Stochastic Reserving

René Dahms

ETH Zurich, Spring 2021

Last update: 26 April 2021

Stochastic Reserving

Stochastic Reserving

René Dahms

ETH Zurich, Spring 2021

Last update: 26 April 2021

Table of contents

4 Linear-Stochastic-Reserving methods

5 Poisson-Model

Table of contents

6 Bootstrap for CLM

7 Mid year reserving

8 CLM: Bayesian & credibility approach

Stochastic Reserving

Table of contents

9 Separation of small and large claims

10 Examples & Trail Exam

Table of contents

Stochastic Reserving

Lecture 1

Introduction

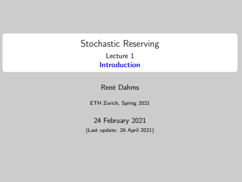
René Dahms

ETH Zurich, Spring 2021

24 February 2021

(Last update: 26 April 2021)

Stochastic Reserving



Stochastic Reserving
Lecture 1
Introduction

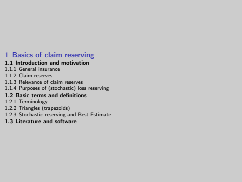
René Dahms

ETH Zurich, Spring 2021

24 February 2021
(Last update: 26 April 2021)

1 Basics of claim reserving

Lecture 1: Table of contents



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

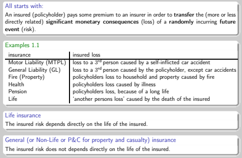
insurance	insured loss
Motor Liability (MTPL)	loss to a 3^{rd} person caused by a self-inflicted car accident
General Liability (GL)	loss to a 3^{rd} person caused by the policyholder, except car accidents
Fire (Property)	policyholders loss to household and property caused by fire
Health	policyholders loss caused by illness
Pension	policyholders loss, because of a long life
Life	'another persons loss' caused by the death of the insured

Life insurance

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

General (or Non-Life or P&C for property and casualty) 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** incurring **future event** (risk).

Examples 1.1

insurance	insured loss
Motor Liability (MTPL)	loss to a 3rd person caused by a self-inflicted car accident
General Liability (GL)	loss to a 3rd person caused by the policyholder, except car accidents
Fire (Property)	policyholders loss to household and property caused by fire
Health	policyholders loss caused by illness
Pension	policyholders loss, because of a long life
Life	'another persons loss' caused by the death of the insured

Life insurance

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

General (or Non-Life or P&C for property and casualty) insurance

The insured risk does not depends 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’

2021-04-26

Stochastic Reserving
Basics of claim reserving
Introduction and motivation

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

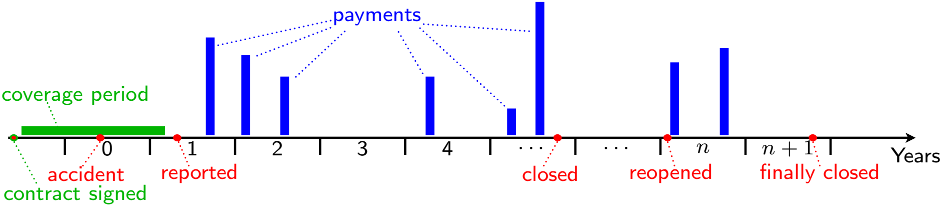
An insurance contract is

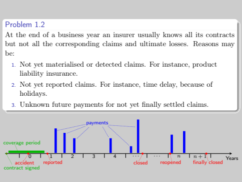
'a contract under which one party (the issuer) accepts **significant** insurance risk from another party (the policyholder) by agreeing to **compensate the policyholder** if a specified **uncertain future** event (the insured event) **adversely affects** the policyholder'

Problem 1.2

At the end of a business year an insurer usually knows all its contracts but not all the corresponding claims and 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.



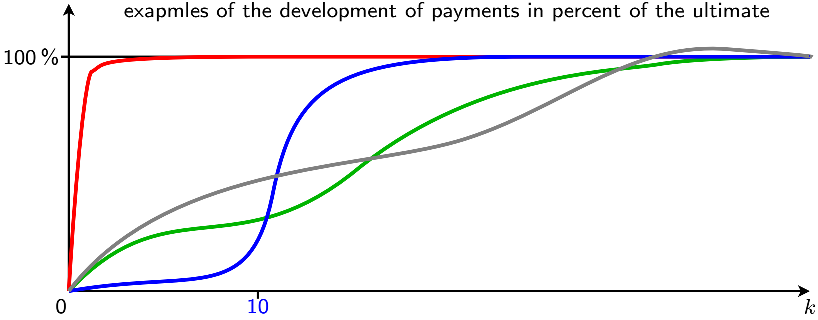


Problem 1.2

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

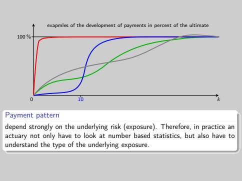
1. Not yet materialised or detected claims. For instance, product liability insurance.
2. Not yet reported claims. For instance, time delay, because of holidays.
3. Unknown future payments for not yet finally settled claims.

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



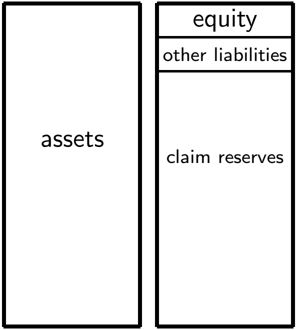
Payment pattern

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 DÃ©cennale in France or Spain
- gray may be madatory accident insurance in Switzerland

Claim reserves are often the most important part of the balance sheet of a general insurer. Moreover, a small changes in the estimate of claim reserves may make the difference between an annual profit or loss.

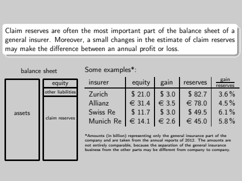
balance sheet



Some examples*:

insurer	equity	gain	reserves	$\frac{\text{gain}}{\text{reserves}}$
Zurich	\$ 21.0	\$ 3.0	\$ 82.7	3.6 %
Allianz	€ 31.4	€ 3.5	€ 78.0	4.5 %
Swiss Re	\$ 11.7	\$ 3.0	\$ 49.5	6.1 %
Munich Re	€ 14.1	€ 2.6	€ 45.0	5.8 %

*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 small changes in the estimate of claim reserves may make the difference between an annual profit or loss.

Some examples*:

insurer	equity	gain	reserves	$\frac{\text{gain}}{\text{reserves}}$
Zurich	\$ 21.0	\$ 3.0	\$ 82.7	3.6 %
Allianz	€ 31.4	€ 3.5	€ 78.0	4.5 %
Swiss Re	\$ 11.7	\$ 3.0	\$ 49.5	6.1 %
Munich Re	€ 14.1	€ 2.6	€ 45.0	5.8 %

*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.

Example: Converium AG

Converium AG was one of the largest reinsurers in the world. At 20^{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 in the world. At 20^{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)

Loss reserving

is an integral part of many processes. For instance:

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

The resulting estimates for claim reserves depend on its 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 the Swiss Solvency Test (SST)
- modelling the value of customers
- . . .

The resulting estimates for claim reserves depend on its 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.

$$\textbf{claim reserves} = \text{case reserves} + \text{IBNR}$$

Definition 1.5 (**I**ncurred **b**ut **n**ot **y**et **r**eported (IBNyR) reserves)

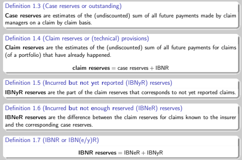
IBNyR reserves are the part of the claim reserves that corresponds to not yet reported claims.

Definition 1.6 (**I**ncurred **b**ut **n**ot **e**nough **r**eserved (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)

$$\textbf{IBNR reserves} = \text{IBNeR} + \text{IBNyR}$$



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.

$$\textbf{claim reserves} = \text{case reserves} + \text{IBNR}$$

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)

$$\textbf{IBNR reserves} = \text{IBNeR} + \text{IBNyR}$$

Provided we take a positive sign for claim reserves IBNyR are non-negative, whereas IBNeR may be positive or negative.

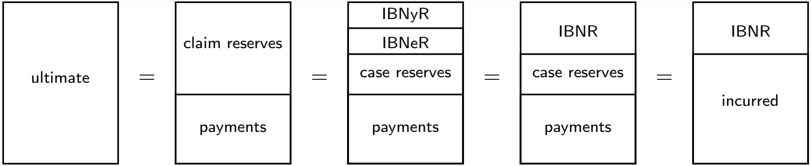
Usually, we will not look at discounted reserves, because discounting (and inflation) disturbs the development of claims and is dealt with separately, i.e. first get undiscounted figures and corresponding payment patterns and then discount.

Definition 1.8 (Incurred (losses) or reported amounts)

$$\textbf{incurred} = \text{payments} + \text{case reserves}$$

Definition 1.9 (Ultimate)

$$\begin{aligned}\textbf{ultimate} &= \text{payments} + \text{claim reserves}\\ &= \text{incurred} + \text{IBNR}\end{aligned}$$



ultimate

=

claim reserves
payments

=

IBNyR
IBNeR
case reserves
payments

=

IBNR
case reserves
payments

=

IBNR
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.
- Precise 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)

$$\textbf{incurred} = \text{payments} + \text{case reserves}$$

Definition 1.9 (Ultimate)

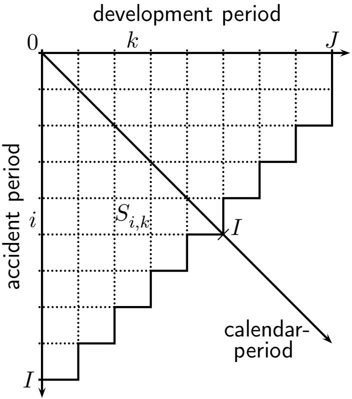
$$\begin{aligned}\textbf{ultimate} &= \text{payments} + \text{claim reserves}\\ &= \text{incurred} + \text{IBNR}\end{aligned}$$

Remarks 1.10

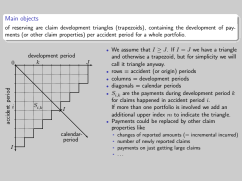
- Payments are often called paid (losses).
- The naming is not consistent within the actuarial world. For instance, actuaries often understand under IBNR only the IBNyR part.
- Precise definitions depend on the accounting standard. For instance, under IFRS 17 one has to discount the cash flows and one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

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.

- 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
  - . . .

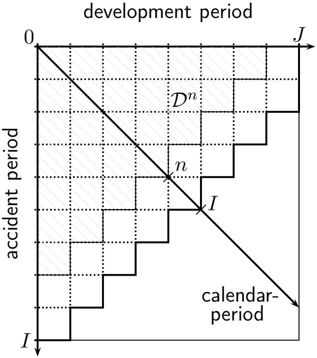
Some actuaries look at those numbers from a different angel:

- 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 months, quarters or years. 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 σ-algebra of all information up to calender period n:

$$\mathcal{D}^n := \sigma\big(S^m_{i,k} \colon 0 \le m \le M,\, 0 \le i \le I,\, 0 \le k \le J \wedge (n-i)\big)$$

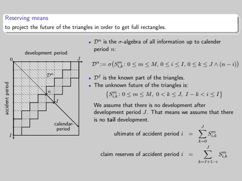
- $\mathcal{D}^I$ is the known part of the triangles.
- The unknown future of the triangles is:

$$\big\{S^m_{i,k} \colon 0 \le m \le M,\ 0 < k \le J,\ I-k < i \le I\big\}$$

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

$$\text{ultimate of accident period } i \;=\; \sum_{k=0}^{J} S^m_{i,k}$$

$$\text{claim reserves of accident period } i \;=\; \sum_{k=I+1-i}^{J} S^m_{i,k}$$



Reserving means
to project the future of the triangles in order to get full rectangles.

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

$$\mathcal{D}^n := \sigma\big(S^m_{i,k} : 0 \le m \le M,\, 0 \le i \le I,\, 0 \le k \le J \wedge (n-i)\big)$$

- $\mathcal{D}^I$ is the known part of the triangles.
- The unknown future of the triangles is:

$$\big\{S^m_{i,k} : 0 \le m \le M,\ 0 < k \le J,\ I-k < i \le I\big\}$$

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

$$\text{ultimate of accident period } i \;=\; \sum_{k=0}^{J} S^m_{i,k}$$

$$\text{claim reserves of accident period } i \;=\; \sum_{k=I+1-i}^{J} S^m_{i,k}$$

On a diagonlal 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 $I = k + i$.

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 $S_{i,k}^m$ 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 1.13 (Best Estimate)

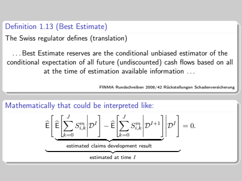
The Swiss regulator defines (translation)

> ... 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:

$$\underbrace{\widehat{\mathsf{E}}\left[\underbrace{\widehat{\mathsf{E}}\left[\sum_{k=0}^{J} S_{i,k}^{m} \middle| \mathcal{D}^{I}\right] - \widehat{\mathsf{E}}\left[\sum_{k=0}^{J} S_{i,k}^{m} \middle| \mathcal{D}^{I+1}\right]}_{\text{estimated claims development result}} \middle| \mathcal{D}^{I}\right]}_{\text{estimated at time } I} = 0.$$



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:

$$\underbrace{\widehat{\mathsf{E}}\left[\underbrace{\widehat{\mathsf{E}}\left[\sum_{k=0}^{J} S_{i,k}^{m}\middle|\mathcal{D}^{I}\right]-\widehat{\mathsf{E}}\left[\sum_{k=0}^{J} S_{i,k}^{m}\middle|\mathcal{D}^{I+1}\right]}_{\text{estimated claims development result}}\middle|\mathcal{D}^{I}\right]}_{\text{estimated at time } I}=0.$$

- 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 then 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 planing framework at time I. Or in other words, we don not expect any profit or loss on already happened claims.

Uncertainty of the Best Estimate

- The Holy Grail of loss reserving is to estimate the ($\mathcal{D}^I$-conditional) distribution of the reserves. Unfortunately, this would require 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 $\mathcal{B}$-conditional mean square error of prediction of the estimate $\widehat{Y}$ of a square integrable random variable Y is defined by

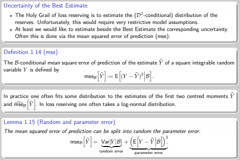
$$\mathsf{mse}_{\mathcal{B}}\left[\widehat{Y}\right] := \mathsf{E}\left[(Y-\widehat{Y})^2\middle|\mathcal{B}\right].$$

In practice one often fits some distribution to the estimates of the first two centred moments $\widehat{Y}$ and $\widehat{\mathsf{mse}}_{\mathcal{B}}\left[\widehat{Y}\right]$. 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:

$$\mathsf{mse}_{\mathcal{B}}\left[\widehat{Y}\right] = \underbrace{\mathsf{Var}[Y|\mathcal{B}]}_{\text{random error}} + \underbrace{\left(\mathsf{E}\left[Y-\widehat{Y}\middle|\mathcal{B}\right]\right)^2}_{\text{parameter error}}.$$



Uncertainty of the Best Estimate

- The Holy Grail of loss reserving is to estimate the ($\mathcal{D}^I$-conditional) distribution of the reserves. Unfortunately, this would require 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 $\mathcal{B}$-conditional mean square error of prediction of the estimate $\widehat{Y}$ of a square integrable random variable Y is defined by

$$\mathrm{mse}_{\mathcal{B}}\left[\widehat{Y}\right] := \mathsf{E}\left[(Y-\widehat{Y})^2\middle|\mathcal{B}\right].$$

In practice one often fits some distribution to the estimates of the first two centred moments $\widehat{Y}$ and $\widehat{\mathrm{mse}}_{\mathcal{B}}\left[\widehat{Y}\right]$. 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:

$$\mathrm{mse}_{\mathcal{B}}\left[\widehat{Y}\right] = \underbrace{\mathsf{Var}[Y|\mathcal{B}]}_{\text{random error}} + \underbrace{\left(\mathsf{E}\left[Y-\widehat{Y}\middle|\mathcal{B}\right]\right)^2}_{\text{parameter error}}.$$

A proof of the split of the mse will be given in Lecture 3.

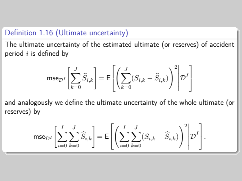
Definition 1.16 (Ultimate uncertainty)

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

$$\mathsf{mse}_{\mathcal{D}^I}\left[\sum_{k=0}^{J}\widehat{S}_{i,k}\right]=\mathsf{E}\left[\left(\sum_{k=0}^{J}(S_{i,k}-\widehat{S}_{i,k})\right)^2\middle|\mathcal{D}^I\right]$$

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

$$\mathsf{mse}_{\mathcal{D}^I}\left[\sum_{i=0}^{I}\sum_{k=0}^{J}\widehat{S}_{i,k}\right]=\mathsf{E}\left[\left(\sum_{i=0}^{I}\sum_{k=0}^{J}(S_{i,k}-\widehat{S}_{i,k})\right)^2\middle|\mathcal{D}^I\right].$$



Definition 1.16 (Ultimate uncertainty)

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

$$\mathsf{mse}_{\mathcal{D}^I}\left[\sum_{k=0}^{J}\widehat{S}_{i,k}\right] = \mathsf{E}\left[\left(\sum_{k=0}^{J}(S_{i,k}-\widehat{S}_{i,k})\right)^2\middle|\mathcal{D}^I\right]$$

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

$$\mathsf{mse}_{\mathcal{D}^I}\left[\sum_{i=0}^{I}\sum_{k=0}^{J}\widehat{S}_{i,k}\right] = \mathsf{E}\left[\left(\sum_{i=0}^{I}\sum_{k=0}^{J}(S_{i,k}-\widehat{S}_{i,k})\right)^2\middle|\mathcal{D}^I\right].$$

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.:

$$\mathsf{CDR}_i^{I+1} := \mathsf{E}\left[\sum_{k=0}^{J} S_{i,k} \middle| \mathcal{D}^I\right] - \mathsf{E}\left[\sum_{k=0}^{J} S_{i,k} \middle| \mathcal{D}^{I+1}\right].$$

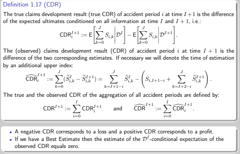
The (observed) claims development result (CDR) of accident period i at time $I+1$ is the difference of the two corresponding estimates. If necessary we will denote the time of estimation by an additional upper index:

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

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

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

- 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 $\mathcal{D}^I$-conditional expectation of the observed CDR equals zero.



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.:

$$\mathsf{CDR}_i^{I+1} := \mathsf{E}\left[\sum_{k=0}^{J} S_{i,k}\middle|\mathcal{D}^I\right] - \mathsf{E}\left[\sum_{k=0}^{J} S_{i,k}\middle|\mathcal{D}^{I+1}\right].$$

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:

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

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

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

- 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 $\mathcal{D}^I$-conditional expectation of the observed CDR equals zero.

For the true CDR we have

$$\mathsf{E}\left[\mathsf{CDR}_i^{I+1}\middle|\mathcal{D}^I\right] = \mathsf{E}\left[\mathsf{E}\left[\sum_{k=I+1-i}^{J} S_{i,k}\middle|\mathcal{D}^I\right] - \left(S_{i,I+1-i} + \mathsf{E}\left[\sum_{k=I+2-i}^{J} S_{i,k}\middle|\mathcal{D}^{I+1}\right]\right)\middle|\mathcal{D}^I\right] = 0.$$

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

$$\widehat{\mathsf{E}}\left[\widehat{\mathsf{CDR}}_i^{I+1}\middle|\mathcal{D}^I\right] = 0.$$

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

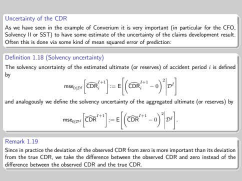
$$\mathsf{mse}_{0|\mathcal{D}^I}\left[\widehat{\mathsf{CDR}}_i^{I+1}\right] := \mathsf{E}\left[\left(\widehat{\mathsf{CDR}}_i^{I+1} - 0\right)^2 \middle| \mathcal{D}^I\right]$$

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

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

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

$$\mathsf{mse}_{0|\mathcal{D}^I}\left[\widehat{\mathsf{CDR}}_i^{I+1}\right] := \mathsf{E}\left[\left(\widehat{\mathsf{CDR}}_i^{I+1} - 0\right)^2 \middle| \mathcal{D}^I\right]$$

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

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

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.

- 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, 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, 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?

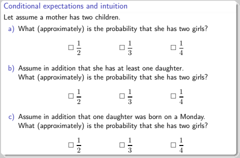
☐ $\frac{1}{2}$ ☐ $\frac{1}{3}$ ☐ $\frac{1}{4}$

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?

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

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

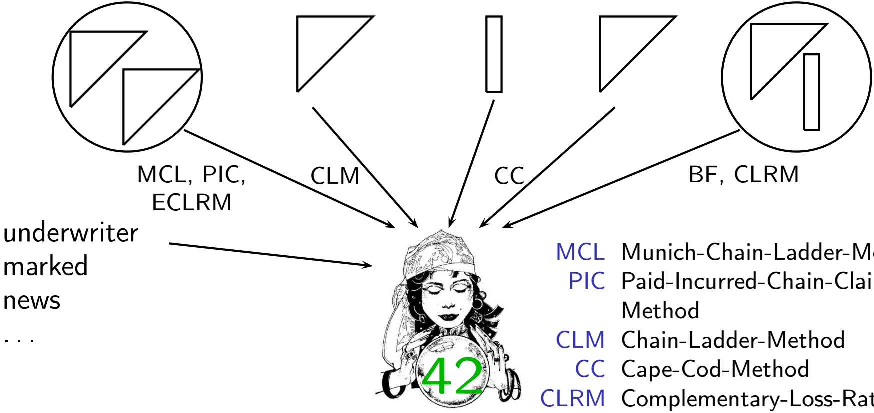
☐ $\frac{1}{2}$ ☐ $\frac{1}{3}$ ☐ $\frac{1}{4}$

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

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

- 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 (conditional) probabilities and expectations.

Reserving in the real world:



MCL, PIC, ECLRM

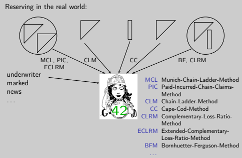
CLM

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
...	



Reserving in the real world:

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
...	

- 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.
- Others combine a triangle and a vector, like the Complementary-Loss-Ratio-Method and the Bornhuetter-Fergueson-Method.
- 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

[1] *Claims Reserving Manual.*
Institute of Actuaries, 2nd revised edition edition, 11 1997.

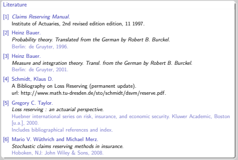
[2] Heinz Bauer.
Probability theory. Translated from the German by Robert B. Burckel.
Berlin: de Gruyter, 1996.

[3] Heinz Bauer.
Measure and integration theory. Transl. from the German by Robert B. Burckel.
Berlin: de Gruyter, 2001.

[4] Schmidt, Klaus D.
A Bibliography on Loss Reserving (permanent update).
url: http://www.math.tu-dresden.de/sto/schmidt/dsvm/reserve.pdf.

[5] Gregory C. Taylor.
Loss reserving : an actuarial perspective.
Huebner international series on risk, insurance, and economic security. Kluwer Academic, Boston [u.a.], 2000.
Includes bibliographical references and index.

[6] Mario V. Wüthrich and Michael Merz.
Stochastic claims reserving methods in insurance.
Hoboken, NJ: John Wiley & Sons, 2008.



Literature

[1] *Claims Reserving Manual.*
Institute of Actuaries, 2nd revised edition edition, 11 1997.

[2] Heinz Bauer.
Probability theory. Translated from the German by Robert B. Burckel.
Berlin: de Gruyter, 1996.

[3] Heinz Bauer.
Measure and integration theory. Transl. from the German by Robert B. Burckel.
Berlin: de Gruyter, 2001.

[4] Schmidt, Klaus D.
A Bibliography on Loss Reserving (permanent update).
url: http://www.math.tu-dresden.de/sto/schmidt/dsvm/reserve.pdf.

[5] Gregory C. Taylor.
Loss reserving : an actuarial perspective.
Huebner international series on risk, insurance, and economic security. Kluwer Academic, Boston [u.a.], 2000.
Includes bibliographical references and index.

[6] Mario V. Wüthrich and Michael Merz.
Stochastic claims reserving methods in insurance.
Hoboken, NJ: John Wiley & Sons, 2008.

- 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)
  - ...

- 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

Lecture 2

Chain-Ladder method

René Dahms

ETH Zurich, Spring 2021

3 March 2021

(Last update: 26 April 2021)

Stochastic Reserving

Stochastic Reserving

Lecture 2

Chain-Ladder method

René Dahms

ETH Zurich, Spring 2021

3 March 2021

(Last update: 26 April 2021)

2 Chain-Ladder-Method (CLM)

Lecture 2: Table of contents

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

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

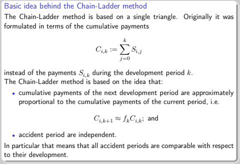
instead of the payments $S_{i,k}$ during the development period k.
The Chain-Ladder method is based on the idea that:

- 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

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

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.

Simple example

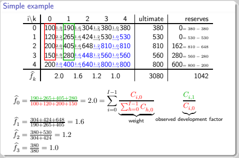
i\k	0	1	2	3	4	ultimate	reserves
0	100 $\xrightarrow{1.9}$	190 $\xrightarrow{1.6}$	304 $\xrightarrow{1.2}$	380 $\xrightarrow{1.0}$	380	380	0 $= 380 - 380$
1	120 $\xrightarrow{2.2}$	265 $\xrightarrow{1.6}$	424 $\xrightarrow{1.2}$	530 $\xrightarrow{1.0}$	530	530	0 $= 530 - 530$
2	200 $\xrightarrow{2.0}$	405 $\xrightarrow{1.6}$	648 $\xrightarrow{1.2}$	810 $\xrightarrow{1.0}$	810	810	162 $= 810 - 648$
3	150 $\xrightarrow{1.9}$	280 $\xrightarrow{1.6}$	448 $\xrightarrow{1.2}$	560 $\xrightarrow{1.0}$	560	560	280 $= 560 - 280$
4	200 $\xrightarrow{2.0}$	400 $\xrightarrow{1.6}$	640 $\xrightarrow{1.2}$	800 $\xrightarrow{1.0}$	800	800	600 $= 800 - 200$
$\widehat{f}_k$	2.0	1.6	1.2	1.0		3080	1042

$$\widehat{f}_0 = \frac{190+265+405+280}{100+120+200+150} = 2.0 = \sum_{i=0}^{I-1} \underbrace{\frac{C_{i,0}}{\sum_{h=0}^{I-1} C_{h,0}}}_{\text{weight}} \underbrace{\frac{C_{i,1}}{C_{i,0}}}_{\text{observed development factor}}$$

$$\widehat{f}_1 = \frac{304+424+648}{190+265+405} = 1.6$$

$$\widehat{f}_2 = \frac{380+530}{304+424} = 1.2$$

$$\widehat{f}_3 = \frac{380}{380} = 1.0$$



Simple example

i\k	0	1	2	3	4	ultimate	reserves
0	100	$\xrightarrow{1.9}$ 190	$\xrightarrow{1.6}$ 304	$\xrightarrow{1.2}$ 380	$\xrightarrow{1.0}$ 380	380	0 $= 380 - 380$
1	120	$\xrightarrow{2.2}$ 265	$\xrightarrow{1.6}$ 424	$\xrightarrow{1.2}$ 530	$\xrightarrow{1.0}$ 530	530	0 $= 530 - 530$
2	200	$\xrightarrow{2.0}$ 405	$\xrightarrow{1.6}$ 648	$\xrightarrow{1.2}$ 810	$\xrightarrow{1.0}$ 810	810	162 $= 810 - 648$
3	150	$\xrightarrow{1.9}$ 280	$\xrightarrow{1.6}$ 448	$\xrightarrow{1.2}$ 560	$\xrightarrow{1.0}$ 560	560	280 $= 560 - 280$
4	200	$\xrightarrow{2.0}$ 400	$\xrightarrow{1.6}$ 640	$\xrightarrow{1.2}$ 800	$\xrightarrow{1.0}$ 800	800	600 $= 800 - 200$
$\widehat{f}_k$		2.0	1.6	1.2	1.0	3080	1042

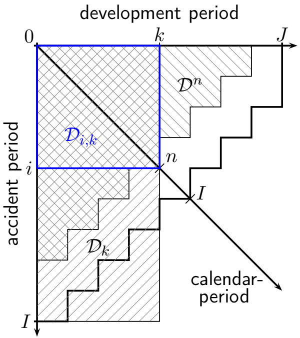
$$\widehat{f}_0 = \frac{190+265+405+280}{100+120+200+150} = 2.0 = \sum_{i=0}^{I-1} \underbrace{\frac{C_{i,0}}{\sum_{h=0}^{I-1} C_{h,0}}}_{\text{weight}} \qquad \underbrace{\frac{C_{i,1}}{C_{i,0}}}_{\text{observed development factor}}$$

$$\widehat{f}_1 = \frac{304+424+648}{190+265+405} = 1.6$$

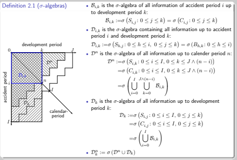
$$\widehat{f}_2 = \frac{380+530}{304+424} = 1.2$$

$$\widehat{f}_3 = \frac{380}{380} = 1.0$$

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:
$$\mathcal{B}_{i,k} := \sigma\left(S_{i,j} \colon 0 \le j \le k\right) = \sigma\left(C_{i,j} \colon 0 \le j \le k\right)$$
- $\mathcal{D}_{i,k}$ is the σ-algebra containing all information up to accident period i and development period k:
$$\mathcal{D}_{i,k} := \sigma\left(S_{h,j} \colon 0 \le h \le i,\ 0 \le j \le k\right) = \sigma\left(B_{h,k} \colon 0 \le h \le i\right)$$
- $\mathcal{D}^n$ is the σ-algebra of all information up to calender period n:
$$\begin{aligned}\mathcal{D}^n :=& \sigma\left(S_{i,k} \colon 0 \le i \le I,\ 0 \le k \le J \wedge (n-i)\right)\\ =& \sigma\left(C_{i,k} \colon 0 \le i \le I,\ 0 \le k \le J \wedge (n-i)\right)\\ =& \sigma\left(\bigcup_{i=0}^{I} \bigcup_{k=0}^{J\wedge(n-i)} \mathcal{B}_{i,k}\right)\end{aligned}$$
- $\mathcal{D}_k$ is the σ-algebra of all information up to development period k:
$$\begin{aligned}\mathcal{D}_k :=& \sigma\left(S_{i,j} \colon 0 \le i \le I,\ 0 \le j \le k\right)\\ =& \sigma\left(C_{i,j} \colon 0 \le i \le I,\ 0 \le j \le k\right)\\ =& \sigma\left(\bigcup_{i=0}^{I} \mathcal{B}_{i,k}\right)\end{aligned}$$
- $\mathcal{D}_k^n := \sigma\left(\mathcal{D}^n \cup \mathcal{D}_k\right)$



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:

$$\mathcal{B}_{i,k} := \sigma\left(S_{i,j} \colon 0 \le j \le k\right) = \sigma\left(C_{i,j} \colon 0 \le j \le k\right)$$

- $\mathcal{D}_{i,k}$ is the σ-algebra containing all information up to accident period i and development period k:

$$\mathcal{D}_{i,k} := \sigma\left(S_{h,j} \colon 0 \le h \le i,\ 0 \le j \le k\right) = \sigma\left(B_{h,k} \colon 0 \le h \le i\right)$$

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

$$\begin{aligned}\mathcal{D}^n :=& \sigma\left(S_{i,k} \colon 0 \le i \le I, 0 \le k \le J \wedge (n-i)\right)\\ =& \sigma\left(C_{i,k} \colon 0 \le i \le I, 0 \le k \le J \wedge (n-i)\right)\\ =& \sigma\left(\bigcup_{i=0}^{I} \bigcup_{k=0}^{J\wedge(n-i)} \mathcal{B}_{i,k}\right)\end{aligned}$$

- $\mathcal{D}_k$ is the σ-algebra of all information up to development period k:

$$\begin{aligned}\mathcal{D}_k :=& \sigma\left(S_{i,j} \colon 0 \le i \le I, 0 \le j \le k\right)\\ =& \sigma\left(C_{i,j} \colon 0 \le i \le I, 0 \le j \le k\right)\\ =& \sigma\left(\bigcup_{i=0}^{I} \mathcal{B}_{i,k}\right)\end{aligned}$$

- $\mathcal{D}_k^n := \sigma\left(\mathcal{D}^n \cup \mathcal{D}_k\right)$

The σ-algebra $\mathcal{D}_k^n$ 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

$$(S_{i_1,k_1} \in \mathcal{D}_k^n \quad \text{and} \quad S_{i_2,k_2} \notin \mathcal{D}_k^n) \quad \text{or} \quad (S_{i_1,k_1} \notin \mathcal{D}_k^n \quad \text{and} \quad S_{i_2,k_2} \in \mathcal{D}_k^n).$$

Assumption 2.A (Mack's Chain-Ladder method)

There exist development factors f_k *and variance parameters* σ_k^2 *such that the cumulative payments*

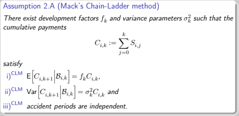
$$C_{i,k} := \sum_{j=0}^{k} S_{i,j}$$

satisfy

i)$^{\mathsf{CLM}}$ $\mathsf{E}\left[C_{i,k+1}\middle|\mathcal{B}_{i,k}\right] = f_k C_{i,k}$,

ii)$^{\mathsf{CLM}}$ $\mathsf{Var}\left[C_{i,k+1}\middle|\mathcal{B}_{i,k}\right] = \sigma_k^2 C_{i,k}$ *and*

iii)$^{\mathsf{CLM}}$ *accident periods are independent.*



Assumption 2.A (Mack's Chain-Ladder method)

There exist development factors f_k and variance parameters σ_k^2 such that the cumulative payments

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

satisfy

i)$^{\text{CLM}}$ $\mathsf{E}\left[C_{i,k+1}\middle|\mathcal{B}_{i,k}\right] = f_k C_{i,k}$,

ii)$^{\text{CLM}}$ $\mathsf{Var}\left[C_{i,k+1}\middle|\mathcal{B}_{i,k}\right] = \sigma_k^2 C_{i,k}$ *and*

iii)$^{\text{CLM}}$ *accident periods are independent.*

If $\mathcal{B}_{i,k}$ are replaced by $\mathcal{D}_k^{i+k}$ then the last assumption about independence is not necessary, i.e. it is enough to assume

$\widetilde{\text{i)}}^{\text{CLM}}$ $\mathsf{E}\left[C_{i,k+1}\middle|\mathcal{D}_k^{i+k}\right] = f_k C_{i,k}$,

$\widetilde{\text{ii)}}^{\text{CLM}}$ $\mathsf{Var}\left[C_{i,k+1}\middle|\mathcal{D}_k^{i+k}\right] = \sigma_k^2 C_{i,k}$.

