Paid to date

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small</td>
<td>500</td>
<td>1300</td>
<td>1800</td>
<td>1800</td>
<td>1800</td>
</tr>
<tr>
<td>Large 1</td>
<td>10</td>
<td>100</td>
<td>500</td>
<td>350</td>
<td>350</td>
</tr>
<tr>
<td>Large 2</td>
<td>0</td>
<td>100</td>
<td>250</td>
<td>950</td>
<td>950</td>
</tr>
</tbody>
</table>

Therefore, we expect the following outcome:

<table>
<thead>
<tr>
<th></th>
<th>AP</th>
<th>IBN(e/y)</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>3100</td>
<td>3100</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>3100</td>
<td>3100</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>2550</td>
<td>3500</td>
<td>550</td>
</tr>
<tr>
<td>3</td>
<td>1500</td>
<td>3100</td>
<td>1600</td>
</tr>
<tr>
<td>4</td>
<td>510</td>
<td>1500</td>
<td>2590</td>
</tr>
<tr>
<td>Total</td>
<td>10760</td>
<td>14500</td>
<td>15500</td>
</tr>
</tbody>
</table>

### Separation of small and large claims

#### How to separate small from large claims

#### Cumulative incurred losses of 100 small claims

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1000</td>
<td>1500</td>
<td>1800</td>
<td>1800</td>
<td>1800</td>
</tr>
<tr>
<td>1</td>
<td>1000</td>
<td>1500</td>
<td>1800</td>
<td>1800</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1000</td>
<td>1500</td>
<td>1800</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1000</td>
</tr>
</tbody>
</table>

#### Cumulative incurred losses of large claim 1

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>300</td>
<td>700</td>
<td>800</td>
<td>350</td>
<td>350</td>
</tr>
<tr>
<td>1</td>
<td>300</td>
<td>700</td>
<td>800</td>
<td>350</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>300</td>
<td>700</td>
<td>800</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>300</td>
</tr>
</tbody>
</table>

#### Cumulative incurred losses of large claim 2

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>500</td>
<td>800</td>
<td>900</td>
<td>950</td>
<td>950</td>
</tr>
<tr>
<td>1</td>
<td>500</td>
<td>800</td>
<td>900</td>
<td>950</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>500</td>
<td>800</td>
<td>900</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>500</td>
</tr>
</tbody>
</table>

#### Cumulative incurred losses of all claims

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1800</td>
<td>3000</td>
<td>3500</td>
<td>3100</td>
<td>3100</td>
</tr>
<tr>
<td>1</td>
<td>1800</td>
<td>3000</td>
<td>3500</td>
<td>3100</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1800</td>
<td>3000</td>
<td>3500</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1800</td>
<td>3000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1800</td>
</tr>
</tbody>
</table>

#### Cumulative payments for 100 small claims

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>500</td>
<td>1300</td>
<td>1800</td>
<td>1800</td>
<td>1800</td>
</tr>
<tr>
<td>1</td>
<td>500</td>
<td>1300</td>
<td>1800</td>
<td>1800</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>500</td>
<td>1300</td>
<td>1800</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>500</td>
</tr>
</tbody>
</table>

#### Cumulative payments for large claim 1

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>10</td>
<td>100</td>
<td>500</td>
<td>350</td>
<td>350</td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td>100</td>
<td>500</td>
<td>350</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>100</td>
<td>500</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10</td>
</tr>
</tbody>
</table>

#### Cumulative payments for large claim 2

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>100</td>
<td>250</td>
<td>950</td>
<td>950</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>100</td>
<td>250</td>
<td>950</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>100</td>
<td>250</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Cumulative payments for all claims

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>510</td>
<td>1500</td>
<td>2550</td>
<td>3100</td>
<td>3100</td>
</tr>
<tr>
<td>1</td>
<td>510</td>
<td>1500</td>
<td>2550</td>
<td>3100</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>510</td>
<td>1500</td>
<td>2550</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>510</td>
<td>1500</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>510</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Small and large by latest information: Classification

- We see that there are no leftovers and no double counting (at any point in time each part is either red or green).
- The classification depends on the estimation date (the colour of each block may change if we look at it one period later). Therefore, accident periods are not comparable, i.e. we don’t have consistency over time.
- The separation is not continuous with respect to the threshold.
First idea is to look at the latest information we have about each claim.
5 Separation of small and large claims

5.2 How to separate small from large claims

5.2.1 Small and large by latest information

Small and large by latest information: Projection

<table>
<thead>
<tr>
<th>small</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>large</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1300</td>
<td>900</td>
<td>400</td>
<td>-450</td>
<td>0</td>
<td>0</td>
<td>500</td>
<td>300</td>
<td>100</td>
<td>50</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1300</td>
<td>900</td>
<td>400</td>
<td>-450</td>
<td>0</td>
<td>1</td>
<td>500</td>
<td>300</td>
<td>100</td>
<td>50</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1000</td>
<td>500</td>
<td>300</td>
<td>-312</td>
<td>0</td>
<td>2</td>
<td>800</td>
<td>700</td>
<td>200</td>
<td>94</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>1000</td>
<td>500</td>
<td>280</td>
<td>-308</td>
<td>0</td>
<td>3</td>
<td>800</td>
<td>700</td>
<td>194</td>
<td>94</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>1300</td>
<td>791</td>
<td>390</td>
<td>-429</td>
<td>0</td>
<td>4</td>
<td>500</td>
<td>385</td>
<td>114</td>
<td>55</td>
<td>0</td>
</tr>
<tr>
<td>(\hat{f}_k)</td>
<td>0.61</td>
<td>0.19</td>
<td>-0.17</td>
<td>0.00</td>
<td>(\hat{f}_k)</td>
<td>0.77</td>
<td>0.13</td>
<td>0.06</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- We see again that accident periods are not comparable, i.e. we don’t have consistency over time.
- We have huge amounts in late development periods within the small triangle, which usually makes projections less stable. The reason for those amounts is the reclassification of a large claim as small in development period 3. 😞 😞
Stochastic Reserving

- Separation of small and large claims
- How to separate small from large claims

We see again that accident periods are not comparable, i.e., we don’t have consistency over time.

We have huge amounts in late development periods within the small triangle, which usually makes projections less stable. The reason for these amounts is the reclassification of a large claim as small in development period 3.
Small and large by latest information: Results

<table>
<thead>
<tr>
<th>AP</th>
<th>expected results</th>
<th>estimated results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ultimate reserves</td>
<td>ultimate reserves</td>
</tr>
<tr>
<td>0</td>
<td>3100 0</td>
<td>3100 0</td>
</tr>
<tr>
<td>1</td>
<td>3100 0</td>
<td>3100 0</td>
</tr>
<tr>
<td>2</td>
<td>3100 550</td>
<td>3283 733</td>
</tr>
<tr>
<td>3</td>
<td>3100 1600</td>
<td>3259 1759</td>
</tr>
<tr>
<td>4</td>
<td>3100 2590</td>
<td>3106 2596</td>
</tr>
<tr>
<td>total</td>
<td>15500 4740</td>
<td>15848 5088</td>
</tr>
</tbody>
</table>

- Under- and overestimation.
- More than 75% of the reserves belong to the large triangle, which is usually less stable.

Conclusion (pros: 1 🎉 versus cons: 1 🙁 and 4 🙁 🙁 🙁)

Do not use the separation method ‘small and large by latest information’ for the estimation of reserves.
### Stochastic Reserving

**Separation of small and large claims**

**How to separate small from large claims**

<table>
<thead>
<tr>
<th>AP</th>
<th>expected results</th>
<th>ultimate reserves</th>
<th>ultimate reserves small res.</th>
<th>large res.</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>3100</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>3100</td>
<td>0</td>
<td>3100</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>3100</td>
<td>550</td>
<td>3283</td>
<td>733</td>
</tr>
<tr>
<td>3</td>
<td>3100</td>
<td>1600</td>
<td>3259</td>
<td>1759</td>
</tr>
<tr>
<td>4</td>
<td>3100</td>
<td>2590</td>
<td>3106</td>
<td>2596</td>
</tr>
<tr>
<td>total</td>
<td>15500</td>
<td>4740</td>
<td>15848</td>
<td>5088</td>
</tr>
</tbody>
</table>

- Under- and overestimation.
- More than 75% of the reserves belong to the large triangle, which is usually less stable.

**Conclusion (pros: 1, versus cons: 2, 3, and 4)**

Do not use the separation method ‘small and large by latest information’ for the estimation of reserves.
Ever and never large by latest information: Classification

\[ \text{claim is large at time } k \iff \max_{j \leq I} (X_j) > \text{threshold} \]

- We see that there are no leftovers and no double counting (at any point in time each part is either red or green).
- The classification depends on the estimation date (the colour of each block may change if we look at it one period later). Therefore, accident periods are not comparable, i.e. we don't have consistency over time.
- The separation is not continuous with respect to the threshold.
In order to get smoother triangles we have to avoid the reclassification of large claims as small. One way to do so is to take all claims as large which have exceeded the threshold at least once.
5. Separation of small and large claims

5.2 How to separate small from large claims

5.2.2 Ever and never large by latest information

**Ever and never large by latest information: Projection**

<table>
<thead>
<tr>
<th>small</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>large</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1000</td>
<td>500</td>
<td>300</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>800</td>
<td>700</td>
<td>200</td>
<td>-400</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1000</td>
<td>500</td>
<td>300</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>800</td>
<td>700</td>
<td>200</td>
<td>-400</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1000</td>
<td>500</td>
<td>300</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>800</td>
<td>700</td>
<td>200</td>
<td>-400</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>1000</td>
<td>500</td>
<td>300</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>800</td>
<td>700</td>
<td>200</td>
<td>-400</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>1300</td>
<td>650</td>
<td>390</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>500</td>
<td>438</td>
<td>125</td>
<td>-250</td>
<td>0</td>
</tr>
</tbody>
</table>

\[ \hat{f}_k \]

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.2</td>
<td>0.00</td>
<td>0.00</td>
<td>0.88</td>
<td>0.13</td>
<td>-0.24</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- We see again that accident periods are not comparable, i.e. we don’t have consistency over time.
- The triangle of small claims is much smoother. 😊 😊
Stochastic Reserving

- Separation of small and large claims
- How to separate small from large claims

We see again that accident periods are not comparable, i.e. we don’t have consistency over time.

The triangle of small claims is much smoother.
5 Separation of small and large claims

5.2 How to separate small from large claims

5.2.2 Ever and never large by latest information

---

**Ever and never large by latest information: Results**

<table>
<thead>
<tr>
<th>AP</th>
<th>expected results</th>
<th>estimated results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ultimate reserves</td>
<td>ultimate reserves</td>
</tr>
<tr>
<td>0</td>
<td>3100 0</td>
<td>3100 0</td>
</tr>
<tr>
<td>1</td>
<td>3100 0</td>
<td>3100 0</td>
</tr>
<tr>
<td>2</td>
<td>3100 550</td>
<td>3100 550</td>
</tr>
<tr>
<td>3</td>
<td>3100 1600</td>
<td>3100 1600</td>
</tr>
<tr>
<td>4</td>
<td>3100 2590</td>
<td>3153 2642</td>
</tr>
<tr>
<td>total</td>
<td>15500 4740</td>
<td>15553 4792</td>
</tr>
</tbody>
</table>

- Under- and overestimation.
- More than 50% of the reserves belong to the large triangle, which is usually less stable.

**Conclusion (pros: 1 😊😊 and 1 😊 versus cons: 1 😖 and 3 😖 😖)***

Do not use the separation method ‘ever and never large by latest information’ for the estimation of reserves.
Stochastic Reserving

Separation of small and large claims

How to separate small from large claims

<table>
<thead>
<tr>
<th>AP</th>
<th>expected results</th>
<th>ultimate reserves</th>
<th>estimated results</th>
<th>small res.</th>
<th>large res.</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>3100</td>
<td>3100</td>
<td>3100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>3100</td>
<td>1600</td>
<td>3100</td>
<td>550</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>3100</td>
<td>1600</td>
<td>3100</td>
<td>1600</td>
<td>500</td>
</tr>
<tr>
<td>4</td>
<td>3100</td>
<td>2590</td>
<td>3153</td>
<td>2642</td>
<td>1830</td>
</tr>
<tr>
<td>total</td>
<td>15500</td>
<td>4740</td>
<td>15553</td>
<td>4792</td>
<td>2330</td>
</tr>
</tbody>
</table>

- Under- and overestimation.
- More than 50% of the reserves belong to the large triangle, which is usually less stable.

Conclusion (pros: 1 and 1 versus cons: 1 and 3)
Do not use the separation method ‘ever and never large by latest information’ for the estimation of reserves.
Small and large now: Classification

claim is large at time $k \iff X_k > \text{threshold}$

- We see that there are no leftovers and no double counting (each part is either red or green).
- The classification does not depend on the estimation date (the colour of each block does not change if we look at it one period later). Therefore, accident periods are comparable, i.e. we have consistency over time.
- The separation is not continuous with respect to the threshold.
Stochastic Reserving

- Separation of small and large claims
- How to separate small from large claims

The separation method ‘ever and never large by latest information’ may stabilise the triangles. But we still have inconsistent accident periods and therefore an under- or overestimation of reserves. In order to get consistent accident periods we could consider a claim as large at time $k$ if it exceeds the threshold at this time.
### Small and large now: Projection

<table>
<thead>
<tr>
<th>small</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1300</td>
<td>200</td>
<td>300</td>
<td>350</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1300</td>
<td>200</td>
<td>300</td>
<td>350</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1300</td>
<td>200</td>
<td>300</td>
<td>350</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>1300</td>
<td>200</td>
<td>300</td>
<td>350</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>1300</td>
<td>200</td>
<td>300</td>
<td>350</td>
<td>0</td>
</tr>
</tbody>
</table>

\[ \hat{f}_k \] | 0.15 | 0.20 | 0.19 | 0.00 |

<table>
<thead>
<tr>
<th>large</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>500</td>
<td>1000</td>
<td>200</td>
<td>-750</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>500</td>
<td>1000</td>
<td>200</td>
<td>-750</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>500</td>
<td>1000</td>
<td>200</td>
<td>-750</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>500</td>
<td>1000</td>
<td>200</td>
<td>-750</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>500</td>
<td>1000</td>
<td>200</td>
<td>-750</td>
<td>0</td>
</tr>
</tbody>
</table>

\[ \hat{f}_k \] | 2.00 | 0.13 | -0.44 | 0.00 |

- We see again that accident periods are comparable, i.e. we have consistency over time.
- We have huge amounts in late development periods, which usually makes projections less stable. The reason for those amounts is the reclassification of a large claim as small. 😞 😞
Stochastic Reserving

Separation of small and large claims

How to separate small from large claims

- We see again that accident periods are comparable, i.e. we have consistency over time.
- We have huge amounts in late development periods, which usually makes projections less stable. The reason for those amounts is the reclassification of a large claim as small.
### Small and large now: Results

<table>
<thead>
<tr>
<th></th>
<th>expected results</th>
<th>estimated results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ultimate reserves</td>
<td>ultimate reserves</td>
</tr>
<tr>
<td>0</td>
<td>3100 0</td>
<td>3100 0</td>
</tr>
<tr>
<td>1</td>
<td>3100 0</td>
<td>3100 0</td>
</tr>
<tr>
<td>2</td>
<td>3100 550</td>
<td>3100 550</td>
</tr>
<tr>
<td>3</td>
<td>3100 1600</td>
<td>3100 1600</td>
</tr>
<tr>
<td>4</td>
<td>3100 2590</td>
<td>3100 2590</td>
</tr>
<tr>
<td>total</td>
<td>15500 4740</td>
<td>15500 4740</td>
</tr>
</tbody>
</table>

- No systematic under- or overestimation.
- Still 40% of the reserves belong to the large triangle, which is usually less stable.

**Conclusion (pros: 3 😊 versus cons: 2 😞 and 1 😞 😞 😞)**

Do not use the separation method ‘small and large now’ for the estimation of reserves.
Stochastic Reserving

Separation of small and large claims

How to separate small from large claims

<table>
<thead>
<tr>
<th>AP</th>
<th>expected results</th>
<th>ultimate reserves</th>
<th>estimated results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>ultimate reserve</td>
<td>small res. large res.</td>
</tr>
<tr>
<td>0</td>
<td>3100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>3100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>3100</td>
<td>550</td>
<td>200</td>
</tr>
<tr>
<td>3</td>
<td>3100</td>
<td>1600</td>
<td>550 850</td>
</tr>
<tr>
<td>4</td>
<td>3100</td>
<td>2590</td>
<td>1640 950</td>
</tr>
<tr>
<td>total</td>
<td>15500</td>
<td>4740</td>
<td>2840 1900</td>
</tr>
</tbody>
</table>

- No systematic under- or overestimation.
- Still 40% of the reserves belong to the large triangle, which is usually less stable.

Conclusion (pros: 3 versus cons: 2 and 1)

Do not use the separation method ‘small and large now’ for the estimation of reserves.
Ever and never large up to now: Classification

claim is large at time $k \iff \max_{j \leq k} (X_j) > \text{threshold}$

- We see that there are no leftovers and no double counting (each part is either red or green).
- The classification does not depend on the estimation date (the colour of each block does not change if we look at it one period later). Therefore, accident periods are comparable, i.e. we have consistency over time.
- The separation is not continuous with respect to the threshold.
Taking the separation method ‘large and small now’ we get consistent accident periods, but lose some stability of the projection. Therefore, lets try to combine the ‘large and small now’ with ‘ever and never large by latest information’. That means we consider a claim as large at time $k$ if it exceeded the threshold at least once up to time $k$. 

\[\text{claim is large at time } k \iff \max_{j \leq k} (X_j) > \text{threshold} \]
### Ever and never large up to now: Projection

<table>
<thead>
<tr>
<th>small</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>large</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1300</td>
<td>200</td>
<td>300</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>500</td>
<td>1000</td>
<td>200</td>
<td>-400</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1300</td>
<td>200</td>
<td>300</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>500</td>
<td>1000</td>
<td>200</td>
<td>-400</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1300</td>
<td>200</td>
<td>300</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>500</td>
<td>1000</td>
<td>200</td>
<td>-400</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>1300</td>
<td>200</td>
<td>300</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>500</td>
<td>1000</td>
<td>200</td>
<td>-400</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>1300</td>
<td>200</td>
<td>300</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>500</td>
<td>1000</td>
<td>200</td>
<td>-400</td>
<td>0</td>
</tr>
</tbody>
</table>

\[ \hat{f}_k \]

| 0.15 | 0.20 | 0.00 | 0.00 |
\[ \hat{f}_k \]

| 2.00 | 0.13 | 0.24 | 0.00 |

- We see again that accident periods are comparable, i.e. we have consistency over time.
- The triangle of small claims is much smoother, in particular for late development periods. But claims that will become large in the future may have huge changes in incurred losses or payments during the time where they are still small. 😞 😊
Stochastic Reserving

Separation of small and large claims

How to separate small from large claims

We see again that accident periods are comparable, i.e. we have consistency over time.

The triangle of small claims is much smoother, in particular for late development periods. But claims that will become large in the future may have huge changes in incurred losses or payments during the time where they are still small.
5 Separation of small and large claims

5.2 How to separate small from large claims

5.2.4 Ever and never large up to now (3/3)

Ever and never large up to now: Results

<table>
<thead>
<tr>
<th>AP</th>
<th>expected results ultimate reserves</th>
<th>estimated results ultimate reserves small res.</th>
<th>large res.</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>3100 0</td>
<td>3100 0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>3100 0</td>
<td>3100 0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>3100 550</td>
<td>3100 550</td>
<td>0 550</td>
</tr>
<tr>
<td>3</td>
<td>3100 1600</td>
<td>3100 1600</td>
<td>500 1100</td>
</tr>
<tr>
<td>4</td>
<td>3100 2590</td>
<td>3100 2590</td>
<td>1290 1300</td>
</tr>
<tr>
<td>total</td>
<td>15500 4740</td>
<td>15500 4740</td>
<td>1790 2950</td>
</tr>
</tbody>
</table>

- No systematic under- or overestimation. 😊
- More than 60% of the reserves belong to the large triangle, which is usually less stable. 😞 😞

Conclusion (pros: 4 😊 versus cons: 2 😞 and 1 😞 😞 😞)

If the threshold is chosen carefully, i.e. if not too much reserves are transferred to the large triangle, we can use the separation method ‘ever and never large up to now’ for the estimation of reserves.
Claims that will become large in the future may have huge changes in incurred losses or payments during the time where they are still small. Therefore, the triangle of small claims may be not so stable as expected.

In order to avoid this behaviour we have to take smaller threshold, which on the other side will transfer more reserves into the triangle of large claims.

---

<table>
<thead>
<tr>
<th></th>
<th>expected results</th>
<th>ultimate reserves</th>
<th>estimated results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ultimate reserves</td>
<td>small res.</td>
<td>large res.</td>
</tr>
<tr>
<td>0</td>
<td>3100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>3100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>3100 550</td>
<td>3100 550</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>3100 1600</td>
<td>3100 1600</td>
<td>500 1100</td>
</tr>
<tr>
<td>4</td>
<td>3100 2590</td>
<td>3100 2590</td>
<td>1290 1300</td>
</tr>
<tr>
<td>total</td>
<td>15500 4740</td>
<td>15500 4740</td>
<td>1790 2950</td>
</tr>
</tbody>
</table>

• No systematic under- or overestimation.
• More than 60% of the reserves belong to the large triangle, which is usually less stable.
Ever large up to now and never large by latest information

- If a claim has huge changes in payments or incurred losses before it exceeds the threshold the first time, it can disturb the triangle of small claims significantly.
- Therefore, ‘ever and never large up to now’ may not lead to smooth enough triangles of small claims and we would like to take all claims out that have exceed the threshold at least once, i.e. we would like to use ‘never large by latest information’.
- But as we have seen ‘ever and never large by latest information’ leads to not comparable accident periods and over- or underestimation of reserves.
- A compromise could be to put all claim that ‘have never been large by latest information’ into the triangle of small claims and all claims that ‘were ever large up to now’ into the triangle of large claims.
- Although this leads to not comparable accident periods within the triangle of small claims as well as leftovers, the corresponding systematic overestimation can often be controlled.
The separation method ‘ever and never large up to now’, which combined the two methods

- ever and never large by latest information
- small and large now

has good properties but may still leave a lot of reserves within the triangle of large claims. One way to get around this is to take the following method.
5 Separation of small and large claims

5.2 How to separate small from large claims

5.2.5 Ever large up to now and never large by latest information (2/4)

Ever large up to now and never large by latest information: Classification

- Claim is large at time $k \iff \max_{j \leq k} (X_j) > \text{threshold}$
- Claim is small at time $k \iff \max_{j \leq I} (X_j) \leq \text{threshold}$

- We have leftovers: Large claims are not counted until they get large for the first time.
- The classification of small claims depends on the estimation date (the colour of each block may change if we look at it one period later). Therefore, accident periods are not comparable, i.e. we don’t have consistency over time. The large triangle is consistent over time.
- The separation is not continuous with respect to the threshold.
Stochastic Reserving

Separation of small and large claims

How to separate small from large claims

We have leftovers: Large claims are not counted until they get large for the first time.

The classification of small claims depends on the estimation date (the colour of each block may change if we look at it one period later). Therefore, accident periods are not comparable, i.e. we don’t have consistency over time. The large triangle is consistent over time.

The separation is not continuous with respect to the threshold.

Ever large up to now and never large by latest information: Classification

Claim is large at time $k$ $\iff$ $\max_{j \leq k} (X_j) >$ threshold

Claim is small at time $k$ $\iff$ $\max_{j \leq k} (X_j) \leq$ threshold

Threshold

Inurred

Paid

2019-02-18

Stochastic Reserving

Separation of small and large claims

How to separate small from large claims
Ever large up to now and never large by latest information:

Projection

<table>
<thead>
<tr>
<th>small</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>large</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1000</td>
<td>500</td>
<td>300</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>500</td>
<td>1000</td>
<td>200</td>
<td>-400</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1000</td>
<td>500</td>
<td>300</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>500</td>
<td>1000</td>
<td>200</td>
<td>-400</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1000</td>
<td>500</td>
<td>300</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>500</td>
<td>1000</td>
<td>200</td>
<td>-400</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>1000</td>
<td>500</td>
<td>300</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>500</td>
<td>1000</td>
<td>200</td>
<td>-400</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>1300</td>
<td>650</td>
<td>390</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>500</td>
<td>1000</td>
<td>200</td>
<td>-400</td>
<td>0</td>
</tr>
</tbody>
</table>

\[ \hat{f}_k = \begin{bmatrix} 0.5 \\ 0.2 \\ 0.00 \\ 0.00 \end{bmatrix} \quad \hat{f}_k = \begin{bmatrix} 2.00 \\ 0.13 \\ 0.24 \\ 0.00 \end{bmatrix} \]

- We see again that accident periods of small claims are not comparable, i.e. we don’t have consistency over time.
- The inconsistency over time leads to a systematic overestimation, because the claims that are not yet large are projected within the small triangle as IBNeR and within the large triangle as IBNyR. Therefore, the overestimation equals

\[
540 = \underbrace{300}_{\text{size at time } k = 0} \cdot \underbrace{1.5 \cdot 1.2}_{\text{cumulative development factors}}
\]

- The triangle of small claims is much smoother. 😊 😊
Separation of small and large claims

How to separate small from large claims

We see again that accident periods of small claims are not comparable, i.e., we don’t have consistency over time. The inconsistency over time leads to a systematic overestimation, because the claims that are not yet large are projected within the small triangle as IBNeR and within the large triangle as IBNyR. Therefore, the overestimation equals

\[ s_{k0} = \sum_{i=0}^{k} \hat{f}_{ki} \]

where each \( \hat{f}_{ki} \) is the cumulative development factor.

The triangle of small claims is much smoother.
Ever large up to now and never large by latest information: 

Results

<table>
<thead>
<tr>
<th>AP</th>
<th>ultimate reserves</th>
<th>ultimate reserves</th>
<th>small res.</th>
<th>large res.</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>3100</td>
<td>3100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>3100</td>
<td>3100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>3100 550</td>
<td>3100 550</td>
<td>0</td>
<td>550</td>
</tr>
<tr>
<td>3</td>
<td>3100 1600</td>
<td>3100 1600</td>
<td>500</td>
<td>1100</td>
</tr>
<tr>
<td>4</td>
<td>3100 2590</td>
<td>3640 3130</td>
<td>1830</td>
<td>1300</td>
</tr>
<tr>
<td>total</td>
<td>15500</td>
<td>16040</td>
<td>2330</td>
<td>2950</td>
</tr>
</tbody>
</table>

- Systematic overestimation, which often can be controlled.
- More than 60% of the reserves belong to the large triangle, which is usually less stable. But, since the small triangle is much more stable, we could increase the threshold and therefore transfer more reserves to the small triangle.

Conclusion (pros: 1 😊😊 and 2 😊 versus cons: 5 😞)

If we can control the systematic overestimation the separation method ‘ever large up to now and never large by latest information’ can be used.
Stochastic Reserving

• Separation of small and large claims

• How to separate small from large claims

<table>
<thead>
<tr>
<th>APR</th>
<th>ultimate reserves</th>
<th>ultimate reserves</th>
<th>small reserves</th>
<th>large reserves</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1000</td>
<td>1000</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>550</td>
<td>3100</td>
<td>500</td>
<td>550</td>
</tr>
<tr>
<td>3</td>
<td>1600</td>
<td>3100</td>
<td>500</td>
<td>1100</td>
</tr>
<tr>
<td>4</td>
<td>2590</td>
<td>3640</td>
<td>1830</td>
<td>1300</td>
</tr>
<tr>
<td>total</td>
<td>15500</td>
<td>4740</td>
<td>2330</td>
<td>2950</td>
</tr>
</tbody>
</table>

- Systematic overestimation, which often can be controlled.
- More than 60% of the reserves belong to the large triangle, which is usually less stable.
- But, since the small triangle is much more stable, we could increase the threshold and therefore transfer more reserves to the small triangle.

Conclusion (pros: 1 and 2 versus cons: 5):
If we can control the systematic overestimation the separation method ‘ever large up to now and never large by latest information’ can be used.
Attritional and excess: Classification

Attritional part at time $k := \min(X_k, \text{threshold})$

Excess part at time $k := X_k - \min(X_k, \text{threshold})$

- We see that there are no leftovers and no double counting (each part is either red or green).
- The classification does not depend on the estimation date (the colour of each block does not change if we look at it one period later). Therefore, accident periods are comparable, i.e. we have consistency over time.
- The separation is continuous with respect to the threshold.
Another method of separation is to split up large claims into a normal (attritional) and an exceptional (excess) part.
5 Separation of small and large claims

5.2 How to separate small from large claims

5.2.6 Attritional and excess

### Attritional and excess: Projection

<table>
<thead>
<tr>
<th>attritional</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>excess</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1700</td>
<td>600</td>
<td>300</td>
<td>-50</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>600</td>
<td>200</td>
<td>-350</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1700</td>
<td>600</td>
<td>300</td>
<td>-50</td>
<td>0</td>
<td>1</td>
<td>100</td>
<td>600</td>
<td>200</td>
<td>-350</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1700</td>
<td>600</td>
<td>300</td>
<td>-50</td>
<td>0</td>
<td>2</td>
<td>100</td>
<td>600</td>
<td>200</td>
<td>-350</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>1700</td>
<td>600</td>
<td>300</td>
<td>-50</td>
<td>0</td>
<td>3</td>
<td>100</td>
<td>600</td>
<td>200</td>
<td>-350</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>1700</td>
<td>600</td>
<td>300</td>
<td>-50</td>
<td>0</td>
<td>4</td>
<td>100</td>
<td>600</td>
<td>200</td>
<td>-350</td>
<td>0</td>
</tr>
</tbody>
</table>

\[ \hat{f}_k = 0.35, 0.13, -0.02, 0.00 \]

\[ \hat{f}_k = 6.00, 0.29, -0.39, 0.00 \]

- We see again that accident periods are comparable, i.e. we have consistency over time.
- The triangle of small claims is much smoother, in particular for late development periods. But claims that will become large in the future may have huge changes in incurred losses or payments during the time where they are still small.
- But the triangle of large claims shows huge development. Therefore, most estimation methods will not work.
- One method that often works for the excess part is ECLRM with additional virtual case reserves \( R_{i,k}^{\text{add}} \) for not yet large claims:

\[
R_{i,k}^{\text{add}} := \left( \hat{N}_{i,J} - N_{i,k} \right) \cdot (\hat{m}_{i,j} - \text{threshold})
\]

number of claims that will become large after time \( k \)  mean ultimate of a large claim
Separation of small and large claims

How to separate small from large claims

<table>
<thead>
<tr>
<th>accident</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>excess</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1700</td>
<td>600</td>
<td>300</td>
<td>-50</td>
<td>0</td>
<td>1</td>
<td>100</td>
<td>500</td>
<td>200</td>
<td>-350</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1700</td>
<td>600</td>
<td>300</td>
<td>-50</td>
<td>0</td>
<td>2</td>
<td>100</td>
<td>500</td>
<td>200</td>
<td>-350</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1700</td>
<td>600</td>
<td>300</td>
<td>-50</td>
<td>0</td>
<td>3</td>
<td>100</td>
<td>500</td>
<td>200</td>
<td>-350</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>1700</td>
<td>600</td>
<td>300</td>
<td>-50</td>
<td>0</td>
<td>4</td>
<td>100</td>
<td>500</td>
<td>200</td>
<td>-350</td>
<td>0</td>
</tr>
</tbody>
</table>

We see again that accident periods are comparable, i.e. we have consistency over time.

- The triangle of small claims is much smoother, in particular for late development periods. But claims that will become large in the future may have huge changes in incurred losses or payments during the time where they are still small.
- But the triangle of large claims shows huge development. Therefore, most estimation methods will not work.

- One method that often works for the excess part is ECLRM with additional virtual case reserves \( R_{add} \) for not yet large claims:

\[
R_{add} = \left( N_{i,J} - N_{i,k} \right) \cdot \frac{\hat{m}_{i,j} - \text{threshold}}{2019-02-18}
\]
Attritional and excess: Results

<table>
<thead>
<tr>
<th>AP</th>
<th>expected results</th>
<th>estimated results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ultimate reserves</td>
<td>ultimate reserves</td>
</tr>
<tr>
<td>0</td>
<td>3100</td>
<td>3100</td>
</tr>
<tr>
<td>1</td>
<td>3100</td>
<td>3100</td>
</tr>
<tr>
<td>2</td>
<td>3100</td>
<td>3100</td>
</tr>
<tr>
<td>3</td>
<td>3100</td>
<td>3100</td>
</tr>
<tr>
<td>4</td>
<td>3100</td>
<td>3100</td>
</tr>
<tr>
<td>total</td>
<td>15500</td>
<td>15500</td>
</tr>
</tbody>
</table>

• No systematic under- or overestimation. 😊
• Less than 33% of the reserves belong to the large triangle, which is usually less stable. 😊

Conclusion (pros: 6 😊 versus cons: 2 😞)

Usually, I prefer the separation method ‘attritional and excess’. But we have to be very careful with the projection of the excess part.
Stochastic Reserving

Separation of small and large claims

How to separate small from large claims

### Attritional and excess: Results

<table>
<thead>
<tr>
<th>AP</th>
<th>expected results</th>
<th>estimated results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ultimate reserves</td>
<td>attritional res.</td>
</tr>
<tr>
<td>0</td>
<td>3500</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>3500</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>3500 550</td>
<td>3500 550</td>
</tr>
<tr>
<td>3</td>
<td>3500 1650</td>
<td>3500 1600</td>
</tr>
<tr>
<td>4</td>
<td>3500 2150</td>
<td>3500 2500</td>
</tr>
<tr>
<td>total</td>
<td>15500 4740</td>
<td>15500 4740</td>
</tr>
</tbody>
</table>

- No systematic under- or overestimation.
- Less than 33% of the reserves belong to the large triangle, which is usually less stable.

### Conclusion (pros: 6 versus cons: 2)

Usually, I prefer the separation method 'attritional and excess'. But we have to be very careful with the projection of the excess part.
5.2 How to separate small from large claims

### 5.2.7 Separation methods summary

<table>
<thead>
<tr>
<th>name</th>
<th>definition of large (th:=threshold)</th>
<th>leftovers or double</th>
<th>consistent accident periods</th>
<th>continuous in threshold</th>
<th>stable projections</th>
<th>under- or overestimation</th>
<th>huge reserves for large claims</th>
</tr>
</thead>
<tbody>
<tr>
<td>large and small by latest information</td>
<td>large at time $k$ $\iff X_I &gt; \text{th}$</td>
<td>☑️ ☑️ ☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☑️ ☑️ ☐</td>
<td>☑️ ☑️ ☑️</td>
</tr>
<tr>
<td>ever and never large by latest information</td>
<td>large at time $k$ $\iff \max_{j \le I}(X_j) &gt; \text{th}$</td>
<td>☑️ ☑️ ☑️</td>
<td>☐</td>
<td>☐</td>
<td>☑️</td>
<td>☐</td>
<td>☑️ ☑️ ☑️</td>
</tr>
<tr>
<td>small and large now</td>
<td>large at time $k$ $\iff X_k &gt; \text{th}$</td>
<td>☑️ ☑️ ☑️</td>
<td>☐</td>
<td>☐</td>
<td>☑️</td>
<td>☐</td>
<td>☑️ ☑️ ☑️</td>
</tr>
<tr>
<td>ever and never large up to now</td>
<td>large at time $k$ $\iff \max_{j \le k}(X_j) &gt; \text{th}$</td>
<td>☑️ ☑️ ☑️</td>
<td>☐</td>
<td>☐</td>
<td>☑️</td>
<td>☐</td>
<td>☑️ ☑️ ☑️</td>
</tr>
<tr>
<td>ever large up to now and never large by latest information</td>
<td>large at time $k$ $\iff \max_{j \le k}(X_j) &gt; \text{th}$</td>
<td>☐ ☐ ☐</td>
<td>☐</td>
<td>☐</td>
<td>☑️ ☑️ ☑️</td>
<td>☑️ ☑️ ☑️</td>
<td>☑️ ☑️ ☑️</td>
</tr>
<tr>
<td>attritional and excess</td>
<td>attritional part $:= \min(X_k, \text{th})$ excess part $:= X_k - \min(X_k, \text{th})$</td>
<td>☑️ ☑️ ☑️</td>
<td>☐</td>
<td>☐</td>
<td>☑️</td>
<td>☐</td>
<td>☑️ ☑️ ☑️</td>
</tr>
</tbody>
</table>
Stochastic Reserving

Separation of small and large claims

How to separate small from large claims

The motivation story (from up to down of the table):

- first idea is to take latest information
- try to get smoother triangles
- try to get consistent accident periods
- try to combine the last two
- try to reduce the amount of reserves within the triangle of large claims
- split up each claims in a ‘good’ and a ‘bad’ part

I prefer the last two separation methods. But under special circumstances, for instance lack of data, it is possible that even the first one is the most suitable method.
5 Separation of small and large claims

5.3 Estimation methods for small and large claims

Estimation methods for small (attritional) claims

- There are no general restrictions to the reserving methods used for small (or attritional) claims.
- Depending on the separation method it might be better to use the paid triangle instead of the incurred triangle, in particular if early development periods are disturbed by future large claims, which usually does not affect the payments as much as the incurred losses.

Estimation methods for large (excess) claims

- Often we have to be very careful with standard methods like CLM and ECLRM, in particular if we don’t have any large claim in early development periods.
- It is not unusual that the triangles of large claims are so unstable that we have to fall back on expert judgement in order to estimate the reserves.

Estimate overall uncertainties

- One way to estimate uncertainties is to couple the estimations of small and large claims, for instance by LSRMs.
- In practice, if we use expert judgement, it is often better to estimate uncertainties on an aggregated level.
Estimation methods for small (attritional) claims

- There are no general restrictions to the reserving methods used for small (or attritional) claims.
- Depending on the separation method it might be better to use the paid triangle instead of the incurred triangle, in particular if early development periods are disturbed by future large claims, which usually does not affect the payments as much as the incurred losses.

Estimation methods for large (excess) claims

- Often we have to be very careful with standard methods like CLM and ECLRM, in particular if we don’t have any large claim in early development periods.
- It is not unusual that the triangles of large claims are so unstable that we have to fall back on expert judgement in order to estimate the reserves.