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, $\mathcal{B}_{i,k}$ could be replaced by $\mathcal{D}_k$, $\mathcal{D}_{i,k}$ or $\mathcal{D}_k^{i+k}$.
- Published by Thomas Mack in 1991, see [22]. 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 [11].

Remark 2.2

- Since accident periods are independent, $\mathcal{B}_{i,k}$ could be replaced by $\mathcal{D}_k$, $\mathcal{D}_{i,k}$ or $\mathcal{D}_k^{i+k}$.
- Published by Thomas Mack in 1991, see [22]. 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 [11].

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* $S_{i,k}$*:*

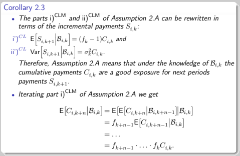
i')^{CL} $\mathsf{E}\left[S_{i,k+1}\middle|\mathcal{B}_{i,k}\right] = (f_k - 1)C_{i,k}$ *and*

ii')^{CL} $\mathsf{Var}\left[S_{i,k+1}\middle|\mathcal{B}_{i,k}\right] = \sigma_k^2 C_{i,k}$.

Therefore, Assumption 2.A means that under the knowledge of $\mathcal{B}_{i,k}$ *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*

$$\begin{aligned}\mathsf{E}\left[C_{i,k+n}\middle|\mathcal{B}_{i,k}\right] &= \mathsf{E}\left[\mathsf{E}\left[C_{i,k+n}\middle|\mathcal{B}_{i,k+n-1}\right]\middle|\mathcal{B}_{i,k}\right] \\ &= f_{k+n-1}\mathsf{E}\left[C_{i,k+n-1}\middle|\mathcal{B}_{i,k}\right] \\ &= \dots \\ &= f_{k+n-1}\cdot\ldots\cdot f_k C_{i,k}.\end{aligned}$$



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')^{CL} $\mathsf{E}\left[S_{i,k+1}\middle|\mathcal{B}_{i,k}\right] = (f_k - 1)C_{i,k}$ *and*

ii')^{CL} $\mathsf{Var}\left[S_{i,k+1}\middle|\mathcal{B}_{i,k}\right] = \sigma_k^2 C_{i,k}$.

Therefore, Assumption 2.A means that under the knowledge of $\mathcal{B}_{i,k}$ *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*

$$\begin{aligned}\mathsf{E}\left[C_{i,k+n}\middle|\mathcal{B}_{i,k}\right] &= \mathsf{E}\left[\mathsf{E}\left[C_{i,k+n}\middle|\mathcal{B}_{i,k+n-1}\right]\middle|\mathcal{B}_{i,k}\right]\\ &= f_{k+n-1}\mathsf{E}\left[C_{i,k+n-1}\middle|\mathcal{B}_{i,k}\right]\\ &= \dots\\ &= f_{k+n-1}\cdot\ldots\cdot f_k C_{i,k}.\end{aligned}$$

Proof of i')$^{\text{CLM}}$:

$$\begin{aligned}\mathsf{E}\left[S_{i,k+1}\middle|\mathcal{B}_{i,k}\right] &= \mathsf{E}\left[C_{i,k+1} - C_{i,k}\middle|\mathcal{B}_{i,k}\right]\\ &= \underbrace{\mathsf{E}\left[C_{i,k+1}\middle|\mathcal{B}_{i,k}\right] - C_{i,k}}_{C_{i,k} \textbf{ is } \mathcal{B}_{i,k} \textbf{ measurable}}\\ &= \underbrace{f_k C_{i,k}}_{\textbf{ii)}^{\textbf{CLM}}} - C_{i,k}\end{aligned}$$

Proof of ii')$^{\text{CLM}}$:

$$\begin{aligned}\mathsf{Var}\left[S_{i,k+1}\middle|\mathcal{B}_{i,k}\right] &= \mathsf{Var}\left[C_{i,k+1} - C_{i,k}\middle|\mathcal{B}_{i,k}\right]\\ &= \underbrace{\mathsf{Var}\left[C_{i,k+1}\middle|\mathcal{B}_{i,k}\right]}_{C_{i,k} \textbf{ is } \mathcal{B}_{i,k} \textbf{ measurable}}\end{aligned}$$

Lemma 2.4 (Chain-Ladder development factors)

Let Assumption 2.A be fulfilled and take arbitrary $\mathcal{D}^I \cap \mathcal{D}_k$*-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*

$$\widehat{f}_k := \sum_{i=0}^{I-1-k} w_{i,k} \frac{C_{i,k+1}}{C_{i,k}} \tag{2.1}$$

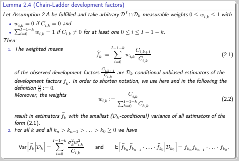
of the observed development factors $\frac{C_{i,k+1}}{C_{i,k}}$ *are* $\mathcal{D}_k$*-conditional unbiased estimators of the development factors* f_k*. In order to shorten notation, we use here and in the following the definition* $\frac{0}{0} := 0$*.*
Moreover, the weights

$$w_{i,k} := \frac{C_{i,k}}{\sum_{h=0}^{I-1-k} C_{h,k}} \tag{2.2}$$

result in estimators $\widehat{f}_k$ *with the smallest (*$\mathcal{D}_k$*-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*

$$\mathrm{Var}\left[\widehat{f}_k \middle| \mathcal{D}_k\right] = \sum_{i=0}^{I-1-k} \frac{\sigma_k^2 w_{i,k}^2}{C_{i,k}} \qquad \text{and} \qquad \mathrm{E}\left[\widehat{f}_{k_n} \widehat{f}_{k_{n-1}} \cdot \ldots \cdot \widehat{f}_{k_0} \middle| \mathcal{D}_{k_0}\right] = f_{k_n} f_{k_{n-1}} \cdot \ldots \cdot f_{k_0}.$$



Lemma 2.4 (Chain-Ladder development factors)

Let Assumption 2.A be fulfilled and take arbitrary $\mathcal{D}^I \cap \mathcal{D}_k$-measurable weights $0 \le w_{i,k} \le 1$ 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* $0 \le i \le I-1-k$.

Then:

1. *The weighted means*
$$\widehat{f}_k := \sum_{i=0}^{I-1-k} w_{i,k} \frac{C_{i,k+1}}{C_{i,k}} \tag{2.1}$$
of the observed development factors $\frac{C_{i,k+1}}{C_{i,k}}$ *are* $\mathcal{D}_k$*-conditional unbiased estimators of the development factors* f_k*. In order to shorten notation, we use here and in the following the definition* $\frac{0}{0} := 0$.
Moreover, the weights
$$w_{i,k} := \frac{C_{i,k}}{\sum_{h=0}^{I-1-k} C_{h,k}} \tag{2.2}$$
result in estimators $\widehat{f}_k$ *with the smallest (*$\mathcal{D}_k$*-conditional) variance of all estimators of the form* (2.1).
2. *For all* k *and all* $k_n > k_{n-1} > \ldots > k_0 \ge 0$ *we have*
$$\mathsf{Var}\left[\widehat{f}_k \middle| \mathcal{D}_k\right] = \sum_{i=0}^{I-1-k} \frac{\sigma_k^2 w_{i,k}^2}{C_{i,k}} \quad \text{and} \quad \mathsf{E}\left[\widehat{f}_{k_n} \widehat{f}_{k_{n-1}} \cdot \ldots \cdot \widehat{f}_{k_0} \middle| \mathcal{D}_{k_0}\right] = f_{k_n} f_{k_{n-1}} \cdot \ldots \cdot f_{k_0}.$$

- **unbiased:**

$$\mathsf{E}\left[\widehat{f}_k \middle| \mathcal{D}_k\right] = \mathsf{E}\left[\sum_{i=0}^{I-1-k} w_{i,k} \frac{C_{i,k+1}}{C_{i,k}} \middle| \mathcal{D}_k\right] = \underbrace{\sum_{i=0}^{I-1-k} w_{i,k} \frac{\mathsf{E}\left[C_{i,k+1} \middle| \mathcal{D}_k\right]}{C_{i,k}}}_{\textbf{measurable with respect to } \mathcal{D}_k} = \underbrace{\sum_{i=0}^{I-k} w_{i,k} \frac{f_k C_{i,k}}{C_{i,k}}}_{\textbf{i)}^{\textbf{CLM}}} = f_k$$

- **minimal variance:** $\mathsf{Var}\left[\widehat{f}_k\right] = \mathsf{E}\left[\mathsf{Var}\left[\widehat{f}_k \middle| \mathcal{D}_k\right]\right] + \mathsf{Var}\left[\mathsf{E}\left[\widehat{f}_k \middle| \mathcal{D}_k\right]\right] = \mathsf{E}\left[\mathsf{Var}\left[\widehat{f}_k \middle| \mathcal{D}_k\right]\right] + 0$

$$\mathsf{Var}\left[\widehat{f}_k \middle| \mathcal{D}_k\right] = \mathsf{Var}\left[\sum_{i=0}^{I-1-k} w_{i,k} \frac{C_{i,k+1}}{C_{i,k}} \middle| \mathcal{D}_k\right] = \underbrace{\sum_{i=0}^{I-1-k} w_{i,k}^2 \frac{\mathsf{Var}\left[C_{i,k+1} \middle| \mathcal{D}_k\right]}{C_{i,k}^2}}_{\textbf{measurable with respect to } \mathcal{D}_k \textbf{ and iii)}^{\textbf{CLM}}} = \underbrace{\sigma_k^2 \sum_{i=0}^{I-1-k} w_{i,k}^2 \frac{1}{C_{i,k}}}_{\textbf{ii)}^{\textbf{CLM}}}$$

Lagrange: minimize $\sum_{i=0}^{I-1-k} w_{i,k}^2 \frac{1}{C_{i,k}} + \lambda \left(1 - \sum_{i=0}^{I-1-k} w_{i,k}\right)$

$$\frac{\partial}{\partial w_{i,k}} \bullet = 2 w_{i,k} \frac{1}{C_{i,k}} - \lambda \quad \Longrightarrow \quad w_{i,k} = \frac{\lambda}{2} C_{i,k} \quad \text{and} \quad \underbrace{\lambda = \frac{2}{\sum_{i=0}^{I-1-k} C_{i,k}}}_{\sum_{i=0}^{I-1-k} w_{i,k} = 1} \quad \Longrightarrow \quad w_{i,k} = \frac{C_{i,k}}{\sum_{h=0}^{I-1-k} C_{h,k}}$$

- **uncorrelated:** $\mathsf{E}\left[\widehat{f}_{k_n} \widehat{f}_{k_{n-1}} \cdot \ldots \cdot \widehat{f}_{k_0} \middle| \mathcal{D}_{k_0}\right] = \mathsf{E}\left[\mathsf{E}\left[\widehat{f}_{k_n} \middle| \mathcal{D}_{k_n}\right] \widehat{f}_{k_{n-1}} \cdot \ldots \cdot \widehat{f}_{k_0} \middle| \mathcal{D}_{k_0}\right]$

$$= f_{k_n} \mathsf{E}\left[\widehat{f}_{k_{n-1}} \cdot \ldots \cdot \widehat{f}_{k_0} \middle| \mathcal{D}_k\right] = \ldots = f_{k_n} f_{k_{n-1}} \cdot \ldots \cdot f_{k_0}$$

Estimator 2.5 (Chain-Ladder Ultimate)

Let Assumption 2.A be fulfilled. Then the estimates

$$\widehat{C}_{i,k} := \widehat{f}_{k-1} \cdot \ldots \cdot \widehat{f}_{I-i} C_{i,I-i}$$

are $\mathcal{D}_{I-i}$-conditional unbiased estimators of $C_{i,k}$, for $I - i < k \leq J$.
In order to shorten notation, we define

$$\widehat{C}_{i,k} := C_{i,k},$$

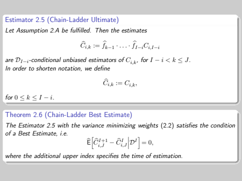
for $0 \leq k \leq I - i$.

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.

$$\widehat{\mathsf{E}}\left[\widehat{C}_{i,J}^{I+1} - \widehat{C}_{i,J}^{I} \middle| \mathcal{D}^{I}\right] = 0,$$

where the additional upper index specifies the time of estimation.



Estimator 2.5 (Chain-Ladder Ultimate)

Let Assumption 2.A be fulfilled. Then the estimates

$$\widehat{C}_{i,k} := \widehat{f}_{k-1} \cdot \ldots \cdot \widehat{f}_{I-i} C_{i,I-i}$$

are $\mathcal{D}_{I-i}$-conditional unbiased estimators of $C_{i,k}$, for $I-i < k \leq J$.
In order to shorten notation, we define

$$\widehat{C}_{i,k} := C_{i,k},$$

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.

$$\widehat{\mathsf{E}}\left[\widehat{C}_{i,J}^{I+1} - \widehat{C}_{i,J}^{I} \middle| \mathcal{D}^I\right] = 0,$$

where the additional upper index specifies the time of estimation.

- **Proof of unbiasedness:**

$$\begin{aligned}
\mathsf{E}\left[\widehat{C}_{i,k}\middle|\mathcal{D}_{I-i}\right] &= \mathsf{E}\left[\widehat{f}_{k-1} \cdot \ldots \cdot \widehat{f}_{I-i} C_{i,I-i}\middle|\mathcal{D}_{I-i}\right] = \mathsf{E}\left[\mathsf{E}\left[\widehat{f}_{k-1}\middle|\mathcal{D}_{k-1}\right]\widehat{f}_{k-2} \cdot \ldots \cdot \widehat{f}_{I-i} C_{i,I-i}\middle|\mathcal{D}_{I-i}\right]\\
&= \mathsf{E}\left[f_{k-1}\widehat{f}_{k-2} \cdot \ldots \cdot \widehat{f}_{I-i} C_{i,I-i}\middle|\mathcal{D}_{I-i}\right] = \ldots = f_{k-1} \cdot \ldots \cdot f_{I-i} C_{i,I-i}\\
&= \mathsf{E}\left[C_{i,k}\middle|\mathcal{D}_{I-i}\right]
\end{aligned}$$

- **Best Estimate:** $\widehat{f}_k^{I+1} := \sum_{i=0}^{I-k} \frac{C_{i,k}}{\sum_{h=0}^{I-k} C_{h,k}} \frac{C_{i,k+1}}{C_{i,k}} = \left(1 - \frac{C_{I-k,k}}{\sum_{h=0}^{I-k} C_{h,k}}\right)\widehat{f}_k^I + \frac{C_{I-k,k}}{\sum_{h=0}^{I-k} C_{h,k}} \frac{C_{I-k,k+1}}{C_{I-k,k}}$

$$\Longrightarrow \qquad \mathsf{E}\left[\widehat{f}_k^{I+1}\middle|\mathcal{D}_k^I\right] = \left(1 - \frac{C_{I-k,k}}{\sum_{h=0}^{I-k} C_{h,k}}\right)\widehat{f}_k^I + \frac{C_{I-k,k}}{\sum_{h=0}^{I-k} C_{h,k}} f_k =: \bar{f}_k$$

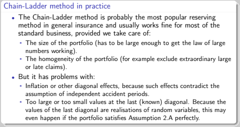
$$\Longrightarrow \qquad \widehat{\mathsf{E}}\left[\widehat{f}_k^{I+1}\middle|\mathcal{D}_k^I\right] = \widehat{f}_k^I$$

$$\begin{aligned}
\mathsf{E}\left[\widehat{C}_{i,J}^{I+1}\middle|\mathcal{D}^I\right] &= \mathsf{E}\left[\widehat{f}_{J-1}^{I+1} \cdot \ldots \cdot \widehat{f}_{I+1-i}^{I+1} C_{i,I+1-i}\middle|\mathcal{D}^I\right]\\
&= \mathsf{E}\left[\mathsf{E}\left[\widehat{f}_{J-1}^{I+1}\middle|\mathcal{D}_{J-1}^I\right]\widehat{f}_{J-2}^{I+1} \cdot \ldots \cdot \widehat{f}_{I+1-i}^{I+1} C_{i,I+1-i}\middle|\mathcal{D}^I\right]\\
&= \mathsf{E}\left[\bar{f}_{J-1}\widehat{f}_{J-2}^{I+1} \cdot \ldots \cdot \widehat{f}_{I+1-i}^{I+1} C_{i,I+1-i}\middle|\mathcal{D}^I\right]\\
&= \ldots = \bar{f}_{J-1} \cdot \ldots \cdot \bar{f}_{I+1-i}\mathsf{E}\left[C_{i,I+1-i}\middle|\mathcal{D}^I\right] = \bar{f}_{J-1} \cdot \ldots \cdot \bar{f}_{I+1-i} f_{I-i} C_{i,I-i}
\end{aligned}$$

$$\Longrightarrow \qquad \widehat{\mathsf{E}}\left[\widehat{C}_{i,J}^{I+1}\middle|\mathcal{D}^I\right] = \widehat{f}_{J-1}^I \cdot \ldots \cdot \widehat{f}_{I-i}^I C_{i,I-i} = \widehat{C}_{i,J}^I = \widehat{\mathsf{E}}\left[\widehat{C}_{i,J}^I\middle|\mathcal{D}^I\right].$$

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.

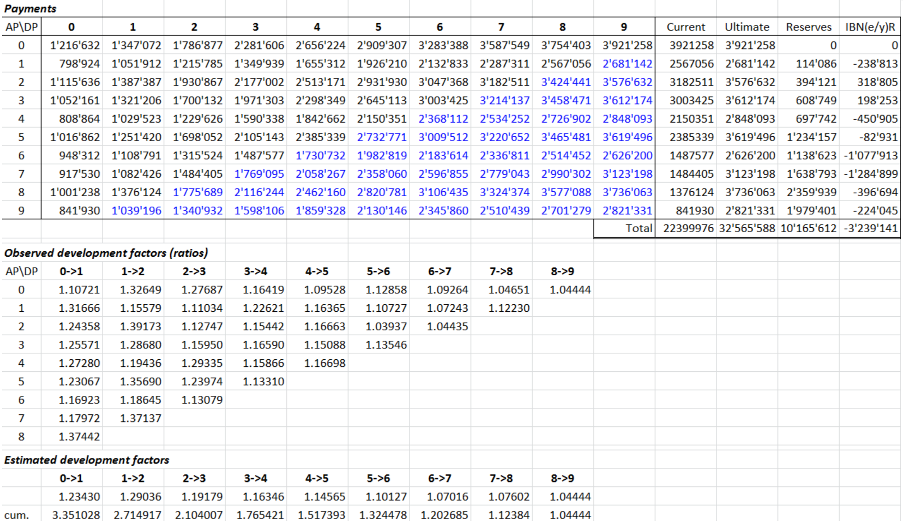


Chain-Ladder method in practice

- 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.

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

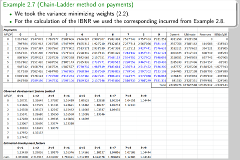
AP\DP	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	Current	Ultimate	Reserves	IBN(e/y)R
0	1'216'632	1'347'072	1'786'877	2'281'606	2'656'224	2'909'307	3'283'388	3'587'549	3'754'403	3'921'258	3921258	3'921'258	0	0
1	798'924	1'051'912	1'215'785	1'349'939	1'655'312	1'926'210	2'132'833	2'287'311	2'567'056	2'681'142	2567056	2'681'142	114'086	-238'813
2	1'115'636	1'387'387	1'930'867	2'177'002	2'513'171	2'931'930	3'047'368	3'182'511	3'424'441	3'576'632	3182511	3'576'632	394'121	318'805
3	1'052'161	1'321'206	1'700'132	1'971'303	2'298'349	2'645'113	3'003'425	3'214'137	3'458'471	3'612'174	3003425	3'612'174	608'749	198'253
4	808'864	1'029'523	1'229'626	1'590'338	1'842'662	2'150'351	2'368'112	2'534'252	2'726'902	2'848'093	2150351	2'848'093	697'742	-450'905
5	1'016'862	1'251'420	1'698'052	2'105'143	2'385'339	2'732'771	3'009'512	3'220'652	3'465'481	3'619'496	2385339	3'619'496	1'234'157	-82'931
6	948'312	1'108'791	1'315'524	1'487'577	1'730'732	1'982'819	2'183'614	2'336'811	2'514'452	2'626'200	1487577	2'626'200	1'138'623	-1'077'913
7	917'530	1'082'426	1'484'405	1'769'095	2'058'267	2'358'060	2'596'855	2'779'043	2'990'302	3'123'198	1484405	3'123'198	1'638'793	-1'284'899
8	1'001'238	1'376'124	1'775'689	2'116'244	2'462'160	2'820'781	3'106'435	3'324'374	3'577'088	3'736'063	1376124	3'736'063	2'359'939	-396'694
9	841'930	1'039'196	1'340'932	1'598'106	1'859'328	2'130'146	2'345'860	2'510'439	2'701'279	2'821'331	841930	2'821'331	1'979'401	-224'045
										Total	22399976	32'565'588	10'165'612	-3'239'141

Observed development factors (ratios)

AP\DP	**0->1**	**1->2**	**2->3**	**3->4**	**4->5**	**5->6**	**6->7**	**7->8**	**8->9**
0	1.10721	1.32649	1.27687	1.16419	1.09528	1.12858	1.09264	1.04651	1.04444
1	1.31666	1.15579	1.11034	1.22621	1.16365	1.10727	1.07243	1.12230	
2	1.24358	1.39173	1.12747	1.15442	1.16663	1.03937	1.04435		
3	1.25571	1.28680	1.15950	1.16590	1.15088	1.13546			
4	1.27280	1.19436	1.29335	1.15866	1.16698				
5	1.23067	1.35690	1.23974	1.13310					
6	1.16923	1.18645	1.13079						
7	1.17972	1.37137							
8	1.37442								

Estimated development factors

	0->1	**1->2**	**2->3**	**3->4**	**4->5**	**5->6**	**6->7**	**7->8**	**8->9**
	1.23430	1.29036	1.19179	1.16346	1.14565	1.10127	1.07016	1.07602	1.04444
cum.	3.351028	2.714917	2.104007	1.765421	1.517393	1.324478	1.202685	1.12384	1.04444



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

AP\DP	0	1	2	3	4	5	6	7	8	9	Current	Ultimate	Reserves	IBN(e/y)R
0	1'216'632	1'347'072	1'786'877	2'281'606	2'656'224	2'909'307	3'283'388	3'587'549	3'754'403	3'921'258	3921258	3'921'258	0	0
1	798'924	1'051'912	1'215'785	1'349'939	1'655'312	1'926'210	2'132'833	2'287'311	2'567'056	2'681'142	2567056	2'681'142	114'086	-238'813
2	1'115'636	1'387'387	1'930'867	2'177'002	2'513'171	2'931'930	3'047'368	3'182'511	3'424'441	3'576'632	3182511	3'576'632	394'121	318'805
3	1'052'161	1'321'206	1'700'132	1'971'303	2'298'349	2'645'113	3'003'425	3'214'137	3'458'471	3'612'174	3003425	3'612'174	608'749	198'253
4	808'864	1'029'523	1'229'626	1'590'338	1'842'662	2'150'351	2'368'112	2'534'252	2'726'902	2'848'093	2150351	2'848'093	697'742	-450'905
5	1'016'862	1'251'420	1'698'052	2'105'143	2'385'339	2'732'771	3'009'512	3'220'652	3'465'481	3'619'496	2385339	3'619'496	1'234'157	-82'931
6	948'312	1'108'791	1'315'524	1'487'577	1'730'732	1'982'819	2'183'614	2'336'811	2'514'452	2'626'200	1487577	2'626'200	1'138'623	-1'077'913
7	917'530	1'082'426	1'484'405	1'769'095	2'058'267	2'358'060	2'596'855	2'779'043	2'990'302	3'123'198	1484405	3'123'198	1'638'793	-1'284'899
8	1'001'238	1'376'124	1'775'689	2'116'244	2'462'160	2'820'781	3'106'435	3'324'374	3'577'088	3'736'063	1376124	3'736'063	2'359'939	-396'094
9	841'930	1'039'196	1'340'932	1'598'106	1'859'328	2'130'146	2'345'860	2'510'439	2'701'279	2'821'331	841930	2'821'331	1'979'401	-224'045
										Total	22399976	32'565'588	10'165'612	-3'239'141

Observed development factors (ratios)

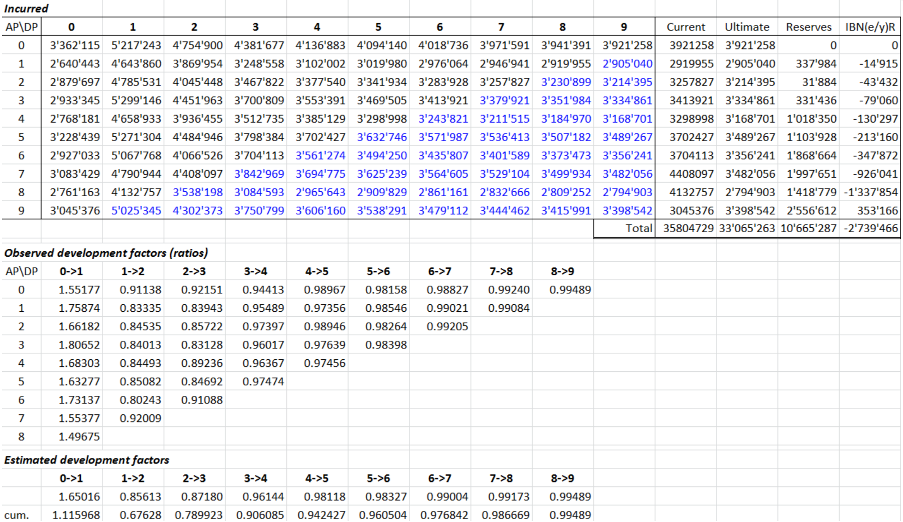
AP\DP	0->1	1->2	2->3	3->4	4->5	5->6	6->7	7->8	8->9
0	1.10721	1.32649	1.27687	1.16419	1.09528	1.12858	1.09264	1.04651	1.04444
1	1.31666	1.15579	1.11034	1.22621	1.16365	1.10727	1.07243	1.12230	
2	1.24358	1.39173	1.12747	1.15442	1.16663	1.03937	1.04435		
3	1.25571	1.28680	1.15950	1.16590	1.15088	1.13546			
4	1.27280	1.19436	1.29335	1.15866	1.16698				
5	1.23067	1.35690	1.23974	1.13310					
6	1.16923	1.18645	1.13079						
7	1.17972	1.37137							
8	1.37442								

Estimated development factors

	0->1	1->2	2->3	3->4	4->5	5->6	6->7	7->8	8->9
	1.23430	1.29036	1.19179	1.16346	1.14565	1.10127	1.07016	1.07602	1.04444
cum.	3.351028	2.714917	2.104007	1.765421	1.517393	1.324478	1.202685	1.12384	1.04444

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.



Incurred

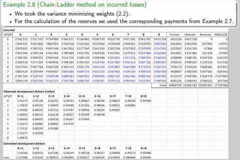
AP\DP	0	1	2	3	4	5	6	7	8	9	Current	Ultimate	Reserves	IBN(e/y)R
0	3'362'115	5'217'243	4'754'900	4'381'677	4'136'883	4'094'140	4'018'736	3'971'591	3'941'391	3'921'258	3921258	3'921'258	0	0
1	2'640'443	4'643'860	3'869'954	3'248'558	3'102'002	3'019'980	2'976'064	2'946'941	2'919'955	2'905'040	2919955	2'905'040	337'984	-14'915
2	2'879'697	4'785'531	4'045'448	3'467'822	3'377'540	3'341'934	3'283'928	3'257'827	3'230'899	3'214'395	3257827	3'214'395	31'884	-43'432
3	2'933'345	5'299'146	4'451'963	3'700'809	3'553'391	3'469'505	3'413'921	3'379'921	3'351'984	3'334'861	3413921	3'334'861	331'436	-79'060
4	2'768'181	4'658'933	3'936'455	3'512'735	3'385'129	3'298'998	3'243'821	3'211'515	3'184'970	3'168'701	3298998	3'168'701	1'018'350	-130'297
5	3'228'439	5'271'304	4'484'946	3'798'384	3'702'427	3'632'746	3'571'987	3'536'413	3'507'182	3'489'267	3702427	3'489'267	1'103'928	-213'160
6	2'927'033	5'067'768	4'066'526	3'704'113	3'561'274	3'494'250	3'435'807	3'401'589	3'373'473	3'356'241	3704113	3'356'241	1'868'664	-347'872
7	3'083'429	4'790'944	4'408'097	3'842'969	3'694'775	3'625'239	3'564'605	3'529'104	3'499'934	3'482'056	4408097	3'482'056	1'997'651	-926'041
8	2'761'163	4'132'757	3'538'198	3'084'593	2'965'643	2'909'829	2'861'161	2'832'666	2'809'252	2'794'903	4132757	2'794'903	1'418'779	-1'337'854
9	3'045'376	5'025'345	4'302'373	3'750'799	3'606'160	3'538'291	3'479'112	3'444'462	3'415'991	3'398'542	3045376	3'398'542	2'556'612	353'166
										Total	35804729	33'065'263	10'665'287	-2'739'466

Observed development factors (ratios)

AP\DP	0->1	1->2	2->3	3->4	4->5	5->6	6->7	7->8	8->9
0	1.55177	0.91138	0.92151	0.94413	0.98967	0.98158	0.98827	0.99240	0.99489
1	1.75874	0.83335	0.83943	0.95489	0.97356	0.98546	0.99021	0.99084	
2	1.66182	0.84535	0.85722	0.97397	0.98946	0.98264	0.99205		
3	1.80652	0.84013	0.83128	0.96017	0.97639	0.98398			
4	1.68303	0.84493	0.89236	0.96367	0.97456				
5	1.63277	0.85082	0.84692	0.97474					
6	1.73137	0.80243	0.91088						
7	1.55377	0.92009							
8	1.49675								

Estimated development factors

	0->1	1->2	2->3	3->4	4->5	5->6	6->7	7->8	8->9
	1.65016	0.85613	0.87180	0.96144	0.98118	0.98327	0.99004	0.99173	0.99489
cum.	1.115968	0.67628	0.789923	0.906085	0.942427	0.960504	0.976842	0.986669	0.99489



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.

Incurred

AP\DP	0	1	2	3	4	5	6	7	8	9	Current	Ultimate	Reserves	IBN(e/y)R
0	3'362'115	5'217'243	4'754'900	4'381'677	4'136'883	4'094'140	4'018'736	3'971'591	3'941'391	3'921'258	3921258	3'921'258	0	0
1	2'640'443	4'643'860	3'869'954	3'248'558	3'102'002	3'019'980	2'976'064	2'946'941	2'919'955	2'905'040	2919955	2'905'040	337'984	-14'915
2	2'879'697	4'785'531	4'045'448	3'467'822	3'377'540	3'341'934	3'283'928	3'257'827	3'230'899	3'214'395	3257827	3'214'395	31'884	-43'432
3	2'933'345	5'299'146	4'451'963	3'700'809	3'553'391	3'469'505	3'413'921	3'379'921	3'351'984	3'334'861	3413921	3'334'861	331'436	-79'060
4	2'768'181	4'658'933	3'936'455	3'512'735	3'385'129	3'298'998	3'243'821	3'211'515	3'184'970	3'168'701	3298998	3'168'701	1'018'350	-130'297
5	3'228'439	5'271'304	4'484'946	3'798'384	3'702'427	3'632'746	3'571'987	3'536'413	3'507'182	3'489'267	3702427	3'489'267	1'103'928	-213'160
6	2'927'033	5'067'768	4'066'526	3'704'113	3'561'274	3'494'250	3'435'807	3'401'589	3'373'473	3'356'241	3704113	3'356'241	1'868'664	-347'872
7	3'083'429	4'790'944	4'408'097	3'842'969	3'694'775	3'625'239	3'564'605	3'529'104	3'499'934	3'482'056	4408097	3'482'056	1'997'651	-926'041
8	2'761'163	4'132'757	3'538'198	3'084'593	2'965'643	2'909'829	2'861'161	2'832'666	2'809'252	2'794'903	4132757	2'794'903	1'418'779	-1'337'854
9	3'045'376	5'025'345	4'302'373	3'750'799	3'606'160	3'538'291	3'479'112	3'444'462	3'415'991	3'398'542	3045376	3'398'542	2'556'612	353'166
										Total	35804729	33'065'263	10'665'287	-2'739'466

Observed development factors (ratios)

AP\DP	0->1	1->2	2->3	3->4	4->5	5->6	6->7	7->8	8->9
0	1.55177	0.91138	0.92151	0.94413	0.98967	0.98158	0.98827	0.99240	0.99489
1	1.75874	0.83335	0.83943	0.95489	0.97356	0.98546	0.99021	0.99084	
2	1.66182	0.84535	0.85722	0.97397	0.98946	0.98264	0.99205		
3	1.80652	0.84013	0.83128	0.96017	0.97639	0.98398			
4	1.68303	0.84493	0.89236	0.96367	0.97456				
5	1.63277	0.85082	0.84692	0.97474					
6	1.73137	0.80243	0.91088						
7	1.55377	0.92009							
8	1.49675								

Estimated development factors

	0->1	1->2	2->3	3->4	4->5	5->6	6->7	7->8	8->9
	1.65016	0.85613	0.87180	0.96144	0.98118	0.98327	0.99004	0.99173	0.99489
cum.	1.115968	0.67628	0.789923	0.906085	0.942427	0.960504	0.976842	0.986669	0.99489

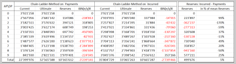
Comparison of the two results

- 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.

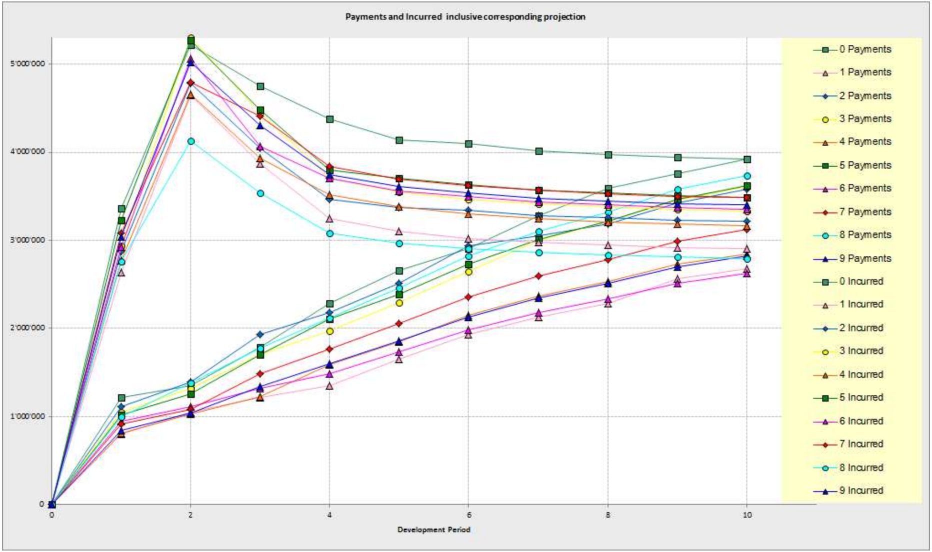
AP\DP	Chain-Ladder-Method on Payments				Chain-Ladder-Method on Incurred				Reserves: Incurred - Payments	
	Current	Ultimate	Reserves	IBN(e/y)R	Current	Ultimate	Reserves	IBN(e/y)R	Reserves	in % of mean Reserves
0	3'921'258	3'921'258	0	0	3'921'258	3'921'258	0	0	0	
1	2'567'056	2'681'142	114'086	-238'813	2'919'955	2'905'040	337'984	-14'915	223'897	99%
2	3'182'511	3'576'632	394'121	318'805	3'257'827	3'214'395	31'884	-43'432	-362'237	-170%
3	3'003'425	3'612'174	608'749	198'253	3'413'921	3'334'861	331'436	-79'060	-277'313	-59%
4	2'150'351	2'848'093	697'742	-450'905	3'298'998	3'168'701	1'018'350	-130'297	320'608	37%
5	2'385'339	3'619'496	1'234'157	-82'931	3'702'427	3'489'267	1'103'928	-213'160	-130'228	-11%
6	1'487'577	2'626'200	1'138'623	-1'077'913	3'704'113	3'356'241	1'868'664	-347'872	730'040	49%
7	1'484'405	3'123'198	1'638'793	-1'284'899	4'408'097	3'482'056	1'997'651	-926'041	358'857	20%
8	1'376'124	3'736'063	2'359'939	-396'694	4'132'757	2'794'903	1'418'779	-1'337'854	-941'160	-50%
9	841'930	2'821'331	1'979'401	-224'045	3'045'376	3'398'542	2'556'612	353'166	577'211	25%
Total	22'399'976	32'565'588	10'165'612	-3'239'141	35'804'729	33'065'263	10'665'287	-2'739'466	499'676	5%

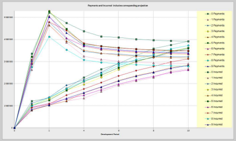
Comparison of the two results

- 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.



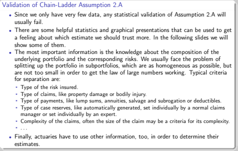
AP\DP	Chain-Ladder-Method on Payments				Chain-Ladder-Method on Incurred				Reserves: Incurred - Payments	
	Current	Ultimate	Reserves	IBN(e/y)R	Current	Ultimate	Reserves	IBN(e/y)R	Reserves	in % of mean Reserves
0	3'921'258	3'921'258	0	0	3'921'258	3'921'258	0	0	0	
1	2'567'056	2'681'142	114'086	-238'813	2'919'955	2'905'040	337'984	-14'915	223'897	99%
2	3'182'511	3'576'632	394'121	318'805	3'257'827	3'214'395	31'884	-43'432	-362'237	-170%
3	3'003'425	3'612'174	608'749	198'253	3'413'921	3'334'861	331'436	-79'060	-277'313	-59%
4	2'150'351	2'848'093	697'742	-450'905	3'298'998	3'168'701	1'018'350	-130'297	320'608	37%
5	2'385'339	3'619'496	1'234'157	-82'931	3'702'427	3'489'267	1'103'928	-213'160	-130'228	-11%
6	1'487'577	2'626'200	1'138'623	-1'077'913	3'704'113	3'356'241	1'868'664	-347'872	730'040	49%
7	1'484'405	3'123'198	1'638'793	-1'284'899	4'408'097	3'482'056	1'997'651	-926'041	358'857	20%
8	1'376'124	3'736'063	2'359'939	-396'694	4'132'757	2'794'903	1'418'779	-1'337'854	-941'160	-50%
9	841'930	2'821'331	1'979'401	-224'045	3'045'376	3'398'542	2'556'612	353'166	577'211	25%
Total	22'399'976	32'565'588	10'165'612	-3'239'141	35'804'729	33'065'263	10'665'287	-2'739'466	499'676	5%


Payments and Incurred inclusive corresponding projection
5'000'000
4'000'000
3'000'000
2'000'000
1'000'000
0
0
2
4
6
8
10
Development Period
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


Payments and Incurred inclusive corresponding projection
5 000 000
4 000 000
3 000 000
2 000 000
1 000 000
0
0
2
4
6
8
10
Development Period
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

Validation of Chain-Ladder Assumption 2.A

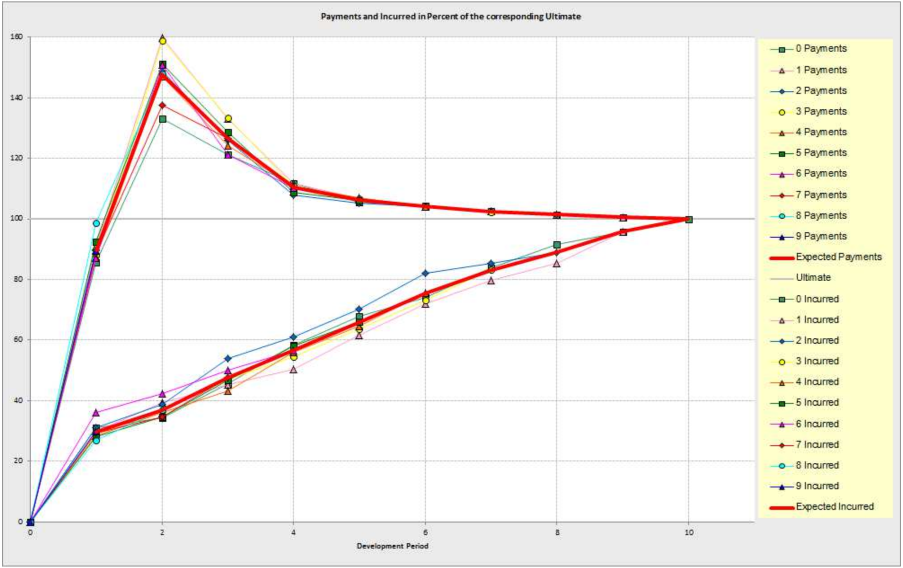
- Since we only have very few data, any statistical validation of Assumption 2.A will usually fail.
- There are some helpful statistics and graphical presentations that can be used to get a feeling about which estimate we should trust more. In the following slides we will show some of them.
- The most 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 estimates.



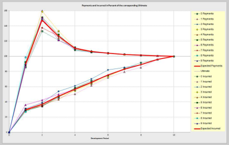
Validation of Chain-Ladder Assumption 2.A

- 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 estimates.

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.

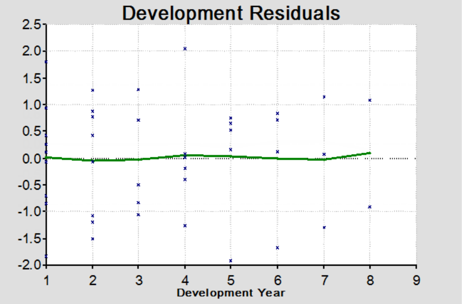


Plot of residuals

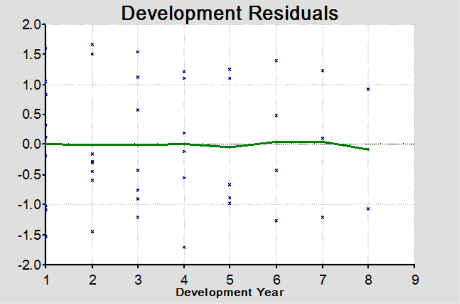
The residuals are defined by

$$\frac{\frac{C_{i,k+1}}{C_{i,k}} - \widehat{f}_k}{\sqrt{\widehat{\mathrm{Var}}\left[\frac{C_{i,k+1}}{C_{i,k}}\middle|\mathcal{D}_k\right]}} = \frac{\frac{C_{i,k+1}}{C_{i,k}} - \widehat{f}_k}{\sqrt{\frac{\widehat{\sigma}_k^2}{C_{i,k}}}}.$$

Payments



Incurred



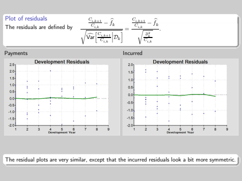
The residual plots are very similar, except that the incurred residuals look a bit more symmetric.

2021-04-26

Stochastic Reserving

└─Chain-Ladder-Method (CLM)

└─Validation and examples (part 1 of 3)



Plot of residuals

The residuals are defined by $\frac{\frac{C_{i,k+1}}{C_{i,k}} - \widehat{f}_k}{\sqrt{\widehat{\mathrm{Var}}\left[\frac{C_{i,k+1}}{C_{i,k}} \middle| \mathcal{D}_k\right]}} = \frac{\frac{C_{i,k+1}}{C_{i,k}} - \widehat{f}_k}{\sqrt{\frac{\widehat{\sigma}_k^2}{C_{i,k}}}}$.

Payments

Incurred

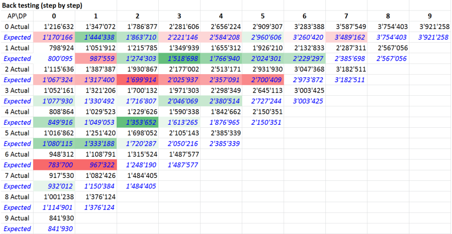
The residual plots are very similar, except that the incurred residuals look a bit more symmetric.

Backtesting step by step

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

$$C_{i,k} \qquad \text{with} \qquad \frac{C_{i,k+1}}{\widehat{f}_k}.$$

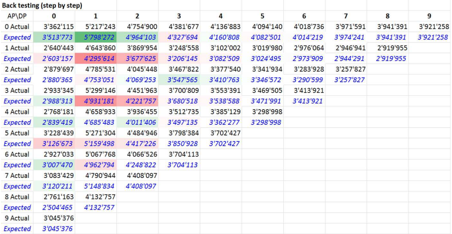
Payments



Back testing (step by step)

AP\DP	0	1	2	3	4	5	6	7	8	9
0 Actual	1'216'632	1'347'072	1'786'877	2'281'606	2'656'224	2'909'307	3'283'388	3'587'549	3'754'403	3'921'258
Expected	*1'170'166*	*1'444'338*	*1'863'710*	*2'221'146*	*2'584'208*	*2'960'606*	*3'260'420*	*3'489'162*	*3'754'403*	*3'921'258*
1 Actual	798'924	1'051'912	1'215'785	1'349'939	1'655'312	1'926'210	2'132'833	2'287'311	2'567'056	
Expected	*800'095*	*987'559*	*1'274'303*	*1'518'698*	*1'766'940*	*2'024'301*	*2'229'297*	*2'385'698*	*2'567'056*	
2 Actual	1'115'636	1'387'387	1'930'867	2'177'002	2'513'171	2'931'930	3'047'368	3'182'511		
Expected	*1'067'324*	*1'317'400*	*1'699'914*	*2'025'937*	*2'357'091*	*2'700'409*	*2'973'872*	*3'182'511*		
3 Actual	1'052'161	1'321'206	1'700'132	1'971'303	2'298'349	2'645'113	3'003'425			
Expected	*1'077'930*	*1'330'492*	*1'716'807*	*2'046'069*	*2'380'514*	*2'727'244*	*3'003'425*			
4 Actual	808'864	1'029'523	1'229'626	1'590'338	1'842'662	2'150'351				
Expected	*849'916*	*1'049'053*	*1'353'652*	*1'613'265*	*1'876'965*	*2'150'351*				
5 Actual	1'016'862	1'251'420	1'698'052	2'105'143	2'385'339					
Expected	*1'080'115*	*1'333'188*	*1'720'287*	*2'050'216*	*2'385'339*					
6 Actual	948'312	1'108'791	1'315'524	1'487'577						
Expected	*783'700*	*967'322*	*1'248'190*	*1'487'577*						
7 Actual	917'530	1'082'426	1'484'405							
Expected	*932'012*	*1'150'384*	*1'484'405*							
8 Actual	1'001'238	1'376'124								
Expected	*1'114'901*	*1'376'124*								
9 Actual	841'930									
Expected	*841'930*									

Incurred



Back testing (step by step)

AP\DP	0	1	2	3	4	5	6	7	8	9
0 Actual	3'362'115	5'217'243	4'754'900	4'381'677	4'136'883	4'094'140	4'018'736	3'971'591	3'941'391	3'921'258
Expected	*3'513'773*	*5'798'272*	*4'964'103*	*4'327'694*	*4'160'808*	*4'082'501*	*4'014'219*	*3'974'241*	*3'941'391*	*3'921'258*
1 Actual	2'640'443	4'643'860	3'869'954	3'248'558	3'102'002	3'019'980	2'976'064	2'946'941	2'919'955	
Expected	*2'603'157*	*4'295'614*	*3'677'625*	*3'206'145*	*3'082'509*	*3'024'495*	*2'973'909*	*2'944'291*	*2'919'955*	
2 Actual	2'879'697	4'785'531	4'045'448	3'467'822	3'377'540	3'341'934	3'283'928	3'257'827		
Expected	*2'880'365*	*4'753'051*	*4'069'253*	*3'547'565*	*3'410'763*	*3'346'572*	*3'290'599*	*3'257'827*		
3 Actual	2'933'345	5'299'146	4'451'963	3'700'809	3'553'391	3'469'505	3'413'921			
Expected	*2'988'313*	*4'931'181*	*4'221'757*	*3'680'518*	*3'538'588*	*3'471'991*	*3'413'921*			
4 Actual	2'768'181	4'658'933	3'936'455	3'512'735	3'385'129	3'298'998				
Expected	*2'839'419*	*4'685'483*	*4'011'406*	*3'497'135*	*3'362'277*	*3'298'998*				
5 Actual	3'228'439	5'271'304	4'484'946	3'798'384	3'702'427					
Expected	*3'126'673*	*5'159'498*	*4'417'226*	*3'850'928*	*3'702'427*					
6 Actual	2'927'033	5'067'768	4'066'526	3'704'113						
Expected	*3'007'470*	*4'962'794*	*4'248'822*	*3'704'113*						
7 Actual	3'083'429	4'790'944	4'408'097							
Expected	*3'120'211*	*5'148'834*	*4'408'097*							
8 Actual	2'761'163	4'132'757								
Expected	*2'504'465*	*4'132'757*								
9 Actual	3'045'376									
Expected	*3'045'376*									

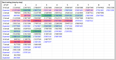
Incurred seems to be a bit more stable, in particular for later development periods.

Backtesting step by step

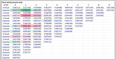
Here we compare the observed values with the one step backwards projected estimate, i.e.

$$C_{i,k} \quad \text{with} \quad \frac{C_{i,k+1}}{\widehat{f}_k}.$$

Payments



Incurred



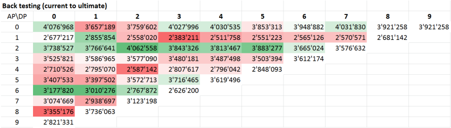
Incurred seems to be a bit more stable, in particular for later development periods.

Backtesting the ultimate

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

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

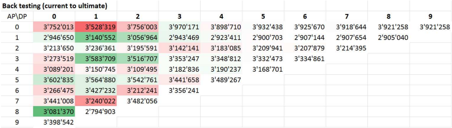
Payments



Back testing (current to ultimate)

AP\DP	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
0	4'076'968	3'657'189	3'759'602	4'027'996	4'030'535	3'853'313	3'948'882	4'031'830	3'921'258	3'921'258
1	2'677'217	2'855'854	2'558'020	2'383'211	2'511'758	2'551'223	2'565'126	2'570'571	2'681'142	
2	3'738'527	3'766'641	4'062'558	3'843'326	3'813'467	3'883'277	3'665'024	3'576'632		
3	3'525'821	3'586'965	3'577'090	3'480'181	3'487'498	3'503'394	3'612'174			
4	2'710'526	2'795'070	2'587'142	2'807'617	2'796'042	2'848'093				
5	3'407'533	3'397'502	3'572'713	3'716'465	3'619'496					
6	3'177'820	3'010'276	2'767'872	2'626'200						
7	3'074'669	2'938'697	3'123'198							
8	3'355'176	3'736'063								
9	2'821'331									

Incurred



Back testing (current to ultimate)

AP\DP	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
0	3'752'013	3'528'319	3'756'003	3'970'171	3'898'710	3'932'438	3'925'670	3'918'644	3'921'258	3'921'258
1	2'946'650	3'140'552	3'056'964	2'943'469	2'923'411	2'900'703	2'907'144	2'907'654	2'905'040	
2	3'213'650	3'236'361	3'195'591	3'142'141	3'183'085	3'209'941	3'207'879	3'214'395		
3	3'273'519	3'583'709	3'516'707	3'353'247	3'348'812	3'332'473	3'334'861			
4	3'089'201	3'150'745	3'109'495	3'182'836	3'190'237	3'168'701				
5	3'602'835	3'564'880	3'542'761	3'441'658	3'489'267					
6	3'266'475	3'427'232	3'212'241	3'356'241						
7	3'441'008	3'240'022	3'482'056							
8	3'081'370	2'794'903								
9	3'398'542									

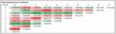
Again, incurred seems to be a bit more stable, in particular for later development periods.

Backtesting the ultimate

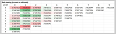
Here we compare the projected ultimate starting at development period k with the one starting at development period $I-i$ (the estimated ultimate), i.e.

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

Payments



Incurred

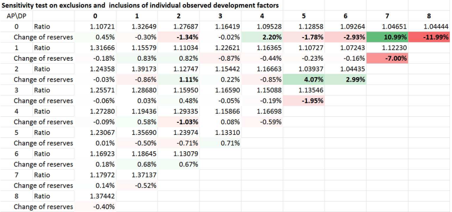


Again, incurred seems to be a bit more stable, in particular for later development periods.

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 $\widehat{f}_k$.

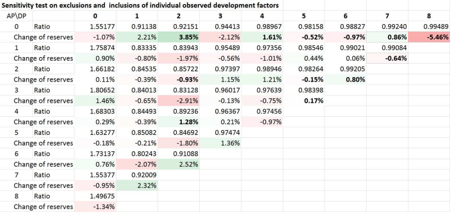
Payments



Sensitivity test on exclusions and inclusions of individual observed development factors

AP\DP		**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
0	Ratio	1.10721	1.32649	1.27687	1.16419	1.09528	1.12858	1.09264	1.04651	1.04444
	Change of reserves	0.45%	-0.30%	**-1.34%**	-0.02%	**2.20%**	**-1.78%**	**-2.93%**	**10.99%**	**-11.99%**
1	Ratio	1.31666	1.15579	1.11034	1.22621	1.16365	1.10727	1.07243	1.12230	
	Change of reserves	-0.18%	0.83%	0.82%	-0.87%	-0.44%	-0.23%	-0.16%	**-7.00%**	
2	Ratio	1.24358	1.39173	1.12747	1.15442	1.16663	1.03937	1.04435		
	Change of reserves	-0.03%	-0.86%	**1.11%**	0.22%	-0.85%	**4.07%**	**2.99%**		
3	Ratio	1.25571	1.28680	1.15950	1.16590	1.15088	1.13546			
	Change of reserves	-0.06%	0.03%	0.48%	-0.05%	-0.19%	**-1.95%**			
4	Ratio	1.27280	1.19436	1.29335	1.15866	1.16698				
	Change of reserves	-0.09%	0.58%	**-1.03%**	0.08%	-0.59%				
5	Ratio	1.23067	1.35690	1.23974	1.13310					
	Change of reserves	0.01%	-0.50%	-0.71%	0.71%					
6	Ratio	1.16923	1.18645	1.13079						
	Change of reserves	0.18%	0.68%	0.67%						
7	Ratio	1.17972	1.37137							
	Change of reserves	0.14%	-0.52%							
8	Ratio	1.37442								
	Change of reserves	-0.40%								

Incurred



Sensitivity test on exclusions and inclusions of individual observed development factors

AP\DP		**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
0	Ratio	1.55177	0.91138	0.92151	0.94413	0.98967	0.98158	0.98827	0.99240	0.99489
	Change of reserves	-1.07%	2.21%	**3.85%**	-2.12%	**1.61%**	**-0.52%**	**-0.97%**	**0.86%**	**-5.46%**
1	Ratio	1.75874	0.83335	0.83943	0.95489	0.97356	0.98546	0.99021	0.99084	
	Change of reserves	0.90%	-0.80%	-1.97%	-0.56%	-1.01%	0.44%	0.06%	**-0.64%**	
2	Ratio	1.66182	0.84535	0.85722	0.97397	0.98946	0.98264	0.99205		
	Change of reserves	0.11%	-0.39%	**-0.93%**	1.15%	1.21%	**-0.15%**	**0.80%**		
3	Ratio	1.80652	0.84013	0.83128	0.96017	0.97639	0.98398			
	Change of reserves	1.46%	-0.65%	-2.91%	-0.13%	-0.75%	**0.17%**			
4	Ratio	1.68303	0.84493	0.89236	0.96367	0.97456				
	Change of reserves	0.29%	-0.39%	**1.28%**	0.21%	-0.97%				
5	Ratio	1.63277	0.85082	0.84692	0.97474					
	Change of reserves	-0.18%	-0.21%	-1.80%	1.36%					
6	Ratio	1.73137	0.80243	0.91088						
	Change of reserves	0.76%	-2.07%	2.52%						
7	Ratio	1.55377	0.92009							
	Change of reserves	-0.95%	2.32%							
8	Ratio	1.49675								
	Change of reserves	-1.34%								

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 $\widehat{f}_k$.

Payments

Incurred

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

Stochastic Reserving

Lecture 3 (Continuation of Lecture 2)

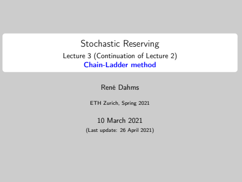
Chain-Ladder method

René Dahms

ETH Zurich, Spring 2021

10 March 2021

(Last update: 26 April 2021)



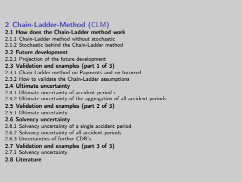
Stochastic Reserving
Lecture 3 (Continuation of Lecture 2)
Chain-Ladder method

René Dahms

ETH Zurich, Spring 2021

10 March 2021
(Last update: 26 April 2021)

2 Chain-Ladder-Method (CLM)



Ultimate uncertainty of a single accident period (repetition)

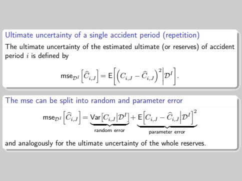
The ultimate uncertainty of the estimated ultimate (or reserves) of accident period i is defined by

$$\mathsf{mse}_{\mathcal{D}^I}\left[\widehat{C}_{i,J}\right] = \mathsf{E}\left[\left(C_{i,J} - \widehat{C}_{i,J}\right)^2 \middle| \mathcal{D}^I\right].$$

The mse can be split into random and parameter error

$$\mathsf{mse}_{\mathcal{D}^I}\left[\widehat{C}_{i,J}\right] = \underbrace{\mathsf{Var}\left[C_{i,J} \middle| \mathcal{D}^I\right]}_{\text{random error}} + \underbrace{\mathsf{E}\left[C_{i,J} - \widehat{C}_{i,J} \middle| \mathcal{D}^I\right]^2}_{\text{parameter error}}$$

and analogously for the ultimate uncertainty of the whole reserves.