Estimate overall uncertainties

- One way to estimate uncertainties is to couple the estimations of small and large claims, for instance by LSRMs.
- In practice, if we use expert judgement, it is often better to estimate uncertainties on an aggregated level.
Bifurcation of large and small losses: Basic idea

The separation methods we have seen up to now do not look at transition of claims from the triangle of small claims to the one of large claims. We now want to try to model these transitions and will follow the notation of U. Riegel [22]. The basic idea is to look separately at:

- the development of small claims conditioned given they are still small at the next period.
- the development of large claims without claims that just exceed the threshold the first time.
- the number of new large claims and their mean expected loss at the time they get large.
Bifurcation of large and small losses: Basic idea

The separation methods we have seen up to now do not look at transitions of claims from the triangle of small claims to the one of large claims. We now want to try to model these transitions and will follow the notation of U. Riegel [22]. The basic idea is to look separately at:

- the development of small claims conditioned given they are still small at the next period.
- the development of large claims without claims that just exceed the threshold the first time.
- the number of new large claims and their mean expected loss at the time they get large.
Bifurcation of large and small losses: Notations

- \( P_{i,k} \) and \( I_{i,k} \) denote the total cumulative payments and incurred losses of all claims of accident period \( i \) at development period \( k \).
- We call a claim large at time \( k \) if its incurred loss did exceed the threshold at least once up to time \( k \) (ever large up to now).
- With \( N_{i,k} \) we denote the number of large claims of accident period \( i \) up to development period \( k \).
- We denote by \( X_{i,\nu,k}^I \) and \( X_{i,\nu,k}^P \) the incurred loss and the cumulative payments, respectively, of the \( \nu \)-th large claim of accident period \( i \) at development periods \( k \).
- \( L_{i,k}^{(j)} := \sum_{\nu=1}^{N_{i,j}} X_{i,\nu,k}^I \) denotes the incurred losses at development period \( k \) of all up to time \( j \) large claims of accident period \( i \).
- \( A_{i,k}^{(j)} := P_{i,k} - \sum_{\nu=1}^{N_{i,j}} X_{i,\nu,k}^P \) are the cumulative payments at development period \( k \) of all claims that are still small at time \( j \).
- The information of incurred losses and payments of small and large claims as well as the individual information of already large claims is denoted by

\[
B_{i,k} := \sigma \left\{ P_{i,j}, I_{i,j}, X_{i,\nu,j}^P, X_{i,\nu,j}^I : j \leq k, \nu \leq N_{i,k} \right\}.
\]
Bifurcation of large and small losses: Notations

- $P_{i,k}$ and $I_{i,k}$ denote the total cumulative payments and incurred losses of all claims of accident period $i$ at development period $k$.
- We call a claim large at time $k$ if its incurred loss did exceed the threshold at least once up to time $k$ (ever large up to now).
- With $N_{i,k}$ we denote the number of large claims of accident period $i$ up to development period $k$.
- We denote by $X^I_{i,ν,k}$ and $X^P_{i,ν,k}$ the incurred loss and the cumulative payments, respectively, of the $ν$-th large claim of accident period $i$ at development period $k$.
- $L^I_{i,k} := \sum_{N_{i,j}ν=1} X^I_{i,ν,k}$ denotes the incurred losses at development period $k$ of all up to time $j$ large claims of accident period $i$.
- $A^P_{i,k} := P_{i,k} - \sum_{N_{i,j}ν=1} X^P_{i,ν,k}$ are the cumulative payments at development period $k$ of all claims that are still small at time $j$.
- The information of incurred losses and payments of small and large claims as well as the individual information of already large claims is denoted by

$$R_{i,k} := \sigma \{ P_{i,j}, I^P_{i,j}, X^I_{i,j}, X^P_{i,j} ; j \leq k, ν \leq N_{i,k} \}$$
5 Separation of small and large claims

5.4 Modelling the transition from small to large (3/4)

Bifurcation of large and small losses: Model (1 of 2)

1. Accident periods as well as individual claims are independent.
2. The number of large claims develop according to CLM, i.e.

\[ \mathbb{E}[N_{i,k+1}|B_{i,k}] = n_k N_{i,k}. \]

3. The cumulative payments of small claims as long as they stay small develop according to CLM, i.e.

\[ \mathbb{E}\left[ A_{i,k+1}^{(k+1)} | B_{i,k}, A_{i,k}^{(k+1)} \right] = a_k A_{i,k}^{(k+1)}. \]

4. The incurred losses of already large claims develop according to CLM, i.e.

\[ \mathbb{E}\left[ L_{i,k+1}^{(k)} | B_{i,k} \right] = l_k L_{i,k}^{(k)}. \]

5. Claims that just became large have a mean incurred loss of \( x_{k+1}^I \) and had mean cumulative payments \( x_k^P \) just before they got large, i.e.

\[ \mathbb{E}[X_{i,\nu,k+1}^I | B_{i,k}] = x_{k+1}^I \quad \text{and} \quad \mathbb{E}[X_{i,\nu,k}^P | B_{i,k}] = x_k^P, \quad \text{for } N_{i,k} < \nu \leq N_{i,k+1}. \]

6. Assumptions on covariances.
Stochastic Reserving

- Separation of small and large claims
- Modelling the transition from small to large

1. We could use other LSRMs instead of CLM. But if so we may have to adapt the covariance conditions and the calculations may become even more complicated.
2. The use of cumulative payments for the small claims is due to the German marked, where, because of the local statutory regulations (HGB), the history of incurred losses is often not of high quality.
3. Except for the additional conditioning for small claims and the wrong upper index for large claims the formulas are almost the same as for LSRMs.
Bifurcation of large and small losses: Model (2 of 2)

We can rewrite the expectations as follows:

2. \( E[N_{i,k+1} | \mathcal{B}_{i,k}] = n_k N_{i,k} \).

3. \( E[A_{i,k+1}^{(k+1)} | \mathcal{B}_{i,k}] = a_k A_{i,k}^{(k)} - \frac{a_k(n_k - 1)x_k P N_{i,k}}{\text{large claims right before becoming large}} \).

4. \( E[L_{i,k+1}^{(k+1)} | \mathcal{B}_{i,k}] = l_k L_{i,k}^{(k)} + \frac{(n_k - 1)x_{k+1} I N_{i,k}}{\text{claims that just have become large}} \).

These formulas look like a LSRM but with up to two development factors per claim property.

Therefore, the same techniques will work and we can derive estimators for the ultimate outcome and for uncertainties.
Bifurcation of large and small losses: Model (2 of 2)

We can rewrite the expectations as follows:

2. \[ E[N_{i,k+1}|B_{i,k}] = n_k N_{i,k} \]

3. \[ E[A_{i,k+1}^{(k+1)}|B_{i,k}, A_k^{(k+1)}|B_{i,k}] = a_k E[A_k^{(k+1)}|B_{i,k}] \]
   \[ = a_k \left( A_{i,k}^{(k)} - E\left[ \sum_{\nu=N_{i,k}+1}^{N_{i,k+1}} X_{i,\nu,k}^P | B_{i,k} \right] \right) \]
   \[ = a_k A_{i,k}^{(k)} - a_k (n_k - 1) x_k^P N_{i,k} \]

4. \[ E[L_{i,k+1}^{(k+1)}|B_{i,k}] = E[L_{i,k+1}^{(k)} + \sum_{\nu=N_{i,k}+1}^{N_{i,k+1}} X_{i,\nu,k+1}^I | B_{i,k}] \]
   \[ = l_k L_{i,k}^{(k)} + E[N_{i,k+1} - N_{i,k} | B_{i,k}] E[X_{i,N_{i,k+1},k} | B_{i,k}] \]
   \[ = l_k L_{i,k}^{(k)} + (n_k - 1) x_k^I N_{i,k} \]
Literature

[23] Ulrich Riegel.
A Bifurcation Approach for Attritional and Large Losses in Chain Ladder Calculations.
Stochastic Reserving

Separation of small and large claims

Literature

[28] Ulrich Riegel.
A Bifurcation Approach for Attritional and Large Losses in Chain Ladder Calculations.
Stochastic Reserving

Lecture 10

Poisson-Model

René Dahms

ETH Zurich, Spring 2019

8 May 2019

(Last update: 18 February 2019)
6 Poisson-Model
6.1 Modelling the number of reported claims
6.2 Projection of the future outcome
6.3 Ultimate uncertainty of the Poisson-Model
6.4 Generalised linear models and reserving
6.5 Literature
Lecture 10: Table of contents

6 Poisson Model
6.1 Modelling the number of reported claims
6.2 Projection of the future outcome
6.3 Ultimate uncertainty of the Poisson Model
6.4 Generalised linear models and reserving
6.5 Literature
Number of occurred claims

- Assume that for each policy a claim occurs during the year with some probability \( p \in (0, 1) \), that we have at most one claim per policy and that claims are independent.

- Then the number of during the year occurred claims \( N \) is Binomial-distributed with parameter \( p \) and \( R \), where the later represents the number of policies, i.e.

\[
P(N = n) = \binom{R}{n} p^n (1 - p)^{R-n} \approx \frac{\mu^n}{n!} e^{-\mu}, \quad \text{with } \mu = Rp
\]

for small \( p \)

- Therefore, we could assume that the number of during a year occurred claims is Poisson-distributed.

- Similar arguments can be applied with the number of during a year reported claims.
• Assume that for each policy a claim occurs during the year with some probability \( p \in (0, 1) \), that we have at most one claim per policy and that claims are independent.

• Then the number of during the year occurred claims \( N \) is Binomial-distributed with parameter \( p \) and \( R \), where the later represents the number of policies, i.e.

\[
P(N = n) = \binom{R}{n} p^n (1 - p)^{R-n} = \frac{R^n}{n!} p^n e^{-\mu}, \quad \text{with } \mu = Rp,
\]

for small \( p \).

• Therefore, we could assume that the number of during a year occurred claims is Poisson-distributed.

• Similar arguments can be applied with the number of during a year reported claims.
Assumption 6.A (Poisson-Model)

Assume that there are parameters $\mu_0, \ldots, \mu_I > 0$ and $\gamma_0, \ldots, \gamma_J > 0$ such that

i) $S_{i,k}$ are independent Poisson-distributed random variables with

$$\mathbb{E}[S_{i,k}] = \gamma_k \mu_i.$$ 

ii) $\sum_{k=0}^J \gamma_k = 1.$

Remark 6.1

- The restriction on $S_{i,k}$ to be an integer is not so restrictive at all. Even for payments we can always argue that they are a multiple of one Cent.
- The Poisson-Model cannot deal with negative claim properties $S_{i,k}$ which is very restrictive, in particular for incurred losses.
- The assumption of independent claim properties $S_{i,k}$ even within the same accident period is also very restrictive.
- The Poisson-Model can deal with incomplete triangles, for which some upper left part is missing.
- In the Poisson-Model we always have $\text{Var}[S_{i,k}] = \mathbb{E}[S_{i,k}] = \gamma_k \mu_i.$
Assumption 6.A (Poisson-Model)
Assume that there are parameters $\mu_0, \ldots, \mu_I > 0$ and $\gamma_0, \ldots, \gamma_J > 0$ such that
1) $S_{i,k}$ are independent Poisson-distributed random variables with $E[S_{i,k}] = \gamma_k \mu_i$.
2) $\sum_{k=0}^{J} \gamma_k = 1$.

Remark 6.1
- The restriction on $S_{i,k}$ to be an integer is not so restrictive at all. Even for payments we can always argue that they are a multiple of one Cent.
- The Poisson-Model cannot deal with negative claim properties $S_{i,k}$ which is very restrictive, in particular for incurred losses.
- The assumption of independent claim properties $S_{i,k}$ even within the same accident period is also very restrictive.
- The Poisson-Model can deal with incomplete triangles, for which some upper left part is missing.
- In the Poisson-Model we always have $\text{Var}(S_{i,k}) = E[S_{i,k}] = \gamma_k \mu_i$. 

Stochastic Reserving

Poisson-Model

Modelling the number of reported claims
Parameters of the Poisson-Model

Since

$$E[C_{i,J}] = \sum_{k=0}^{J} E[S_{i,k}] = \sum_{k=0}^{J} \mu_i \gamma_k = \mu_i$$

the parameter

$$\mu_i$$ represents the expected ultimate outcome of accident period $i$ and

$$\gamma_k$$ represents the expected fraction of the ultimate outcome that have manifested (or will manifest) itself during development period $k$ (reporting or cashflow pattern).
Parameters of the Poisson-Model

Since

\[ E(C_{i,J}) = \sum_{k=0}^{J} E(S_{i,k}) = \sum_{k=0}^{J} \mu_i \gamma_k = \mu_i \]

the parameter \( \mu_i \) represents the expected ultimate outcome of accident period \( i \) and
\( \gamma_k \) represents the expected fraction of the ultimate outcome that have manifested (or will manifest) itself during development period \( k \) (reporting or cashflow pattern).
6 Poisson-Model

6.2 Projection of the future outcome

Probability of the observed triangle

\[ P \left( (S_{i,k})_{i+k \leq I} = (x_{i+k})_{i+k \leq I} \right) = \prod_{i+k \leq I} \frac{(\mu_i \gamma_k)^{x_{i,k}}}{x_{i,k}!} e^{-\mu_i \gamma_k}. \]

Maximum likelihood (ML) for the Poisson-Model

The maximum likelihood estimators for the parameters are those \( \hat{\mu}_i \) and \( \hat{\gamma}_k \) for which the probability of the observed triangle is maximal. In order to get shorter formulas we will maximize the logarithm of the probability. Therefore, we set its partial derivatives with respect to each parameter to zero and try to solve the resulting system of linear equations:

\[ 0 = \frac{\partial \log P \left( (S_{i,k})_{i+k \leq I} \right)}{\partial \mu_i} = (I-i)^{\wedge} J \sum_{k=0}^{I} \frac{S_{i,k}}{\mu_i} - \gamma_k \iff \mu_i \sum_{k=0}^{I} \gamma_k = \sum_{k=0}^{I} S_{i,k} = C_{i,(I-i)^{\wedge} J} \]

\[ 0 = \frac{\partial \log P \left( (S_{i,k})_{i+k \leq I} \right)}{\partial \gamma_k} = \sum_{i=0}^{I-k} \frac{S_{i,k}}{\gamma_k} - \mu_i \iff \gamma_k \sum_{i=0}^{I-k} \mu_i = \sum_{i=0}^{I-k} S_{i,k}. \quad (6.1) \]

We denote the solution (if it exists) by \( \hat{\mu}_i \) and \( \hat{\gamma}_k \).
One can prove that if the observed data are not too strange then there exists an unique solution of (6.1), which represents a maximum. An example for ‘too strange’ is $S_{i,k} = 0$ for all observed accident and development periods.
Estimator 6.2 (for the future outcome within the Poisson-Model)

\[ \hat{S}_{i,k}^{\text{Poi}} := \hat{E}[S_{i,k}] := \hat{\mu}_i \hat{\gamma}_k \]

\[ \hat{C}_{i,J}^{\text{Poi}} := \hat{E}[C_{i,J}\mid D^I] := C_{i,I-i} + \sum_{k=I-i+1}^{J} \hat{S}_{i,k}^{\text{Poi}}. \]

Theorem 6.3 (Poisson-Model vs. Chain-Ladder method)

Assume that there exists an unique positive solution of (6.1). Then

\[ \hat{S}_{i,k}^{\text{Poi}} = \hat{S}_{i,k}^{\text{CLM}}, \]

where \( \hat{S}_{i,k}^{\text{CLM}} \) denotes the Chain-Ladder-projection corresponding to the variance minimizing weights.
Stochastic Reserving

\[ \text{Poisson-Model} \]

Projection of the future outcome

Lemma

\[
\begin{align*}
I - k & \quad \sum_{i=0}^{I-k} C_{i,k} = \sum_{i=0}^{I-k} \hat{\mu}_i \sum_{j=0}^{k} \hat{\gamma}_j \\
I - k & \quad \sum_{i=0}^{I-k} C_{i,k-1} = \sum_{i=0}^{I-k} (C_{i,k} - S_{i,k}) = \sum_{i=0}^{I-k} \hat{\mu}_i \sum_{j=0}^{k-1} \hat{\gamma}_j,
\end{align*}\]

for all \( i + k \leq I \)

Proof of the above lemma (by induction):

Start with \( k = J \)

\[
\begin{align*}
\sum_{i=0}^{I-J} C_{i,J} & = \sum_{i=0}^{I-J} \sum_{j=0}^{J} S_{i,j} = \sum_{i=0}^{I-J} \hat{\mu}_i \sum_{j=0}^{J} \hat{\gamma}_j.
\end{align*}\]

Now assume that the lemma is true for some \( k > 0 \) then we get

\[
\begin{align*}
\sum_{i=0}^{I-(k-1)} C_{i,k-1} & = \sum_{i=0}^{I-k} C_{i,k-1} + C_{I-(k-1),k-1} = \sum_{i=0}^{I-k} C_{i,k} - \sum_{j=0}^{k-1} S_{i,k} + \sum_{j=0}^{k-1} S_{I-(k-1),j} \\
& = \sum_{i=0}^{I-k} \hat{\mu}_i \sum_{j=0}^{k} \hat{\gamma}_j - \hat{\gamma}_k \sum_{i=0}^{I-k} \hat{\mu}_i + \hat{\mu}_{I-(k-1)} \sum_{j=0}^{k-1} \hat{\gamma}_k = \sum_{i=0}^{I-k} \sum_{j=0}^{k} \hat{\mu}_i \hat{\gamma}_j.
\end{align*}\]

Proof of Theorem 6.3:

\[
\begin{align*}
\widehat{C}_{i,J}^\text{Poi} & = C_{i,I-i} + \hat{\mu}_i \sum_{k=I-i+1}^{J} \hat{\gamma}_k = C_{i,I-i} + \frac{C_{i,I-i}}{\sum_{k=0}^{I-i} \hat{\gamma}_k} \sum_{k=I-i+1}^{J} \hat{\gamma}_k = C_{i,I-i} \left( 1 + \frac{\sum_{k=I-i+1}^{J} \hat{\gamma}_k}{\sum_{k=0}^{I-i} \hat{\gamma}_k} \right) \\
\text{Estimator 6.2} & = C_{i,I-i} \frac{\sum_{k=0}^{J} \hat{\gamma}_k}{\sum_{k=I-i}^{J} \hat{\gamma}_k} = C_{i,I-i} \frac{\sum_{k=0}^{J} \hat{\gamma}_k}{\sum_{k=0}^{I-i+1} \hat{\gamma}_k} \cdot \ldots \cdot C_{i,I-i} \frac{\sum_{k=0}^{J-1} \hat{\gamma}_k}{\sum_{k=0}^{I-i+1} \hat{\gamma}_k} = C_{i,I-i} \frac{\sum_{h=0}^{I-(I-i+1)} C_{i,I-i+1}}{\sum_{h=0}^{I-i} C_{i,I-i}} \ldots \frac{\sum_{h=0}^{I-J} C_{i,J}}{\sum_{h=0}^{I-i} C_{i,I-i}}
\end{align*}\]

above lemma

above lemma
Corollary 6.4 (Poisson-Model vs. Chain-Ladder method)

Taking CLM as LSRM with the variance minimizing weights we have

\[
\hat{S}_{i,k}^{\text{CLM}} = \hat{E}[S_{i,k}] = \hat{f}_{k-1}(1 + \hat{f}_{k-2}) \cdots (1 + \hat{f}_{I-i})C_{i,I-i}.
\]

Combining this with Estimator 6.2 and Theorem 6.3 we get

\[
\hat{\gamma}_k = \frac{\hat{S}_{i,k}^{\text{Poi}}}{\hat{\mu}_i} = \frac{\hat{S}_{i,k}^{\text{CLM}}}{\hat{\mu}_i} = \frac{\hat{S}_{i,k}^{\text{CLM}}}{\hat{C}_{i,J}^{\text{CLM}}} = \frac{\hat{f}_{k-1}(1 + \hat{f}_{k-2}) \cdots (1 + \hat{f}_{I-i})C_{i,I-i}}{(1 + \hat{f}_{J-1}) \cdots (1 + \hat{f}_{I-i})C_{i,I-i}}
\]

and

\[
\hat{f}_k = \frac{\hat{\gamma}_{k+1}}{\sum_{j=0}^{k} \hat{\gamma}_j} = \frac{\hat{\gamma}_{k+1}}{1 - \sum_{j=k+1}^{J} \hat{\gamma}_j}.
\]
Proof of the last statement: From the proof of Theorem 6.3 we know that

\[ 1 + \hat{f}_k = \frac{\sum_{j=0}^{k+1} \hat{\gamma}_j}{\sum_{j=0}^{k} \hat{\gamma}_j} . \]

From this we compute

\[ \hat{f}_k = \frac{\sum_{j=0}^{k+1} \hat{\gamma}_j}{\sum_{j=0}^{k} \hat{\gamma}_j} - 1 = \frac{\hat{\gamma}_{k+1}}{\sum_{j=0}^{k} \hat{\gamma}_j} = \frac{\hat{\gamma}_{k+1}}{1 - \sum_{j=k+1}^{J} \hat{\gamma}_j} . \]
Ultimate uncertainty

\[
mse \left[ \sum_{i=0}^{I} \hat{C}_{i,J}^{Poi} \right] = E \left[ \left( \sum_{i+k>l} \left( S_{i,k} - \hat{S}_{i,k}^{Poi} \right) \right)^2 \right]
\]

\[
= E \left[ \left( \sum_{i+k>l} \left( S_{i,k} - E[S_{i,k}] \right) - \sum_{i+k>l} \left( \hat{S}_{i,k}^{Poi} - E[\hat{S}_{i,k}^{Poi}] \right) \right)^2 \right]
\]

\[
\approx E \left[ \left( \sum_{i+k>l} \left( S_{i,k} - E[S_{i,k}] \right) - \sum_{i+k>l} \left( \hat{S}_{i,k}^{Poi} - E[\hat{S}_{i,k}^{Poi}] \right) \right)^2 \right]
\]

\[
= E \left[ \left( \sum_{i+k>l} \left( S_{i,k} - E[S_{i,k}] \right) \right)^2 \right] + E \left[ \left( \sum_{i+k>l} \left( \hat{S}_{i,k}^{Poi} - E[\hat{S}_{i,k}^{Poi}] \right) \right)^2 \right]
\]

\[
- 2E \left[ \left( \sum_{i+k>l} \left( S_{i,k} - E[S_{i,k}] \right) \right) \left( \sum_{i+k>l} \left( \hat{S}_{i,k}^{Poi} - E[\hat{S}_{i,k}^{Poi}] \right) \right) \right]
\]

\[
= \text{Var} \left[ \sum_{i+k>l} S_{i,k} \right] + \text{Var} \left[ \sum_{i+k>l} \hat{S}_{i,k}^{Poi} \right] - 0
\]

\[
\text{random error} \quad \text{parameter error} \quad \text{independence of past and future}
\]
Ultimate uncertainty

\[ \text{mse} \left[ \sum_{i=0}^{\infty} \hat{C}_{Poi, i} \right] = \mathbb{E} \left( \sum_{i > I} \left( S_{i,k} - \mathbb{E}[S_{i,k}] \right)^2 \right) \]

\[ \approx \mathbb{E} \left( \sum_{i > I} \left( S_{i,k} - \mathbb{E}[S_{i,k}] \right)^2 \right) + \mathbb{E} \left( \sum_{i > I} \left( \hat{S}_{Poi, i,k} - \mathbb{E}[\hat{S}_{Poi, i,k}] \right)^2 \right) - 2 \mathbb{E} \left( \sum_{i > I} \left( S_{i,k} - \mathbb{E}[S_{i,k}] \right) \left( \hat{S}_{Poi, i,k} - \mathbb{E}[\hat{S}_{Poi, i,k}] \right) \right) \]

\[ = \text{Var} \left[ \sum_{i > I} X_{i,k} \right] = \text{Var} \left[ \sum_{i > I} S_{i,k} \right] - \text{independence of past and future} \]
Random error