Ultimate uncertainty of a single accident period (repetition)

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

$$\mathsf{mse}_{\mathcal{D}^I}\left[\widehat{C}_{i,J}\right] = \mathsf{E}\left[\left(C_{i,J} - \widehat{C}_{i,J}\right)^2 \middle| \mathcal{D}^I\right].$$

The mse can be split into random and parameter error

$$\mathsf{mse}_{\mathcal{D}^I}\left[\widehat{C}_{i,J}\right] = \underbrace{\mathsf{Var}\left[C_{i,J} \middle| \mathcal{D}^I\right]}_{\text{random error}} + \underbrace{\mathsf{E}\left[C_{i,J} - \widehat{C}_{i,J} \middle| \mathcal{D}^I\right]^2}_{\text{parameter error}}$$

and analogously for the ultimate uncertainty of the whole reserves.

$$\begin{aligned}\mathsf{Var}\left[C_{i,J} \middle| \mathcal{D}^I\right] &= \mathsf{Var}\left[C_{i,J} - \widehat{C}_{i,J} \middle| \mathcal{D}^I\right] \\ &= \underbrace{\mathsf{E}\left[\left(C_{i,J} - \widehat{C}_{i,J}\right)^2 \middle| \mathcal{D}^I\right]}_{=\mathsf{mse}_{\mathcal{D}^I}\left[\widehat{C}_{i,J}\right]} - \mathsf{E}\left[C_{i,J} - \widehat{C}_{i,J} \middle| \mathcal{D}^I\right]^2\end{aligned}$$

Taylor approximation of the mse (introduced by Ancus Röhr in [12])

Lets look at the (multi-linear) functional:

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

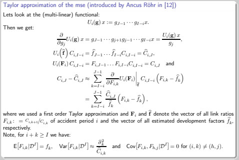
Then we get:

$$\begin{aligned}
\frac{\partial}{\partial g_j} U_i(\mathbf{g})\, x &= g_{J-1} \cdots g_{j+1} g_{j-1} \cdots g_{I-i} x = \frac{U_i(\mathbf{g})\, x}{g_j}, \\
U_i\left(\widehat{\mathbf{f}}\right) C_{i,I-i} &= \widehat{f}_{J-1} \ldots \widehat{f}_{I-i} C_{i,I-i} = \widehat{C}_{i,J}, \\
U_i(\mathbf{F}_i)\, C_{i,I-i} &= F_{i,J-1} \ldots F_{i,I-i} C_{i,I-i} = C_{i,J} \quad \text{and} \\
C_{i,J} - \widehat{C}_{i,J} &\approx \sum_{k=I-i}^{J-1} \frac{\partial}{\partial F_{i,k}} U_i(\mathbf{F}_i) \bigg|_{\widehat{\mathbf{f}}} C_{i,I-i} \left(F_{i,k} - \widehat{f}_k \right) \\
&= \sum_{k=I-i}^{J-1} \frac{\widehat{C}_{i,J}}{\widehat{f}_k} \left(F_{i,k} - \widehat{f}_k \right),
\end{aligned}$$

where we used a first order Taylor approximation and $\mathbf{F}_i$ and $\widehat{\mathbf{f}}$ 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 $\widehat{f}_k$, respectively.

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

$$\mathsf{E}\left[F_{i,k} | \mathcal{D}^I\right] = f_k, \quad \mathsf{Var}\left[F_{i,k} | \mathcal{D}^I\right] \approx \frac{\widehat{\sigma}_k^2}{\widehat{C}_{i,k}} \quad \text{and} \quad \mathsf{Cov}\left[F_{i,k}, F_{h,j} | \mathcal{D}^I\right] = 0 \text{ for } (i,k) \neq (h,j).$$



Taylor approximation of the mse (introduced by Ancus Röhr in [12])

Lets look at the (multi-linear) functional:

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

Then we get:

$$\begin{aligned} \frac{\partial}{\partial g_j} U_i(\mathbf{g})\, x &= g_{J-1} \cdots g_{j+1} g_{j-1} \cdots g_{I-i} x = \frac{U_i(\mathbf{g})\, x}{g_j}, \\ U_i\left(\widehat{\mathbf{f}}\right) C_{i,I-i} &= \widehat{f}_{J-1} \ldots \widehat{f}_{I-i} C_{i,I-i} = \widehat{C}_{i,J}, \\ U_i(\mathbf{F}_i)\, C_{i,I-i} &= F_{i,J-1} \ldots F_{i,I-i} C_{i,I-i} = C_{i,J} \quad \text{and} \\ C_{i,J} - \widehat{C}_{i,J} &\approx \sum_{k=I-i}^{J-1} \frac{\partial}{\partial F_{i,k}} U_i(\mathbf{F}_i)\Big|_{\widehat{\mathbf{f}}} C_{i,I-i} \left(F_{i,k} - \widehat{f}_k\right) \\ &= \sum_{k=I-i}^{J-1} \frac{\widehat{C}_{i,J}}{\widehat{f}_k} \left(F_{i,k} - \widehat{f}_k\right), \end{aligned}$$

where we used a first order Taylor approximation and $\mathbf{F}_i$ and $\widehat{\mathbf{f}}$ 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 $\widehat{f}_k$, respectively.

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

$$\mathsf{E}\left[F_{i,k} | \mathcal{D}^I\right] = f_k, \quad \mathsf{Var}\left[F_{i,k} | \mathcal{D}^I\right] \approx \frac{\widehat{\sigma}_k^2}{\widehat{C}_{i,k}} \quad \text{and} \quad \mathsf{Cov}\left[F_{i,k}, F_{h,j} | \mathcal{D}^I\right] = 0 \text{ for } (i,k) \neq (h,j).$$

Sine $F_{i,k} = {}^{C_{i,k+1}}/C_{i,k}$, we get for $i + k \geq I$

$$\mathsf{E}\Big[F_{i,k} \Big| \mathcal{D}^I\Big] = \mathsf{E}\Big[\mathsf{E}\Big[F_{i,k} \Big| \mathcal{D}_k^{i+k}\Big] \Big| \mathcal{D}^I\Big] = \mathsf{E}\left[\mathsf{E}\left[\frac{C_{i,k+1}}{C_{i,k}} \middle| \mathcal{D}_k^{i+k}\right] \middle| \mathcal{D}^I\right] = \mathsf{E}\Big[f_k \Big| \mathcal{D}^I\Big] = f_k$$

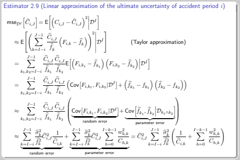
$$\begin{aligned} \mathsf{Var}\Big[F_{i,k} \Big| \mathcal{D}^I\Big] &= \mathsf{E}\Big[\mathsf{Var}\Big[F_{i,k} \Big| \mathcal{D}_k^{i+k}\Big] \Big| \mathcal{D}^I\Big] + \mathsf{Var}\Big[\mathsf{E}\Big[F_{i,k} \Big| \mathcal{D}_k^{i+k}\Big] \Big| \mathcal{D}^I\Big] \\ &= \mathsf{E}\left[\frac{\mathsf{Var}\Big[C_{i,k+1} \Big| \mathcal{D}_k^{i+k}\Big]}{C_{i,k}^2} \middle| \mathcal{D}^I\right] + 0 = \mathsf{E}\left[\frac{\sigma_k^2 C_{i,k}}{C_{i,k}^2} \middle| \mathcal{D}^I\right] \approx \frac{\widehat{\sigma}_k^2}{\widehat{C}_{i,k}}. \end{aligned}$$

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

$$\begin{aligned} \mathsf{Cov}\Big[F_{i,k}, F_{h,j} \Big| \mathcal{D}^I\Big] &= \mathsf{E}\Big[\mathsf{Cov}\Big[F_{i,k}, F_{h,j} \Big| \mathcal{D}_j^{h+j}\Big] \Big| \mathcal{D}^I\Big] + \mathsf{Cov}\Big[\mathsf{E}\Big[F_{i,k} \Big| \mathcal{D}_j^{h+j}\Big], \mathsf{E}\Big[F_{h,j} \Big| \mathcal{D}_j^{h+j}\Big] \Big| \mathcal{D}^I\Big] \\ &= 0 + \mathsf{Cov}\Big[F_{i,k}, f_j \Big| \mathcal{D}^I\Big] = 0 \end{aligned}$$

Estimator 2.9 (Linear approximation of the ultimate uncertainty of accident period i)

$$\begin{aligned}
\text{mse}_{\mathcal{D}^I}\left[\widehat{C}_{i,J}\right] &= \mathsf{E}\left[\left(C_{i,J}-\widehat{C}_{i,J}\right)^2\middle|\mathcal{D}^I\right] \\
&\approx \mathsf{E}\left[\left(\sum_{k=I-i}^{J-1}\frac{\widehat{C}_{i,J}}{\widehat{f}_k}\left(F_{i,k}-\widehat{f}_k\right)\right)^2\middle|\mathcal{D}^I\right] \qquad \text{(Taylor approximation)} \\
&= \sum_{k_1,k_2=I-i}^{J-1}\frac{\widehat{C}_{i,J}}{\widehat{f}_{k_1}}\frac{\widehat{C}_{i,J}}{\widehat{f}_{k_2}}\mathsf{E}\left[\left(F_{i,k_1}-\widehat{f}_{k_1}\right)\left(F_{i,k_2}-\widehat{f}_{k_2}\right)\middle|\mathcal{D}^I\right] \\
&= \sum_{k_1,k_2=I-i}^{J-1}\frac{\widehat{C}_{i,J}}{\widehat{f}_{k_1}}\frac{\widehat{C}_{i,J}}{\widehat{f}_{k_2}}\left(\text{Cov}\left[F_{i,k_1},F_{i,k_2}|\mathcal{D}^I\right]+\left(\widehat{f}_{k_1}-f_{k_1}\right)\left(\widehat{f}_{k_2}-f_{k_2}\right)\right) \\
&\approx \sum_{k_1,k_2=I-i}^{J-1}\frac{\widehat{C}_{i,J}}{\widehat{f}_{k_1}}\frac{\widehat{C}_{i,J}}{\widehat{f}_{k_2}}\left(\underbrace{\text{Cov}\left[F_{i,k_1},F_{i,k_2}|\mathcal{D}^I\right]}_{\text{random error}}+\underbrace{\text{Cov}\left[\widehat{f}_{k_1},\widehat{f}_{k_2}\middle|\mathcal{D}_{k_1\wedge k_2}\right]}_{\text{parameter error}}\right) \\
&\approx \underbrace{\sum_{k=I-i}^{J-1}\frac{\widehat{\sigma}_k^2}{\widehat{f}_k^2}\widehat{C}_{i,J}^2\frac{1}{\widehat{C}_{i,k}}}_{\text{random error}}+\underbrace{\sum_{k=I-i}^{J-1}\frac{\widehat{\sigma}_k^2}{\widehat{f}_k^2}\widehat{C}_{i,J}^2\sum_{h=0}^{I-k-1}\frac{w_{h,k}^2}{C_{h,k}}}_{\text{parameter error}}=\widehat{C}_{i,J}^2\sum_{k=I-i}^{J-1}\frac{\widehat{\sigma}_k^2}{\widehat{f}_k^2}\left(\frac{1}{\widehat{C}_{i,k}}+\sum_{h=0}^{I-k-1}\frac{w_{h,k}^2}{C_{h,k}}\right)
\end{aligned}$$



Estimator 2.9 (Linear approximation of the ultimate uncertainty of accident period i)

$$\begin{aligned}
\mathsf{mse}_{\mathcal{D}^I}\left[\widehat{C}_{i,J}\right] &= \mathsf{E}\left[\left(C_{i,J}-\widehat{C}_{i,J}\right)^2\middle|\mathcal{D}^I\right]\\
&\approx \mathsf{E}\left[\left(\sum_{k=I-i}^{J-1}\frac{\widehat{C}_{i,J}}{\widehat{f}_k}\left(F_{i,k}-\widehat{f}_k\right)\right)^2\middle|\mathcal{D}^I\right] \qquad \text{(Taylor approximation)}\\
&= \sum_{k_1,k_2=I-i}^{J-1}\frac{\widehat{C}_{i,J}}{\widehat{f}_{k_1}}\frac{\widehat{C}_{i,J}}{\widehat{f}_{k_2}}\mathsf{E}\left[\left(F_{i,k_1}-\widehat{f}_{k_1}\right)\left(F_{i,k_2}-\widehat{f}_{k_2}\right)\middle|\mathcal{D}^I\right]\\
&= \sum_{k_1,k_2=I-i}^{J-1}\frac{\widehat{C}_{i,J}}{\widehat{f}_{k_1}}\frac{\widehat{C}_{i,J}}{\widehat{f}_{k_2}}\left(\mathsf{Cov}\left[F_{i,k_1},F_{i,k_2}|\mathcal{D}^I\right]+\left(\widehat{f}_{k_1}-f_{k_1}\right)\left(\widehat{f}_{k_2}-f_{k_2}\right)\right)\\
&\approx \sum_{k_1,k_2=I-i}^{J-1}\frac{\widehat{C}_{i,J}}{\widehat{f}_{k_1}}\frac{\widehat{C}_{i,J}}{\widehat{f}_{k_2}}\left(\underbrace{\mathsf{Cov}\left[F_{i,k_1},F_{i,k_2}|\mathcal{D}^I\right]}_{\text{random error}}+\underbrace{\mathsf{Cov}\left[\widehat{f}_{k_1},\widehat{f}_{k_2}\middle|\mathcal{D}_{k_1\wedge k_2}\right]}_{\text{parameter error}}\right)\\
&\approx \underbrace{\sum_{k=I-i}^{J-1}\frac{\widehat{\sigma}_k^2}{\widehat{f}_k^2}\widehat{C}_{i,J}^2\frac{1}{\widehat{C}_{i,k}}}_{\text{random error}}+\underbrace{\sum_{k=I-i}^{J-1}\frac{\widehat{\sigma}_k^2}{\widehat{f}_k^2}\widehat{C}_{i,J}^2\sum_{h=0}^{I-k-1}\frac{w_{h,k}^2}{C_{h,k}}}_{\text{parameter error}} = \widehat{C}_{i,J}^2\sum_{k=I-i}^{J-1}\frac{\widehat{\sigma}_k^2}{\widehat{f}_k^2}\left(\frac{1}{\widehat{C}_{i,k}}+\sum_{h=0}^{I-k-1}\frac{w_{h,k}^2}{C_{h,k}}\right)
\end{aligned}$$

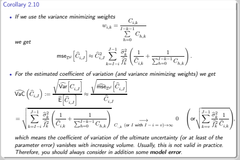
$$\begin{aligned}
&\mathsf{E}\left[\left(F_{i,k_1}-\widehat{f}_{k_1}\right)\left(F_{i,k_2}-\widehat{f}_{k_2}\right)\middle|\mathcal{D}^I\right] = \mathsf{E}\left[\left(\left(F_{i,k_1}-f_{k_1}\right)-\left(\widehat{f}_{k_1}-f_{k_1}\right)\right)\left(\left(F_{i,k_2}-f_{k_2}\right)-\left(\widehat{f}_{k_2}-f_{k_2}\right)\right)\middle|\mathcal{D}^I\right]\\
&\quad= \mathsf{E}\left[\left(F_{i,k_1}-f_{k_1}\right)\left(F_{i,k_2}-f_{k_2}\right)\middle|\mathcal{D}^I\right]-\mathsf{E}\left[\left(F_{i,k_1}-f_{k_1}\right)\left(\widehat{f}_{k_2}-f_{k_2}\right)\middle|\mathcal{D}^I\right]\\
&\qquad-\mathsf{E}\left[\left(\widehat{f}_{k_1}-f_{k_1}\right)\left(F_{i,k_2}-f_{k_2}\right)\middle|\mathcal{D}^I\right]+\mathsf{E}\left[\left(\widehat{f}_{k_1}-f_{k_1}\right)\left(\widehat{f}_{k_2}-f_{k_2}\right)\middle|\mathcal{D}^I\right]\\
&\quad= \mathsf{Cov}\left[F_{i,k_1},F_{i,k_2}\middle|\mathcal{D}^I\right]-\mathsf{E}\left[F_{i,k_1}-f_{k_1}\middle|\mathcal{D}^I\right]\left(\widehat{f}_{k_2}-f_{k_2}\right)\\
&\qquad-\mathsf{E}\left[F_{i,k_2}-f_{k_2}\middle|\mathcal{D}^I\right]\left(\widehat{f}_{k_1}-f_{k_1}\right)+\left(\widehat{f}_{k_1}-f_{k_1}\right)\left(\widehat{f}_{k_2}-f_{k_2}\right)\\
&\quad= \mathsf{Cov}\left[F_{i,k_1},F_{i,k_2}\middle|\mathcal{D}^I\right]-0-0+\left(\widehat{f}_{k_1}-f_{k_1}\right)\left(\widehat{f}_{k_2}-f_{k_2}\right)
\end{aligned}$$

For $k_1 < k_2$ we have $\widehat{f}_{k_1} \in \mathcal{D}_{k_2}$ and $F_{i,k_1} \in \mathcal{D}^I_{k_2}$. This leads to

$$\begin{aligned}
\mathsf{Cov}\left[\widehat{f}_{k_1},\widehat{f}_{k_2}\middle|\mathcal{D}_{k_1\wedge k_2}\right] &= \mathsf{E}\left[\mathsf{Cov}\left[\widehat{f}_{k_1},\widehat{f}_{k_2}\middle|\mathcal{D}_{k_2}\right]\middle|\mathcal{D}_{k_1\wedge k_2}\right]+\mathsf{Cov}\left[\mathsf{E}\left[\widehat{f}_{k_1}\middle|\mathcal{D}_{k_2}\right],\mathsf{E}\left[\widehat{f}_{k_2}\middle|\mathcal{D}_{k_2}\right]\middle|\mathcal{D}_{k_1\wedge k_2}\right]\\
&= 0+\mathsf{Cov}\left[\widehat{f}_{k_1},f_{k_2}\middle|\mathcal{D}_{k_1\wedge k_2}\right]=0\\
\mathsf{Cov}\left[F_{i,k_1},F_{i,k_2}\middle|\mathcal{D}^I\right] &= \mathsf{E}\left[\mathsf{Cov}\left[F_{i,k_1},F_{i,k_2}\middle|\mathcal{D}^I_{k_2}\right]\middle|\mathcal{D}^I\right]+\mathsf{Cov}\left[\mathsf{E}\left[F_{i,k_1}\middle|\mathcal{D}^I_{k_2}\right],\mathsf{E}\left[F_{i,k_2}\middle|\mathcal{D}^I_{k_2}\right]\middle|\mathcal{D}^I\right]\\
&= 0+\mathsf{Cov}\left[F_{i,k_1},f_{k_2}\middle|\mathcal{D}^I\right]=0
\end{aligned}$$

Corollary 2.10

- *If we use the variance minimizing weights*
$$w_{i,k} = \frac{C_{i,k}}{\sum\limits_{h=0}^{I-k-1} C_{h,k}}$$
we get
$$\mathsf{mse}_{\mathcal{D}^I}\left[\widehat{C}_{i,J}\right] \approx \widehat{C}_{i,J}^2 \sum_{k=I-i}^{J-1} \frac{\widehat{\sigma}_k^2}{\widehat{f}_k^2}\left(\frac{1}{\widehat{C}_{i,k}} + \frac{1}{\sum_{h=0}^{I-k-1} C_{h,k}}\right).$$
- *For the estimated coefficient of variation (and variance minimizing weights) we get*
$$\widehat{\mathsf{VaC}}\left(\widehat{C}_{i,J}\right) := \frac{\sqrt{\widehat{\mathsf{Var}}\left[C_{i,J}\right]}}{\widehat{\mathsf{E}}\left[\widehat{C}_{i,J}\right]} \approx \frac{\sqrt{\widehat{\mathsf{mse}}_{\mathcal{D}^I}\left[\widehat{C}_{i,J}\right]}}{\widehat{C}_{i,J}}$$
$$= \sqrt{\sum_{k=I-i}^{J-1} \frac{\widehat{\sigma}_k^2}{\widehat{f}_k^2}\left(\frac{1}{\widehat{C}_{i,k}} + \frac{1}{\sum_{h=0}^{I-k-1} C_{h,k}}\right)} \xrightarrow[C_{\cdot,k}\ \text{(or } I \text{ with } I-i=v)\to\infty]{} 0 \quad \left(\text{or} \sqrt{\sum_{k=v}^{J-1} \frac{\widehat{\sigma}_k^2}{\widehat{f}_k^2}\frac{1}{\widehat{C}_{i,k}}}\right),$$
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****.*



Corollary 2.10

- *If we use the variance minimizing weights*
$$w_{i,k} = \frac{C_{i,k}}{\sum\limits_{h=0}^{I-k-1} C_{h,k}}$$
we get
$$\mathsf{mse}_{\mathcal{D}^I}\left[\widehat{C}_{i,J}\right] \approx \widehat{C}_{i,J}^2 \sum_{k=I-i}^{J-1} \frac{\widehat{\sigma}_k^2}{\widehat{f}_k^2} \left(\frac{1}{\widehat{C}_{i,k}} + \frac{1}{\sum_{h=0}^{I-k-1} C_{h,k}}\right).$$
- *For the estimated coefficient of variation (and variance minimizing weights) we get*
$$\widehat{\mathsf{VaC}}\left(\widehat{C}_{i,J}\right) := \frac{\sqrt{\widehat{\mathsf{Var}}\left[C_{i,J}\right]}}{\widehat{\mathsf{E}}\left[\widehat{C}_{i,J}\right]} \approx \frac{\sqrt{\widehat{\mathsf{mse}}_{\mathcal{D}^I}\left[\widehat{C}_{i,J}\right]}}{\widehat{C}_{i,J}}$$
$$= \sqrt{\sum_{k=I-i}^{J-1} \frac{\widehat{\sigma}_k^2}{\widehat{f}_k^2} \left(\frac{1}{\widehat{C}_{i,k}} + \frac{1}{\sum_{h=0}^{I-k-1} C_{h,k}}\right)} \underset{C_{\cdot,k}\ (\text{or } I \text{ with } I-i=v)\to\infty}{\longrightarrow} 0 \quad \left(\text{or} \sqrt{\sum_{k=v}^{J-1} \frac{\widehat{\sigma}_k^2}{\widehat{f}_k^2} \frac{1}{\widehat{C}_{i,k}}}\right),$$
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 last observed 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 a Taylor approximation you can directly estimate the random and the parameter error like Mack has done in the original approach, see [22]. The result is the same.*
- *For the process error we have made five approximations:*
  - *Taylor approximation,*
  - $\mathrm{Var}\left[{}^1\!/_{C_{i,k}} | \mathcal{D}^I\right] \approx {}^1\!/_{C_{i,k}}$,
  - $C_{i,k} \approx \widehat{C}_{i,k}$,
  - $f_k \approx \widehat{f}_k$, *and*
  - $\sigma_k^2 \approx \widehat{\sigma}_k^2$.

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

Corollary 2.11

- *Instead of using a Taylor approximation you can directly estimate the random and the parameter error like Mack has done in the original approach, see [22]. The result is the same.*
- *For the process error we have made five approximations:*
    - *Taylor approximation,*
    - $\mathsf{Var}\left[{}^1\!/C_{i,k}\middle|\mathcal{D}^I\right] \approx {}^1\!/C_{i,k}$,
    - $C_{i,k} \approx \widehat{C}_{i,k}$,
    - $f_k \approx \widehat{f}_k$, *and*
    - $\sigma_k^2 \approx \widehat{\sigma}_k^2$.

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

Original estimation of the random error:

$$
\begin{aligned}
\mathsf{Var}\left[C_{i,J}\middle|\mathcal{D}^I\right] &= \underbrace{\mathsf{Var}\left[C_{i,J}\middle|\mathcal{B}_{i,I-i}\right]}_{\textbf{iii)}^{\textbf{CLM}}} \\
&= \mathsf{Var}\left[\mathsf{E}\left[C_{i,J}\middle|\mathcal{B}_{i,J-1}\right]\middle|\mathcal{B}_{i,I-i}\right] + \mathsf{E}\left[\mathsf{Var}\left[C_{i,J}\middle|\mathcal{B}_{i,J-1}\right]\middle|\mathcal{B}_{i,I-i}\right] \\
&= \underbrace{\mathsf{Var}\left[f_{J-1}C_{i,J-1}\middle|\mathcal{B}_{i,I-i}\right]}_{\textbf{i)}^{\textbf{CLM}}} + \underbrace{\mathsf{E}\left[\sigma_{J-1}^2 C_{i,J-1}\middle|\mathcal{B}_{i,I-i}\right]}_{\textbf{ii)}^{\textbf{CLM}}} \\
&= f_{J-1}^2\mathsf{Var}\left[C_{i,J-1}\middle|\mathcal{B}_{i,I-i}\right] + \sigma_{J-1}^2 \underbrace{\prod_{j=I-i}^{J-2} f_j\, C_{i,I-i}}_{\textbf{Corollary } 2.3} \\
= \cdots &= \sum_{k=I-i}^{J-1} \prod_{j=k+1}^{J-1} f_j^2\, \sigma_k^2 \prod_{j=I-i}^{k-1} f_j\, C_{i,I-i} \\
&= \sum_{k=I-i}^{J-1} \frac{\sigma_k^2}{f_k^2\, \prod_{j=I-i}^{k-1} f_j\, C_{i.I-i}} \left(\prod_{j=I-i}^{J-1} f_j\, C_{i,I-i}\right)^2 \\
&\approx \widehat{C}_{i,J}^2 \sum_{k=I-i}^{J-1} \frac{\widehat{\sigma}_k^2}{\widehat{f}_k^2\, \prod_{j=I-i}^{k-1} \widehat{f}_j\, C_{i.I-i}} \\
&= \widehat{C}_{i,J}^2 \sum_{k=I-i}^{J-1} \frac{\widehat{\sigma}_k^2}{\widehat{f}_k^2 \widehat{C}_{i,k}}
\end{aligned}
$$

Estimator 2.12 (Variance parameter)

Let Assumption 2.A be fulfilled. Then

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

with

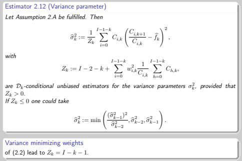
$$Z_k := I - 2 - k + \sum_{i=0}^{I-1-k} w_{i,k}^2 \frac{1}{C_{i,k}} \sum_{h=0}^{I-1-k} C_{h,k},$$

are $\mathcal{D}_k$-conditional unbiased estimators for the variance parameters σ_k^2, provided that $Z_k > 0$.
If $Z_k \leq 0$ one could take

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

Variance minimizing weights

of (2.2) lead to $Z_k = I - k - 1$.



Estimator 2.12 (Variance parameter)

Let Assumption 2.A be fulfilled. Then

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

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 $\mathcal{D}_k$-conditional unbiased estimators for the variance parameters σ_k^2, provided that $Z_k > 0$.
If $Z_k \leq 0$ one could take

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

Variance minimizing weights

of (2.2) lead to $Z_k = I - k - 1$.

Unbiasedness:

$$\mathsf{E}\left[C_{i,k}\left(\frac{C_{i,k+1}}{C_{i,k}} - \widehat{f}_k\right)^2 \middle| \mathcal{D}_k\right] = C_{i,k}\mathsf{Var}\left[\frac{C_{i,k+1}}{C_{i,k}} - \widehat{f}_k \middle| \mathcal{D}_k\right]$$

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

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

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

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

$$= \sigma_k^2\left(1 - 2w_{i,k} + C_{i,k}\sum_{h=0}^{I-1-k}\frac{w_{h,k}^2}{C_{h,k}}\right) \underbrace{\Longrightarrow}_{\textbf{change order of summation}} \sum_{i=0}^{I-1-k}\mathsf{E}\left[C_{i,k}\left(\frac{C_{i,k+1}}{C_{i,k}} - \widehat{f}_k\right)^2 \middle| \mathcal{D}_k\right] = Z_k$$

Taking the variance minimizing weights we get

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

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:

$$\mathrm{mse}_{\mathcal{D}^I}\left[\sum_{i=0}^{I}\widehat{C}_{i,J}\right] = \underbrace{\mathrm{Var}\left[\sum_{i=I-J+1}^{I} C_{i,J}\middle|\mathcal{D}^I\right]}_{\text{random error}} + \underbrace{\mathrm{E}\left[\sum_{i=I-J+1}^{I}\left(C_{i,J}-\widehat{C}_{i,J}\right)\middle|\mathcal{D}^I\right]^2}_{\text{squared parameter estimation error}}.$$

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.

2021-04-26

Stochastic Reserving
Chain-Ladder-Method (CLM)
Ultimate uncertainty

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:

$$\text{mse}_{\mathcal{D}^I}\left[\sum_{i=0}^{I}\widehat{C}_{i,J}\right] = \underbrace{\text{Var}\left[\sum_{i=I-J+1}^{I} C_{i,J}\middle|\mathcal{D}^I\right]}_{\text{random error}} + \underbrace{\text{E}\left[\sum_{i=I-J+1}^{I}\left(C_{i,J}-\widehat{C}_{i,J}\right)\middle|\mathcal{D}^I\right]^2}_{\text{squared parameter estimation error}}.$$

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)

$$\mathsf{mse}_{\mathcal{D}^I}\left[\sum_{i=0}^{I}\widehat{C}_{i,J}\right]=\mathsf{E}\left[\left(\sum_{i=0}^{I}\left(C_{i,J}-\widehat{C}_{i,J}\right)\right)^2\middle|\mathcal{D}^I\right]$$

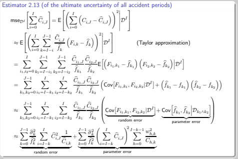
$$\approx\mathsf{E}\left[\left(\sum_{i=0}^{I}\sum_{k=I-i}^{J-1}\frac{\widehat{C}_{i,J}}{\widehat{f}_k}\left(F_{i,k}-\widehat{f}_k\right)\right)^2\middle|\mathcal{D}^I\right]\qquad\text{(Taylor approximation)}$$

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

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

$$\approx\sum_{k_1,k_2=0}^{J-1}\sum_{i_1=I-k_1}^{I}\sum_{i_2=I-k_2}^{I}\frac{\widehat{C}_{i_1,J}}{\widehat{f}_{k_1}}\frac{\widehat{C}_{i_2,J}}{\widehat{f}_{k_2}}\left(\underbrace{\mathsf{Cov}\left[F_{i_1,k_1},F_{i_2,k_2}|\mathcal{D}^I\right]}_{\text{random error}}+\underbrace{\mathsf{Cov}\left[\widehat{f}_{k_1},\widehat{f}_{k_2}\middle|\mathcal{D}_{k_1\wedge k_2}\right]}_{\text{parameter error}}\right)$$

$$\approx\underbrace{\sum_{k=0}^{J-1}\frac{\widehat{\sigma}_k^2}{\widehat{f}_k^2}\sum_{i=I-k}^{I}\widehat{C}_{i,J}^2\frac{1}{\widehat{C}_{i,k}}}_{\text{random error}}+\underbrace{\sum_{k=0}^{J-1}\frac{\widehat{\sigma}_k^2}{\widehat{f}_k^2}\left(\sum_{i=I-k}^{I}\widehat{C}_{i,J}\right)^2\sum_{h=0}^{I-k-1}\frac{w_{h,k}^2}{C_{h,k}}}_{\text{parameter error}}$$



Estimator 2.13 (of the ultimate uncertainty of all accident periods)

$$
\begin{aligned}
\text{mse}_{\mathcal{D}^I}\left[\sum_{i=0}^{I}\widehat{C}_{i,J}\right] &= \mathsf{E}\left[\left(\sum_{i=0}^{I}\left(C_{i,J}-\widehat{C}_{i,J}\right)\right)^2\middle|\mathcal{D}^I\right] \\
&\approx \mathsf{E}\left[\left(\sum_{i=0}^{I}\sum_{k=I-i}^{J-1}\frac{\widehat{C}_{i,J}}{\widehat{f}_k}\left(F_{i,k}-\widehat{f}_k\right)\right)^2\middle|\mathcal{D}^I\right] \qquad \text{(Taylor approximation)} \\
&= \sum_{i_1,i_2=0}^{I}\sum_{k_1=I-i_1}^{J-1}\sum_{k_2=I-i_2}^{J-1}\frac{\widehat{C}_{i_1,J}}{\widehat{f}_{k_1}}\frac{\widehat{C}_{i_2,J}}{\widehat{f}_{k_2}}\mathsf{E}\left[\left(F_{i_1,k_1}-\widehat{f}_{k_1}\right)\left(F_{i_2,k_2}-\widehat{f}_{k_2}\right)\middle|\mathcal{D}^I\right] \\
&= \sum_{k_1,k_2=0}^{J-1}\sum_{i_1=I-k_1}^{I}\sum_{i_2=I-k_2}^{I}\frac{\widehat{C}_{i_1,J}}{\widehat{f}_{k_1}}\frac{\widehat{C}_{i_2,J}}{\widehat{f}_{k_2}}\left(\mathsf{Cov}[F_{i_1,k_1},F_{i_2,k_2}|\mathcal{D}^I]+\left(\widehat{f}_{k_1}-f_{k_1}\right)\left(\widehat{f}_{k_2}-f_{k_2}\right)\right) \\
&\approx \sum_{k_1,k_2=0}^{J-1}\sum_{i_1=I-k_1}^{I}\sum_{i_2=I-k_2}^{I}\frac{\widehat{C}_{i_1,J}}{\widehat{f}_{k_1}}\frac{\widehat{C}_{i_2,J}}{\widehat{f}_{k_2}}\left(\underbrace{\mathsf{Cov}[F_{i_1,k_1},F_{i_2,k_2}|\mathcal{D}^I]}_{\text{random error}}+\underbrace{\mathsf{Cov}\left[\widehat{f}_{k_1},\widehat{f}_{k_2}\middle|\mathcal{D}_{k_1\wedge k_2}\right]}_{\text{parameter error}}\right) \\
&\approx \underbrace{\sum_{k=0}^{J-1}\frac{\widehat{\sigma}_k^2}{\widehat{f}_k^2}\sum_{i=I-k}^{I}\widehat{C}_{i,J}^2\frac{1}{\widehat{C}_{i,k}}}_{\text{random error}}+\underbrace{\sum_{k=0}^{J-1}\frac{\widehat{\sigma}_k^2}{\widehat{f}_k^2}\left(\sum_{i=I-k}^{I}\widehat{C}_{i,J}\right)^2\sum_{h=0}^{I-k-1}\frac{w_{h,k}^2}{C_{h,k}}}_{\text{parameter error}}
\end{aligned}
$$

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

$$
\begin{aligned}
&\mathsf{Cov}\left[F_{i_1,k_1},F_{i_2,k_2}\middle|\mathcal{D}^I\right] \\
&\qquad = \mathsf{Cov}\left[\mathsf{E}\left[F_{i_1,k_1}\middle|\mathcal{D}^I_{k_1\vee k_2}\right],\mathsf{E}\left[F_{i_2,k_2}\middle|\mathcal{D}^I_{k_1\vee k_2}\right]\middle|\mathcal{D}^I\right]+\mathsf{E}\left[\mathsf{Cov}\left[F_{i_1,k_1},F_{i_2,k_2}\middle|\mathcal{D}^I_{k_1\vee k_2}\right]\middle|\mathcal{D}^I\right] \\
&\qquad = 0+\mathbf{1}_{i_1=i_2}\mathbf{1}_{k_1=k_2}\sigma_{k_1}^2\,\mathsf{E}\left[\frac{1}{C_{i_1,k_1}}\middle|\mathcal{D}^I\right]\approx\mathbf{1}_{i_1=i_2}\mathbf{1}_{k_1=k_2}\widehat{\sigma}_{k_1}^2\,\frac{1}{\widehat{C}_{i_1,k_1}}
\end{aligned}
$$

$$
\mathsf{Cov}\left[\widehat{f}_{k_1},\widehat{f}_{k_2}\middle|\mathcal{D}_{k_1\wedge k_2}\right]=\mathbf{1}_{k_1=k_2}\sigma_{k_1}^2\sum_{h=0}^{I-k_1-1}\frac{w_{h,k_1}^2}{C_{h,k}}.
$$

Therefore,

$$
\begin{aligned}
&\text{mse}_{\mathcal{D}^I}\left[\widehat{C}_{i,J}\right] \\
&\approx \sum_{k_1,k_2=0}^{J-1}\sum_{i_1=I-k_1}^{I}\sum_{i_2=I-k_2}^{I}\frac{\widehat{C}_{i_1,J}}{\widehat{f}_{k_1}}\frac{\widehat{C}_{i_2,J}}{\widehat{f}_{k_2}}\widehat{\sigma}_{k_1}^2\mathbf{1}_{k_1=k_2}\left(\mathbf{1}_{i_1=i_2}\frac{1}{\widehat{C}_{i_1,k_2}}+\sum_{h=0}^{I-k_1-1}\frac{w_{h,k_1}^2}{C_{h,k_1}}\right) \\
&= \sum_{k=0}^{J-1}\frac{\widehat{\sigma}_k^2}{\widehat{f}_k^2}\left(\sum_{i=I-k}^{I}\widehat{C}_{i,J}^2\frac{1}{\widehat{C}_{i,k}}+\left(\sum_{i=I-k}^{I}\widehat{C}_{i,J}\right)^2\sum_{h=0}^{I-k-1}\frac{w_{h,k}^2}{C_{h,k}}\right)
\end{aligned}
$$

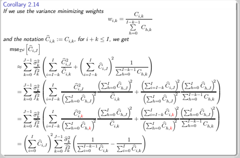
Corollary 2.14

If we use the variance minimizing weights

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

and the notation $\widehat{C}_{i,k} := C_{i,k}$*, for* $i + k \leq I$*, we get*

$$\begin{aligned}
&\text{mse}_{\mathcal{D}^I}\left[\widehat{C}_{i,J}\right] \\
&\approx \sum_{k=0}^{J-1} \frac{\widehat{\sigma}_k^2}{\widehat{f}_k^2} \left(\sum_{i=I-k}^{I} \frac{\widehat{C}_{i,J}^2}{\widehat{C}_{i,k}} + \left(\sum_{i=I-k}^{I} \widehat{C}_{i,J} \right)^2 \frac{1}{\sum_{h=0}^{I-k-1} C_{h,k}} \right) \\
&= \sum_{k=0}^{J-1} \frac{\widehat{\sigma}_k^2}{\widehat{f}_k^2} \left(\sum_{i=I-k}^{I} \frac{\widehat{C}_{i,J}^2}{\left(\sum_{h=0}^{I} \widehat{C}_{h,J}\right)^2} \frac{\left(\sum_{h=0}^{I} \widehat{C}_{h,J}\right)^2}{\widehat{C}_{i,k}} + \frac{\left(\sum_{i=I-k}^{I} \widehat{C}_{i,J}\right)^2}{\left(\sum_{h=0}^{I} \widehat{C}_{h,J}\right)^2} \frac{\left(\sum_{h=0}^{I} \widehat{C}_{h,J}\right)^2}{\sum_{h=0}^{I-k-1} C_{h,k}} \right) \\
&= \sum_{k=0}^{J-1} \frac{\widehat{\sigma}_k^2}{\widehat{f}_k^2} \left(\sum_{i=I-k}^{I} \frac{\widehat{C}_{i,k}^2}{\left(\sum_{h=0}^{I} \widehat{C}_{h,k}\right)^2} \frac{\left(\sum_{h=0}^{I} \widehat{C}_{h,J}\right)^2}{\widehat{C}_{i,k}} + \frac{\left(\sum_{i=I-k}^{I} \widehat{C}_{i,k}\right)^2}{\left(\sum_{h=0}^{I} \widehat{C}_{h,k}\right)^2} \frac{\left(\sum_{h=0}^{I} \widehat{C}_{h,J}\right)^2}{\sum_{h=0}^{I-k-1} C_{h,k}} \right) \\
&= \left(\sum_{i=0}^{I} \widehat{C}_{i,J} \right)^2 \sum_{k=0}^{J-1} \frac{\widehat{\sigma}_k^2}{\widehat{f}_k^2} \left(\frac{1}{\sum_{i=0}^{I-k-1} \widehat{C}_{i,k}} - \frac{1}{\sum_{i=0}^{I} \widehat{C}_{i,k}} \right)
\end{aligned}$$



Corollary 2.14

If we use the variance minimizing weights

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

and the notation $\widehat{C}_{i,k} := C_{i,k}$, *for* $i+k \leq I$, *we get*

$$\begin{aligned}
&\mathsf{mse}_{\mathcal{D}^I}\left[\widehat{C}_{i,J}\right] \\
&\approx \sum_{k=0}^{J-1} \frac{\widehat{\sigma}_k^2}{\widehat{f}_k^2} \left(\sum_{i=I-k}^{I} \frac{\widehat{C}_{i,J}^2}{\widehat{C}_{i,k}} + \left(\sum_{i=I-k}^{I} \widehat{C}_{i,J} \right)^2 \frac{1}{\sum_{h=0}^{I-k-1} C_{h,k}} \right) \\
&= \sum_{k=0}^{J-1} \frac{\widehat{\sigma}_k^2}{\widehat{f}_k^2} \left(\sum_{i=I-k}^{I} \frac{\widehat{C}_{i,J}^2}{\left(\sum_{h=0}^{I} \widehat{C}_{h,J}\right)^2} \frac{\left(\sum_{h=0}^{I} \widehat{C}_{h,J}\right)^2}{\widehat{C}_{i,k}} + \frac{\left(\sum_{i=I-k}^{I} \widehat{C}_{i,J}\right)^2}{\left(\sum_{h=0}^{I} \widehat{C}_{h,J}\right)^2} \frac{\left(\sum_{h=0}^{I} \widehat{C}_{h,J}\right)^2}{\sum_{h=0}^{I-k-1} C_{h,k}} \right) \\
&= \sum_{k=0}^{J-1} \frac{\widehat{\sigma}_k^2}{\widehat{f}_k^2} \left(\sum_{i=I-k}^{I} \frac{\widehat{C}_{i,k}^2}{\left(\sum_{h=0}^{I} \widehat{C}_{h,k}\right)^2} \frac{\left(\sum_{h=0}^{I} \widehat{C}_{h,J}\right)^2}{\widehat{C}_{i,k}} + \frac{\left(\sum_{i=I-k}^{I} \widehat{C}_{i,k}\right)^2}{\left(\sum_{h=0}^{I} \widehat{C}_{h,k}\right)^2} \frac{\left(\sum_{h=0}^{I} \widehat{C}_{h,J}\right)^2}{\sum_{h=0}^{I-k-1} C_{h,k}} \right) \\
&= \left(\sum_{i=0}^{I} \widehat{C}_{i,J} \right)^2 \sum_{k=0}^{J-1} \frac{\widehat{\sigma}_k^2}{\widehat{f}_k^2} \left(\frac{1}{\sum_{i=0}^{I-k-1} \widehat{C}_{i,k}} - \frac{1}{\sum_{i=0}^{I} \widehat{C}_{i,k}} \right)
\end{aligned}$$

For each $k < J$ we have

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

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

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

Finally,

$$\begin{aligned}
&\sum_{i=I-k}^{I} \frac{\widehat{C}_{i,k}^2}{\left(\sum_{h=0}^{I} \widehat{C}_{h,k}\right)^2} \frac{1}{\widehat{C}_{i,k}} + \frac{\left(\sum_{i=I-k}^{I} \widehat{C}_{i,k}\right)^2}{\left(\sum_{h=0}^{I} \widehat{C}_{h,k}\right)^2} \frac{1}{\sum_{h=0}^{I-k-1} C_{h,k}} \\
&\qquad = \frac{\sum_{i=I-k}^{I} \widehat{C}_{i,k} \sum_{i=0}^{I-k-1} \widehat{C}_{i,k} + \left(\sum_{i=I-k}^{I} \widehat{C}_{i,k}\right)^2}{\left(\sum_{i=0}^{I} \widehat{C}_{i,k}\right)^2 \sum_{i=0}^{I-k-1} C_{i,k}} \\
&\qquad = \frac{\sum_{i=I-k}^{I} \widehat{C}_{i,k} \left(\sum_{i=0}^{I-k-1} \widehat{C}_{i,k} + \sum_{i=I-k}^{I} \widehat{C}_{i,k}\right)}{\left(\sum_{i=0}^{I} \widehat{C}_{i,k}\right)^2 \sum_{i=0}^{I-k-1} C_{i,k}} \\
&\qquad = \frac{\sum_{i=I-k}^{I} \widehat{C}_{i,k}}{\sum_{h=0}^{I} \widehat{C}_{h,k} \sum_{h=0}^{I-k-1} C_{h,k}} = \frac{1}{\sum_{i=0}^{I-k-1} \widehat{C}_{i,k}} - \frac{1}{\sum_{i=0}^{I} \widehat{C}_{i,k}}
\end{aligned}$$

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 $\widehat{C}^m_{i,J}$, $0 \le m \le M$, *be estimates of the same (unknown) ultimate. Then*

$$\sum_{m=0}^{M} \underbrace{\min\left(\frac{\widehat{C}^m_{i,J}}{C^m_{i,I-i}}, \frac{C^m_{i,I-i}}{\widehat{C}^m_{i,J}}\right)\left(\sum_{l=0}^{M}\min\left(\frac{\widehat{C}^l_{i,J}}{C^l_{i,I-i}}, \frac{C^l_{i,I-i}}{\widehat{C}^l_{i,J}}\right)\right)^{-1}}_{\text{mixing weights}} \widehat{C}^m_{i,J}$$

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

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

$$\sum_{m=0}^{M} \underbrace{\min\left(\frac{\widehat{C}^m_{i,J}}{C^m_{i,I-i}}, \frac{C^m_{i,I-i}}{\widehat{C}^m_{i,J}}\right)\left(\sum_{l=0}^{M} \min\left(\frac{\widehat{C}^l_{i,J}}{C^l_{i,I-i}}, \frac{C^l_{i,I-i}}{\widehat{C}^l_{i,J}}\right)\right)^{-1}}_{\text{mixing weights}} \widehat{C}^m_{i,J}$$

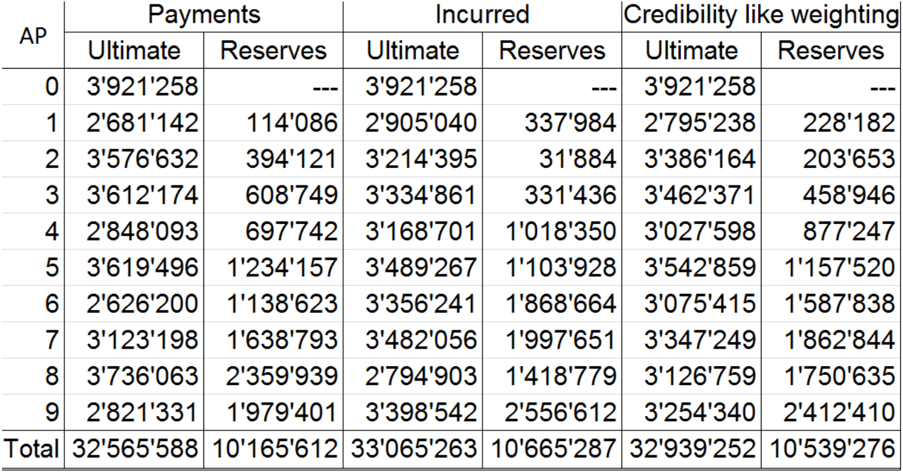
is a credibility like weighted mean of these estimates.

Remark 2.16 (Credibility like weighted ultimate uncertainty)

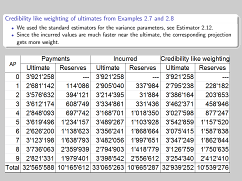
We will see later, see Section 4, that it is possible to transfer the weighting of ultimates to the corresponding ultimate uncertainties.

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.



AP	Payments		Incurred		Credibility like weighting	
	Ultimate	Reserves	Ultimate	Reserves	Ultimate	Reserves
0	3'921'258	---	3'921'258	---	3'921'258	---
1	2'681'142	114'086	2'905'040	337'984	2'795'238	228'182
2	3'576'632	394'121	3'214'395	31'884	3'386'164	203'653
3	3'612'174	608'749	3'334'861	331'436	3'462'371	458'946
4	2'848'093	697'742	3'168'701	1'018'350	3'027'598	877'247
5	3'619'496	1'234'157	3'489'267	1'103'928	3'542'859	1'157'520
6	2'626'200	1'138'623	3'356'241	1'868'664	3'075'415	1'587'838
7	3'123'198	1'638'793	3'482'056	1'997'651	3'347'249	1'862'844
8	3'736'063	2'359'939	2'794'903	1'418'779	3'126'759	1'750'635
9	2'821'331	1'979'401	3'398'542	2'556'612	3'254'340	2'412'410
Total	32'565'588	10'165'612	33'065'263	10'665'287	32'939'252	10'539'276

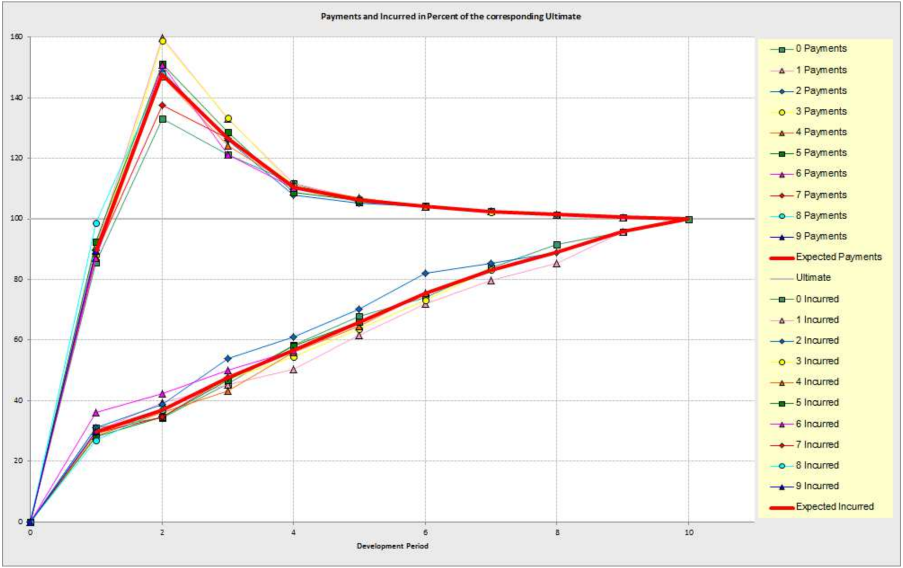


Credibility like weighting of ultimates from Examples 2.7 and 2.8

- 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.

AP	Payments		Incurred		Credibility like weighting	
	Ultimate	Reserves	Ultimate	Reserves	Ultimate	Reserves
0	3'921'258	---	3'921'258	---	3'921'258	---
1	2'681'142	114'086	2'905'040	337'984	2'795'238	228'182
2	3'576'632	394'121	3'214'395	31'884	3'386'164	203'653
3	3'612'174	608'749	3'334'861	331'436	3'462'371	458'946
4	2'848'093	697'742	3'168'701	1'018'350	3'027'598	877'247
5	3'619'496	1'234'157	3'489'267	1'103'928	3'542'859	1'157'520
6	2'626'200	1'138'623	3'356'241	1'868'664	3'075'415	1'587'838
7	3'123'198	1'638'793	3'482'056	1'997'651	3'347'249	1'862'844
8	3'736'063	2'359'939	2'794'903	1'418'779	3'126'759	1'750'635
9	2'821'331	1'979'401	3'398'542	2'556'612	3'254'340	2'412'410
Total	32'565'588	10'165'612	33'065'263	10'665'287	32'939'252	10'539'276

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.



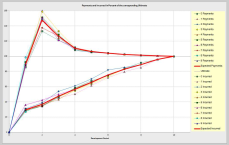
2021-04-26

Stochastic Reserving

Chain-Ladder-Method (CLM)

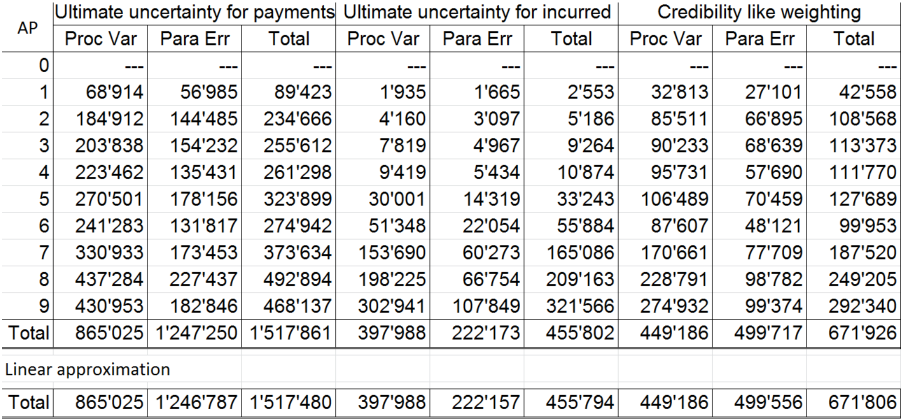
Validation and examples (part 2 of 3)

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.



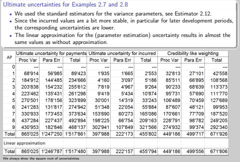
Ultimate uncertainties for Examples 2.7 and 2.8

- 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 as without approximation.



AP	Ultimate uncertainty for payments			Ultimate uncertainty for incurred			Credibility like weighting		
	Proc Var	Para Err	Total	Proc Var	Para Err	Total	Proc Var	Para Err	Total
0	---	---	---	---	---	---	---	---	---
1	68'914	56'985	89'423	1'935	1'665	2'553	32'813	27'101	42'558
2	184'912	144'485	234'666	4'160	3'097	5'186	85'511	66'895	108'568
3	203'838	154'232	255'612	7'819	4'967	9'264	90'233	68'639	113'373
4	223'462	135'431	261'298	9'419	5'434	10'874	95'731	57'690	111'770
5	270'501	178'156	323'899	30'001	14'319	33'243	106'489	70'459	127'689
6	241'283	131'817	274'942	51'348	22'054	55'884	87'607	48'121	99'953
7	330'933	173'453	373'634	153'690	60'273	165'086	170'661	77'709	187'520
8	437'284	227'437	492'894	198'225	66'754	209'163	228'791	98'782	249'205
9	430'953	182'846	468'137	302'941	107'849	321'566	274'932	99'374	292'340
Total	865'025	1'247'250	1'517'861	397'988	222'173	455'802	449'186	499'717	671'926
Linear approximation									
Total	865'025	1'246'787	1'517'480	397'988	222'157	455'794	449'186	499'556	671'806

We always show the square root of uncertainties.



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 as without approximation.

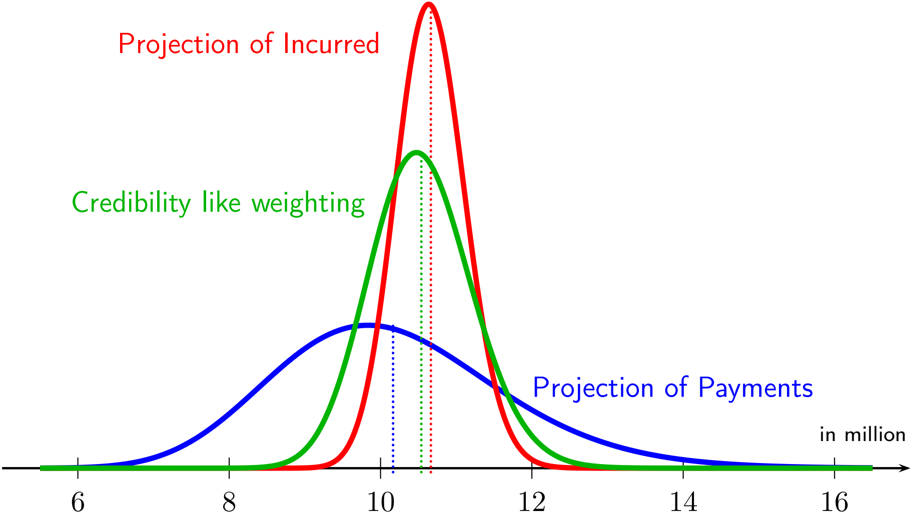
AP	Ultimate uncertainty for payments			Ultimate uncertainty for incurred			Credibility like weighting		
	Proc Var	Para Err	Total	Proc Var	Para Err	Total	Proc Var	Para Err	Total
0	---	---	---	---	---	---	---	---	---
1	68'914	56'985	89'423	1'935	1'665	2'553	32'813	27'101	42'558
2	184'912	144'485	234'666	4'160	3'097	5'186	85'511	66'895	108'568
3	203'838	154'232	255'612	7'819	4'967	9'264	90'233	68'639	113'373
4	223'462	135'431	261'298	9'419	5'434	10'874	95'731	57'690	111'770
5	270'501	178'156	323'899	30'001	14'319	33'243	106'489	70'459	127'689
6	241'283	131'817	274'942	51'348	22'054	55'884	87'607	48'121	99'953
7	330'933	173'453	373'634	153'690	60'273	165'086	170'661	77'709	187'520
8	437'284	227'437	492'894	198'225	66'754	209'163	228'791	98'782	249'205
9	430'953	182'846	468'137	302'941	107'849	321'566	274'932	99'374	292'340
Total	865'025	1'247'250	1'517'861	397'988	222'173	455'802	449'186	499'717	671'926
Linear approximation									
Total	865'025	1'246'787	1'517'480	397'988	222'157	455'794	449'186	499'556	671'806

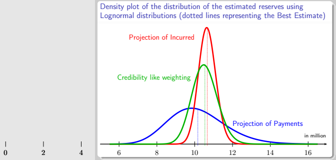
We always show the square root of uncertainties.

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

One can derive estimators for the ultimate uncertainty without a first order Taylor approximation, see [21]. In practice, the resulting figures are almost alike.

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 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?

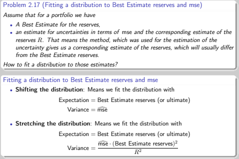
Fitting a distribution to Best Estimate reserves and mse

- **Shifting the distribution**: Means we fit the distribution with

$$\text{Expectation} = \text{Best Estimate reserves (or ultimate)}$$
$$\text{Variance} = \widehat{\text{mse}}$$

- **Stretching the distribution**: Means we fit the distribution with

$$\text{Expectation} = \text{Best Estimate reserves (or ultimate)}$$
$$\text{Variance} = \frac{\widehat{\text{mse}} \cdot (\text{Best Estimate reserves})^2}{R^2}$$



Problem 2.17 (Fitting a distribution to Best Estimate reserves and 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 mse

- **Shifting the distribution**: Means we fit the distribution with

$$\begin{aligned}\text{Expectation} &= \text{Best Estimate reserves (or ultimate)}\\ \text{Variance} &= \widehat{\text{mse}}\end{aligned}$$

- **Stretching the distribution**: Means we fit the distribution with

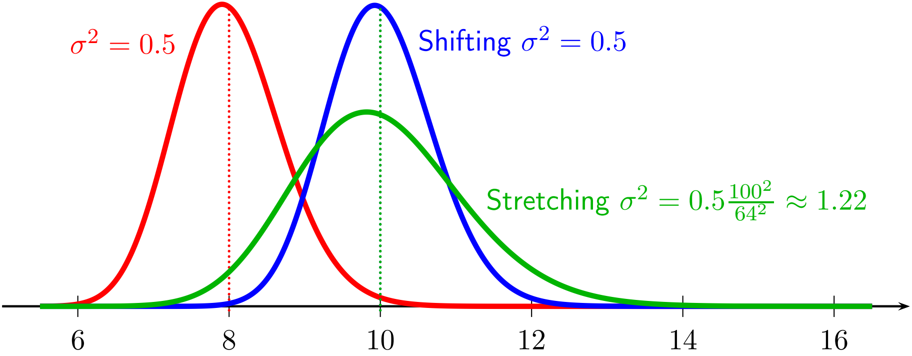
$$\begin{aligned}\text{Expectation} &= \text{Best Estimate reserves (or ultimate)}\\ \text{Variance} &= \frac{\widehat{\text{mse}} \cdot (\text{Best Estimate reserves})^2}{R^2}\end{aligned}$$

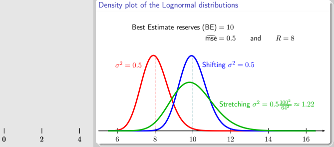
I prefer the stretching, as long as it leads to plausible results, which in particular is not the case if $R \approx 0$.

Density plot of the Lognormal distributions

Best Estimate reserves (BE) $= 10$

$\widehat{\text{mse}} = 0.5$ and $R = 8$





Density plot of the Lognormal distributions

Best Estimate reserves (BE) $= 10$

$\widehat{\text{mse}} = 0.5$ and $R = 8$

Stretching means to keep the coefficient of variation $\frac{\sqrt{\text{Variance}}}{\text{Expectation}}$ constant.

Stochastic Reserving

Lecture 4 (Continuation of Lecture 2)

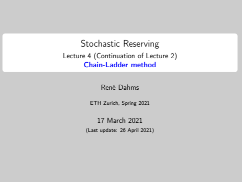
Chain-Ladder method

René Dahms

ETH Zurich, Spring 2021

17 March 2021

(Last update: 26 April 2021)



Stochastic Reserving

Lecture 4 (Continuation of Lecture 2)

Chain-Ladder method

René Dahms

ETH Zurich, Spring 2021

17 March 2021

(Last update: 26 April 2021)

2 Chain-Ladder-Method (CLM)

2021-04-26

Stochastic Reserving

Chain-Ladder-Method (CLM)

Validation and examples (part 2 of 3)

Lecture 4: Table of contents

Claims development result and solvency uncertianty (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 $I+1$:

$$\widehat{\mathsf{CDR}}_i^{I+1} := \widehat{C}_{i,J}^{I} - \widehat{C}_{i,J}^{I+1}.$$

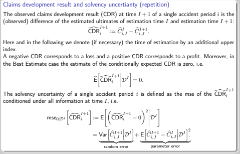
Here and in the following we denote (if necessary) the time of estimation by an additional upper index.