Since all $S_{i,k}$ are independent we get

$$\text{Var}\left[\sum_{i+k>I} S_{i,k}\right] = \sum_{i+k>I} \text{Var}[S_{i,k}] = \sum_{i+k>I} \gamma_k \mu_i \approx \sum_{i+k>I} \hat{\gamma}_k \hat{\mu}_i.$$ 

Parameter error

In order to analyse the parameter error we use the following Taylor expansion:

$$\ln(z) \approx \ln(z_0) + \frac{1}{z_0} (z - z_0) \quad \text{for } z_0 = 1 \text{ and } z = \frac{\hat{\gamma}_k \hat{\mu}_i}{\gamma_k \mu_i}.$$ 

Therefore, we get

$$\hat{\gamma}_k \hat{\mu}_i \approx \gamma_k \mu_i \left(\ln(\hat{\gamma}_k \hat{\mu}_i) - \ln(\gamma_k \mu_i) + 1\right).$$

Finally, taking the covariance it follows

$$\text{Cov}[\hat{\gamma}_{k_1} \hat{\mu}_{i_1}, \hat{\gamma}_{k_2} \hat{\mu}_{i_2}] \approx \gamma_{k_1} \mu_{i_1} \gamma_{k_2} \mu_{i_2} \text{Cov}[\ln(\hat{\gamma}_{k_1} \hat{\mu}_{i_1}), \ln(\hat{\gamma}_{k_2} \hat{\mu}_{i_2})] \approx \hat{\gamma}_{k_1} \hat{\mu}_{i_1} \hat{\gamma}_{k_2} \hat{\mu}_{i_2} \text{Cov}[\ln(\hat{\gamma}_{k_1} \hat{\mu}_{i_1}), \ln(\hat{\gamma}_{k_2} \hat{\mu}_{i_2})].$$

The last covariance term can be estimated by the inverse of the Fisher information matrix $\mathbf{I}$

$$\text{Cov}[\hat{\gamma}_{k_1} \hat{\mu}_{i_1}, \hat{\gamma}_{k_2} \hat{\mu}_{i_2}] \approx \hat{\gamma}_{k_1} \hat{\mu}_{i_1} \hat{\gamma}_{k_2} \hat{\mu}_{i_2} (\mathbf{I}^{-1})_{(i_1,k_1),(i_2,k_2)}. $$
Random error
Since all $\gamma_{i,k}$ are independent we get
\[
\text{Var}\left[\sum_{i+k>\alpha} S_{i,k}\right] = \sum_{i+k>\alpha} \gamma_{i,k} = \sum_{i+k>\alpha} \hat{\gamma}_{i,k}.
\]

Parameter error
In order to analyse the parameter error we use the following Taylor expansion:
\[
\ln(z) \approx \ln(z_0) + \frac{z_0}{z_0}(z - z_0)
\]
for $z_0 = 1$ and $z = \hat{\gamma}_{i,k}\hat{\mu}_i$.

Therefore, we get
\[
\hat{\gamma}_{i,k}\hat{\mu}_i \approx \gamma_{i,k}\mu_i \left(\ln(\hat{\gamma}_{i,k}\hat{\mu}_i) - \ln(\gamma_{i,k}\mu_i) + 1\right).
\]

Finally, taking the covariance it follows
\[
\text{Cov}[\hat{\gamma}_{i_1,k_1}\hat{\mu}_{i_1}, \hat{\gamma}_{i_2,k_2}\hat{\mu}_{i_2}] = \gamma_{i_1,k_1}\mu_{i_1}\gamma_{i_2,k_2}\mu_{i_2}\text{Cov}\left[\ln(\hat{\gamma}_{i_1,k_1}\hat{\mu}_{i_1}), \ln(\hat{\gamma}_{i_2,k_2}\hat{\mu}_{i_2})\right]
\]
\[
\approx \gamma_{i_1,k_1}\mu_{i_1}\gamma_{i_2,k_2}\mu_{i_2}\text{Cov}\left[\ln(\gamma_{i_1,k_1}\mu_{i_1}), \ln(\gamma_{i_2,k_2}\mu_{i_2})\right].
\]

The last covariance term can be estimated by the inverse of the Fisher information matrix $I$:
\[
\text{Cov}[\hat{\gamma}_{i_1,k_1}\hat{\mu}_{i_1}, \hat{\gamma}_{i_2,k_2}\hat{\mu}_{i_2}] \approx \gamma_{i_1,k_1}\mu_{i_1}\gamma_{i_2,k_2}\mu_{i_2}(I^{-1})_{(i_1,k_1),(i_2,k_2)}.
\]
Estimator 6.5 (of the ultimate uncertainty)

$$\hat{\text{mse}} \left[ \sum_{i=0}^{I} \hat{C}_{i,J}^{\text{Poi}} \right] \approx \sum_{i+k>I} \hat{\gamma}_k \hat{\mu}_i + \sum_{i_1+k_1,i_2+k_2>I} \hat{\gamma}_{k_1} \hat{\mu}_{i_1} \hat{\gamma}_{k_2} \hat{\mu}_{i_2} \left( I^{-1} \right)_{(i_1,k_1),(i_2,k_2)}.$$

random error \hspace{1cm} parameter error

Remark 6.6

- The ultimate uncertainty of a single accident period $i$ can be estimated by

$$\hat{\text{mse}} \left[ \hat{C}_{i,J}^{\text{Poi}} \right] \approx \sum_{k=I-i+1}^{J} \hat{\gamma}_k \hat{\mu}_i + \sum_{k_1,k_2=I-i+1}^{J} \hat{\gamma}_{k_1} \hat{\mu}_{i_1} \hat{\gamma}_{k_2} \hat{\mu}_{i_2} \left( I^{-1} \right)_{(i_1,k_1),(i,k_2)}.$$

random error \hspace{1cm} parameter error

- Using the Fisher information matrix for the estimation of the parameter error is a standard approach in the theory of generalised linear models (GLMs). An introduction to generalised linear models can be found in [23].
- The inverse of the Fisher information matrix is a standard output of most GLM-software.
Estimator 6.5 (of the ultimate uncertainty)

\[
\hat{\text{mse}} \left[ \hat{C}_{\text{Poi}} \right] \approx \sum_{i} + \sum_{k > I} \hat{\gamma}_k \hat{\mu}_i \text{random error} + \sum_{k_1, k_2 > I} \hat{\gamma}_{k_1} \hat{\mu}_{i1} \hat{\gamma}_{k_2} \hat{\mu}_{i2} \text{parameter error}.
\]

Remark 6.6

- The ultimate uncertainty of a single accident period \(i\) can be estimated by

\[
\hat{\text{mse}} \left[ \hat{C}_{\text{Poi}} \right] \approx \sum_{i} \hat{\gamma}_k \hat{\mu}_i \text{random error} + \sum_{k_1, k_2 > I} \hat{\gamma}_{k_1} \hat{\mu}_{i1} \hat{\gamma}_{k_2} \hat{\mu}_{i2} \text{parameter error}.
\]

- Using the Fisher information matrix for the estimation of the parameter error is a standard approach in the theory of generalised linear models (GLMs). An introduction to generalised linear models can be found in [23].

- The inverse of the Fisher information matrix is a standard output of most GLM software.
The Poisson-Model as generalised linear model (GLM)

In order to deal with GLMs it is not necessary to know the underlying distribution exactly. It is enough to assume that it belongs to the ‘exponential dispersion family’ and that all $S_{i,k}$ are independent with

$$E[S_{i,k}] = \text{Var}[S_{i,k}] = \gamma_k \mu_i.$$

Overdispersed Poisson-Model

The restriction on the variance to be equal to the expectation can be softened by taking

$$E[S_{i,k}] = \gamma_k \mu_i \quad \text{and} \quad \text{Var}[S_{i,k}] = \varphi_k \gamma_k \mu_i,$$

where $\varphi_k > 0$ is called the dispersion parameter. The estimates for the future development are the same as for the Poisson-Model, but the estimates for the ultimate uncertainty will change.
The Poisson-Model as generalised linear model (GLM)
In order to deal with GLMs it is not necessary to know the underlying distribution exactly. It is enough to assume that it belongs to the ‘exponential dispersion family’ and that all $S_{i,k}$ are independent with

$$E[S_{i,k}] = Var[S_{i,k}] = \gamma_k \mu_i.$$ 

Overdispersed Poisson-Model
The restriction on the variance to be equal to the expectation can be softened by taking

$$E[S_{i,k}] = \gamma_k \mu_i \quad \text{and} \quad Var[S_{i,k}] = \phi_k \gamma_k \mu_i,$$

where $\phi_k > 0$ is called the dispersion parameter. The estimates for the future development are the same as for the Poisson-Model, but the estimates for the ultimate uncertainty will change.
GLMs in general

In general we could assume that

\[
E[S_{i,k}] = x_{i,k} \quad \text{and} \quad \text{Var}[S_{i,k}] = \frac{\varphi_{i,k}}{\omega_{i,k}} V(x_{i,k}),
\]

where

- \( \varphi_{i,k} > 0 \) are the dispersion parameters (unknown),
- \( \omega_{i,k} > 0 \) are known weights in order to compensate for different volumes and
- \( V(\cdot) \) is an appropriate variance function.
GLMs in general

In general we could assume that

\[
\mathbb{E}[S_{i,k}] = x_{i,k} \quad \text{and} \quad \mathbb{V}[S_{i,k}] = \frac{\omega_{i,k}}{\psi_{i,k}} \cdot V(x_{i,k}),
\]

where

- \( \psi_{i,k} > 0 \) are the dispersion parameters (unknown),
- \( \omega_{i,k} > 0 \) are known weights in order to compensate for different volumes and
- \( V(\cdot) \) is an appropriate variance function.
Literature


*Generalised Linear Models.*

Stochastic Reserving

Poisson-Model

Literature

Generalized Linear Models.
Stochastic Reserving

Lecture 11
Bootstrapping

René Dahms
ETH Zurich, Spring 2019

15 May 2019
(Last update: 10 February 2019)
7 Bootstrap for CLM

7.1 Motivation
7.2 Chain-Ladder method and bootstrapping, variant 1
7.3 Bootstrapping Chain-Ladder step by step, variant 1
7.4 Chain-Ladder method and bootstrapping, variant 2
7.5 Bootstrapping Chain-Ladder step by step, variant 2
7.6 Possible problems with bootstrapping
7.7 Parametric vs. non-parametric bootstrap
7.8 Literature
Stochastic Reserving

Lecture 11: Table of contents

7 Bootstrap for CLM
7.1 Motivation
7.2 Chain-Ladder method and bootstrapping, variant 1
7.3 Bootstrapping Chain-Ladder step by step, variant 1
7.4 Chain-Ladder method and bootstrapping, variant 2
7.5 Bootstrapping Chain-Ladder step by step, variant 2
7.6 Possible problems with bootstrapping
7.7 Parametric vs. non-parametric bootstrap
7.8 Literature
Approximation by the empirical distribution (resampling)

- Let \( g(\Phi) \) be a (bounded) real function depending on the random vector \( \Phi = (\Phi_m)_{0 \leq m \leq M} \).
- We are interested in the distribution \( P \) of \( g \).
- **Resampling**: Assume we know the distribution of \( \Phi \) then we could sample an independent sequence \((\varphi^n)_0 \leq n \leq N = (\varphi^m)_{0 \leq m \leq M}\) and approximate \( P \) by the empirical distribution

\[
P^{\text{emp}}(g \leq x) := \frac{\text{number of } \varphi^n \text{ with } g(\varphi^n) \leq x}{N + 1}.
\]

- Unfortunately, instead of the distribution of \( \Phi \) we often only know a single realisation \((\varphi_m)_{0 \leq m \leq M}\).
• If the function $g$ is ‘nice enough’ it is well known that the empirical distribution converges to $P$ in some sense.
Basic idea behind bootstrapping

- **parametric bootstrap**: 
  * make an assumption about the distribution family for $\Phi$ 
  * use the observation $(\varphi_m)_{0 \leq m \leq M}$ to estimate the corresponding parameters 
  * resample

- **non-parametric bootstrap**: 
  Use the empirical distribution $P_M$ generated by resampling the observation $(\varphi_m)_{0 \leq m \leq M}$, i.e.

$$P_M(g \leq x) := \frac{\text{number of vectors } \pi \in \{0, 1, \ldots, M\}^{M+1} \text{ with } g((\varphi_{\pi_m})_{0 \leq m \leq M}) \leq x}{(M + 1)^{M+1}}.$$
Basic idea behind bootstrapping

- **parametric bootstrap**:
  - make an assumption about the distribution family for \( \Phi \)
  - use the observation \((\phi_m)_{0 \leq m \leq M}\) to estimate the corresponding parameters
  - resample

- **non-parametric bootstrap**:
  - Use the empirical distribution \( P_M \) generated by resampling the observation \((\phi_m)_{0 \leq m \leq M}\), i.e.
  \[
  P_M(g \leq x) := \text{number of vectors } \pi \in \{0, 1, \ldots, M\}^{M+1} \text{ with } g((\phi_m)_{0 \leq m \leq M}) \leq x
  \]

**Motivation**

**Parametric bootstrap**:

- Which distribution family should we take?
- In order to be able to estimate the parameters of the underlying distribution of \( \Phi \) based on a single observation we have to make some additional assumptions, for instance that the components of the vector \( \Phi \) are independent identical distributed.

**Non-parametric bootstrap**:

- At least we have to assume that the components of \( \Phi \) are identical distributed.
- In most cases the number \((M + 1)^{M+1}\) is too large. So we do not calculate the sum method all such vectors \( \pi \) but approximate it by resampling (sampling with replacement).
- In some cases it is possible to prove that \( P_M \) converges in some sense to \( P \) as \( M \) goes to infinity. For instance, if \( \Phi \) has independent identical distributed bounded components and

\[
g((\Phi_m)_{0 \leq m \leq M}) := \frac{1}{\sqrt{M + 1}} \sum_{m=0}^{M} \Phi_m.
\]
The flying words ‘to bootstrap’ comes from

‘to pull oneself up by one’s bootstraps’

In our case we want to get the whole distribution by the observation of one realisation.

How to combine the Chain-Ladder method and bootstrapping

We have to find random variables, which

- can be assumed to be i.i.d.
- define the reserves.
The flying words ‘to bootstrap’ come from ‘to pull oneself up by one’s bootstraps.’ In our case we want to get the whole distribution by the observation of one realization.

How to combine the Chain-Ladder method and bootstrapping

We have to find random variables, which
• can be assumed to be i.i.d.
• define the reserves.
Recapitulation

Let $C_{i,k} := \sum_{j=0}^{k} S_{i,j}$. If we have

\begin{align*}
&i')^{\text{CLM}} \quad \mathbb{E}\left[S_{i,k+1 \mid B_{i,k}}\right] = f_k C_{i,k}, \\
&ii')^{\text{CLM}} \quad \text{Var}\left[S_{i,k+1 \mid B_{i,k}}\right] = \sigma_k^2 C_{i,k} \quad \text{and} \\
&iii')^{\text{CLM}} \quad \text{accident periods are independent.}
\end{align*}

Then

\[
\widehat{S}_{i,k} := \hat{f}_{k-1} (1 + \hat{f}_{k-2}) \cdots (1 + \hat{f}_{I-i}) C_{i,I-i} \quad \text{with} \quad \hat{f}_k := \sum_{i=0}^{I-1-k} \frac{C_{i,k}}{\sum_{h=0}^{i-1-k} C_{h,k}} \frac{S_{i,k+1}}{C_{i,k}}
\]

are $D_{I-i}$-conditional unbiased estimators of $S_{i,k}$, for $I-i < k \leq J$.

Therefore, we get

\[
S_{i,k+1} = f_k C_{i,k} + \sqrt{\sigma_k^2 C_{i,k}} \frac{S_{i,k+1} - f_k C_{i,k}}{\sqrt{\sigma_k^2 C_{i,k}}} =: \Phi_{i,k}
\]

where $\Phi_{i,k}$ have mean zero and variance one.

We can look at $S_{i,k}$ as function of $\Phi := \left(\Phi_{i,k}\right)_{i+k<I, k<J}$ and some starting values, for instance $(C_{i,0})_{0 \leq i \leq I}$. 
Recapitulation

Let $C_{i,k} := \sum_{j=0}^{k} S_{i,j}$. If we have

1. $CLM \ E[S_{i,k} + 1 \mid B_{i,k}] = f_k C_{i,k}$,
2. $CLM \ Var[S_{i,k} + 1 \mid B_{i,k}] = \sigma^2_k C_{i,k}$ and
3. $CLM$ accident periods are independent.

Then

$$\hat{S}_{i,k} = \hat{f}_k - 1 \cdot (1 + \hat{f}_k - 2)^{\cdot \cdot \cdot} \cdot (1 + \hat{f}_{I-i}) C_{i,I} - i$$

are $D_{I-i}$-conditional unbiased estimators of $S_{i,k}$, for $I-i < k \leq J$.

Therefore, we get

$$S_{i,k} = \hat{f}_k C_{i,k} + \sqrt{\sigma^2_k C_{i,k}} \Phi_{i,k}$$

where $\Phi_{i,k}$ have mean zero and variance one.

We can look at $S_{i,k}$ as function of $\Phi := (\Phi_{i,k})_{i=0, \ldots, I; k=I-i, \ldots, J}$ and some starting values, for instance $(C_{i,0})_{0 \leq i \leq I}$.