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.

$$\widehat{\mathsf{E}}\left[\widehat{\mathsf{CDR}}_i^{I+1}\middle|\mathcal{D}^I\right] = 0.$$

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

$$\begin{aligned}\mathsf{mse}_{0|\mathcal{D}^I}\left[\widehat{\mathsf{CDR}}_i^{I+1}\right] &:= \mathsf{E}\left[\left(\widehat{\mathsf{CDR}}_i^{I+1} - 0\right)^2\middle|\mathcal{D}^I\right] \\ &= \underbrace{\mathsf{Var}\left[\widehat{C}_{i,J}^{I+1}\middle|\mathcal{D}^I\right]}_{\text{random error}} + \underbrace{\mathsf{E}\left[\widehat{C}_{i,J}^{I+1} - \widehat{C}_{i,J}^{I}\middle|\mathcal{D}^I\right]^2.}_{\text{parameter error}}\end{aligned}$$



Claims development result and solvency uncertianty (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 $I+1$:

$$\widehat{\mathsf{CDR}}_i^{I+1} := \widehat{C}_{i,J}^{I} - \widehat{C}_{i,J}^{I+1}.$$

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.

$$\widehat{\mathsf{E}}\left[\widehat{\mathsf{CDR}}_i^{I+1}\middle|\mathcal{D}^I\right] = 0.$$

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

$$\begin{aligned}\mathsf{mse}_{0|\mathcal{D}^I}\left[\widehat{\mathsf{CDR}}_i^{I+1}\right] &:= \mathsf{E}\left[\left(\widehat{\mathsf{CDR}}_i^{I+1} - 0\right)^2\middle|\mathcal{D}^I\right]\\ &= \underbrace{\mathsf{Var}\left[\widehat{C}_{i,J}^{I+1}\middle|\mathcal{D}^I\right]}_{\text{random error}} + \underbrace{\mathsf{E}\left[\widehat{C}_{i,J}^{I+1} - \widehat{C}_{i,J}^{I}\middle|\mathcal{D}^I\right]^2}_{\text{parameter error}}.\end{aligned}$$

Assumption 2.B (Consistent estimates over time)

In order to have consistent estimates at times I *and* $I+1$ *we assume that there exist* $\mathcal{D}^I \cap \mathcal{D}_k$*-measurable weights* $0 \leq w_{I-k,k}^{I+1} \leq 1$ *with*

- $C_{I-k,k} = 0$ *implies* $w_{I-k,k}^{I+1} = 0$,
- $w_{i,k}^{I+1} := (1 - w_{I-k,k}^{I+1}) w_{i,k}^{I}$, *for* $0 \leq i \leq I-1-k$.

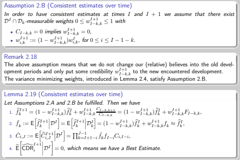
Remark 2.18

The above assumption means that we do not change our (relative) believes into the old development periods and only put some credibility $w_{I-k,k}^{I+1}$ 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. $\widehat{f}_k^{I+1} = (1 - w_{I-k,k}^{I+1}) \widehat{f}_k^{I} + w_{I-k,k}^{I+1} \frac{C_{I-k,k+1}}{C_{I-k,k}} = (1 - w_{I-k,k}^{I+1}) \widehat{f}_k^{I} + w_{I-k,k}^{I+1} F_{I-k,k}$,
2. $\bar{f}_k := \mathsf{E}\left[\widehat{f}_k^{I+1} \middle| \mathcal{D}^I\right] = \mathsf{E}\left[\widehat{f}_k^{I+1} \middle| \mathcal{D}_k^I\right] = (1 - w_{I-k,k}^{I+1}) \widehat{f}_k^{I} + w_{I-k,k}^{I+1} f_k \approx \widehat{f}_k^{I}$,
3. $\bar{C}_{i,J} := \mathsf{E}\left[\widehat{C}_{i,J}^{I+1} \middle| \mathcal{D}^I\right] = \prod_{k=I+1-i}^{J-1} \bar{f}_k f_{I-i} C_{i,I-i}$,
4. $\widehat{\mathsf{E}}\left[\widehat{\mathsf{CDR}}_i^{I+1} \middle| \mathcal{D}^I\right] = 0$, *which means we have a Best Estimate.*



Assumption 2.B (Consistent estimates over time)

In order to have consistent estimates at times I and $I+1$ we assume that there exist $\mathcal{D}^I \cap \mathcal{D}_k$-measurable weights $0 \le w_{I-k,k}^{I+1} \le 1$ with

- $C_{I-k,k} = 0$ *implies* $w_{I-k,k}^{I+1} = 0$,
- $w_{i,k}^{I+1} := (1 - w_{I-k,k}^{I+1}) w_{i,k}^I$, *for* $0 \le i \le I-1-k$.

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 $w_{I-k,k}^{I+1}$ 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. $\widehat{f}_k^{I+1} = (1 - w_{I-k,k}^{I+1})\widehat{f}_k^I + w_{I-k,k}^{I+1} \frac{C_{I-k,k+1}}{C_{I-k,k}} = (1 - w_{I-k,k}^{I+1})\widehat{f}_k^I + w_{I-k,k}^{I+1} F_{I-k,k}$,
2. $\bar{f}_k := \mathsf{E}\left[\widehat{f}_k^{I+1} \middle| \mathcal{D}^I\right] = \mathsf{E}\left[\widehat{f}_k^{I+1} \middle| \mathcal{D}_k^I\right] = (1 - w_{I-k,k}^{I+1})\widehat{f}_k^I + w_{I-k,k}^{I+1} f_k \approx \widehat{f}_k^I$,
3. $\bar{C}_{i,J} := \mathsf{E}\left[\widehat{C}_{i,J}^{I+1} \middle| \mathcal{D}^I\right] = \prod_{k=I+1-i}^{J-1} \bar{f}_k f_{I-i} C_{i,I-i}$,
4. $\widehat{\mathsf{E}}\left[\widehat{\mathsf{CDR}}_i^{I+1} \middle| \mathcal{D}^I\right] = 0$, *which means we have a Best Estimate.*

1. $$\widehat{f}_k^{I+1} := \sum_{i=0}^{I-k} w_{i,k}^{I+1} \frac{C_{i,k+1}}{C_{i,k}} = (1 - w_{I-k,k}^{I+1}) \sum_{i=0}^{I-k-1} w_{i,k}^I \frac{C_{i,k+1}}{C_{i,k}} + w_{I-k,k}^{I+1} \frac{C_{I-k,k+1}}{C_{I-k,k}}$$
$$= (1 - w_{I-k,k}^{I+1})\widehat{f}_k^I + w_{I-k,k}^{I+1} \frac{C_{I-k,k+1}}{C_{I-k,k}} = (1 - w_{I-k,k}^{I+1})\widehat{f}_k^I + w_{I-k,k}^{I+1} F_{I-k,}$$

2. $$\implies \quad \mathsf{E}\left[\widehat{f}_k^{I+1} \middle| \mathcal{D}^I\right] = \mathsf{E}\left[\widehat{f}_k^{I+1} \middle| \mathcal{D}_k^I\right] = (1 - w_{I-k,k}^{I+1})\widehat{f}_k^I + w_{I-k,k}^{I+1} f_k$$
$$\implies \quad \widehat{\bar{f}}_k := \widehat{f}_k^I$$

3. $$\mathsf{E}\left[\widehat{C}_{i,J}^{I+1} \middle| \mathcal{D}^I\right] = \mathsf{E}\left[\prod_{k=I+1-i}^{J-1} \widehat{f}_k^{I+1} C_{i,I+1-i} \middle| \mathcal{D}^I\right] = \mathsf{E}\left[\mathsf{E}\left[\widehat{f}_{J-1}^{I+1} \middle| \mathcal{D}_{J-1}^I\right] \prod_{k=I+1-i}^{J-2} \widehat{f}_k^{I+1} C_{i,I+1-i} \middle| \mathcal{D}^I\right]$$
$$= \bar{f}_{J-1} \mathsf{E}\left[\prod_{k=I+1-i}^{J-2} \widehat{f}_k^{I+1} C_{i,I+1-i} \middle| \mathcal{D}^I\right] = \ldots = \prod_{k=I+1-i}^{J-1} \bar{f}_k \mathsf{E}\left[C_{i,I+1-i} \middle| \mathcal{D}^I\right]$$
$$= \prod_{k=I+1-i}^{J-1} \bar{f}_k \mathsf{E}[C_{i,I+1-i} | \mathcal{B}_{i,I-i}] = \prod_{k=I+1-i}^{J-1} \bar{f}_k f_{I-i} C_{i,I-i}$$

4. $$\mathsf{E}\left[\widehat{\mathsf{CDR}}_i^{I+1} \middle| \mathcal{D}^I\right] = \mathsf{E}\left[\widehat{C}_{i,J}^I \middle| \mathcal{D}^I\right] - \mathsf{E}\left[\widehat{C}_{i,J}^{I+1} \middle| \mathcal{D}^I\right] = \prod_{k=I-i}^{J-1} \widehat{f}_k^I C_{i,I-i} - \prod_{k=I+1-i}^{J-1} \bar{f}_k f_{I-i} C_{i,I-i}$$
$$\approx \prod_{k=I-i}^{J-1} \widehat{f}_k^I C_{i,I-i} - \prod_{k=I+1-i}^{J-1} \widehat{f}_k^I \widehat{f}_{I-i}^I C_{i,I-i} = 0$$

Taylor approximation of next years estimates

Recall the (multi-linear) functional:

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

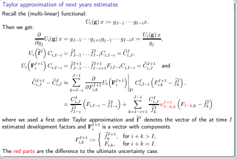
Then we get:

$$\begin{aligned}
\frac{\partial}{\partial g_j} U_i(\mathbf{g})\, x &= g_{J-1} \cdots g_{j+1} g_{j-1} \cdots g_{I-i} x = \frac{U_i(\mathbf{g})\, x}{g_j}, \\
U_i\left(\widehat{\mathbf{f}}^I\right) C_{i,I-i} &= \widehat{f}^I_{J-1} \ldots \widehat{f}^I_{I-i} C_{i,I-i} = \widehat{C}^I_{i,J}, \\
U_i\left(\mathbf{F}_i^{I+1}\right) C_{i,I-i} &= \widehat{f}^{I+1}_{J-1} \ldots \widehat{f}^{I+1}_{I-i+1} F_{i,I-i} C_{i,I-i} = \widehat{C}^{I+1}_{i,J} \quad \text{and} \\
\widehat{C}^{I+1}_{i,J} - \widehat{C}^I_{i,J} &\approx \sum_{k=I-i}^{J-1} \frac{\partial}{\partial F^{I+1}_{i,k}} U_i\left(\mathbf{F}_i^{I+1}\right)\Bigg|_{\widehat{\mathbf{f}}^I} C^I_{i,I-i} \left(F^{I+1}_{i,k} - \widehat{f}^I_k\right), \\
&= \frac{C^I_{i,J}}{\widehat{f}^I_{I-i}} \left(F_{i,I-i} - \widehat{f}^I_{I-i}\right) + \sum_{k=I-i+1}^{J-1} \frac{C^I_{i,J}}{\widehat{f}^I_k} w^{I+1}_{I-k,k} \left(F_{I-k,k} - \widehat{f}^I_k\right)
\end{aligned}$$

where we used a first order Taylor approximation and $\widehat{\mathbf{f}}^I$ denotes the vector of the at time I estimated development factors and $\mathbf{F}_i^{I+1}$ is a vector with components

$$F^{I+1}_{i,k} := \begin{cases} \widehat{f}^{I+1}_k, & \text{for } i+k > I, \\ F_{i,k}, & \text{for } i+k = I. \end{cases}$$

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



Taylor approximation of next years estimates

Recall the (multi-linear) functional:

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

Then we get:

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

$$U_i\left(\widehat{\mathbf{f}}^I\right)C_{i,I-i} = \widehat{f}^I_{J-1}\dots\widehat{f}^I_{I-i}C_{i,I-i} = \widehat{C}^I_{i,J},$$

$$U_i\left(\mathbf{F}^{I+1}_i\right)C_{i,I-i} = \widehat{f}^{I+1}_{J-1}\dots\widehat{f}^{I+1}_{I-i+1}F_{i,I-i}C_{i,I-i} = \widehat{C}^{I+1}_{i,J} \quad \text{and}$$

$$\begin{aligned}\widehat{C}^{I+1}_{i,J} - \widehat{C}^I_{i,J} &\approx \sum_{k=I-i}^{J-1}\frac{\partial}{\partial F^{I+1}_{i,k}}U_i\left(\mathbf{F}^{I+1}_i\right)\Big|_{\widehat{\mathbf{f}}^I}C^I_{i,I-i}\left(F^{I+1}_{i,k} - \widehat{f}^I_k\right),\\ &= \frac{C^I_{i,J}}{\widehat{f}^I_{I-i}}\left(F_{i,I-i} - \widehat{f}^I_{I-i}\right) + \sum_{k=I-i+1}^{J-1}\frac{C^I_{i,J}}{\widehat{f}^I_k}w^{I+1}_{I-k,k}\left(F_{I-k,k} - \widehat{f}^I_k\right)\end{aligned}$$

where we used a first order Taylor approximation and $\widehat{\mathbf{f}}^I$ denotes the vector of the at time I estimated development factors and $\mathbf{F}^{I+1}_i$ is a vector with components

$$F^{I+1}_{i,k} := \begin{cases}\widehat{f}^{I+1}_k, & \text{for } i+k>I,\\ F_{i,k}, & \text{for } i+k=I.\end{cases}$$

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

For $k = I - i$ we get

$$F^{I+1}_{i,I-i} - \widehat{f}^I_{I-i} = F_{i,I-i} - \widehat{f}^I_{I-i}$$

and for $k > I - i$ it is

$$F^{I+1}_{i,k} - \widehat{f}^I_k = \widehat{f}^{I+1}_k - \widehat{f}^I_k = (1 - w^{I+1}_{I-k,k})\widehat{f}^I_k + w^{I+1}_{I-k,k}F_{I-k,k} - \widehat{f}^I_k = w^{I+1}_{I-k,k}\left(F_{I-k,k} - \widehat{f}^I_k\right)$$

Linear approximation of the CDR

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

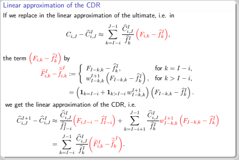
$$C_{i,J} - \widehat{C}_{i,J}^{I} \approx \sum_{k=I-i}^{J-1} \frac{\widehat{C}_{i,J}^{I}}{\widehat{f}_k^{I}} \left(F_{i,k} - \widehat{f}_k^{I} \right),$$

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

$$\widetilde{F}_{i,k}^{I} - \widetilde{\widehat{f}}_{i,k}^{I} := \begin{cases} F_{I-k,k} - \widehat{f}_k^{I}, & \text{for } k = I-i, \\ w_{I-k,k}^{I+1} \left(F_{I-k,k} - \widehat{f}_k^{I} \right), & \text{for } k > I-i, \end{cases}$$
$$= \left(\mathbf{1}_{k=I-i} + \mathbf{1}_{k>I-i}\, w_{I-k,k}^{I+1} \right) \left(F_{I-k,k} - \widehat{f}_k^{I} \right).$$

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

$$\widehat{C}_{i,J}^{I+1} - \widehat{C}_{i,J}^{I} \approx \frac{\widehat{C}_{i,J}^{I}}{\widehat{f}_{I-i}^{I}} \left(F_{i,I-i} - \widehat{f}_{I-i}^{I} \right) + \sum_{k=I-i+1}^{J-1} \frac{\widehat{C}_{i,J}^{I}}{\widehat{f}_k^{I}} w_{I-k,k}^{I+1} \left(F_{I-k,k} - \widehat{f}_k^{I} \right)$$
$$= \sum_{k=I-i}^{J-1} \frac{\widehat{C}_{i,J}^{I}}{\widehat{f}_k^{I}} \left(\widetilde{F}_{i,k}^{I} - \widetilde{\widehat{f}}_k^{I} \right).$$



Linear approximation of the CDR

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

$$C_{i,J} - \widehat{C}^I_{i,J} \approx \sum_{k=I-i}^{J-1} \frac{\widehat{C}^I_{i,J}}{\widehat{f}^I_k} \left(F_{i,k} - \widehat{f}^I_k\right),$$

the term $\left(F_{i,k} - \widehat{f}^I_k\right)$ by

$$\begin{aligned} \widetilde{F}^I_{i,k} - \widetilde{\widehat{f}}^I_{i,k} &:= \begin{cases} F_{I-k,k} - \widehat{f}^I_k, & \text{for } k = I-i, \\ w^{I+1}_{I-k,k} \left(F_{I-k,k} - \widehat{f}^I_k\right), & \text{for } k > I-i, \end{cases} \\ &= \left(\mathbf{1}_{k=I-i} + \mathbf{1}_{k>I-i}\, w^{I+1}_{I-k,k}\right) \left(F_{I-k,k} - \widehat{f}^I_k\right). \end{aligned}$$

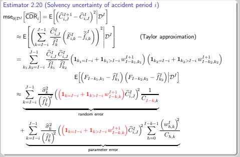
we get the linear approximation of the CDR, i.e.

$$\begin{aligned} \widehat{C}^{I+1}_{i,J} - \widehat{C}^I_{i,J} &\approx \frac{\widehat{C}^I_{i,J}}{\widehat{f}^I_{I-i}} \left(F_{i,I-i} - \widehat{f}^I_{I-i}\right) + \sum_{k=I-i+1}^{J-1} \frac{\widehat{C}^I_{i,J}}{\widehat{f}^I_k} w^{I+1}_{I-k,k} \left(F_{I-k,k} - \widehat{f}^I_k\right) \\ &= \sum_{k=I-i}^{J-1} \frac{\widehat{C}^I_{i,J}}{\widehat{f}^I_k} \left(\widetilde{F}^I_{i,k} - \widetilde{\widehat{f}}^I_k\right). \end{aligned}$$

The term $\widetilde{F}^I_{i,k} - \widetilde{\widehat{f}}^I_{i,k}$ depends on the accident period i only via the indicator functions $\mathbf{1}_{k=I-i}$ and $\mathbf{1}_{k>I-i}$.

Estimator 2.20 (Solvency uncertainty of accident period i)

$$
\begin{aligned}
\mathrm{mse}_{0|\mathcal{D}^I}\left[\widehat{\mathrm{CDR}}_i\right] &= \mathrm{E}\left[\left(\widehat{C}_{i,J}^{I+1} - \widehat{C}_{i,J}^{I}\right)^2 \middle| \mathcal{D}^I\right] \\
&\approx \mathrm{E}\left[\left(\sum_{k=I-i}^{J-1} \frac{\widehat{C}_{i,J}^{I}}{\widehat{f}_k^I}\left(\widetilde{F}_{i,k}^I - \widetilde{\widehat{f}}_{i,k}^{I}\right)\right)^2 \middle| \mathcal{D}^I\right] \qquad \text{(Taylor approximation)} \\
&= \sum_{k_1,k_2=I-i}^{J-1} \frac{\widehat{C}_{i,J}^{I}}{\widehat{f}_{k_1}^I}\frac{\widehat{C}_{i,J}^{I}}{\widehat{f}_{k_2}^I}\left(\mathbf{1}_{k_1=I-i} + \mathbf{1}_{k_1>I-i}\, w_{I-k_1,k_1}^{I+1}\right)\left(\mathbf{1}_{k_2=I-i} + \mathbf{1}_{k_2>I-i}\, w_{I-k_2,k_2}^{I+1}\right) \\
&\qquad \mathrm{E}\left[\left(F_{I-k_1,k_1} - \widehat{f}_{k_1}^I\right)\left(F_{I-k_2,k_2} - \widehat{f}_{k_2}^I\right)\middle|\mathcal{D}^I\right] \\
&\approx \underbrace{\sum_{k=I-i}^{J-1} \frac{\widehat{\sigma}_k^2}{\left(\widehat{f}_k^I\right)^2}\left(\left(\mathbf{1}_{k=I-i} + \mathbf{1}_{k>I-i}\, w_{I-k,k}^{I+1}\right)\widehat{C}_{i,J}^{I}\right)^2 \frac{1}{C_{I-k,k}}}_{\text{random error}} \\
&\quad + \underbrace{\sum_{k=I-i}^{J-1} \frac{\widehat{\sigma}_k^2}{\left(\widehat{f}_k^I\right)^2}\left(\left(\mathbf{1}_{k=I-i} + \mathbf{1}_{k>I-i}\, w_{I-k,k}^{I+1}\right)\widehat{C}_{i,J}^{I}\right)^2 \sum_{h=0}^{I-k-1} \frac{\left(w_{h,k}^I\right)^2}{C_{h,k}}}_{\text{parameter error}}
\end{aligned}
$$



Estimator 2.20 (Solvency uncertainty of accident period i)

$$
\begin{aligned}
\mathsf{mse}_{0|\mathcal{D}^I}\left[\widehat{\mathsf{CDR}}_i\right] &= \mathsf{E}\left[\left(\widehat{C}_{i,J}^{I+1}-\widehat{C}_{i,J}^{I}\right)^2\middle|\mathcal{D}^I\right]\\
&\approx \mathsf{E}\left[\left(\sum_{k=I-i}^{J-1}\frac{\widehat{C}_{i,J}^I}{\widehat{f}_k^I}\left(\widetilde{F}_{i,k}^I-\widetilde{\widehat{f}}_{i,k}^I\right)\right)^2\middle|\mathcal{D}^I\right] \qquad \text{(Taylor approximation)}\\
&= \sum_{k_1,k_2=I-i}^{J-1}\frac{\widehat{C}_{i,J}^I}{\widehat{f}_{k_1}^I}\frac{\widehat{C}_{i,J}^I}{\widehat{f}_{k_2}^I}\left(\mathbf{1}_{k_1=I-i}+\mathbf{1}_{k_1>I-i}w_{I-k_1,k_1}^{I+1}\right)\left(\mathbf{1}_{k_2=I-i}+\mathbf{1}_{k_2>I-i}w_{I-k_2,k_2}^{I+1}\right)\\
&\qquad \mathsf{E}\left[\left(F_{I-k_1,k_1}-\widehat{f}_{k_1}^I\right)\left(F_{I-k_2,k_2}-\widehat{f}_{k_2}^I\right)\middle|\mathcal{D}^I\right]\\
&\approx \underbrace{\sum_{k=I-i}^{J-1}\frac{\widehat{\sigma}_k^2}{\left(\widehat{f}_k^I\right)^2}\left(\left(\mathbf{1}_{k=I-i}+\mathbf{1}_{k>I-i}w_{I-k,k}^{I+1}\right)\widehat{C}_{i,J}^I\right)^2\frac{1}{C_{I-k,k}}}_{\text{random error}}\\
&\quad+\underbrace{\sum_{k=I-i}^{J-1}\frac{\widehat{\sigma}_k^2}{\left(\widehat{f}_k^I\right)^2}\left(\left(\mathbf{1}_{k=I-i}+\mathbf{1}_{k>I-i}w_{I-k,k}^{I+1}\right)\widehat{C}_{i,J}^I\right)^2\sum_{h=0}^{I-k-1}\frac{\left(w_{h,k}^I\right)^2}{C_{h,k}}}_{\text{parameter error}}
\end{aligned}
$$

From the derivation of the ultimate uncertainty we already know

$$
\begin{aligned}
\mathsf{E}\left[\left(F_{I-k_1,k_1}-\widehat{f}_{k_1}^I\right)\left(F_{I-k_2,k_2}^I-\widehat{f}_{k_2}^I\right)\middle|\mathcal{D}^I\right] &= \mathsf{Cov}\left[F_{I-k_1,k_1},F_{I-k_2,k_2}\middle|\mathcal{D}^I\right]+\left(\widehat{f}_{k_1}^I-f_{k_1}\right)\left(\widehat{f}_{k_2}^I-f_{k_2}\right)\\
&\approx \mathsf{Cov}\left[F_{I-k_1,k_1},F_{I-k_2,k_2}\middle|\mathcal{D}^I\right]+\mathsf{Cov}\left[\widehat{f}_{k_1}^I,\widehat{f}_{k_2}^I\middle|\mathcal{D}_{k_1\wedge k_2}\right]\\
&\approx \mathbf{1}_{k_1=k_2}\left(\frac{\widehat{\sigma}_{k_1}^2}{C_{I-k_1,k_1}}+\sum_{h=0}^{I-k_1-1}\frac{\widehat{\sigma}_{k_1}^2\left(w_{h,k}^I\right)^2}{C_{h,k_1}}\right).
\end{aligned}
$$

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

$$
\begin{aligned}
&\mathsf{mse}_{0|\mathcal{D}^I}\left[\widehat{\mathsf{CDR}}_i\right]\\
&\approx \sum_{k_1,k_2=0}^{J-1}\frac{\widehat{C}_{i,J}^I}{\widehat{f}_{k_1}^I}\frac{\widehat{C}_{i,J}^I}{\widehat{f}_{k_2}^I}\left(\mathbf{1}_{k_1=I-i}+\mathbf{1}_{k_1>I-i}w_{I-k_1,k_1}^{I+1}\right)\left(\mathbf{1}_{k_2=I-i}+\mathbf{1}_{k_2>I-i}w_{I-k_2,k_2}^{I+1}\right)\\
&\qquad\qquad \mathbf{1}_{k_1=k_2}\left(\frac{\widehat{\sigma}_{k_1}^2}{C_{I-k_1,k_1}}+\sum_{h=0}^{I-k_1-1}\frac{\widehat{\sigma}_{k_1}^2\left(w_{h,k_1}^I\right)^2}{C_{h,k_1}}\right)\\
&= \sum_{k=I-i}^{J-1}\left(\frac{\widehat{C}_{i,J}^I}{\widehat{f}_k^I}\right)^2\left(\mathbf{1}_{k=I-i}+\mathbf{1}_{k>I-i}w_{I-k,k}^{I+1}\right)^2\left(\frac{\widehat{\sigma}_k^2}{C_{I-k,k}}+\sum_{h=0}^{I-k-1}\frac{\widehat{\sigma}_k^2\left(w_{h,k}^I\right)^2}{C_{h,k}}\right)
\end{aligned}
$$

Ultimate uncertainty for accident period i

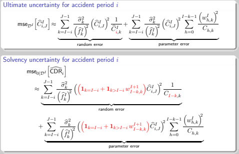
$$\mathsf{mse}_{\mathcal{D}^I}\left[\widehat{C}^I_{i,J}\right] \approx \underbrace{\sum_{k=I-i}^{J-1} \frac{\widehat{\sigma}^2_k}{\left(\widehat{f}^I_k\right)^2} \left(\widehat{C}^I_{i,J}\right)^2 \frac{1}{\widehat{C}^I_{i,k}}}_{\text{random error}} + \underbrace{\sum_{k=I-i}^{J-1} \frac{\widehat{\sigma}^2_k}{\left(\widehat{f}^I_k\right)^2} \left(\widehat{C}^I_{i,J}\right)^2 \sum_{h=0}^{I-k-1} \frac{\left(w^I_{h,k}\right)^2}{C_{h,k}}}_{\text{parameter error}}$$

Solvency uncertainty for accident period i

$$\mathsf{mse}_{0|\mathcal{D}^I}\left[\widehat{\mathsf{CDR}}_i\right]$$

$$\approx \underbrace{\sum_{k=I-i}^{J-1} \frac{\widehat{\sigma}^2_k}{\left(\widehat{f}^I_k\right)^2} \left(\left(\mathbf{1}_{k=I-i} + \mathbf{1}_{k>I-i}\, w^{I+1}_{I-k,k}\right)\widehat{C}^I_{i,J}\right)^2 \frac{1}{C_{I-k,k}}}_{\text{random error}}$$

$$+ \underbrace{\sum_{k=I-i}^{J-1} \frac{\widehat{\sigma}^2_k}{\left(\widehat{f}^I_k\right)^2} \left(\left(\mathbf{1}_{k=I-i} + \mathbf{1}_{k>I-i}\, w^{I+1}_{I-k,k}\right)\widehat{C}^I_{i,J}\right)^2 \sum_{h=0}^{I-k-1} \frac{\left(w^I_{h,k}\right)^2}{C_{h,k}}}_{\text{parameter error}}$$



Ultimate uncertainty for accident period i

$$\mathsf{mse}_{\mathcal{D}^I}\left[\widehat{C}^I_{i,J}\right] \approx \underbrace{\sum_{k=I-i}^{J-1} \frac{\widehat{\sigma}^2_k}{\left(\widehat{f}^I_k\right)^2} \left(\widehat{C}^I_{i,J}\right)^2 \frac{1}{\widehat{C}^I_{i,k}}}_{\text{random error}} + \underbrace{\sum_{k=I-i}^{J-1} \frac{\widehat{\sigma}^2_k}{\left(\widehat{f}^I_k\right)^2} \left(\widehat{C}^I_{i,J}\right)^2 \sum_{h=0}^{I-k-1} \frac{\left(w^I_{h,k}\right)^2}{C_{h,k}}}_{\text{parameter error}}$$

Solvency uncertainty for accident period i

$$\begin{aligned} &\mathsf{mse}_{0|\mathcal{D}^I}\left[\widehat{\mathsf{CDR}}_i\right] \\ &\approx \underbrace{\sum_{k=I-i}^{J-1} \frac{\widehat{\sigma}^2_k}{\left(\widehat{f}^I_k\right)^2} \left(\left(\mathbf{1}_{k=I-i} + \mathbf{1}_{k>I-i} w^{I+1}_{I-k,k}\right)\widehat{C}^I_{i,J}\right)^2 \frac{1}{C_{I-k,k}}}_{\text{random error}} \\ &+ \underbrace{\sum_{k=I-i}^{J-1} \frac{\widehat{\sigma}^2_k}{\left(\widehat{f}^I_k\right)^2} \left(\left(\mathbf{1}_{k=I-i} + \mathbf{1}_{k>I-i} w^{I+1}_{I-k,k}\right)\widehat{C}^I_{i,J}\right)^2 \sum_{h=0}^{I-k-1} \frac{\left(w^I_{h,k}\right)^2}{C_{h,k}}}_{\text{parameter error}} \end{aligned}$$

It almost looks like a simple multiplication by the factor

$$\left(\mathbf{1}_{k=I-i} + \mathbf{1}_{k>I-i} w^{I+1}_{I-k,k}\right),$$

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

Corollary 2.21

If we use the variance minimizing weights

$$w_{i,k}^{I} = \frac{C_{i,k}}{\sum\limits_{h=0}^{I-k-1} C_{h,k}} \qquad \text{and} \qquad w_{i,k}^{I+1} = \frac{C_{i,k}}{\sum\limits_{h=0}^{I-k} C_{h,k}}$$

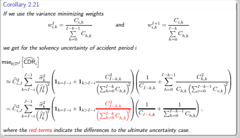
we get for the solvency uncertainty of accident period i

$$\mathrm{mse}_{0|\mathcal{D}^I}\left[\widehat{\mathrm{CDR}}_i\right]$$

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

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

where the red terms indicate the differences to the ultimate uncertainty case.



Corollary 2.21

If we use the variance minimizing weights

$$w_{i,k}^{I} = \frac{C_{i,k}}{\sum\limits_{h=0}^{I-k-1} C_{h,k}} \qquad \text{and} \qquad w_{i,k}^{I+1} = \frac{C_{i,k}}{\sum\limits_{h=0}^{I-k} C_{h,k}}$$

we get for the solvency uncertainty of accident period i

$$\begin{aligned}
&\mathsf{mse}_{0|\mathcal{D}^I}\left[\widehat{\mathsf{CDR}}_i\right] \\
&\approx \widehat{C}_{i,J}^2 \sum_{k=I-i}^{J-1} \frac{\widehat{\sigma}_k^2}{\left(\widehat{f}_k^I\right)^2} \left(\mathbf{1}_{k=I-i} + \mathbf{1}_{k>I-i} \frac{C_{I-k,k}^2}{\left(\sum_{h=0}^{I-k} C_{h,k}\right)^2}\right) \left(\frac{1}{C_{I-k,k}} + \sum_{h=0}^{I-k-1} \frac{C_{h,k}^2}{C_{h,k}\left(\sum_{v=0}^{I-k-1} C_{v,k}\right)^2}\right) \\
&= \widehat{C}_{i,J}^2 \sum_{k=I-i}^{J-1} \frac{\widehat{\sigma}_k^2}{\left(\widehat{f}_k^I\right)^2} \left(\mathbf{1}_{k=I-i} + \mathbf{1}_{k>I-i} \frac{C_{I-k,k}^2}{\left(\sum_{h=0}^{I-k} C_{h,k}\right)^2}\right) \left(\frac{1}{C_{I-k,k}} + \frac{1}{\sum_{h=0}^{I-k-1} C_{h,k}}\right),
\end{aligned}$$

where the red terms indicate the differences to the ultimate uncertainty case.

Dependent accident periods

Since $\widetilde{F}_{i,k}^{I}$ and $\widehat{\widetilde{f}}_k^{I}$ depend on $F_{I-k,k} = {}^{C_{I-k,k+1}}/_{C_{I-k,k}}$, for all i, the $\widehat{\mathrm{CDR}}_i^{I+1}$, $i \leq I$, 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:

Dependent accident periods

Since $\widetilde{F}^{I}_{i,k}$ and $\widetilde{\widehat{f}}_{k}^{I}$ depend on $F_{I-k,k} = {}^{C_{I-k,k+1}}/_{C_{I-k,k}}$, for all i, the $\widehat{\mathrm{CDR}}_{i}^{I+1}$, $i \leq I$, 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)

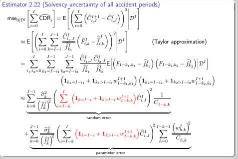
$$\mathsf{mse}_{0|\mathcal{D}^I}\left[\sum_{i=0}^{I}\widehat{\mathsf{CDR}}_i\right] = \mathsf{E}\left[\left(\sum_{i=0}^{I}\left(\widehat{C}_{i,J}^{I+1}-\widehat{C}_{i,J}^{I}\right)\right)^2\middle|\mathcal{D}^I\right]$$

$$\approx \mathsf{E}\left[\left(\sum_{i=0}^{I}\sum_{k=I-i}^{J-1}\frac{\widehat{C}_{i_1,J}^{I}}{\widehat{f}_{k_1}^{I}}\left(\widetilde{F}_{i,k}^{I}-\widetilde{\widehat{f}}_{i,k}^{I}\right)\right)^2\middle|\mathcal{D}^I\right] \qquad \text{(Taylor approximation)}$$

$$= \sum_{i_1,i_2=0}^{I}\sum_{k_1=I-i_1}^{J-1}\sum_{k_2=I-i_2}^{J-1}\frac{\widehat{C}_{i_1,J}^{I}}{\widehat{f}_{k_1}^{I}}\frac{\widehat{C}_{i_2,J}^{I}}{\widehat{f}_{k_2}^{I}}\mathsf{E}\left[\left(F_{I-k_1,k_1}-\widehat{f}_{k_1}^{I}\right)\left(F_{I-k_2,k_2}-\widehat{f}_{k_2}^{I}\right)\middle|\mathcal{D}^I\right]$$
$$\left(\mathbf{1}_{k_1=I-i_1}+\mathbf{1}_{k_1>I-i_1}w_{I-k_1,k_1}^{I+1}\right)\left(\mathbf{1}_{k_2=I-i_2}+\mathbf{1}_{k_2>I-i_2}w_{I-k_2,k_2}^{I+1}\right)$$

$$\approx \underbrace{\sum_{k=0}^{J-1}\frac{\widehat{\sigma}_k^2}{\left(\widehat{f}_k^I\right)^2}\left(\sum_{i=I-k}^{I}\left(\mathbf{1}_{k=I-i}+\mathbf{1}_{k>I-i}\,w_{I-k,k}^{I+1}\right)\widehat{C}_{i,J}^{I}\right)^2\frac{1}{C_{I-k,k}}}_{\text{random error}}$$

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



Estimator 2.22 (Solvency uncertainty of all accident periods)

$$\begin{aligned}
\mathsf{mse}_{0|\mathcal{D}^I}\left[\sum_{i=0}^{I}\widehat{\mathsf{CDR}}_i\right] &= \mathsf{E}\left[\left(\sum_{i=0}^{I}\left(\widehat{C}_{i,J}^{I+1}-\widehat{C}_{i,J}^{I}\right)\right)^2\middle|\mathcal{D}^I\right]\\
&\approx \mathsf{E}\left[\left(\sum_{i=0}^{I}\sum_{k=I-i}^{J-1}\frac{\widehat{C}_{i_1,J}^{I}}{\widehat{f}_{k_1}^{I}}\left(\widetilde{F}_{i,k}^{I}-\widetilde{\widehat{f}}_{i,k}^{I}\right)\right)^2\middle|\mathcal{D}^I\right] \qquad \text{(Taylor approximation)}\\
&= \sum_{i_1,i_2=0}^{I}\sum_{k_1=I-i_1}^{J-1}\sum_{k_2=I-i_2}^{J-1}\frac{\widehat{C}_{i_1,J}^{I}}{\widehat{f}_{k_1}^{I}}\frac{\widehat{C}_{i_2,J}^{I}}{\widehat{f}_{k_2}^{I}}\mathsf{E}\left[\left(F_{I-k_1,k_1}-\widehat{f}_{k_1}^{I}\right)\left(F_{I-k_2,k_2}-\widehat{f}_{k_2}^{I}\right)\middle|\mathcal{D}^I\right]\\
&\qquad \left(\mathbf{1}_{k_1=I-i_1}+\mathbf{1}_{k_1>I-i_1}w_{I-k_1,k_1}^{I+1}\right)\left(\mathbf{1}_{k_2=I-i_2}+\mathbf{1}_{k_2>I-i_2}w_{I-k_2,k_2}^{I+1}\right)\\
&\approx \underbrace{\sum_{k=0}^{J-1}\frac{\widehat{\sigma}_k^2}{\left(\widehat{f}_k^I\right)^2}\left(\sum_{i=I-k}^{I}\left(\mathbf{1}_{k=I-i}+\mathbf{1}_{k>I-i}\,w_{I-k,k}^{I+1}\right)\widehat{C}_{i,J}^{I}\right)^2\frac{1}{C_{I-k,k}}}_{\text{random error}}\\
&\quad +\underbrace{\sum_{k=0}^{J-1}\frac{\widehat{\sigma}_k^2}{\left(\widehat{f}_k^I\right)^2}\left(\sum_{i=I-k}^{I}\left(\mathbf{1}_{k=I-i}+\mathbf{1}_{k>I-i}\,w_{I-k,k}^{I+1}\right)\widehat{C}_{i,J}^{I}\right)^2\sum_{h=0}^{I-k-1}\frac{\left(w_{h,k}^I\right)^2}{C_{h,k}}}_{\text{parameter error}}
\end{aligned}$$

From the single accident period case we know

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

Therefore, we get

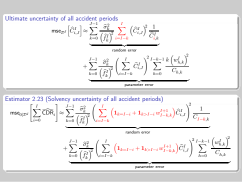
$$\begin{aligned}
&\mathsf{mse}_{0|\mathcal{D}^I}\left[\sum_{i=0}^{I}\widehat{\mathsf{CDR}}_i\right]\\
&\quad\approx \sum_{i_1,i_2=0}^{I}\sum_{k_1=I-i_1}^{J-1}\sum_{k_2=I-i_2}^{J-1}\frac{\widehat{C}_{i_1,J}^{I}}{\widehat{f}_{k_1}^{I}}\frac{\widehat{C}_{i_2,J}^{I}}{\widehat{f}_{k_2}^{I}}\mathbf{1}_{k_1=k_2}\left(\frac{\widehat{\sigma}_{k_1}^2}{C_{I-k_1,k_1}}+\sum_{h=0}^{I-k_1-1}\frac{\widehat{\sigma}_{k_1}^2\left(w_{h,k_1}^I\right)^2}{C_{h,k_1}}\right)\\
&\qquad\qquad \left(\mathbf{1}_{k_1=I-i_1}+\mathbf{1}_{k_1>I-i_1}w_{I-k_1,k_1}^{I+1}\right)\left(\mathbf{1}_{k_2=I-i_2}+\mathbf{1}_{k_2>I-i_2}w_{I-k_2,k_2}^{I+1}\right)\\
&\quad= \sum_{k_1,k_2=0}^{J-1}\sum_{i_1=I-k_1}^{I}\sum_{i_2=I-k_2}^{I}\frac{\widehat{C}_{i_1,J}^{I}}{\widehat{f}_{k_1}^{I}}\frac{\widehat{C}_{i_2,J}^{I}}{\widehat{f}_{k_2}^{I}}\mathbf{1}_{k_1=k_2}\left(\frac{\widehat{\sigma}_{k_1}^2}{C_{I-k_1,k_1}}+\sum_{h=0}^{I-k_1-1}\frac{\widehat{\sigma}_{k_1}^2\left(w_{h,k_1}^I\right)^2}{C_{h,k_1}}\right)\\
&\qquad\qquad \left(\mathbf{1}_{k_1=I-i_1}+\mathbf{1}_{k_1>I-i_1}w_{I-k_1,k_1}^{I+1}\right)\left(\mathbf{1}_{k_2=I-i_2}+\mathbf{1}_{k_2>I-i_2}w_{I-k_2,k_2}^{I+1}\right)\\
&\quad= \sum_{k=0}^{J-1}\frac{\widehat{\sigma}_k^2}{\left(\widehat{f}_k^I\right)^2}\left(\sum_{i=I-k}^{I}\left(\mathbf{1}_{k=I-i}+\mathbf{1}_{k>I-i}w_{I-k,k}^{I+1}\right)\widehat{C}_{i,J}^{I}\right)^2\left(\frac{1}{C_{I-k,k}}+\sum_{h=0}^{I-k-1}\frac{\left(w_{h,k}^I\right)^2}{C_{h,k}}\right)
\end{aligned}$$

Ultimate uncertainty of all accident periods

$$\mathrm{mse}_{\mathcal{D}^I}\left[\widehat{C}_{i,J}^I\right] \approx \underbrace{\sum_{k=0}^{J-1} \frac{\widehat{\sigma}_k^2}{\left(\widehat{f}_k^I\right)^2} \sum_{i=I-k}^{I} \left(\widehat{C}_{i,J}^I\right)^2 \frac{1}{\widehat{C}_{i,k}^I}}_{\text{random error}}$$
$$+ \underbrace{\sum_{k=0}^{J-1} \frac{\widehat{\sigma}_k^2}{\left(\widehat{f}_k^I\right)^2} \left(\sum_{i=I-k}^{I} \widehat{C}_{i,J}^I\right)^2 \sum_{h=0}^{I-k-1} \frac{k\left(w_{h,k}^I\right)^2}{C_{h,k}}}_{\text{parameter error}}$$

Estimator 2.23 (Solvency uncertainty of all accident periods)

$$\mathrm{mse}_{0|\mathcal{D}^I}\left[\sum_{i=0}^{I} \widehat{\mathrm{CDR}}_i\right] \approx \underbrace{\sum_{k=0}^{J-1} \frac{\widehat{\sigma}_k^2}{\left(\widehat{f}_k^I\right)^2} \left(\sum_{i=I-k}^{I} \left(\mathbf{1}_{k=I-i} + \mathbf{1}_{k>I-i}\, w_{I-k,k}^{I+1}\right)\widehat{C}_{i,J}^I\right)^2 \frac{1}{C_{I-k,k}}}_{\text{random error}}$$
$$+ \underbrace{\sum_{k=0}^{J-1} \frac{\widehat{\sigma}_k^2}{\left(\widehat{f}_k^I\right)^2} \left(\sum_{i=I-k}^{I} \left(\mathbf{1}_{k=I-i} + \mathbf{1}_{k>I-i}\, w_{I-k,k}^{I+1}\right)\widehat{C}_{i,J}^I\right)^2 \sum_{h=0}^{I-k-1} \frac{\left(w_{h,k}^I\right)^2}{C_{h,k}}}_{\text{parameter error}}$$



Ultimate uncertainty of all accident periods

$$\mathrm{mse}_{\mathcal{D}^I}\left[\widehat{C}^I_{i,J}\right] \approx \underbrace{\sum_{k=0}^{J-1} \frac{\widehat{\sigma}^2_k}{\left(\widehat{f}^I_k\right)^2} \sum_{i=I-k}^{I} \left(\widehat{C}^I_{i,J}\right)^2 \frac{1}{\widehat{C}^I_{i,k}}}_{\text{random error}}$$

$$+ \underbrace{\sum_{k=0}^{J-1} \frac{\widehat{\sigma}^2_k}{\left(\widehat{f}^I_k\right)^2} \left(\sum_{i=I-k}^{I} \widehat{C}^I_{i,J}\right)^2 \sum_{h=0}^{I-k-1} \frac{k\left(w^I_{h,k}\right)^2}{C_{h,k}}}_{\text{parameter error}}$$

Estimator 2.23 (Solvency uncertainty of all accident periods)

$$\mathrm{mse}_{0|\mathcal{D}^I}\left[\sum_{i=0}^{I} \widehat{\mathrm{CDR}}_i\right] \approx \underbrace{\sum_{k=0}^{J-1} \frac{\widehat{\sigma}^2_k}{\left(\widehat{f}^I_k\right)^2} \left(\sum_{i=I-k}^{I} \left(\mathbf{1}_{k=I-i} + \mathbf{1}_{k>I-i}\, w^{I+1}_{I-k,k}\right) \widehat{C}^I_{i,J}\right)^2 \frac{1}{C_{I-k,k}}}_{\text{random error}}$$

$$+ \underbrace{\sum_{k=0}^{J-1} \frac{\widehat{\sigma}^2_k}{\left(\widehat{f}^I_k\right)^2} \left(\sum_{i=I-k}^{I} \left(\mathbf{1}_{k=I-i} + \mathbf{1}_{k>I-i}\, w^{I+1}_{I-k,k}\right) \widehat{C}^I_{i,J}\right)^2 \sum_{h=0}^{I-k-1} \frac{\left(w^I_{h,k}\right)^2}{C_{h,k}}}_{\text{parameter error}}$$

Corollary 2.24

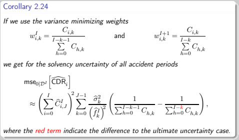
If we use the variance minimizing weights

$$w_{i,k}^{I} = \frac{C_{i,k}}{\sum\limits_{h=0}^{I-k-1} C_{h,k}} \qquad \text{and} \qquad w_{i,k}^{I+1} = \frac{C_{i,k}}{\sum\limits_{h=0}^{I-k} C_{h,k}}$$

we get for the solvency uncertainty of all accident periods

$$\begin{aligned} &\mathrm{mse}_{0|\mathcal{D}^I}\left[\widehat{\mathrm{CDR}}_i\right] \\ &\approx \left(\sum_{i=0}^{I} \widehat{C}_{i,J}^{I}\right)^2 \sum_{k=0}^{J-1} \frac{\widehat{\sigma}_k^2}{\left(\widehat{f}_k^I\right)^2} \left(\frac{1}{\sum_{h=0}^{I-k-1} C_{h,k}} - \frac{1}{\sum_{h=0}^{I{\color{red}-k}} C_{h,k}}\right), \end{aligned}$$

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



Corollary 2.24

If we use the variance minimizing weights

$$w_{i,k}^{I} = \frac{C_{i,k}}{\sum\limits_{h=0}^{I-k-1} C_{h,k}} \qquad \text{and} \qquad w_{i,k}^{I+1} = \frac{C_{i,k}}{\sum\limits_{h=0}^{I-k} C_{h,k}}$$

we get for the solvency uncertainty of all accident periods

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

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

$$\begin{aligned}
\sum_{i=I-k}^{I} \left(\mathbf{1}_{k=I-i} + \mathbf{1}_{k>I-i} w_{I-k,k}^{I+1}\right) \widehat{C}_{i,J}^{I} &= \widehat{C}_{I-k,J}^{I} + \sum_{i=I-k+1}^{I} \frac{C_{I-k,k}}{\sum_{h=0}^{I-k} C_{h,k}} \widehat{C}_{i,J}^{I} \\
&= \widehat{C}_{I-k,J}^{I} + \sum_{i=I-k+1}^{I} \frac{\widehat{C}_{I-k,J}^{I}}{\sum_{h=0}^{I-k} \widehat{C}_{h,J}^{I}} \widehat{C}_{i,J}^{I} \\
&= \widehat{C}_{I-k,J}^{I} \left(1 + \frac{\sum_{i=I-k+1}^{I} \widehat{C}_{i,J}^{I}}{\sum_{h=0}^{I-k} \widehat{C}_{h,J}^{I}}\right) = \sum_{i=0}^{I} \widehat{C}_{i,J}^{I} \frac{C_{I-k,k}}{\sum_{h=0}^{I-k} \widehat{C}_{h,J}^{I}}
\end{aligned}$$

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

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

Estimation at time $n \geq I$

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

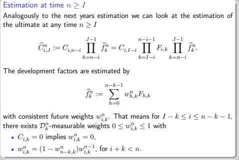
$$\widehat{C}^n_{i,J} := C_{i,n-i} \prod_{k=n-i}^{J-1} \widehat{f}^n_k = C_{i,I-i} \prod_{k=I-i}^{n-i-1} F_{i,k} \prod_{k=n-i}^{J-1} \widehat{f}^n_k.$$

The development factors are estimated by

$$\widehat{f}^n_k := \sum_{h=0}^{n-k-1} w^n_{h,k} F_{h,k}$$

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

- $C_{i,k} = 0$ implies $w^n_{i,k} = 0$,
- $w^n_{i,k} = (1 - w^n_{n-k,k}) w^{n-1}_{i,k}$, 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 $n \geq I$

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

The development factors are estimated by

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

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 $i + k < n$.

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

Since formulas will get very tedious (see for instance [12]), 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

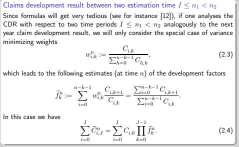
$$w_{i,k}^{n} := \frac{C_{i,k}}{\sum_{h=0}^{n-k-1} C_{h,k}}, \tag{2.3}$$

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

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

In this case we have

$$\sum_{i=0}^{I} \widehat{C}_{i,J}^n = \sum_{i=0}^{I} C_{i,0} \prod_{k=0}^{J-1} \widehat{f}_k^n. \tag{2.4}$$



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

Since formulas will get very tedious (see for instance [12]), 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

$$w_{i,k}^n := \frac{C_{i,k}}{\sum_{h=0}^{n-k-1} C_{h,k}}, \tag{2.3}$$

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

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

In this case we have

$$\sum_{i=0}^{I} \widehat{C}_{i,J}^n = \sum_{i=0}^{I} C_{i,0} \prod_{k=0}^{J-1} \widehat{f}_k^n. \tag{2.4}$$

For each $k > 0$ we have

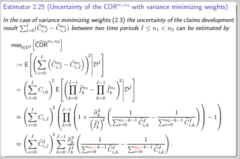
$$\begin{aligned}
\sum_{i=0}^{I} \widehat{C}_{i,k+1}^n &= \sum_{i=0}^{n-k-1} C_{i,k+1} + \sum_{i=n-k}^{I} \widehat{f}_k^n \widehat{C}_{i,k}^n \\
&= \frac{\sum_{i=0}^{n-k-1} C_{i,k+1}}{\sum_{i=0}^{n-k-1} C_{i,k}} \sum_{i=0}^{n-k-1} C_{i,k} + \widehat{f}_k^n \sum_{i=n-k}^{I} \widehat{C}_{i,k}^n \\
&= \widehat{f}_k^n \sum_{i=0}^{n-k-1} C_{i,k} + \widehat{f}_k^n \sum_{i=n-k}^{I} \widehat{C}_{i,k}^n \\
&= \widehat{f}_k^n \sum_{i=0}^{I} \widehat{C}_{i,k}^n,
\end{aligned}$$

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* $\sum_{i=0}^{I}(\widehat{C}_{i,J}^{n_2}-\widehat{C}_{i,J}^{n_1})$ *between two time periods* $I \leq n_1 < n_2$ *can be estimated by*

$$\begin{aligned}
&\mathrm{mse}_{0|\mathcal{D}^I}\left[\widehat{\mathrm{CDR}}^{n_1,n_2}\right] \\
&:= \mathsf{E}\left[\left(\sum_{i=0}^{I}\left(\widehat{C}_{i,J}^{n_2}-\widehat{C}_{i,J}^{n_1}\right)\right)^2\middle|\mathcal{D}^I\right] \\
&= \left(\sum_{i=0}^{I} C_{i,0}\right)^2 \mathsf{E}\left[\left(\prod_{k=0}^{J-1}\widehat{f}_k^{n_2}-\prod_{k=0}^{J-1}\widehat{f}_k^{n_1}\right)^2\middle|\mathcal{D}^I\right] \\
&\approx \left(\sum_{i=0}^{I} C_{i,J}\right)^2 \left(\prod_{k=0}^{J-1}\left(1+\frac{\widehat{\sigma}_k^2}{\left(\widehat{f}_k^I\right)^2}\left(\frac{1}{\sum_{i=0}^{n_1-k-1}\widehat{C}_{i,k}^I}-\frac{1}{\sum_{i=0}^{n_2-k-1}\widehat{C}_{i,k}^I}\right)\right)-1\right) \\
&\approx \left(\sum_{i=0}^{I}\widehat{C}_{i,J}^I\right)^2 \sum_{k=0}^{J-1}\frac{\widehat{\sigma}_k^2}{\widehat{f}_k^I}\left(\frac{1}{\sum_{i=0}^{n_1-k-1}\widehat{C}_{i,k}^I}-\frac{1}{\sum_{i=0}^{n_2-k-1}\widehat{C}_{i,k}^I}\right).
\end{aligned}$$



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 $\sum_{i=0}^{I}(\widehat{C}_{i,J}^{n_2} - \widehat{C}_{i,J}^{n_1})$ between two time periods $I \leq n_1 < n_2$ can be estimated by

$$\begin{aligned}
&\mathrm{mse}_{0|\mathcal{D}^I}\left[\widehat{\mathrm{CDR}}^{n_1,n_2}\right] \\
&:= \mathsf{E}\left[\left(\sum_{i=0}^{I}\left(\widehat{C}_{i,J}^{n_2} - \widehat{C}_{i,J}^{n_1}\right)\right)^2\middle|\mathcal{D}^I\right] \\
&= \left(\sum_{i=0}^{I} C_{i,0}\right)^2 \mathsf{E}\left[\left(\prod_{k=0}^{J-1}\widehat{f}_k^{n_2} - \prod_{k=0}^{J-1}\widehat{f}_k^{n_1}\right)^2\middle|\mathcal{D}^I\right] \\
&\approx \left(\sum_{i=0}^{I} C_{i,J}\right)^2 \left(\prod_{k=0}^{J-1}\left(1 + \frac{\widehat{\sigma}_k^2}{\left(\widehat{f}_k^I\right)^2}\left(\frac{1}{\sum_{i=0}^{n_1-k-1}\widehat{C}_{i,k}^I} - \frac{1}{\sum_{i=0}^{n_2-k-1}\widehat{C}_{i,k}^I}\right)\right) - 1\right) \\
&\approx \left(\sum_{i=0}^{I} \widehat{C}_{i,J}^I\right)^2 \sum_{k=0}^{J-1}\frac{\widehat{\sigma}_k^2}{\widehat{f}_k^I}\left(\frac{1}{\sum_{i=0}^{n_1-k-1}\widehat{C}_{i,k}^I} - \frac{1}{\sum_{i=0}^{n_2-k-1}\widehat{C}_{i,k}^I}\right).
\end{aligned}$$

We will start with some very brave approximations

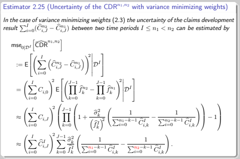
$$\begin{aligned}
\mathrm{mse}_{0|\mathcal{D}^I}\left[\widehat{\mathrm{CDR}}^{n_1,n_2}\right] &= \left(\sum_{i=0}^{I} C_{i,0}\right)^2 \mathsf{E}\left[\left(\prod_{k=0}^{J-1}\widehat{f}_k^{n_2} - \prod_{k=0}^{J-1}\widehat{f}_k^{n_1}\right)^2\middle|\mathcal{D}^I\right] \\
&\approx \left(\sum_{i=0}^{I} C_{i,0}\right)^2 \left(\prod_{k=0}^{J-1}\mathsf{E}\left[\left(\widehat{f}_k^{n_2}\right)^2\middle|\mathcal{D}^{n_1}\right] - \prod_{k=0}^{J-1}\mathsf{E}\left[\left(\widehat{f}_k^{n_1}\right)^2\middle|\mathcal{D}^{n_1}\right]\right) \\
&\approx \left(\sum_{i=0}^{I} C_{i,0}\right)^2 \left(\prod_{k=0}^{J-1}\mathsf{E}\left[\left(\widehat{f}_k^{n_2}\right)^2\middle|\mathcal{D}^{n_1}\right] - \prod_{k=0}^{J-1}\left(\widehat{f}_k^{I}\right)^2\right)
\end{aligned}$$

In order to estimate the remaining expectations of the square of $\widehat{f}_k^{n_2}$, we will look at the corresponding variance and expectation of $\widehat{f}_k^{n_2}$ and always replace all future weights $w_{i,k}^{n_2}$ by their estimates at time I, i.e. by

$$\widehat{w}_{i,k}^{n_2} := \frac{\widehat{C}_{i,k}^I}{\sum_{h=0}^{n_2-k-1}\widehat{C}_{h,k}^I}.$$

We get

$$\begin{aligned}
\mathsf{E}\left[\widehat{f}_k^{n_2}\middle|\mathcal{D}^{n_1}\right] &= \mathsf{E}\left[\sum_{i=0}^{n_2-k-1} w_{i,k}^{n_2}\frac{C_{i,k+1}}{C_{i,k}}\middle|\mathcal{D}^{n_1}\right] \approx \mathsf{E}\left[\sum_{i=0}^{n_2-k-1} \widehat{w}_{i,k}^{n_2}\frac{C_{i,k+1}}{C_{i,k}}\middle|\mathcal{D}^{n_1}\right] = \sum_{i=0}^{n_1-k-1}\widehat{w}_{i,k}^{n_2}\frac{C_{i,k+1}}{C_{i,k}} + \sum_{i=n_1-k}^{n_2-k-1}\widehat{w}_{i,k}^{n_2} f_k \\
&= \sum_{i=0}^{n_1-k-1}\frac{\widehat{C}_{i,k}^I}{\sum_{h=0}^{n_2-k-1}\widehat{C}_{h,k}^I}\frac{C_{i,k+1}}{C_{i,k}} + \frac{\sum_{h=n_1-k}^{n_2-k-1}\widehat{C}_{h,k}^I}{\sum_{h=0}^{n_2-k-1}\widehat{C}_{h,k}^I} f_k \approx \frac{\sum_{h=0}^{n_1-k-1}\widehat{C}_{h,k}^I}{\sum_{h=0}^{n_2-k-1}\widehat{C}_{h,k}^I}\widehat{f}_k^{n_1} + \frac{\sum_{h=n_1-k}^{n_2-k-1}\widehat{C}_{h,k}^I}{\sum_{h=0}^{n_2-k-1}\widehat{C}_{h,k}^I} f_k \\
&\approx \widehat{f}_k^{n_1} + \frac{\sum_{h=n_1-k}^{n_2-k-1}\widehat{C}_{h,k}^I}{\sum_{h=0}^{n_2-k-1}\widehat{C}_{h,k}^I}\left(f_k - \widehat{f}_k^{n_1}\right) =: \widehat{f}_k^{n_1} + \Omega_k^{n_1,n_2}\left(f_k - \widehat{f}_k^{n_1}\right)
\end{aligned}$$



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* $\sum_{i=0}^{I}(\widehat C_{i,J}^{n_2} - \widehat C_{i,J}^{n_1})$ *between two time periods* $I \le n_1 < n_2$ *can be estimated by*

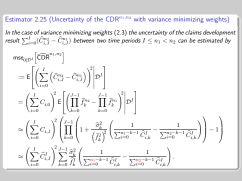
$$\begin{aligned}
&\mathsf{mse}_{0|\mathcal{D}^I}\left[\widehat{\mathsf{CDR}}^{n_1,n_2}\right] \\
&:= \mathsf{E}\left[\left(\sum_{i=0}^{I}\left(\widehat C_{i,J}^{n_2} - \widehat C_{i,J}^{n_1}\right)\right)^2\middle|\mathcal{D}^I\right] \\
&= \left(\sum_{i=0}^{I} C_{i,0}\right)^2 \mathsf{E}\left[\left(\prod_{k=0}^{J-1}\widehat f_k^{n_2} - \prod_{k=0}^{J-1}\widehat f_k^{n_1}\right)^2\middle|\mathcal{D}^I\right] \\
&\approx \left(\sum_{i=0}^{I} C_{i,J}\right)^2 \left(\prod_{k=0}^{J-1}\left(1 + \frac{\widehat\sigma_k^2}{\left(\widehat f_k^I\right)^2}\left(\frac{1}{\sum_{i=0}^{n_1-k-1}\widehat C_{i,k}^I} - \frac{1}{\sum_{i=0}^{n_2-k-1}\widehat C_{i,k}^I}\right)\right) - 1\right) \\
&\approx \left(\sum_{i=0}^{I} \widehat C_{i,J}^I\right)^2 \sum_{k=0}^{J-1}\frac{\widehat\sigma_k^2}{\widehat f_k^I}\left(\frac{1}{\sum_{i=0}^{n_1-k-1}\widehat C_{i,k}^I} - \frac{1}{\sum_{i=0}^{n_2-k-1}\widehat C_{i,k}^I}\right).
\end{aligned}$$

This leads to

$$\begin{aligned}
\left(\mathsf{E}\left[\widehat f_k^{n_2}\middle|\mathcal{D}^{n_1}\right]\right)^2 &= \left(\widehat f_k^{n_1}\right)^2 + 2\Omega_k^{n_1,n_2}\widehat f_k^{n_1}(f_k - \widehat f_k^{n_1}) + \left(\Omega_k^{n_1,n_2}\right)^2(f_k - \widehat f_k^{n_1})^2 \\
&\approx \left(\widehat f_k^{n_1}\right)^2 + 0 + \left(\Omega_k^{n_1,n_2}\right)^2 \mathsf{Var}\left[\widehat f_k^{n_1}\middle|\mathcal{D}_k\right] \approx \left(\widehat f_k^{I}\right)^2 + \widehat\sigma_k^2 \frac{\left(\sum_{i=n_1-k}^{n_2-k-1}\widehat C_{i,k}^I\right)^2}{\sum_{i=0}^{n_1-k-1}\widehat C_{i,k}^I\left(\sum_{i=0}^{n_2-k-1}\widehat C_{i,k}^I\right)^2},
\end{aligned}$$