• In the last formula we still have some unknown parameters, i.e. $f_k$ and $\sigma^2_k$.
• $\Phi_{i,k}$ are the Pearson residuals.
• Some of the Pearson residuals have to be ignored, because they cannot have the same distribution like all other residuals. For instance:
  - $\Phi_{I,J-1}$ in the case where $I = J$, because it is equal to zero (deterministically).
  - all $\Phi_{i,k}$ for all development periods $k$ where we know that all claims will be closed. For those $k$ the residuals are deterministic and equal to zero.
Step 1: Chain-Ladder method

claim property: $S_{i,k}$
estimated development factors: $\hat{f}_k$
estimated variance parameters: $\hat{\sigma}_k^2$

<table>
<thead>
<tr>
<th>$S_{i,k}$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>ultimate</th>
<th>reserve</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>0</td>
<td>250</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>300</td>
<td>190</td>
<td>88</td>
<td>0</td>
<td>578</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>100</td>
<td>85</td>
<td>37</td>
<td>0</td>
<td>222</td>
<td>37</td>
</tr>
<tr>
<td>3</td>
<td>200</td>
<td>150</td>
<td>70</td>
<td>0</td>
<td>420</td>
<td>220</td>
</tr>
</tbody>
</table>

| $\hat{f}_k$ | 0.75 | 0.2 | 0 | 1470 | 257 |
| $\hat{\sigma}_k^2$ | 6.67 | 0.70 | 0.09 |

Step 2: Residuals

Pearson residuals inclusive variance adjustment:

$$\varphi_{i,k} := \frac{S_{i,k+1} - \hat{f}_k C_{i,k}}{\sqrt{\hat{\sigma}_k^2 C_{i,k}}} \sqrt{\frac{I - k}{I - k - 1}}$$
correction by the empirical mean:

$$\varphi^*_{i,k} := \varphi_{i,k} - \frac{1}{I(I-1)/2 - 1} \sum_{i+k<I, k<I-1} \varphi_{i,k}$$
$$= \varphi_{i,k} - 0.24$$
Step 1: Chain-Ladder method

Claim property: \( S_{i,k} \)

Estimated development factors: \( \hat{f}_k \)

Estimated variance parameters: \( \hat{\sigma}_k^2 \)

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>0</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>0</td>
<td>250</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>190</td>
<td>88</td>
<td>0</td>
<td>578</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>85</td>
<td>37</td>
<td>0</td>
<td>222</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>150</td>
<td>70</td>
<td>0</td>
<td>420</td>
<td>220</td>
</tr>
</tbody>
</table>

\( \hat{f}_k = 0.75, 0.2, 0 \)

\( \hat{\sigma}_k^2 = 6.67, 0.70, 0.09 \)

Step 2: Residuals

Pearson residuals inclusive variance adjustment:

\[ \varphi_{i,k} = \frac{S_{i,k} + 1 - \hat{f}_k C_{i,k}}{\sqrt{\hat{\sigma}_k^2 C_{i,k}}} \]

Correction by the empirical mean:

\[ \varphi_{i,k}^* = \varphi_{i,k} - \frac{1}{I_k} \frac{1}{I_k - 1} \sum_{i=0}^{I_k - 1} \varphi_{i,k} = \varphi_{i,k} - 0.24 \]

- No residuals for development periods \( k \), where \( \hat{\sigma}_k^2 \) has to be approximated.
- Although \( \Phi_{i,k} \) has zero mean and variance equal to one, its estimate \( \varphi_{i,k} \) doesn’t. The reason for this is that we do not know the parameters \( f_k \) and \( \sigma_k \) and use some estimators instead. Let’s assume we know \( \sigma_k^2 \) and take the variance minimizing weights \( w_{i,k} \):

\[
\text{Var} \left[ \frac{S_{i,k+1} - \hat{f}_k C_{i,k}}{\sqrt{\sigma_k^2 C_{i,k}}} \right] = \frac{1}{\sigma_k^2 C_{i,k}} \left( \text{Var} [S_{i,k+1} | D_k] - 2C_{i,k} \text{Cov} [S_{i,k+1}, \hat{f}_k | D_k] + C_{i,k}^2 \text{Var} [\hat{f}_k | D_k] \right)
\]

\[
= \frac{1}{\sigma_k^2 C_{i,k}} \left( \sigma_k^2 C_{i,k} - 2C_{i,k} \text{Cov} [S_{i,k+1}, \varphi_{i,k} S_{i,k+1} / C_{i,k} | D_k] + C_{i,k}^2 \sum_{h=0}^{I_k - 1} \frac{w_{h,k}^2}{w_{h,k}} \text{Var} [S_{h,k+1} | D_k] \right)
\]

\[
= \left( 1 - 2 \frac{C_{i,k}}{\sum_{h=0}^{I_k - 1} C_{h,k}} \right) + \frac{C_{i,k}}{\sum_{h=0}^{I_k - 1} C_{h,k}} = \left( 1 - \frac{C_{i,k}}{\sum_{h=0}^{I_k - 1} C_{h,k}} \right) < 1
\]

\[ w_{i,k} = \frac{C_{i,k}}{\sum_{h=0}^{I_k - 1} C_{h,k}} \]

Therefore, we could take

\[
\left( 1 - \frac{C_{i,k}}{\sum_{h=0}^{I_k - 1} C_{h,k}} \right)^{-1}
\]

as variance adjustment factor. But since we don’t know the variance parameter \( \sigma_k^2 \) we take \( \sqrt{\frac{I_k}{I_k - 1}} \) instead. This insures that

\[
\frac{1}{I_k - 1} \sum_{i=0}^{I_k - 1} \text{Var} [\varphi_{i,k} | D_k] = 1 \quad \text{provided we know } \sigma_k^2.
\]
Step 3: Resampled residuals (non-parametric bootstrap)

set of residuals:

\{ -1.28, -1.00, 0.28, 0.95, 1.05 \}

\[
\begin{array}{c|ccc}
\varphi_{i,k} & 0 & 1 & 2 & 3 \\
\hline
0 & 0.28 & -1.00 & 1.05 \\
1 & -1.00 & 0.95 & -1.00 \\
2 & 1.05 & 0.28 & -1.28 \\
3 & -1.28 & 0.28 & 0.95 \\
\end{array}
\]

Step 4a: Resampled triangle and Chain-Ladder method without process variance

\[
\begin{align*}
S_{i,0}^* &:= S_{i,0} \\
S_{i,k+1}^* &:= \hat{f}_k C_{i,k}^* + \sqrt{\hat{\sigma}_k^2 C_{i,k}^*} \varphi_{i,k}^*, \quad i + k \leq I \\
S_{i,k+1}^* &:= \hat{f}_k^* C_{i,k}^*, \quad i + k > I
\end{align*}
\]

\[
\begin{array}{c|cccc|cc}
S_{i,k}^* & 0 & 1 & 2 & 3 & \text{ultimate} & \text{reserve} \\
\hline
0 & 100 & 82 & 25 & 4 & 211 & 0 \\
1 & 300 & 184 & 114 & 13 & 611 & 13 \\
2 & 100 & 100 & 42 & 5 & 247 & 47 \\
3 & 200 & 146 & 72 & 9 & 428 & 228 \\
\end{array}
\]

\[
\begin{array}{c|ccc}
\hat{f}_k^* & 0.73 & 0.21 & 0.02 \\
\end{array}
\]

Step 4b: Resampled triangle and Chain-Ladder method with process variance

\[
\begin{align*}
S_{i,0}^* &:= S_{i,0} \\
S_{i,k+1}^* &:= \hat{f}_k C_{i,k}^* + \sqrt{\hat{\sigma}_k^2 C_{i,k}^*} \varphi_{i,k}^*, \quad i + k \leq I \\
S_{i,k+1}^* &:= \hat{f}_k^* C_{i,k}^*, \quad i + k > I
\end{align*}
\]

\[
\begin{array}{c|cccc|cc}
S_{i,k}^* & 0 & 1 & 2 & 3 & \text{ultimate} & \text{reserve} \\
\hline
0 & 100 & 82 & 25 & 4 & 211 & 0 \\
1 & 300 & 184 & 114 & 6 & 604 & 6 \\
2 & 100 & 100 & 45 & -1 & 244 & 44 \\
3 & 200 & 103 & 77 & 15 & 394 & 194 \\
\end{array}
\]

\[
\begin{array}{c|ccc}
\hat{f}_k^* & 0.73 & 0.21 & 0.03 \\
\end{array}
\]
In the case of parametric bootstrap we use the residuals in order to fit a distribution and use this distribution to get the resampled residuals.
Step 5: Repeat steps 3 and 4 and collect the resulting reserves

Reserves without process variance (sorted):
\{145, 146, 148, 156, 156, 157, 159, 165, 166, 167, 168, 168, \ldots, 345, 346, 347, 347, 347, 349, 351, 352, 354, 355, 357, 357\}

Reserves with process variance (sorted):

empirical distribution function

\[
\begin{align*}
\sigma^2 &= 1724 \\
\mu &= 256
\end{align*}
\]

empirical density

\[
\begin{align*}
\sigma^2 &= 1724 \\
\mu &= 256
\end{align*}
\]
Step 5: Repeat steps 3 and 4 and collect the resulting reserves

Reserves without process variance (sorted):

\{145, 146, 148, 156, 156, 157, 159, 165, 166, 167, 168, 168, \ldots, 345, 346, 347, 347, 347, 349, 351, 352, 354, 355, 357, 357, \ldots\}

Empirical distribution function

100 140 180 220 260 300 340 380 420

\[\sigma^2 = 1724\] \[\mu = 256\]

Reserves with process variance (sorted):


Empirical distribution function

100 140 180 220 260 300 340 380 420

\[\sigma^2 = 3197\] \[\mu = 257\]

- without process variance: \(\sigma^2\) represents the squared parameter estimation error
- with process variance: \(\sigma^2\) represents the sum of the squared parameter estimation error and the process variance
- in this example the mean of the empirical distribution is almost equal to the Best Estimate of the Chain-Ladder method
- Other variants of bootstrap methods are
  - other starting values, for instance the last known diagonal,
  - \(S^*_{i,k+1} := \hat{f}^* C_{i,k} + \sqrt{\hat{\sigma}^2 C_{i,k}} \varphi^*_{i,k}\)
Recapitulation: (methoddispersed) Poisson-Model

If we have

i) $\text{Poi} \ S_{i,k}$ are independent random variables,

ii) $\text{Poi} \ \text{the distribution of } S_{i,k} \ \text{belongs to the exponential dispersion family and}$

iii) $\text{Poi} \ \text{Var} \left[ S_{i,k} \right] = \vartheta_k \text{E} \left[ S_{i,k} \right] = \vartheta_k \gamma_k \mu_i.$

Then $\hat{S}_{i,k} := \hat{\gamma}_k \hat{\mu}_i$, where $\hat{\gamma}_k$ and $\hat{\mu}_i$ solve (6.1) and $\sum_{k=0}^{J} \hat{\gamma}_k = 1$, are unbiased estimators of $S_{i,k}$, for $I - i < k \leq J$.

Therefore, we get

$$S_{i,k} = \gamma_k \mu_i + \sqrt{\gamma_k \mu_i} \frac{S_{i,k} - \gamma_k \mu_i}{\sqrt{\gamma_k \mu_i}},$$

where $\Phi_{i,k}$ have mean zero and variance $\vartheta_k$.

We can look at $S_{i,k}$ as function of $\Phi := \left( \Phi_{i,k} \right)_{i+k \leq I, i<I, k \leq J}.$
Recapitulation: (methoddispersed) Poisson-Model

If we have

1) \( Poi \) \( S_{i,k} \) are independent random variables,

2) \( Poi \) the distribution of \( S_{i,k} \) belongs to the exponential dispersion family and

3) \( Var[S_{i,k}] = \mu_i \gamma_k \mu_i \).

Then \( \hat{S}_{i,k} := \hat{\gamma}_k \hat{\mu}_i \), where \( \hat{\gamma}_k \) and \( \hat{\mu}_i \) solve (6.1) and \( \sum_k \hat{\gamma}_k = 1 \), are unbiased estimators of \( S_{i,k} \) for \( i < k \leq j \).

Therefore, we get

\[ S_{i,k} = \gamma_k \mu_i + \sqrt{\gamma_k \mu_i} \Phi_{i,k} \]

where \( \Phi_{i,k} \) have mean zero and variance \( \theta_k \).

We can look at \( S_{i,k} \) as function of \( \Phi := \{ \Phi_{i,k} | i < k \leq j \} \).

- In the last formula we still have some unknown parameters, i.e. \( \gamma_k \), \( \mu_i \) and \( \theta_k \).
7 Bootstrap for CLM

7.5 Bootstrapping Chain-Ladder step by step, variant 2

Step 1: Chain-Ladder method (Poisson-Model)

<table>
<thead>
<tr>
<th>$S_{i,k}$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>ultimate ($\hat{\mu}_i$)</th>
<th>reserve</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>100</td>
<td>100</td>
<td>50</td>
<td>0</td>
<td>250</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>300</td>
<td>190</td>
<td>88</td>
<td>0</td>
<td>578</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>100</td>
<td>85</td>
<td>37</td>
<td>0</td>
<td>222</td>
<td>37</td>
</tr>
<tr>
<td>3</td>
<td>200</td>
<td>150</td>
<td>70</td>
<td>0</td>
<td>420</td>
<td>220</td>
</tr>
</tbody>
</table>

| $\hat{\gamma}_k$ | 0.48 | 0.36 | 0.17 | 0.00 | 1470 | 257 |

Step 2: Residuals

Residuals inclusive variance adjustment:

$$\varphi_{i,k} := \frac{S_{i,k} - \hat{\gamma}_k \hat{\mu}_i}{\sqrt{\hat{\gamma}_k \hat{\mu}_i}} \sqrt{\frac{I - k}{\sum_{h=0}^{I-k} (S_{h,k} - \hat{\gamma}_k \hat{\mu}_h)^2}} \approx \sqrt{\vartheta_k}$$

Correction by the empirical mean:

$$\varphi_{i,k}^* := \varphi_{i,k} - \frac{1}{I(I+1)/2 - 2 \sum_{i+k\leq I, i< I, k< I} \varphi_{i,k}}$$

$$= \varphi_{i,k} - 0.04$$
• No residual for \((i, k) = (I, 0)\), because it is equal to zero (deterministically).
• Although \(\Phi_{i,k}\) has zero mean and variance equal to \(\vartheta_k\), its estimate \(\varphi_{i,k}\) doesn’t. The reason for this is that we do not know the parameters \(f_k\) and \(\sigma_k\) and use some estimators instead. The variance adjustment

\[
\sqrt{\hat{\vartheta}_k} = \sqrt{\frac{I - k}{\sum_{h=0}^{I-k} \left( \frac{S_{h,k} - \hat{\gamma}_k \hat{\mu}_h}{\hat{\gamma}_k \hat{\mu}_h} \right)^2}}
\]

ensures that the empirical variance equals one, i.e. that

\[
\frac{1}{I - k} \sum_{i=0}^{I-k} (\varphi_{i,k}^* - 0)^2 = 1.
\]
Step 3: Resampled residuals (non-parametric bootstrap)

set of residuals:

\{-1.52, -1.19, -0.82, -0.51, 0.60, 1.09, 1.14, 1.22\}

\[
\begin{array}{cccc|c}
\varphi_{i,k}^{*} & 0 & 1 & 2 & 3 \\
0 & -0.51 & -1.19 & -1.52 & 1.09 \\
1 & 1.22 & 1.09 & -1.19 & -0.82 \\
2 & -1.52 & -0.51 & 1.22 & -1.19 \\
3 & 1.22 & -0.51 & 1.09 & 1.14 \\
\end{array}
\]

Step 4a: Resampled triangle and Chain-Ladder method without process variance

\[
S_{i,k}^{*} := \hat{\gamma}_k \hat{\mu}_i + \sqrt{\hat{\vartheta}_k \hat{\gamma}_k \hat{\mu}_i \varphi_{i,k}^{*}}, \quad i + k \leq I, \\
S_{i,k}^{*} := \hat{\gamma}_k^{*} \hat{\mu}_i^{*}, \quad i + k > I
\]

\[
\begin{array}{cccc|c|c}
S_{i,k}^{*} & 0 & 1 & 2 & 3 & \text{ult.} (\hat{\mu}_i^{*}) & \text{reserve} \\
0 & 107 & 69 & 22 & 0 & 198 & 0 \\
1 & 233 & 235 & 73 & 0 & 541 & 0 \\
2 & 72 & 71 & 21 & 0 & 164 & 21 \\
3 & 237 & 216 & 67 & 0 & 520 & 283 \\
\gamma_k^{*} & 0.44 & 0.36 & 0.11 & 0.00 & 1423 & 304 \\
\end{array}
\]

Step 4b: Resampled triangle and Chain-Ladder method with process variance

\[
S_{i,k}^{*} := \hat{\gamma}_k \hat{\mu}_i + \sqrt{\hat{\vartheta}_k \hat{\gamma}_k \hat{\mu}_i \varphi_{i,k}^{*}}, \quad i + k \leq I, \\
S_{i,k}^{*} := \hat{\gamma}_k^{*} \hat{\mu}_i^{*} + \sqrt{\hat{\vartheta}_k \hat{\gamma}_k \hat{\mu}_i \varphi_{i,k}^{*}}, \quad i + k > I
\]

\[
\begin{array}{cccc|cc|c|c}
S_{i,k}^{*} & 0 & 1 & 2 & 3 & \text{ultimate} & \text{reserve} \\
0 & 107 & 69 & 22 & 0 & 198 & 0 \\
1 & 233 & 235 & 73 & 0 & 541 & 0 \\
2 & 72 & 71 & 50 & 0 & 193 & 50 \\
3 & 237 & 200 & 87 & 0 & 525 & 288 \\
\gamma_k^{*} & 0.44 & 0.36 & 0.11 & 0.00 & 1456 & 337 \\
\end{array}
\]
In the case of parametric bootstrap we use the residuals in order to fit a distribution and use this distribution to get the resampled residuals.
Step 5: Repeat steps 3 and 4 and collect the resulting reserves

Reserves without process variance (sorted):
\{139, 153, 154, 155, 157, 157, 161, 162, 164, 165, 166, 166, \ldots , 378, 380, 381, 384, 385, 386, 387, 388, 389, 400, 402, 447\}

Reserves with process variance (sorted):
\{86, 92, 95, 97, 99, 101, 102, 105, 109, 111, 117, 117, \ldots , 420, 427, 428, 430, 432, 449, 451, 459, 460, 466, 472, 481\}

Empirical distribution function

\[\begin{array}{c}
s^2 = 2488 \\
\mu = 259
\end{array}\]

Empirical density

\[\begin{array}{c}
s^2 = 2488 \\
\mu = 259
\end{array}\]
Step 5: Repeat steps 3 and 4 and collect the resulting reserves without process variance (sorted):

\{139, 153, 154, 155, 157, 157, 161, 162, 164, 165, 166, 166, 378, 380, 381, 384, 385, 386, 387, 388, 389, 400, 402, 447\}

Step 5: Repeat steps 3 and 4 and collect the resulting reserves with process variance (sorted):

\{86, 92, 95, 97, 99, 101, 102, 105, 109, 111, 117, 117, 420, 427, 428, 430, 432, 449, 451, 459, 460, 466, 472, 481\}

- without process variance: $\sigma^2$ represents the squared parameter estimation error
- with process variance: $\sigma^2$ represents the sum of the squared parameter estimation error and the process variance
- in this example the mean of the empirical distribution is almost equal to the Best Estimate of the Chain-Ladder method
- Another version of bootstrapping is to take

$$S_{i,k}^* := \hat{\gamma}_k \hat{\mu}_i + \sqrt{\hat{\varrho}_k \hat{\gamma}_k \hat{\mu}_i \hat{\varphi}_{i,k}}, \quad i + k > I.$$
Possible problems with bootstrapping

- Following the bootstrap idea strictly would imply that instead applying the standard Chain-Ladder method automatically we had to hire some experienced reserving actuaries and let them estimate the reserves for each resampled triangle.
- If the mean of the resampled empirical distribution is not equal to the Best-Estimate we have to rescale
  * each resampled outcome individually or
  * the resampled empirical distribution
- Exclude non-random areas otherwise the resulting variance will be too small. For example, if we know that all claims will be settled after 10 years we should exclude all residuals (all deterministic and equal to zero) after development year 10.
- We may exclude resampled triangles which are not possible. For instance, if we have payments without subrogation then we know that all payments will be non-negative. Therefore, we may exclude resampled triangles with negative entries.
In the case of the last bullet point it could even happen that the cumulative payments get negative.
Bootstrapped probabilities (inclusive process variance) of both variants

Variant 1 (ResQ output)

Variant 2 (ResQ output)

The bootstrapped distribution using variant 1 looks a bit too symmetric. Therefore, I would prefer variant 2 in this case.
Stochastic Reserving

Bootstrap for CLM

Possible problems with bootstrapping

The bootstrapped distribution using variant 1 looks a bit too symmetric. Therefore, I would prefer variant 2 in this case.
Parametric bootstrap

- We can resample triangles with extreme behaviour even if we only observe very small residuals. 😊
- We have to make an assumption about the distribution of the reserves. 😞

Non-parametric bootstrap

- If we only observe very small residuals the bootstrapped empirical distribution may be too ‘nice’. We may underestimate uncertainties. 😞
- We do not have to make an assumption about the distribution of the residuals. 😊
Up to now there is no proof that either of the presented bootstrapping variants converge in some sense to the real distribution of the reserves. On the contrary there are empirical studies, where

- a Poisson distribution was chosen to generate a triangle
- the resulting bootstrap distribution and the real distribution of the reserves has been compared

The results indicate, that the uncertainty may be underestimated by bootstrapping.
Literature

[25] Efron, B. and Tibshirani, R.J.
   *An Introduction to the Bootstrap.*

[26] England, P.D. and Verrall, R.J.
   Analytic and bootstrap estimates of prediction errors in claims reserving.

[27] England, P.D. and Verrall, R.J.
   Stochastic claims reserving in general insurance.
Stochastic Reserving

Bootstrap for CLM

Literature

[25] Efron, B. and Tibshirani, R.J.
   An Introduction to the Bootstrap.

[26] England, P.D. and Verrall, R.J.
   Analytic and bootstrap estimates of prediction errors in claims reserving.

[27] England, P.D. and Verrall, R.J.
   Stochastic claims reserving in general insurance.
Stochastic Reserving

Lecture 12

CLM: Bayesian & credibility approach

René Dahms

ETH Zurich, Spring 2019

22 May 2019

(Last update: 18 February 2019)
8 CLM: Bayesian & credibility approach
8.1 A Bayesian approach to the Chain-Ladder method
8.2 A credibility approach to the Chain-Ladder method
8.3 Example
8.4 Literature
Recapitulation of the Chain-Ladder method

Let \( C_{i,k} := \sum_{j=0}^{k} S_{i,j} \). If we have

\begin{align*}
\text{i)} & \quad \text{CLM} \quad \mathbb{E}[C_{i,k+1} | B_{i,k}] = f_k C_{i,k}, \\
\text{ii)} & \quad \text{CLM} \quad \text{Var}[C_{i,k+1} | B_{i,k}] = \sigma_k^2 C_{i,k} \quad \text{and} \\
\text{iii)} & \quad \text{CLM} \quad \text{accident periods are independent.}
\end{align*}

Then \( \hat{C}_{i,k+1} := \hat{f}_k \cdot \ldots \cdot \hat{f}_{I-i} C_{i,I-i} \) with

\[
\hat{f}_k := \sum_{i=0}^{I-1-k} \frac{C_{i,k}}{\sum_{h=0}^{I-1-k} C_{h,k}} \frac{C_{i,k+1}}{C_{i,k}}
\]

are \( D_{I-i} \)-conditional unbiased estimators of \( C_{i,k} \), for \( I - i \leq k < J \).

But

this is only true if we assume that the development factors \( f_k \) are fixed. We now want to look at the Chain-Ladder method where they are assumed to be realisations of random variables \( \varphi_k \) with \( \mathbb{E}[\varphi_k] = f_k \). We denote by

\[
\varphi := (\varphi_0, \ldots, \varphi_{J-1})
\]

the corresponding collections of all random development factors.
Recapitulation of the Chain-Ladder method

Let \( C_{i,k} := \sum_{j=0}^{C_{i,j}} \). If we have

1) CLM \( \text{E}[C_{i,k+1}|B_{i,k}] = f_k C_{i,k} \)
2) CLM \( \text{Var}[C_{i,k+1}|B_{i,k}] = \sigma^2_k C_{i,k} \) and
3) CLM事故期间是独立的.

Then \( \hat{C}_{i,k+1} = \hat{f}_k \cdot \ldots \cdot \hat{f}_I \cdot C_{i,I} \) with \( \hat{f}_k := I^{-1} - k \sum_{i=0}^{C_{i,k}} \frac{C_{i,k}}{I^k} \) are \( B_{i,k} \)-conditional unbiased estimators of \( C_{i,k} \), for \( I \leq k < J \).

But this is only true if we assume that the development factors \( f_k \) are fixed. We now want to look at the Chain-Ladder method where they are assumed to be realizations of random variables \( \varphi_k \) with \( \text{E}[\varphi_k] = f_k \). We denote by

\[ \varphi = (\varphi_0, \ldots, \varphi_J) \]

the corresponding collections of all random development factors.

Note, everything will stay correct if we replace \( B_{i,k} \) with \( D_{i,k} \) and skip the independence assumption.
Assumption 8.A (Bayesian Chain-Ladder method)

We assume that

1) $\text{Bay} \quad \mathbb{E}[C_{i,k+1} | \varphi, B_{i,k}] = \varphi_k C_{i,k},$

2) $\text{Bay} \quad \text{Var}[C_{i,k+1} | \varphi, B_{i,k}] = \sigma_k^2(\varphi) C_{i,k},$

3) $\text{Bay} \quad \text{conditional given } \varphi \text{ the accident periods are independent and}$

4) $\text{Bay} \quad \text{For any selection } u_k \in \{1, \varphi_k, \varphi^2_k, \sigma_k^2(\varphi)\} \text{ we have}$

$$\mathbb{E}[u_0 \cdot \ldots \cdot u_{J-1} | D] = \mathbb{E}[u_0 | D] \cdot \ldots \cdot \mathbb{E}[u_{J-1} | D],$$

where $D$ is any claim information $D^n, D_k, D^n_k$ or $D^n \cap D_k$.

Remark 8.1

• We assume that the variance parameters $\sigma_k^2$ may depend on the random development factors $\varphi$.

• Conditionally given $\varphi$ we have a standard Chain-Ladder method with development factors $\varphi_k$ and variance parameters $\sigma_k^2(\varphi)$. 
Assumption 8.A (Bayesian Chain-Ladder method)

We assume that

1) \( \mathbb{E}[C_{i,k}^{i+k} | \Phi, B_{i,k}] = \phi_k C_{i,k} \),
2) \( \text{Var}[C_{i,k}^{i+k} | \Phi, B_{i,k}] = \sigma_k^2(\phi) C_{i,k} \),
3) \( \text{conditional given } \phi \) the accident periods are independent and
4) \( \text{For any selection } u_k \in \{1, \phi_k, \phi_k^2, \sigma_k^2(\phi)\} \) we have

\[
\mathbb{E}[u_0 \cdot \ldots \cdot u_{J-1} | \mathcal{D}] = \mathbb{E}[u_0 | \mathcal{D}] \cdot \ldots \cdot \mathbb{E}[u_{J-1} | \mathcal{D}],
\]

where \( \mathcal{D} \) is any claim information \( \mathcal{D}^i, \mathcal{D}^j, \mathcal{D}^i \cap \mathcal{D}^j \).

Remark 8.1

• We assume that the variance parameters \( \sigma_k^2 \) may depend on the random development factors \( \phi \).
• Conditionally given \( \phi \) we have a standard Chain-Ladder method with development factors \( \phi_k \) and variance parameters \( \sigma_k^2(\phi) \).

Note, everything will stay correct if we replace \( B_{i,k} \) with \( D_{k}^{i+k} \) and skip the independence assumption.
Definition 8.2 (Bayes estimators)
Let $Z$ be a random variable and $\mathcal{D}$ some $\sigma$-algebra (for instance the information contained in some observations). The Bayes estimator $Z^{Bay}$ of $Z$ given $\mathcal{D}$ is defined by

$$Z^{Bay} := E[Z|\mathcal{D}].$$

Corollary 8.3
If $Z^2$ is integrable then the Bayes estimator is the $\mathcal{D}$-measurable estimator that minimizes the conditionally, given $\mathcal{D}$, mean squared error of prediction, i.e.

$$Z^{Bay} = \arg\min_{\widehat{Z}} E[(Z - \widehat{Z})^2|\mathcal{D}].$$

Estimator 8.4 (of the future outcome)
Under Assumption 8.4 we get

$$C_{i,k+1}^{Bay} := E[C_{i,k+1}|\mathcal{D}^I] = E[\varphi_k|\mathcal{D}^I] \cdot \ldots \cdot E[\varphi_{I-i}|\mathcal{D}^I] C_{i,I-i} =: \varphi_k^{Bay} \cdot \ldots \cdot \varphi_{I-i}^{Bay} C_{i,I-i}$$
Definition 8.2 (Bayes estimators)
Let $Z$ be a random variable and $D$ some $\sigma$-algebra (for instance the information contained in some observations). The Bayes estimator $Z_{\text{Bay}}$ of $Z$ given $D$ is defined by

$$Z_{\text{Bay}} := E[Z|D].$$

Corollary 8.3
If $Z$ is integrable then the Bayes estimator is the $D$-measurable estimator that minimizes the conditionally, given $D$, mean squared error of prediction, i.e.

$$Z_{\text{Bay}} = \arg\min_{\hat{Z}} E[(Z - \hat{Z})^2 | D].$$

Estimator 8.4 (of the future outcome)
Under Assumption 8.A we get

$$C_{i,k+1}^{\text{Bay}} := E[C_{i,k+1} | D] = E \left[ E[C_{i,k+1} | \varphi, D] | D \right] = E \left[ \varphi_k \cdot \ldots \cdot \varphi_{I-i} C_{i,I-i} | D \right].$$

The corollary is true, because the conditional expectation is the orthogonal projection onto the subspace of all $D$ measurable functions (within the space of all square integrable functions).

Proof of Estimator 8.4:

$$C_{i,k+1}^{\text{Bay}} = E \left[ C_{i,k+1} | D \right] = E \left[ E \left[ C_{i,k+1} | \varphi, D \right] | D \right]$$

$$= E \left[ \varphi_k \cdot \ldots \cdot \varphi_{I-i} C_{i,I-i} | D \right].$$

standard CLM for fixed development factors

$$= E \left[ \varphi_k | D \right] \cdot \ldots \cdot E \left[ \varphi_{I-i} | D \right] C_{i,I-i}.$$
Ultimate uncertainty in the Bayesian case

For the mean squared error of prediction of the ultimate outcome we get

\[
\text{mse}_{DI} \left[ \sum_{i=0}^{I} C_{i,J}^{Bay} \right] = E \left[ E \left[ \left( \sum_{i=0}^{I} \left( C_{i,J} - C_{i,J}^{Bay} \right) \right)^2 \mid \varphi, DI \right] \right] DI
\]

\[
= E \left[ \left( \sum_{i=0}^{I} \left( C_{i,J} - E[C_{i,J} \mid \varphi, DI] \right) \right)^2 \mid \varphi, DI \right] DI
\]

\[
= E \sum_{i=0}^{I} \text{Var}[C_{i,J} \mid \varphi, DI] + \left( \sum_{i=0}^{I} \left( E[C_{i,J} \mid \varphi, DI] - C_{i,J}^{Bay} \right) \right)^2 DI
\]

\[
= \sum_{i=0}^{I} E[\text{Var}[C_{i,J} \mid \varphi, DI] \mid DI] + E \left[ \left( \sum_{i=0}^{I} \left( E[C_{i,J} \mid \varphi, DI] - C_{i,J}^{Bay} \right) \right)^2 \right] DI
\]

\[
\text{random error} \quad \text{parameter error}
\]
For the mean squared error of prediction of the ultimate outcome we get

\[ \text{mse} \left[ \sum_{i=0}^{C_{i,J}} \right] = \mathbb{E} \left[ \left( \sum_{i=0}^{C_{i,J}} (C_{i,J} - C_{Bay_i,J}) \right)^2 \right] \]

\[ = \mathbb{E} \left[ \left( \sum_{i=0}^{C_{i,J}} (C_{i,J} - \mathbb{E}[C_{i,J} | \phi, D_I]) + \left( \sum_{i=0}^{C_{Bay_i,J}} (C_{Bay_i,J} - \mathbb{E}[C_{i,J} | \phi, D_I]) \right) \right)^2 \right] \]

\[ = \sum_{i=0}^{C_{i,J}} \mathbb{E} \left[ \text{var}(C_{i,J} | \phi, D_I) \right] + \mathbb{E} \left[ \left( \sum_{i=0}^{C_{Bay_i,J}} (C_{Bay_i,J} - \mathbb{E}[C_{i,J} | \phi, D_I]) \right)^2 \right] \]
Derivation of the random error

\[
\begin{align*}
E[\text{Var}[C_{i,j} | \varphi, D^I] | D^I] &= \sum_{k=I-i}^{J-1} E \left[ \prod_{j=k+1}^{J-1} \varphi_j^2 \sigma_k^2(\varphi) \prod_{j=I-i}^{k-1} \varphi_j \bigg| D^I \right] C_{i,I-i} \\
&= \sum_{k=I-i}^{J-1} \prod_{j=k+1}^{J-1} E[\varphi_j^2 | D^I] E[\sigma_k^2(\varphi) | D^I] \prod_{j=I-i}^{k-1} E[\varphi_j | D^I] C_{i,I-i}. \\
\end{align*}
\]

iv) Bay

Derivation of the parameter error

\[
\begin{align*}
E \left[ \left( \sum_{i=0}^{I} \left( E[C_{i,j} | \varphi, D^I] - C_{i,j}^{\text{Bay}} \right) \right)^2 \bigg| D^I \right] &= \sum_{i_1,i_2=0}^{I} C_{i_1,I-i_1} C_{i_2,I-i_2} \text{Cov} \left[ \prod_{k=I-i_1}^{J-1} \varphi_k, \prod_{k=I-i_2}^{J-1} \varphi_k \bigg| D^I \right] \\
&= \sum_{i_1,i_2=0}^{I} C_{i_1,I-i_1} C_{i_2,I-i_2} \prod_{k=I-(i_1 \lor i_2)}^{I-(i_1 \land i_2) - 1} E[\varphi_k | D^I] \left( \prod_{k=I-(i_1 \land i_2)}^{J-1} E[\varphi_k^2 | D^I] - \prod_{k=I-(i_1 \lor i_2)}^{J-1} E[\varphi_k | D^I]^2 \right).
\end{align*}
\]
A Bayesian approach to the Chain-Ladder method

Parameter error:

\[
\begin{align*}
E \left[ \left( \sum_{i=0}^{I} \left( E \left[ C_{i,J} \mid D^I \right] - C_{i,J}^{Bay} \right) \right)^2 \right] & = \text{Var} \left[ \sum_{i=0}^{I} E \left[ C_{i,J} \mid D^I \right] D^I \right] \\
& \quad \quad \quad + \text{E} \left[ E \left[ C_{i,J} \mid D^I \right] D^I \right] = C_{i,J}^{Bay}
\end{align*}
\]

\[
= \sum_{i_1, i_2=0}^{I} \text{Cov} \left[ C_{i_1,I-i_1}, C_{i_2,I-i_2} \prod_{k=I-i_1}^{J-1} \varphi_k, \prod_{k=I-i_2}^{J-1} \varphi_k \mid D^I \right]
\]

\[
= \sum_{i_1, i_2=0}^{I} C_{i_1,I-i_1} C_{i_2,I-i_2} \text{Cov} \left[ \prod_{k=I-i_1}^{J-1} \varphi_k, \prod_{k=I-i_2}^{J-1} \varphi_k \mid D^I \right]
\]

\[
= \sum_{i_1, i_2=0}^{I} C_{i_1,I-i_1} C_{i_2,I-i_2} \left( E \left[ \prod_{k=I-i_1}^{J-1} \varphi_k \mid D^I \right] - E \left[ \prod_{k=I-i_1}^{J-1} \varphi_k \mid D^I \right] \right) \left( E \left[ \prod_{k=I-i_2}^{J-1} \varphi_k \mid D^I \right] - E \left[ \prod_{k=I-i_2}^{J-1} \varphi_k \mid D^I \right] \right)
\]

\[
= \sum_{i_1, i_2=0}^{I} C_{i_1,I-i_1} C_{i_2,I-i_2} \prod_{k=I-(i_1 \land i_2)}^{I-(i_1 \lor i_2)-1} E \left[ \varphi_k \mid D^I \right] \left( E \left[ \prod_{k=I-(i_1 \land i_2)}^{J-1} \varphi_k \mid D^I \right] - E \left[ \prod_{k=I-(i_1 \land i_2)}^{J-1} \varphi_k \mid D^I \right] \right)
\]

Note: Although accident periods are independent given \(D^I\) and \(\varphi\) they are usually not independent given \(D^I\).
Problem 8.5

*We still have to estimate*

\[ E[\varphi_k | D^I], \quad E[\varphi_k^2 | D^I] \quad \text{and} \quad E[\sigma_k^2(\varphi) | D^I]. \]

**Distribution based models**

On solution is to make an assumption on the joint distribution of \((C_{i,k})_{i+k \leq I}\) and \(\varphi\) and than calculate the a posteriori distribution of \(\varphi\) given \(D^I\), which then can be used to calculate the missing objects.

**Credibility approximation**

Another way is to look only at estimators \(\hat{F}_k^{Cred}\), which depends in an affine way on the observations \(F_{i,k} = \frac{C_{i,k+1}}{C_{i,k}}\).
Problem 8.5
We still have to estimate \( E[\varphi_k|D_I] \), \( E[\varphi_2|D_I] \) and \( E[\sigma^2_k(\varphi)|D_I] \).

Distribution based models
On solution is to make an assumption on the joint distribution of \((C_i,k)_{i+k\leq I}\) and \(\varphi\) and then calculate the a posteriori distribution of \(\varphi\) given \(D_I\), which then can be used to calculate the missing objects.

Credibility approximation
Another way is to look only at estimators \( \hat{F}^{\text{Cred}}_i \), which depends in an affine way on the observations \( F_i,k = C_i,k + 1 \).

- Note, we know \( E[\varphi_k] = f_k \), but usually \( E[\varphi_k|D_I] \neq f_k \).
- Even if we have a good model for the joint distribution of \((C_i,k)_{i+k\leq I}\) and \(\varphi\), the calculation of posteriori distributions is very hard, since we have only very few data.
- Looking at the credibility estimator instead of the Bayesian estimator means to look at the a orthogonal projection on the affine subspace of \(D_I\) generated by the link ratios \( F_i,k \) instead of the projection onto \(D_I\) itself.
Definition 8.6 (Credibility estimators of the development factors)

\[ \hat{F}_k^{\text{Cred}} := \arg\min_{\hat{\varphi}} \mathbb{E} \left[ (\varphi_k - \hat{\varphi})^2 \bigg| \mathcal{D}_I \right] \]

Theorem 8.7 (Credibility estimator for the development factors)

Let Assumption 8.A be fulfilled. Then

- the credibility estimators of the development factors are given by

\[ F_k^{\text{Cred}} = \alpha_k \hat{f}_k^{\text{CLM}} + (1 - \alpha_k) f_k, \quad \text{with} \quad \alpha_k := \frac{\sum_{i=0}^{I-k-1} C_{i,k}}{\sum_{i=0}^{I-k-1} C_{i,k} + \sigma_k^2 \tau_k^2}, \]

where \( f_k := \mathbb{E}[\varphi_k], \sigma_k^2 := \mathbb{E}[\sigma_k^2(\varphi)], \tau_k^2 := \text{Var}[\varphi_k] \) and

\[ \hat{f}_k^{\text{CLM}} := \sum_{i=0}^{I-k-1} \frac{C_{i,k}}{\sum_{h=0}^{I-k-1} \frac{C_{h,k}}{C_{i,k}}} \frac{C_{i,k+1}}{C_{i,k}}. \]

- the corresponding mean squared error of prediction is given by

\[ \text{mse}_{\mathcal{D}_k} \left[ F_k^{\text{Cred}} \right] := \mathbb{E} \left[ (\varphi_k - F_k^{\text{Cred}})^2 \bigg| \mathcal{D}_k \right] = \alpha_k \frac{\sigma_k^2 \tau_k^2}{\sum_{i=0}^{I-k-1} C_{i,k}} = (1 - \alpha_k) \tau_k^2. \]
Stochastic Reserving

**CLM: Bayesian & credibility approach**

A credibility approach to the Chain-Ladder method

- Conditionally given $D_k$, the random variables $F_{i,k} = \frac{C_{i,k}+1}{C_{i,k}}$, $i = 0, \ldots, I - k - 1$, fulfill the assumptions of the Bühlmann and Straub model (see [28, Section 4.2]). The first part of the theorem is the well known credibility estimator of Bühlmann and Straub and the second part is the corresponding mean square error of prediction (see [28, Chapter 4]).

- The case $\tau_k^2 \to \infty$, i.e. $\alpha_k = 1$, is called the non-informative priors. It corresponds to the standard Chain-Ladder method introduced in Section 2.

- Since $F_{k}^{cred}$ still depends on the unknown expectation $f_k = \mathbb{E}[\varphi_k]$ we don’t mark it with a hat like other estimator.
Estimator 8.8 (Credibility estimator of the future development)

\[ \hat{C}^{\text{Cred}}_{i,k} := \hat{F}^{\text{Cred}}_{k-1} \cdot \ldots \cdot \hat{F}^{\text{Cred}}_{I-i} C_{i,I-i}, \quad \text{for } i + k > I. \]

Estimation of the structural parameters \( f_k, \sigma^2_k, \) and \( \tau^2_k \), see [28, Section 4.8]

Either ask experts or if we have several similar portfolios \( C^m_{i,k}, 0 \leq m \leq M \), we can take

\[ \hat{F}^{m,\text{Cred}}_k := \hat{\alpha}_k f^{m,\text{CLM}}_k + (1 - \hat{\alpha}_k) \hat{f}_k \quad \text{and} \quad \hat{C}^{m,\text{Cred}}_{i,k} := \hat{F}^{m,\text{Cred}}_{k-1} \cdot \ldots \cdot \hat{F}^{m,\text{Cred}}_{I-i} C^m_{i,I-i} \]

with

\[ \hat{f}_k := \begin{cases} \frac{\sum_{m=0}^M \hat{\alpha}_k f^{m,\text{CLM}}_k}{\sum_{m=0}^M \hat{\alpha}_k}, & \text{if } \sum_{m=0}^M \hat{\alpha}_k <> 0, \\ \hat{f}^{\text{tot,CLM}}_k, & \text{otherwise}, \end{cases} \]

\[ \hat{\alpha}_k := \frac{\omega^m_{\cdot,k}}{\omega^m_{\cdot,k} + \frac{\sigma^2_k}{\tau^2_k}} \quad (:= 0, \text{ if } \tau^2_k = 0), \]

\[ \hat{\tau}^2_k := \max \left\{ 0; c_k \left( \frac{M+1}{M} \sum_{m=0}^M \frac{\omega^m_{\cdot,k}}{\omega^m_{\cdot,k}} \left( \hat{f}^{m,\text{CLM}}_k - \hat{f}^{\text{tot,CLM}}_k \right)^2 - \frac{(M+1)\sigma^2_k}{\omega^m_{\cdot,k}} \right) \right\}, \]

\[ \hat{\sigma}^2_k := \frac{1}{M+1} \sum_{m=0}^M \frac{1}{I - k - 1} \sum_{i=0}^{I-k-1} C^m_{i,k} \left( \frac{C^m_{i,k+1}}{C^m_{i,k}} - \hat{f}^{m,\text{CLM}}_k \right)^2, \]

\[ \omega^m_{\cdot,k} := \sum_{i=0}^{I-k-1} C^m_{i,k} \quad \text{and} \quad \omega^m_{\cdot,k} := \sum_{m=0}^M \omega^m_{\cdot,k}. \]
- In the case of non-informative priors, i.e. \( \tau_k^2 \rightarrow \infty \), the estimators of the future development are the same as for the standard Chain-Ladder method introduced in Section 2.
- \( \hat{f}_{tot, CLM} = \frac{\omega_{i,k}}{\omega_{i,k}} \) are the standard estimates of the development factors of the combined portfolio \( \sum_{m=0}^{M} C_{i,k}^m \).
- The factors \( c_k \) are normalizing factors that makes the estimators \( \hat{\tau}_k \) unbiased (conditioned \( \hat{\tau}_k > 0 \)).
Estimator 8.9 (of the ultimate uncertainty)

Let Assumption 8.A be fulfilled. Then the ultimate uncertainty is given by

$$\text{mse}_{D^I} \left[ \sum_{i=0}^{I} \hat{C}_{i,J}^{\text{Cred}} \right] = \sum_{i=0}^{I} \sum_{k=I-i}^{J-1} \prod_{j=k+1}^{J-1} \mathbb{E} \left[ \varphi_j^2 \mid D^I \right] \mathbb{E} \left[ \sigma_k^2 \mid D^I \right] \prod_{j=I-i}^{k-1} \mathbb{E} \left[ \varphi_j \mid D^I \right] C_{i,I-i}$$

$$+ \sum_{i_1,i_2=0}^{I} C_{i_1,I-i_1} C_{i_2,I-i_2} \mathbb{E} \left[ \left( \prod_{k=I-i_1}^{J-1} \hat{F}_k^{\text{Cred}} - \prod_{k=I-i_1}^{J-1} \varphi_k \right) \left( \prod_{k=I-i_2}^{J-1} \hat{F}_k^{\text{Cred}} - \prod_{k=I-i_2}^{J-1} \varphi_k \right) \mid D^I \right]$$

Remark 8.10 (connection to the standard CLM)

In the case of non-informative priors, i.e. $\tau_k^2 \to \infty$, the random error is slightly bigger than in the standard CLM case, whereas the parameter error is the same.
First, like in the Bayesian case, we decompose the mse
\[
\text{mse}_\mathcal{D} \left[ \sum_{i=0}^{I} C_{i,J}^{\text{cred}} \right] = \sum_{i=0}^{I} \text{E} \left[ \text{Var} \left[ C_{i,J} \mid \varphi, \mathcal{D}^I \right] \right] \mathcal{D}^I + \text{E} \left[ \left( \sum_{i=0}^{I} \left( \text{E} \left[ C_{i,J} \mid \varphi, \mathcal{D}^I \right] - C_{i,J}^{\text{cred}} \right) \right)^2 \right] \mathcal{D}^I.
\]

The random error is the same like in the Bayesian case and for the second term we take the summation out of the expectation.

In order to estimate it we take
\[
\text{E} \left[ \sigma_k^2 (\varphi) \mid \mathcal{D}^I \right] \approx \sigma_k^2 \quad \text{and} \quad \text{E} \left[ \varphi_k \mid \mathcal{D}^I \right] \approx \hat{\alpha}_k C_{i,J}^{\text{cred}}.
\]

Moreover, we estimate
\[
\text{E} \left[ \varphi_j^2 \mid \mathcal{D}^I \right] = \left( \varphi_j - \text{E} \left[ \varphi_j \mid \mathcal{D}^I \right] \right)^2 + \left( \text{E} \left[ \varphi_k \mid \mathcal{D}^I \right] \right)^2 \approx \hat{\alpha}_j \frac{\hat{\sigma}_j^2}{\sum_{i=0}^{I-j-1} C_{i,j}} + (\hat{F}_j^{\text{cred}})^2.
\]

Finally, we compute
\[
\text{E} \left[ \left( \frac{J-1}{I-i_1} \prod_{k=I-i_1}^{J-1} F_k^{\text{cred}} - \frac{J-1}{I-i_1} \prod_{k=I-i_1}^{J-1} \varphi_k \right) \left( \frac{J-1}{I-i_2} \prod_{k=I-i_2}^{J-1} F_k^{\text{cred}} - \prod_{k=I-i_2}^{J-1} \varphi_k \right) \mid \mathcal{D}^I \right]
\]
\[
\approx \text{E} \left[ \left( \frac{J-1}{I-i_1} \prod_{k=I-i_1}^{J-1} F_k^{\text{Bay}} - \prod_{k=I-i_1}^{J-1} \varphi_k \right) \left( \frac{J-1}{I-i_2} \prod_{k=I-i_2}^{J-1} F_k^{\text{Bay}} - \prod_{k=I-i_2}^{J-1} \varphi_k \right) \mid \mathcal{D}^I \right] = \text{Cov} \left[ \frac{J-1}{I-i_1} \prod_{k=I-i_1}^{J-1} \varphi_k, \prod_{k=I-i_2}^{J-1} \varphi_k \mid \mathcal{D}^I \right]
\]
\[
= \prod_{k=I-(i_1 \lor i_2)}^{I-(i_1 \land i_2)-1} \text{E} \left[ \varphi_k \mid \mathcal{D}^I \right] \left( \frac{J-1}{I-(i_1 \land i_2)} \prod_{k=I-(i_1 \land i_2)}^{J-1} \text{E} \left[ \varphi_k \mid \mathcal{D}^I \right] \right) - \prod_{k=I-(i_1 \land i_2)}^{J-1} \text{E} \left[ \varphi_k \mid \mathcal{D}^I \right]^2
\]
and replace all unknown parameters by they estimates and take the factors \( \hat{\alpha}_i C_{i,J}^{\text{cred}} \) and \( \hat{\alpha}_1 C_{i_1,J}^{\text{cred}} \) out.
Pricing of similar subportfolios

- In [27] an example of a portfolio was discussed that consists of six subportfolios, ‘BU A’…‘BU F’. Results and figures are copied from this article.
- For reserving we would usually combine all six of them to get the law of large numbers more volume to get working.
- But in pricing we need individual premiums for each subportfolio.
- One way to do so is to use the introduced credibility reserving.

<table>
<thead>
<tr>
<th>BU</th>
<th>reserves</th>
<th>√mse</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CLM</td>
<td>Cred</td>
</tr>
<tr>
<td>A</td>
<td>486</td>
<td>504</td>
</tr>
<tr>
<td>B</td>
<td>235</td>
<td>244</td>
</tr>
<tr>
<td>C</td>
<td>701</td>
<td>517</td>
</tr>
<tr>
<td>D</td>
<td>1029</td>
<td>899</td>
</tr>
<tr>
<td>E</td>
<td>495</td>
<td>621</td>
</tr>
<tr>
<td>F</td>
<td>40</td>
<td>25</td>
</tr>
<tr>
<td>sum</td>
<td>2987</td>
<td>2810</td>
</tr>
</tbody>
</table>

For overall CLM

<table>
<thead>
<tr>
<th>overall CLM</th>
<th>2746</th>
<th>1418</th>
</tr>
</thead>
<tbody>
<tr>
<td>LSRM</td>
<td>2987</td>
<td>1353</td>
</tr>
</tbody>
</table>

For LSRM we coupled the individual Chain-Ladder projections by $R_{i,k}^{m_1,m_2} := \sqrt{C_{i,k}^{m_1} C_{i,k}^{m_2}}$. 
- The total reserves differ only by 6%, but per subportfolio the differences are much larger (up to 46%).
- The CLM reserves for the combined portfolio are even smaller.
- The mse of the combined portfolios is about 25% larger than the sum of the individual ones. This may be a hint that the estimated reserves of the subportfolios are correlated.
- The LSRM leads to almost the same results as the overall CLM.
- In the file ‘Example_Cor_Dll.xlsx’ (or ‘Example_Cor_ActiveX.xlsx’), see Example 147, the CLM and the LSRM estimates are (re)calculated. The presented figures for CLM, which are taken from the original article [27], differ slightly from the recalculated once, because of rounding effects.
Correlation of the estimated reserves

<table>
<thead>
<tr>
<th>BU</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1.00</td>
<td>-0.15</td>
<td>0.01</td>
<td>0.23</td>
<td>-0.17</td>
<td>0.26</td>
</tr>
<tr>
<td>B</td>
<td>-0.15</td>
<td>1.00</td>
<td>0.03</td>
<td>0.13</td>
<td>-0.03</td>
<td>-0.00</td>
</tr>
<tr>
<td>C</td>
<td>0.01</td>
<td>0.03</td>
<td>1.00</td>
<td>0.04</td>
<td>0.06</td>
<td>-0.05</td>
</tr>
<tr>
<td>D</td>
<td>0.23</td>
<td>0.13</td>
<td>0.04</td>
<td>1.00</td>
<td>-0.05</td>
<td>0.09</td>
</tr>
<tr>
<td>E</td>
<td>-0.17</td>
<td>-0.03</td>
<td>0.06</td>
<td>-0.05</td>
<td>1.00</td>
<td>0.03</td>
</tr>
<tr>
<td>F</td>
<td>0.26</td>
<td>-0.00</td>
<td>-0.05</td>
<td>0.09</td>
<td>0.03</td>
<td>1.00</td>
</tr>
</tbody>
</table>

We see that at least the estimated reserves for subportfolio BU A are correlated to the others.
We see that at least the estimated reserves for subportfolio BU A are correlated to the others.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1.00</td>
<td>-0.15</td>
<td>0.01</td>
<td>0.23</td>
<td>-0.17</td>
<td>0.26</td>
</tr>
<tr>
<td>B</td>
<td>-0.15</td>
<td>1.00</td>
<td>0.03</td>
<td>0.13</td>
<td>-0.03</td>
<td>0.00</td>
</tr>
<tr>
<td>C</td>
<td>0.01</td>
<td>0.03</td>
<td>1.00</td>
<td>0.04</td>
<td>0.06</td>
<td>0.05</td>
</tr>
<tr>
<td>D</td>
<td>0.23</td>
<td>0.13</td>
<td>0.04</td>
<td>1.00</td>
<td>-0.05</td>
<td>0.09</td>
</tr>
<tr>
<td>E</td>
<td>-0.17</td>
<td>-0.03</td>
<td>0.06</td>
<td>-0.05</td>
<td>1.00</td>
<td>0.03</td>
</tr>
<tr>
<td>F</td>
<td>0.26</td>
<td>-0.00</td>
<td>-0.05</td>
<td>0.09</td>
<td>0.03</td>
<td>1.00</td>
</tr>
</tbody>
</table>
Comparison of the estimated development pattern (1/2)

The individual CLM development pattern are smoothed by the credibility approach:
Comparison of the estimated development pattern (1/2)
The individual CLM development pattern are smoothed by the credibility approach.
Comparison of the estimated development pattern (2/2)

The credibility approach shifts the individual CLM development pattern into the direction of the overall CLM pattern:
Comparison of the estimated development pattern (2/2)
The credibility approach shifts the individual CLM development pattern into the direction of the overall CLM pattern.
Literature

[28] Aloise Gisler and Mario V. Wüthrich.  
Credibility for the Chain Ladder Reserving Method.  

[29] Bühlmann, H. and Gisler, A.  
_A Course in Credibility Theory and its Applications._  
Universitext, Springer Verlag, 2005.
Literature


9 Mid year reserving

9.1 Problem of mid-year reserving

9.2 Methods for mid-year reserving
9.2.1 Splitting or shifting of development periods
9.2.2 Extrapolation of the last diagonal
9.2.3 Shifting accident periods
9.2.4 Splitting of accident periods
9.2.5 Separating semesters
9.2.6 Separating the youngest semester

9.3 Conclusion

9.4 Literature
Stochastic Reserving

Lecture 13: Table of contents
### Chain-Ladder method at year end

<table>
<thead>
<tr>
<th>accident years (periods)</th>
<th>development month (periods)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>12 (0) 24 (1) 36 (2)</td>
</tr>
<tr>
<td></td>
<td>f₀ = 2.5</td>
</tr>
<tr>
<td></td>
<td>f₁ = 1.4</td>
</tr>
<tr>
<td>0</td>
<td>100 → 250 → 350</td>
</tr>
<tr>
<td>1</td>
<td>200 500 700</td>
</tr>
<tr>
<td>2</td>
<td>260 650 910</td>
</tr>
</tbody>
</table>

### Chain-Ladder method at mid-year

<table>
<thead>
<tr>
<th>accident years (periods)</th>
<th>development month (periods)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>12 24 36</td>
</tr>
<tr>
<td></td>
<td>0 100 250 350</td>
</tr>
<tr>
<td>1</td>
<td>1 200 500 650</td>
</tr>
<tr>
<td>2</td>
<td>2 260 455 75</td>
</tr>
</tbody>
</table>

### Chain-Ladder assumptions (Mack [21]):

- \( \mathbb{E}[C_{i,k+1}|B_{i,k}] = f_k C_{i,k} \)
- \( \text{Var}[C_{i,k+1}|B_{i,k}] = \sigma_k^2 C_{i,k} \)
- independent accident years (periods)

### Additional comments:
- additional semester of experience
- new cells are incomplete

\[ \Rightarrow \text{years are not comparable} \]
\[ \Rightarrow \text{Chain-Ladder will not work.} \]
### Stochastic Reserving

#### Mid year reserving

- **Problem of mid-year reserving**

#### Chain-Ladder method at year end

<table>
<thead>
<tr>
<th>Accident Years</th>
<th>Development Month (Periods)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>12 (0)</td>
</tr>
<tr>
<td>1</td>
<td>24 (1)</td>
</tr>
<tr>
<td>2</td>
<td>36 (2)</td>
</tr>
</tbody>
</table>

#### Chain-Ladder assumptions (Mack [21]):

- $\text{E}[C_{i,k+1}|B_{i,k}] = f_k C_{i,k}$
- $\text{Var}[C_{i,k+1}|B_{i,k}] = \sigma^2_k C_{i,k}$
- independent accident years (periods)

#### Chain-Ladder method at mid-year

<table>
<thead>
<tr>
<th>Accident Years</th>
<th>Development Month (Periods)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>12 (0)</td>
</tr>
<tr>
<td>1</td>
<td>24 (1)</td>
</tr>
<tr>
<td>2</td>
<td>36 (2)</td>
</tr>
</tbody>
</table>

- additional semester of experience
- new cells are incomplete
- years are not comparable
- Chain-Ladder will not work.
Problems

- forecast or closing
  * If the method produces estimates for a closing the second semester of the latest accident year is missing for a forecast estimate.
  * If the method produces estimates for a forecast the estimated ultimate for the latest accident year contains the estimate for the second semester, which has to be eliminated for a mid-year closing.

- generalisation to other dates during the year

- consistency at year end

- usability:
  * discussion of the claims development result
  * comparability of observed development factors
  * comparability of estimated development factors
  * estimation error (ultimate and solvency uncertainties)
  * additional workload
If we have estimates for a forecast of the next year end closing then the estimated ultimate for the latest accident year contains the corresponding second semester. Usually, this has to be eliminated from the estimates if we want to use it for a mid-year closing (under USGAAP, PAA under IFRS 17 and many other accounting standards). Such an elimination is not always easy. Often one looks in the history to get an ‘first to second semester ratio’ which is then applied at the forecast estimate of the latest accident year.

But one has to be careful. For instance, assume we expect one large claim per accident year. What do we do at end of June if

- we already observed one large claim for the latest accident year?
  - We should not transfer any part of this large claim into the second semester!
  - Should we account for the possibility of another large claim via IBNe/yR?

- we have not observed any large claim for the latest accident year?
  - How much of the IBNe/yR for large claims should we take into account for the first semester?
Assume we have complete data for each semester.

<table>
<thead>
<tr>
<th>Year</th>
<th>12</th>
<th>24</th>
<th>36</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>100</td>
<td>250</td>
<td>350</td>
</tr>
<tr>
<td>1</td>
<td>200</td>
<td>500</td>
<td>650</td>
</tr>
<tr>
<td>2</td>
<td>260</td>
<td>455</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>75</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[9\text{ Mid year reserving} \]

9.1 Problem of mid-year reserving (3/4)
Assume we have complete data for each semester

For the numerical example we took for each accident semester the following non-random development pattern

<table>
<thead>
<tr>
<th>development month</th>
<th>6</th>
<th>12</th>
<th>18</th>
<th>24</th>
<th>30</th>
<th>36</th>
<th>42</th>
</tr>
</thead>
<tbody>
<tr>
<td>cumulative</td>
<td>25</td>
<td>75</td>
<td>100</td>
<td>150</td>
<td>175</td>
<td>175</td>
<td>175</td>
</tr>
<tr>
<td>incremental</td>
<td>25</td>
<td>50</td>
<td>25</td>
<td>50</td>
<td>25</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

and accident semester volumes

<table>
<thead>
<tr>
<th>accident semester</th>
<th>1H 0</th>
<th>2H 0</th>
<th>1H 1</th>
<th>2H 1</th>
<th>1H 2</th>
<th>2H 2</th>
<th>1H 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>volume</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2.6</td>
<td>2.6</td>
<td>3</td>
</tr>
</tbody>
</table>

We take this easy and non-random example in order to illustrate issues and possible solutions. A more realistic example with random data would make it much harder to understand the effects. Moreover, we cannot expect that a method will work fine in practice, if it fails (to some degree) for such an easy example.
Problem 9.1 (Mid-year reserving)

What can we do at the end of the first semester in order to estimate reserves that correspond to the reserves at year end, which are estimated by Chain-Ladder on the basis of the 12x12 triangle (12 accident months within one accident period and the same for development periods)?
Stochastic Reserving

Mid year reserving

Problem of mid-year reserving

Problem 9.1 (Mid-year reserving)

What can we do at the end of the first semester in order to estimate reserves that correspond to the reserves at year end, which are estimated by Chain-Ladder on the basis of the 12x12 triangle (12 accident months within an accident period and the same for development periods)?
### Step by step

<table>
<thead>
<tr>
<th></th>
<th>12</th>
<th>24</th>
<th>36</th>
<th>48</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>25</td>
<td>75</td>
<td>100</td>
<td>150</td>
</tr>
<tr>
<td>1</td>
<td>25</td>
<td>75</td>
<td>100</td>
<td>150</td>
</tr>
<tr>
<td>2</td>
<td>25</td>
<td>75</td>
<td>100</td>
<td>150</td>
</tr>
<tr>
<td>3</td>
<td>25</td>
<td>75</td>
<td>100</td>
<td>150</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>24</th>
<th>36</th>
<th>48</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>100</td>
<td>150</td>
<td>200</td>
</tr>
<tr>
<td>1</td>
<td>100</td>
<td>150</td>
<td>200</td>
</tr>
<tr>
<td>2</td>
<td>100</td>
<td>150</td>
<td>200</td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>150</td>
<td>200</td>
</tr>
</tbody>
</table>

### Results

- **Ultimate:** PY = 1960, CY = 1050, Total = 3010 ✓ for forecasts
- **Reserves:** PY = 505, CY = 975, Total = 1480
- The development factors in the third triangle are the products of two corresponding development factors of the second, i.e.

\[
7 = \frac{7}{4}, \quad \frac{13}{7} = \frac{10}{7} \cdot \frac{13}{10}, \quad \frac{14}{13} = \frac{14}{13} \cdot 1.
\]
From an ultimate point of view, it does not matter if we look at development periods

- 6, 12, 18, 24 . . . , or
- 6, 18, 30 . . .
Properties

- results in a forecast
- easy to generalise to other dates during the year 😊
- it is consistent with the yearly Chain-Ladder at year end, because shifting and splitting results in the same (estimated) ultimates 😊
- usability:
  - claims development result can be discussed 😊
  - observed and estimated development factor can only be compared if we use split development periods, but this goes along with much larger triangles
  - although, in theory the estimated prediction errors are the same for split and shifted data in practice often less values for split data are observed 😞
  - split data triangles can get very huge, for instance for a forecast at the end of November 😞
Denote by $C_{i,k}$ the cumulative values for accident year $i$ at the end of development semester $k$ and by $C_{i,k}^*$ the cumulative values for accident year $i$ at the end of development year $k$. Moreover, let
\[
B_{i,k} := \sigma \left( C_{i,j}, 0 \leq j \leq k \right) \quad \text{and} \quad B_{i,k}^* := \sigma \left( C_{i,j}^*, 0 \leq j \leq k \right)
\]
the corresponding information of the past. Then we have
\[
C_{i,k}^* = C_{i,2k+1} \quad \text{and} \quad B_{i,k}^* \subseteq B_{i,2k+1}.
\]
Assume that the semester data $C_{i,k}$ satisfies the Chain-Ladder assumptions, i.e.
\begin{itemize}
  \item $E[C_{i,k+1} | B_{i,k}] = f_k C_{i,k}$
  \item $\text{Var}[C_{i,k+1} | B_{i,k}] = \sigma_k^2 C_{i,k}$
  \item accident years are independent.
\end{itemize}
Then $C_{i,k}^*$ satisfies the Chain-Ladder assumptions, too:
\begin{itemize}
  \item $E[C_{i,k+1} | B_{i,k}^*] = E\left[ E[C_{i,2(k+1)+1} | B_{i,2k+1}] | B_{i,k}^* \right] = E\left[ f_{2k+2} f_{2k+1} C_{i,2k+1} | B_{i,k}^* \right] = f_{2k+2} f_{2k+1} C_{i,k}^*$
  \item $\text{Var}[C_{i,k+1} | B_{i,k}^*] = E\left[ \text{Var}[C_{i,2(k+1)+1} | B_{i,2k+2}] | B_{i,k}^* \right] + \text{Var}\left[ E[C_{i,2(k+1)+1} | B_{i,2k+2}] | B_{i,k}^* \right]$
    \[= E\left[ \sigma_{2k+2}^2 C_{i,2k+2} | B_{i,k}^* \right] + \text{Var}\left[ f_{2k+2} C_{i,2k+2} | B_{i,k}^* \right] = \sigma_{2k+2}^2 E\left[ C_{i,2k+2} | B_{i,k+1} \right] | B_{i,k}^* \right] + \text{Var}\left[ f_{2k+2} C_{i,2k+2} | B_{i,k}^* \right]$
    \[= \sigma_{2k+2}^2 f_{2k+1} C_{i,k}^* + f_{2k+2}^2 \left( 0 + \sigma_{2k+2}^2 C_{i,k}^* \right) = \sigma_{2k+2}^2 f_{2k+1} + f_{2k+2}^2 \sigma_{2k+1}^2 \]
  \item accident years are independent.
\end{itemize}
But in practice one often observes $\sigma_{k}^2 > \sigma_{2k+2}^2 f_{2k+1} + f_{2k+2}^2 \sigma_{2k+1}^2$
### Step by step

<table>
<thead>
<tr>
<th></th>
<th>6</th>
<th>12</th>
<th>18</th>
<th>24</th>
<th>30</th>
<th>36</th>
<th>42</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>25</td>
<td>75</td>
<td>100</td>
<td>150</td>
<td>175</td>
<td>175</td>
<td>175</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>75</td>
<td>100</td>
<td>150</td>
<td>175</td>
<td>175</td>
<td>175</td>
</tr>
<tr>
<td>1</td>
<td>50</td>
<td>150</td>
<td>200</td>
<td>300</td>
<td>350</td>
<td>700</td>
<td></td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>150</td>
<td>200</td>
<td>300</td>
<td>350</td>
<td>700</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>65</td>
<td>195</td>
<td>260</td>
<td>455</td>
<td>650</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>65</td>
<td>195</td>
<td>260</td>
<td>455</td>
<td>650</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>75</td>
<td>300</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Results

- **Ultimate**: PY = 1960, CY = 1050, Total = 3010 ✓ for forecasts
- **Reserves**: PY = 505, CY = 975, Total = 1480
- The (estimated) development in the third picture are the product of the (estimated) development factors of the triangle with split development periods.
Stochastic Reserving

Mid year reserving

Methods for mid-year reserving

Results:
- Ultimate: PY = 1960, CY = 1050, Total = 3010 ✓ for forecasts
- Reserves: PY = 505, CY = 975, Total = 1480
- The (estimated) development in the third picture are the product of the (estimated) development factors of the triangle with split development periods.
Properties

- results in a forecast
- easy to generalise to other dates during the year
- it is consistent (end in some way optimal) with the yearly Chain Ladder at year end
- usability:
  - claims development result can be discussed
  - observed and estimated development factor of the last picture are the same as at year end
  - since ultimates are the same as for split or shifted development periods, the same estimates for prediction errors can be used
  - the estimated development factors are almost the best predictions of the corresponding estimates of the following year end closure
  - not so easy to implement with standard reserving software
Using the same notation like in the case of split development periods we get

\[
\hat{f}_k = \frac{\sum_{i=0}^{I-k} \hat{C}_{i,k+1}}{\sum_{i=0}^{I-k} C^*_{i,k}} = \frac{\sum_{i=0}^{I-k-1} C^*_{i,k+1} + \hat{f}_{hy} C_{I-k,2k+2}}{\sum_{i=0}^{I-k} C^*_{i,k}} \\
= \frac{\sum_{i=0}^{I-k-1} \hat{f}_{ye} C^*_{i,k}}{\sum_{i=0}^{I-k} C^*_{i,k}} + \left(1 - \frac{\sum_{i=0}^{I-k-1} C^*_{i,k}}{\sum_{i=0}^{I-k} C^*_{i,k}}\right) \frac{C_{I-k,2k+2}}{C_{I-k,2k+1}} \hat{f}_{hy} 2k+2.
\]

That means the estimated development factors \(\hat{f}_k\) are a weighted mean of the estimated development factors \(\hat{f}_{ye}\) from last year end closing and the newly observed development \(\frac{C_{I-k,2k+2}}{C_{I-k,2k+1}}\) multiplied by the estimated development of the second half year \(\hat{f}_{hy}\).

Moreover, one can show that the weights \(\frac{\sum_{i=0}^{I-k-1} C^*_{i,k}}{\sum_{i=0}^{I-k} C^*_{i,k}}\) are almost the best weights \(\alpha_k\) in order to forecast the estimated development factors \(\hat{f}_{k+1}\) of the next year end closing, i.e. \(\alpha_k\) that minimize (see [29] for details)

\[
E \left[ \left( (1 - \alpha_k) \hat{f}_{ye} + \alpha_k \frac{C_{I-k,2k+2}}{C_{I-k,2k+1}} \hat{f}_{hy} 2k+2 - \hat{f}_{k+1} \right) ^2 \right] C_{i,j} \text{ known at end of June}.
\]
### Step by step

<table>
<thead>
<tr>
<th>Step</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>25</td>
<td>50</td>
<td>65</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td></td>
<td>75</td>
<td>150</td>
<td>195</td>
<td>260</td>
<td></td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>200</td>
<td>260</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>25</td>
<td>25</td>
<td>125</td>
<td>215</td>
<td></td>
</tr>
<tr>
<td></td>
<td>75</td>
<td>75</td>
<td>350</td>
<td>560</td>
<td></td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>100</td>
<td>175</td>
<td>840</td>
<td></td>
</tr>
<tr>
<td></td>
<td>150</td>
<td>150</td>
<td>175</td>
<td></td>
<td>805</td>
</tr>
<tr>
<td>2</td>
<td>25</td>
<td>50</td>
<td>66</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td></td>
<td>75</td>
<td>150</td>
<td>195</td>
<td>260</td>
<td></td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>200</td>
<td>260</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>25</td>
<td>25</td>
<td>125</td>
<td>215</td>
<td></td>
</tr>
<tr>
<td></td>
<td>75</td>
<td>75</td>
<td>175</td>
<td>1084</td>
<td></td>
</tr>
<tr>
<td></td>
<td>75</td>
<td>175</td>
<td>175</td>
<td>980</td>
<td></td>
</tr>
</tbody>
</table>

### Results

- **Ultimate**: “PY” = 1505, “CY” = 980, Total = 2514 (2485)
- **Reserves**: “PY” = 245, “CY” = 710, Total = 994 (955)
- correct values in red
- should give estimates for closings, but only if ‘volumes are stable’
Stochastic Reserving

- Mid year reserving
- Methods for mid-year reserving

Results:
- Ultimate: "PY" = 1505, "CY" = 980, Total = 2514 (2485)
- Reserves: "PY" = 245, "CY" = 710, Total = 994 (955)
- Correct values in red
- Should give estimates for closings, but only if 'volumes are stable'
Properties

- results in closing figures, but only if ‘volumes are stable’ 😞
- easy to generalise to other dates during the year 😊
- it is not consistent with the yearly Chain Ladder at year end 😞
- usability:
  - a discussion of the claims development result is almost impossible 😞 😞
  - observed and estimated development factors at mid year and at year end are not alike 😞
  - estimation errors can be estimated by the standard formulas 😊
  - may be useful in a merger and acquisition process at mid year, if no other information except for triangles are available
The method is inconsistent with the yearly Chain-Ladder for the same reasons as the method of split accident years, see section 4.
9 Mid year reserving

9.2 Methods for mid-year reserving

9.2.4 Splitting of accident periods

Step by step

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>25</td>
<td>75</td>
<td>100</td>
<td>150</td>
</tr>
<tr>
<td>1</td>
<td>25</td>
<td>75</td>
<td>100</td>
<td>150</td>
</tr>
<tr>
<td>2</td>
<td>25</td>
<td>75</td>
<td>100</td>
<td>150</td>
</tr>
<tr>
<td>3</td>
<td>25</td>
<td>75</td>
<td>100</td>
<td>150</td>
</tr>
</tbody>
</table>

Results

- **Ultimate**: PY = 1960, CY = 525, Total = 2485 ✓ for closings
- **Reserves**: PY = 505, CY = 450, Total = 955
Stochastic Reserving

- Mid year reserving

- Methods for mid-year reserving

Results:

- Ultimate: PY = 1960, CY = 525, Total = 2485 ✓ for closings
- Reserves: PY = 515, CY = 450, Total = 955
Properties

- results in closing figures
- easy to generalise to other dates during the year 😊
- it is, except for strange situation, not consistent with the yearly Chain Ladder at year end 😞
- Usability:
  - claims development result can be discussed 😊
  - observed and estimated development factor can only be compared if we always use the same split, but this comes along much larger triangles
  - uncertainties can be estimated by standard formulas (not by the original formulas of Mack [21] and Merz-Wüthrich [17], but by the formulas of Röhr [30])
  - split data triangles can get very huge, for instance for a forecast at the end of November 😞
Denote by $C_{i,k}$ the cumulative values for accident semester $i$ at the end of development semester $k$ and by $C^*_{i,k}$ the cumulative values for accident year $i$ at the end of development year $k$. Moreover, let

$$B_{i,k} := \sigma \left( C_{i,j}, 0 \leq j \leq k \right) \quad \text{and} \quad B^*_{i,k} := \sigma \left( C^*_{i,j}, 0 \leq j \leq k \right)$$

the corresponding information of the past. Then we have

$$C^*_{i,k} = C_{2i,2k+1} + C_{2i+1,2k} \quad \text{and} \quad B^*_{i,k} \subseteq \sigma \left( B_{2i,2k+1} \cup B_{2i+1,2k} \right).$$

Assume that $C_{i,k}$ and $C^*_{i,k}$ satisfy the Chain-Ladder assumptions, i.e.

- $E \left[ C_{i,k+1} \bigg| B_{i,k} \right] = f_k C_{i,k}$
- $\text{Var} \left[ C_{i,k+1} \bigg| B_{i,k} \right] = \sigma_k^2 C_{i,k}$
- accident semester are independent.

- $E \left[ C^*_{i,k+1} \bigg| B^*_{i,k} \right] = g_k C^*_{i,k}$
- $\text{Var} \left[ C^*_{i,k+1} \bigg| B^*_{i,k} \right] = \tau_k^2 C^*_{i,k}$
- accident years are independent.

Then we get:

$$g_k \left( C_{2i,2k+1} + C_{2i+1,2k} \right) = g_k C^*_{i,k} = E \left[ C_{2i,2(k+1)+1} + C_{2i+1,2(k+1)} \bigg| B^*_{i,k} \right]$$

$$= E \left[ E \left[ C_{2i,2(k+1)+1} + C_{2i+1,2(k+1)} \bigg| B_{2i,2k+1}, B_{2i+1,2k} \right] \bigg| B^*_{i,k} \right]$$

$$= E \left[ f_{2k+2} f_{2k+1} C_{2i,2k+1} + f_{2k+1} f_{2k} C_{2i+1,2k} \bigg| B^*_{i,k} \right]. \quad (9.1)$$

Therefore, it follows

$$0 = (g_k - f_{2k+2} f_{2k+1}) E \left[ C_{2i,2k+1} \bigg| D \right] + (g_k - f_{2k+1} f_{2k}) E \left[ C_{2i+1,2k} \bigg| D \right], \quad (9.2)$$

for each $\sigma$-algebra $D \subseteq B^*_{i,k}$. 
Moreover multiplying (9.1) by \((C_{2i,2k+1} + C_{2i+1,2k})\), resorting the terms and using (9.2) we get and

\[
0 = (g_k - f_{2k+2}f_{2k+1}) \text{Var}[C_{2i,2k+1}\mid \mathcal{D}] + (g_k - f_{2k+1}f_{2k}) \text{Var}[C_{2i+1,2k}\mid \mathcal{D}],
\]

which is only possible if

- \(\text{Var}[C_{i,k}] = 0\), which means that there is no randomness,
- \(f_{2k+2} = f_{2k}\), which in practice implies \(f_{2k+2} = f_{2k+1} = f_{2k} = 1\), or
- 

\[
\frac{\text{Var}[C_{2i,2k+1}\mid \mathcal{D}]}{f_{2k} \text{Var}[C_{2i+1,2k}\mid \mathcal{D}]} = \frac{g_k - f_{2k+1}f_{2k}}{f_{2k}(g_k - f_{2k+2}f_{2k+1})} = \frac{\mathbb{E}[C_{2i,2k+1}\mid \mathcal{D}]}{f_{2k} \mathbb{E}[C_{2i+1,2k}\mid \mathcal{D}]} = \frac{\mathbb{E}[C_{2i,0}]}{\mathbb{E}[C_{2i+1,0}]},
\]

where the last equation is true, because the second term is independent of the \(\sigma\)-algebra \(\mathcal{D}\) and we can take the trivial \(\sigma\)-algebra. This means, first and second semesters are alike (not only in expectation, but also in expectation conditioned to all information of the past) up to a fixed factor.

All these cases are very strange circumstances.
Methods for mid-year reserving

- Results in closing figures
- Easy to generalise to other dates during the year
- It is, except for strange situations, not consistent with the yearly Chain Ladder at year end
- Usability:
  - Claims development result can be discussed
  - Observed and estimated development factor can only be compared if we always use the same split, but this comes along much larger triangles
  - Uncertainties can be estimated by standard formulas (not by the original formulas of Mack [21] and Merz-Wüthrich [17], but by the formulas of Röhr [30])
  - Split data triangles can get very huge, for instance for a forecast at the end of November
9 Mid year reserving

9.2 Methods for mid-year reserving

9.2.5 Separating semesters

Step by step

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>25</td>
<td>50</td>
<td>65</td>
<td>75</td>
</tr>
<tr>
<td>24</td>
<td>75</td>
<td>200</td>
<td>260</td>
<td>75</td>
</tr>
<tr>
<td>36</td>
<td>100</td>
<td>500</td>
<td>455</td>
<td>75</td>
</tr>
<tr>
<td>48</td>
<td>150</td>
<td>200</td>
<td>195</td>
<td>75</td>
</tr>
</tbody>
</table>

Step 1:

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>25</td>
<td>50</td>
<td>65</td>
<td>100</td>
</tr>
<tr>
<td>18</td>
<td>75</td>
<td>150</td>
<td>195</td>
<td>100</td>
</tr>
<tr>
<td>30</td>
<td>100</td>
<td>300</td>
<td>260</td>
<td>175</td>
</tr>
<tr>
<td>42</td>
<td>150</td>
<td>350</td>
<td>455</td>
<td>175</td>
</tr>
</tbody>
</table>

Step 2:

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>25</td>
<td>50</td>
<td>65</td>
<td>100</td>
</tr>
<tr>
<td>18</td>
<td>75</td>
<td>150</td>
<td>195</td>
<td>100</td>
</tr>
<tr>
<td>30</td>
<td>100</td>
<td>300</td>
<td>260</td>
<td>175</td>
</tr>
<tr>
<td>42</td>
<td>150</td>
<td>350</td>
<td>455</td>
<td>175</td>
</tr>
</tbody>
</table>

Results

- **Ultimate**: PY = 1960, CY = 525, Total = 2485 ✓ for closings
- **Reserves**: PY = 505, CY = 450, Total = 955
Stochastic Reserving

- Mid year reserving
- Methods for mid-year reserving

Results
- Ultimate: PY = 1960, CY = 525, Total = 2485 ✓ for closings
- Reserves: PY = 515, CY = 450, Total = 955
Properties

- results in closing figures
- easy to generalise to other dates during the year 😊
- even at year end you will get different reserves looking at accident year or accident semesters 😞
- usability:
  - claims development result can be discussed 😊
  - observed and estimated development factors at mid year and at year end are only comparable if we always use separated data
  - standard formulas for estimating uncertainties will not work, because they cannot reflect dependencies (which in addition have to be specified) between the triangles 😞
  - we may end up with a lot of triangles, for instance at the end of November 😞
Properties

- Results in closing figures
- Easy to generalise to other dates during the year
- Even at year end you will get different reserves looking at accident year or accident semesters
- Usability:
  - Claims development result can be discussed
  - Observed and estimated development factors at mid year and at year end are only comparable if we always use separated data
  - Standard formulas for estimating uncertainties will not work, because they cannot reflect dependencies (which in addition have to be specified)
  - We may end up with a lot of triangles, for instance at the end of November

Stochastic Reserving
- Mid year reserving
- Methods for mid-year reserving
9.2.6 Separating the youngest semester

Step by step

Results

- **Ultimate**: PY = 1960, CY = 525, Total = 2485 ✓ for closings
- **Reserves**: PY = 505, CY = 450, Total = 955
Stochastic Reserving

- Mid year reserving
- Methods for mid-year reserving

Results

- Ultimate: PY = 1960, CY = 525, Total = 2485 ✓ for closings
- Reserves: PY = 515, CY = 450, Total = 955

2019-02-18
Properties

- resulting in closing figures
- easy to generalise for other dates during the year 😊
- at year end both triangles are the same and equal to the yearly triangle 😞

usability:

* claims development results can be discussed 😊
* observed and estimated development factors for prior years at mid year and at year end are comparable 😊
* standard formulas for estimating uncertainties will not work, because they cannot reflect dependencies (which in addition have to be specified) between the triangles 😞
* not so easy to implement with standard reserving software
Properties

- resulting in closing figures
- easy to generalise for other dates during the year
- at year end both triangles are the same and equal to the yearly triangle
- usability:
  - claims development results can be discussed
  - observed and estimated development factors for prior years at mid year and at year end are comparable
  - standard formulas for estimating uncertainties will not work, because they cannot reflect dependencies (which in addition have to be specified) between the triangles
- not so easy to implement with standard reserving software
### 9.3 Conclusion

<table>
<thead>
<tr>
<th>method</th>
<th>forecast or closing</th>
<th>generalisation to other dates</th>
<th>methodology consistent with year end</th>
<th>estimates consistent at year end</th>
<th>analysis of claims development result</th>
<th>dev. factors comparable to year end</th>
<th>estimation of uncertainties</th>
<th>workload</th>
</tr>
</thead>
<tbody>
<tr>
<td>splitting development years</td>
<td>forecast</td>
<td>(✓)</td>
<td>?</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>can get huge</td>
</tr>
<tr>
<td>shifting development periods</td>
<td>forecast</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>extrapolating last diagonal</td>
<td>forecast</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>shifting accident periods</td>
<td>closing</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>splitting accident years</td>
<td>closing</td>
<td>(✓)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>can get huge</td>
</tr>
<tr>
<td>separation semesters</td>
<td>closing</td>
<td>(✓)</td>
<td>?</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>can get huge</td>
</tr>
<tr>
<td>separating youngest semesters</td>
<td>closing</td>
<td>(✓)</td>
<td>?</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
(✓) stands for ‘yes, but’ and refers to possible huge triangles, cannot be implemented (easily) in standard reserving software or other reasons

My favourite for mid-year closings is the separation of the youngest semester, because

– estimated ultimates (CDR), estimated development factors as well as observed development factors are comparable with year end figures based on yearly triangles
– with some tricks it can be implemented in most standard reserving software
– uncertainties should anyway be estimated separately

My favourite for forecasts is the shifting of development periods, because

– it can be implemented in most standard reserving software, which is not the case for the (correct) extrapolation of the last diagonal
Literature

Chain-ladder method and midyear loss reserving.

Chain Ladder and Error Propagation.
Literature

[36] René Dahms.
Chain-ladder method and midyear loss reserving.

Chain Ladder and Error Propagation.