For the variance we get

$$\begin{aligned}
\mathsf{Var}\left[\widehat f_k^{n_2}\middle|\mathcal{D}^{n_1}\right] &= \mathsf{Var}\left[\sum_{i=0}^{n_2-k-1} w_{i,k}^{n_2}\frac{C_{i,k+1}}{C_{i,k}}\middle|\mathcal{D}^{n_1}\right] \approx \mathsf{Var}\left[\sum_{i=0}^{n_2-k-1} \widehat w_{i,k}^{n_2}\frac{C_{i,k+1}}{C_{i,k}}\middle|\mathcal{D}^{n_1}\right] \\
&= \sum_{i=n_1-k}^{n_2-k-1}\left(\widehat w_{i,k}^{n_2}\right)^2 \mathsf{Var}\left[\frac{C_{i,k+1}}{C_{i,k}}\middle|\mathcal{D}^{n_1}\right] \\
&= \sum_{i=n_1-k}^{n_2-k-1}\left(\widehat w_{i,k}^{n_2}\right)^2 \left(\mathsf{Var}\left[\mathsf{E}\left[\frac{C_{i,k+1}}{C_{i,k}}\middle|\mathcal{D}_k^{n_1}\right]\middle|\mathcal{D}^{n_1}\right] + \mathsf{E}\left[\mathsf{Var}\left[\frac{C_{i,k+1}}{C_{i,k}}\middle|\mathcal{D}_k^{n_1}\right]\middle|\mathcal{D}^{n_1}\right]\right) \\
&= \sum_{i=n_1-k}^{n_2-k-1}\left(\widehat w_{i,k}^{n_2}\right)^2 \left(0 + \mathsf{E}\left[\frac{\sigma_k^2}{C_{i,k}}\middle|\mathcal{D}^{n_1}\right]\right) \approx \sum_{i=n_1-k}^{n_2-k-1}\left(\widehat w_{i,k}^{n_2}\right)^2 \frac{\widehat\sigma_k^2}{\widehat C_{i,k}^I} = \widehat\sigma_k^2 \frac{\sum_{i=n_1-k}^{n_2-k-1}\widehat C_{i,k}^I}{\left(\sum_{i=0}^{n_2-k-1}\widehat C_{i,k}^I\right)^2}.
\end{aligned}$$



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 $\sum_{i=0}^{I}(\widehat{C}_{i,J}^{n_2} - \widehat{C}_{i,J}^{n_1})$ between two time periods $I \leq n_1 < n_2$ can be estimated by

$$\begin{aligned}
&\mathsf{mse}_{0|\mathcal{D}^I}\left[\widehat{\mathsf{CDR}}^{n_1,n_2}\right] \\
&:= \mathsf{E}\left[\left(\sum_{i=0}^{I}\left(\widehat{C}_{i,J}^{n_2} - \widehat{C}_{i,J}^{n_1}\right)\right)^2 \middle| \mathcal{D}^I\right] \\
&= \left(\sum_{i=0}^{I} C_{i,0}\right)^2 \mathsf{E}\left[\left(\prod_{k=0}^{J-1}\widehat{f}_k^{n_2} - \prod_{k=0}^{J-1}\widehat{f}_k^{n_1}\right)^2 \middle| \mathcal{D}^I\right] \\
&\approx \left(\sum_{i=0}^{I} C_{i,J}\right)^2 \left(\prod_{k=0}^{J-1}\left(1 + \frac{\widehat{\sigma}_k^2}{\left(\widehat{f}_k^I\right)^2}\left(\frac{1}{\sum_{i=0}^{n_1-k-1}\widehat{C}_{i,k}^I} - \frac{1}{\sum_{i=0}^{n_2-k-1}\widehat{C}_{i,k}^I}\right)\right) - 1\right) \\
&\approx \left(\sum_{i=0}^{I} \widehat{C}_{i,J}^I\right)^2 \sum_{k=0}^{J-1}\frac{\widehat{\sigma}_k^2}{\widehat{f}_k^I}\left(\frac{1}{\sum_{i=0}^{n_1-k-1}\widehat{C}_{i,k}^I} - \frac{1}{\sum_{i=0}^{n_2-k-1}\widehat{C}_{i,k}^I}\right).
\end{aligned}$$

Both estimates together lead to

$$\begin{aligned}
\mathsf{E}\left[\left(\widehat{f}_k^{n_2}\right)^2 \middle| \mathcal{D}^{n_1}\right] &\approx \left(\widehat{f}_k^I\right)^2 + \widehat{\sigma}_k^2\left(\frac{\left(\sum_{i=n_1-k}^{n_2-k-1}\widehat{C}_{i,k}^I\right)^2}{\sum_{i=0}^{n_1-k-1}\widehat{C}_{i,k}^I\left(\sum_{i=0}^{n_2-k-1}\widehat{C}_{i,k}^I\right)^2} + \frac{\sum_{i=n_1-k}^{n_2-k-1}\widehat{C}_{i,k}^I}{\left(\sum_{i=0}^{n_2-k-1}\widehat{C}_{i,k}^I\right)^2}\right) \\
&= \left(\widehat{f}_k^I\right)^2 + \widehat{\sigma}_k^2 \frac{\sum_{i=n_1-k}^{n_2-k-1}\widehat{C}_{i,k}^I}{\sum_{i=0}^{n_1-k-1}\widehat{C}_{i,k}^I \sum_{i=0}^{n_2-k-1}\widehat{C}_{i,k}^I} \\
&= \left(\widehat{f}_k^I\right)^2 + \widehat{\sigma}_k^2\left(\frac{1}{\sum_{i=0}^{n_1-k-1}\widehat{C}_{i,k}^I} - \frac{1}{\sum_{i=0}^{n_2-k-1}\widehat{C}_{i,k}^I}\right)
\end{aligned}$$

Combining all we get

$$\begin{aligned}
\mathsf{mse}_{0|\mathcal{D}^I}\left[\widehat{\mathsf{CDR}}^{n_1,n_2}\right] &\approx \left(\sum_{i=0}^{I} C_{i,0}\right)^2\left(\prod_{k=0}^{J-1}\left(\left(\widehat{f}_k^I\right)^2 + \widehat{\sigma}_k^2\left(\frac{1}{\sum_{i=0}^{n_1-k-1}\widehat{C}_{i,k}^I} - \frac{1}{\sum_{i=0}^{n_2-k-1}\widehat{C}_{i,k}^I}\right)\right) - \prod_{k=0}^{J-1}\left(\widehat{f}_k^I\right)^2\right) \\
&= \left(\sum_{i=0}^{I} C_{i,J}\right)^2\left(\prod_{k=0}^{J-1}\left(1 + \frac{\widehat{\sigma}_k^2}{\left(\widehat{f}_k^I\right)^2}\left(\frac{1}{\sum_{i=0}^{n_1-k-1}\widehat{C}_{i,k}^I} - \frac{1}{\sum_{i=0}^{n_2-k-1}\widehat{C}_{i,k}^I}\right)\right) - 1\right) \\
&\approx \left(\sum_{i=0}^{I} \widehat{C}_{i,J}^I\right)^2 \sum_{k=0}^{J-1}\frac{\widehat{\sigma}_k^2}{\widehat{f}_k^I}\left(\frac{1}{\sum_{i=0}^{n_1-k-1}\widehat{C}_{i,k}^I} - \frac{1}{\sum_{i=0}^{n_2-k-1}\widehat{C}_{i,k}^I}\right),
\end{aligned}$$

where we used in the last step a Taylor approximation in $\widehat{\sigma}_k^2$ at zero.

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 [12] 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

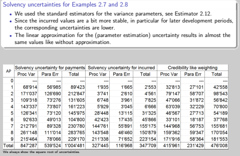
- All summation over accident periods stop at I, but we skipped $\wedge I$ in order to keep the formulas a bit simpler.
- the red parts are the differences to our estimators for the solvency and ultimate uncertainty, i.e.
  - If we take $n_2 = I + 1$ and $n_1 = I$ Estimator 2.25 leads to the same formulas 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 [12] 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.

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.

AP	Solvency uncertainty for payments			Solvency uncertainty for incurred			Credibility like weighting		
	Proc Var	Para Err	Total	Proc Var	Para Err	Total	Proc Var	Para Err	Total
0	---	---	---	---	---	---	---	---	---
1	68'914	56'985	89'423	1'935	1'665	2'553	32'813	27'101	42'558
2	171'037	126'690	212'847	3'741	2'610	4'561	79'147	58'707	98'543
3	109'318	73'276	131'605	6'748	3'961	7'825	47'066	31'872	56'842
4	143'337	73'807	161'223	5'929	3'045	6'666	63'039	32'229	70'800
5	126'341	73'120	145'975	28'448	13'115	31'325	46'567	27'713	54'189
6	92'633	49'013	104'800	42'423	17'435	45'866	33'101	18'187	37'768
7	212'791	89'328	230'780	144'761	55'891	155'175	144'968	56'753	155'681
8	261'148	111'014	283'765	143'548	46'460	150'879	159'362	59'347	170'054
9	215'464	78'066	229'170	211'338	71'652	223'154	171'916	58'364	181'553
Total	847'287	539'524	1'004'481	327'445	116'968	347'709	415'961	231'429	476'008

We always show the square root of uncertainties.



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.

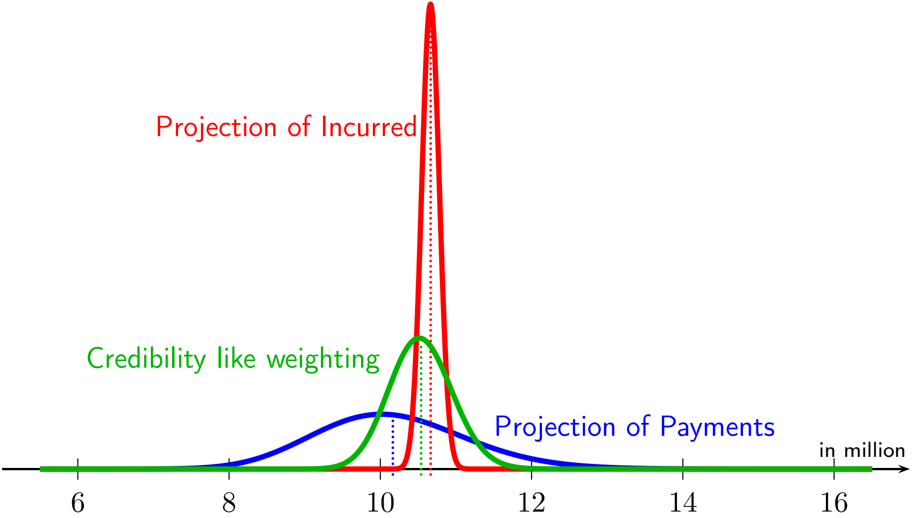
AP	Solvency uncertainty for payments			Solvency uncertainty for incurred			Credibility like weighting		
	Proc Var	Para Err	Total	Proc Var	Para Err	Total	Proc Var	Para Err	Total
0	---	---	---	---	---	---	---	---	---
1	68'914	56'985	89'423	1'935	1'665	2'553	32'813	27'101	42'558
2	171'037	126'690	212'847	3'741	2'610	4'561	79'147	58'707	98'543
3	109'318	73'276	131'605	6'748	3'961	7'825	47'066	31'872	56'842
4	143'337	73'807	161'223	5'929	3'045	6'666	63'039	32'229	70'800
5	126'341	73'120	145'975	28'448	13'115	31'325	46'567	27'713	54'189
6	92'633	49'013	104'800	42'423	17'435	45'866	33'101	18'187	37'768
7	212'791	89'328	230'780	144'761	55'891	155'175	144'968	56'753	155'681
8	261'148	111'014	283'765	143'548	46'460	150'879	159'362	59'347	170'054
9	215'464	78'066	229'170	211'338	71'652	223'154	171'916	58'364	181'553
Total	847'287	539'524	1'004'481	327'445	116'968	347'709	415'961	231'429	476'008

We always show the square root of uncertainties.

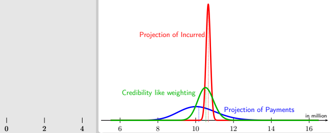
The uncertainty of the weighing has been calculated using a LSRM coupling of both CLM via the exposure $R^{0,1}_{i,k} = R^{1,0}_{i,k} := R^{0,0}_{i,k} + R^{1,1}_{i,k}$, see Section 4.

One can derive estimators for uncertainties without a first order Taylor approximation, see [21]. In practice, the resulting figures are almost alike.

Density plot of the distribution of the CDR using Lognormal distributions (dotted lines representing the Best Estimate)



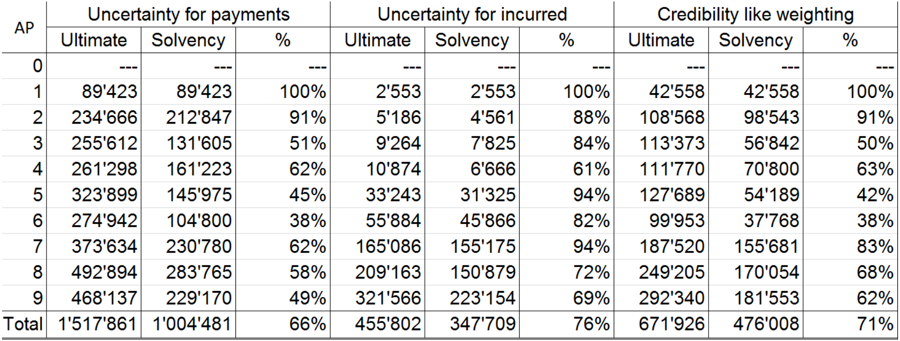
Density plot of the distribution of the CDR using Lognormal distributions (dotted lines representing the Best Estimate)



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 $R^{0,1}_{i,k} = R^{1,0}_{i,k} := R^{0,0}_{i,k} + R^{1,1}_{i,k}$, see Section 4.

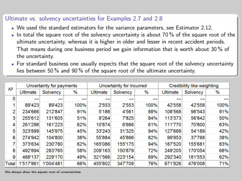
Ultimate vs. solvency uncertainties for Examples 2.7 and 2.8

- 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.



AP	Uncertainty for payments			Uncertainty for incurred			Credibility like weighting		
	Ultimate	Solvency	%	Ultimate	Solvency	%	Ultimate	Solvency	%
0	---	---	---	---	---	---	---	---	---
1	89'423	89'423	100%	2'553	2'553	100%	42'558	42'558	100%
2	234'666	212'847	91%	5'186	4'561	88%	108'568	98'543	91%
3	255'612	131'605	51%	9'264	7'825	84%	113'373	56'842	50%
4	261'298	161'223	62%	10'874	6'666	61%	111'770	70'800	63%
5	323'899	145'975	45%	33'243	31'325	94%	127'689	54'189	42%
6	274'942	104'800	38%	55'884	45'866	82%	99'953	37'768	38%
7	373'634	230'780	62%	165'086	155'175	94%	187'520	155'681	83%
8	492'894	283'765	58%	209'163	150'879	72%	249'205	170'054	68%
9	468'137	229'170	49%	321'566	223'154	69%	292'340	181'553	62%
Total	1'517'861	1'004'481	66%	455'802	347'709	76%	671'926	476'008	71%

We always show the square root of uncertainties.



Ultimate vs. solvency uncertainties for Examples 2.7 and 2.8

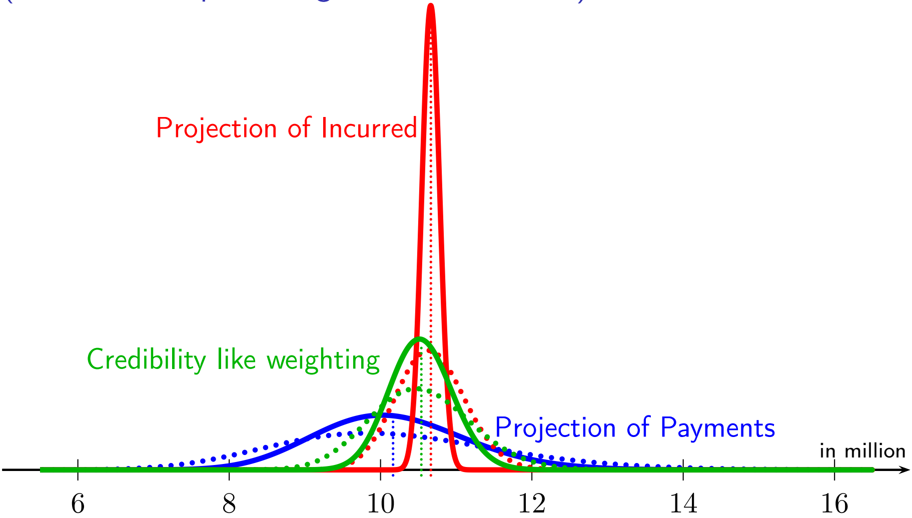
- We used the standard estimators for the variance parameters, see Estimator 2.12.
- In total the square root of the solvency uncertainty is about 70 % of the square root of the ultimate uncertainty, whereas it is higher in older and lesser in recent accident periods. That means during one business period 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.

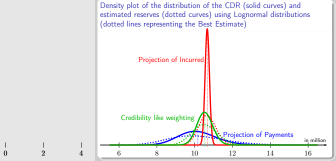
AP	Uncertainty for payments			Uncertainty for incurred			Credibility like weighting		
	Ultimate	Solvency	%	Ultimate	Solvency	%	Ultimate	Solvency	%
0	---	---	---	---	---	---	---	---	---
1	89'423	89'423	100%	2'553	2'553	100%	42'558	42'558	100%
2	234'666	212'847	91%	5'186	4'561	88%	108'568	98'543	91%
3	255'612	131'605	51%	9'264	7'825	84%	113'373	56'842	50%
4	261'298	161'223	62%	10'874	6'666	61%	111'770	70'800	63%
5	323'899	145'975	45%	33'243	31'325	94%	127'689	54'189	42%
6	274'942	104'800	38%	55'884	45'866	82%	99'953	37'768	38%
7	373'634	230'780	62%	165'086	155'175	94%	187'520	155'681	83%
8	492'894	283'765	58%	209'163	150'879	72%	249'205	170'054	68%
9	468'137	229'170	49%	321'566	223'154	69%	292'340	181'553	62%
Total	1'517'861	1'004'481	66%	455'802	347'709	76%	671'926	476'008	71%

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)





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

[7] Christian Braun.
The prediction error of the chain ladder method applied to correlated run-off triangles.
Astin Bull., 34(2):399–423, 2004.

[8] René Dahms.
Linear stochastic reserving methods.
Astin Bull., 42(1):1–34, 2012.

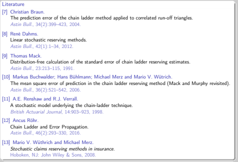
[9] Thomas Mack.
Distribution-free calculation of the standard error of chain ladder reserving estimates.
Astin Bull., 23:213–115, 1991.

[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).
Astin Bull., 36(2):521–542, 2006.

[11] A.E. Renshaw and R.J. Verrall.
A stochastic model underlying the chain-ladder technique.
British Actuarial Journal, 14:903–923, 1998.

[12] Ancus Röhr.
Chain Ladder and Error Propagation.
Astin Bull., 46(2):293–330, 2016.

[13] Mario V. Wüthrich and Michael Merz.
Stochastic claims reserving methods in insurance.
Hoboken, NJ: John Wiley & Sons, 2008.



Literature

[7] Christian Braun.
The prediction error of the chain ladder method applied to correlated run-off triangles.
Astin Bull., 34(2):399–423, 2004.

[8] René Dahms.
Linear stochastic reserving methods.
Astin Bull., 42(1):1–34, 2012.

[9] Thomas Mack.
Distribution-free calculation of the standard error of chain ladder reserving estimates.
Astin Bull., 23:213–115, 1991.

[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).
Astin Bull., 36(2):521–542, 2006.

[11] A.E. Renshaw and R.J. Verrall.
A stochastic model underlying the chain-ladder technique.
British Actuarial Journal, 14:903–923, 1998.

[12] Ancus Röhr.
Chain Ladder and Error Propagation.
Astin Bull., 46(2):293–330, 2016.

[13] Mario V. Wüthrich and Michael Merz.
Stochastic claims reserving methods in insurance.
Hoboken, NJ: John Wiley & Sons, 2008.

Stochastic Reserving

Lecture 5

Other Reserving Methods

René Dahms

ETH Zurich, Spring 2021

24 March 2021

(Last update: 26 April 2021)

Stochastic Reserving

Lecture 5

Other Reserving Methods

René Dahms

ETH Zurich, Spring 2021

24 March 2021
(Last update: 26 April 2021)

3 Other classical reserving methods

Lecture 5: Table of contents

Basic idea behind the Complementary-Loss-Ration method

The Complementary-Loss-Ration method is based on a single triangle and a 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 period 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 a 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 period are independent.

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

Simple example

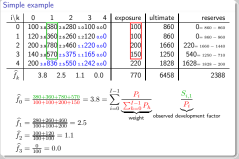
i\k	0	1	2	3	4	exposure	ultimate	reserves
0	100 $_{3.8}$	380 $_{2.8}$	280 $_{1.0}$	100 $_{0.0}$	0	100	860	0 $= 860 - 860$
1	120 $_{3.6}$	360 $_{2.6}$	260 $_{1.2}$	120 $_{0.0}$	0	100	860	0 $= 860 - 860$
2	200 $_{3.9}$	780 $_{2.3}$	460 $_{1.1}$	220 $_{0.0}$	0	200	1660	220 $= 1660 - 1440$
3	140 $_{3.8}$	570 $_{2.5}$	375 $_{1.1}$	165 $_{0.0}$	0	150	1250	540 $= 1250 - 710$
4	200 $_{3.8}$	836 $_{2.5}$	550 $_{1.1}$	242 $_{0.0}$	0	220	1828	1628 $= 1828 - 200$
$\widehat{f}_k$	3.8	2.5	1.1	0.0		770	6458	2388

$$\widehat{f}_0 = \frac{380+360+780+570}{100+100+200+150} = 3.8 = \sum_{i=0}^{I-1} \underbrace{\frac{P_i}{\sum_{h=0}^{I-1} P_h}}_{\text{weight}} \underbrace{\frac{S_{i,1}}{P_i}}_{\text{observed development factor}}$$

$$\widehat{f}_1 = \frac{280+260+460}{100+100+200} = 2.5$$

$$\widehat{f}_2 = \frac{100+120}{100+100} = 1.1$$

$$\widehat{f}_3 = \frac{0}{100} = 0.0$$



Simple example

i\k	0	1	2	3	4	exposure	ultimate	reserves
0	100	$_{3.8}$380	$_{2.8}$280	$_{1.0}$100	$_{0.0}$0	100	860	0$_{= 860 - 860}$
1	120	$_{3.6}$360	$_{2.6}$260	$_{1.2}$120	$_{0.0}$0	100	860	0$_{= 860 - 860}$
2	200	$_{3.9}$780	$_{2.3}$460	$_{1.1}$220	$_{0.0}$0	200	1660	220$_{= 1660 - 1440}$
3	140	$_{3.8}$570	$_{2.5}$375	$_{1.1}$165	$_{0.0}$0	150	1250	540$_{= 1250 - 710}$
4	200	$_{3.8}$836	$_{2.5}$550	$_{1.1}$242	$_{0.0}$0	220	1828	1628$_{= 1828 - 200}$
$\widehat{f}_k$		3.8	2.5	1.1	0.0	770	6458	2388

$$\widehat{f}_0 = \frac{380+360+780+570}{100+100+200+150} = 3.8 = \sum_{i=0}^{I-1} \underbrace{\frac{P_i}{\sum_{h=0}^{I-1} P_h}}_{\text{weight}} \underbrace{\frac{S_{i,1}}{P_i}}_{\text{observed development factor}}$$

$$\widehat{f}_1 = \frac{280+260+460}{100+100+200} = 2.5$$

$$\widehat{f}_2 = \frac{100+120}{100+100} = 1.1$$

$$\widehat{f}_3 = \frac{0}{100} = 0.0$$

Assumption 3.A (CLRM)

There exist exposures P_i*, development factors* f_k *and variance parameters* σ_k^2 *such that*

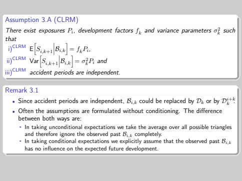
i)$^{\text{CLRM}}$ $\mathsf{E}\left[S_{i,k+1}\middle|\mathcal{B}_{i,k}\right] = f_k P_i,$

ii)$^{\text{CLRM}}$ $\mathsf{Var}\left[S_{i,k+1}\middle|\mathcal{B}_{i,k}\right] = \sigma_k^2 P_i$ *and*

iii)$^{\text{CLRM}}$ *accident periods are independent.*

Remark 3.1

- Since accident periods are independent, $\mathcal{B}_{i,k}$ could be replaced by $\mathcal{D}_k$ or by $\mathcal{D}_k^{i+k}$.
- 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 $\mathcal{B}_{i,k}$ completely.
  - In taking conditional expectations we explicitly assume that the observed past $\mathcal{B}_{i,k}$ has no influence on the expected future development.



Assumption 3.A (CLRM)

There exist exposures P_i, development factors f_k and variance parameters σ_k^2 such that

i)$^{\text{CLRM}}$ $\mathsf{E}\left[S_{i,k+1}\middle|\mathcal{B}_{i,k}\right] = f_k P_i$,
ii)$^{\text{CLRM}}$ $\mathsf{Var}\left[S_{i,k+1}\middle|\mathcal{B}_{i,k}\right] = \sigma_k^2 P_i$ *and*
iii)$^{\text{CLRM}}$ *accident periods are independent.*

Remark 3.1

- Since accident periods are independent, $\mathcal{B}_{i,k}$ could be replaced by $\mathcal{D}_k$ or by $\mathcal{D}_k^{i+k}$.
- 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 $\mathcal{B}_{i,k}$ completely.
  - In taking conditional expectations we explicitly assume that the observed past $\mathcal{B}_{i,k}$ 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 $\mathcal{D}_k$*-conditionally unbiased estimators* $\widehat{f}_k$ *of* f_k *the estimator*

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

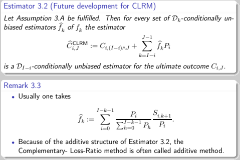
is a $\mathcal{D}_{I-i}$*-conditionally unbiased estimator for the ultimate outcome* $C_{i,J}$.

Remark 3.3

- Usually one takes

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

- 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 $\mathcal{D}_k$-conditionally unbiased estimators $\widehat{f}_k$ of f_k the estimator

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

is a $\mathcal{D}_{I-i}$-conditionally unbiased estimator for the ultimate outcome $C_{i,J}$.

Remark 3.3

- Usually one takes

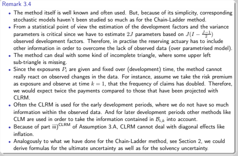
$$\widehat{f}_k := \sum_{i=0}^{I-k-1} \frac{P_i}{\sum_{h=0}^{I-k-1} P_h} \frac{S_{i,k+1}}{P_i}.$$

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

$$\begin{aligned}
\mathsf{E}\left[\widehat{C}_{i,J}^{\mathsf{CLRM}} \middle| \mathcal{D}_{I-i}\right] &= C_{i,(I-i)\wedge J} + \sum_{k=I-i}^{J-1} \mathsf{E}\left[\widehat{f}_k \middle| \mathcal{D}_{I-i}\right] P_i \\
&= C_{i,(I-i)\wedge J} + \sum_{k=I-i}^{J-1} \mathsf{E}\left[\mathsf{E}\left[\widehat{f}_k \middle| \mathcal{D}_k\right] \middle| \mathcal{D}_{I-i}\right] P_i \\
&= C_{i,(I-i)\wedge J} + \sum_{k=I-i}^{J-1} f_k P_i \\
&= C_{i,(I-i)\wedge J} + \sum_{k=I-i}^{J-1} \underbrace{\mathsf{E}\left[\mathsf{E}\left[S_{i,k+1} \middle| \mathcal{D}_k\right] \middle| \mathcal{D}_{I-i}\right]}_{\text{i)}^{\mathsf{CLRM}}} \\
&= \mathsf{E}\left[C_{i,J} \middle| \mathcal{D}_{I-i}\right]
\end{aligned}$$

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 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 $\mathcal{B}_{i,k}$ into account.
- Because of part iii)$^{\text{CLRM}}$ 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 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 $\mathcal{B}_{i,k}$ into account.
- Because of part iii)$^{\text{CLRM}}$ 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.

We will analyse uncertainties in the more general setup of Linear Stochastic Reserving methods later in section 4.

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 [15]. 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.

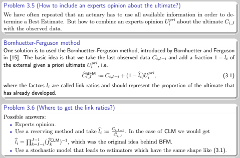
$$\widehat{C}_{i,J}^{\mathsf{BFM}} := C_{i,I-i} + (1 - \widehat{l}_i) U_i^{pri}, \tag{3.1}$$

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 reasoning method and take $\widehat{l}_i := \frac{C_{i,I-i}}{\widehat{C}_{i,J}}$. In the case of CLM we would get $\widehat{l}_i = \prod_{k=I-i}^{J-1} (\widehat{f}_k^{\mathsf{CLM}})^{-1}$, 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 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 [15]. 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.

$$\widehat{C}_{i,J}^{\mathsf{BFM}} := C_{i,I-i} + (1 - \widehat{l}_i)U_i^{pri}, \tag{3.1}$$

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 $\widehat{l}_i := \frac{C_{i,I-i}}{\widehat{C}_{i,J}}$. In the case of CLM we would get $\widehat{l}_i = \prod_{k=I-i}^{J-1}(\widehat{f}_k^{\mathsf{CLM}})^{-1}$, which was the original idea behind BFM.
- Use a stochastic model that leads to estimators which have the same shape like (3.1).

Roughly spoken, we take the a priori ultimate and replace the already developed part by the observated data.

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 $l_i := \frac{C_{i,I-i}}{\widehat{C}_{i,J}}$ and if all link ratios $0 \leq l_i \leq 1$ then $C_{i,J}^{\mathsf{BFM}}$ could be looked at as credibility like weighted average of the a priori ultimate U_i^{pri} and the estimated ultimate $\widehat{C}_{i,J}$ with credibility weights $(1 - l_i)$ and l_i, respectively:

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

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 $l_{I-J} = 1$, 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 $l_i := \frac{C_{i,I-i}}{\widehat{C}_{i,J}}$ and if all link ratios $0 \leq l_i \leq 1$ then $C_{i,J}^{\mathsf{BFM}}$ could be looked at as credibility like weighted average of the a priori ultimate U_i^{pri} and the estimated ultimate $\widehat{C}_{i,J}$ with credibility weights $(1 - l_i)$ and l_i, respectively:

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

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

Remark 3.8 (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)

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

Defining the link ratios via

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

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 [18, Section 6.6].

Remark 3.8 (BFM as Complementary-Loss-Ratio method)

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

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

Defining the link ratios via

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

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 [18, 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 priory ultimate

$$U_i^{\mathsf{BHM}\,pri} := \widehat{l}_i \widehat{C}_{i,J}^{\mathsf{CLM}} + (1 - \widehat{l}_i) U_i^{pri} = C_{i,I-i} + (1 - \widehat{l}_i) U_i^{pri} = \widehat{C}_{i,J}^{\mathsf{BFM}},$$

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

Then we get the estimate

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

Remark 3.10

Connection between BHM, BFM and CLM

- BHM was independently developed by Benktander, see [14], and Hovinen, see [16].
- The BHM is a twice iterated BFM with Chain-Ladder link ratios.
- Iterating BFM further will finally lead to the CLM Best Estimate, see [17].

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 priory ultimate

$$U_i^{\mathsf{BHM}\,pri} := \widehat{l}_i \widehat{C}_{i,J}^{\mathsf{CLM}} + (1-\widehat{l}_i)U_i^{pri} = C_{i,I-i} + (1-\widehat{l}_i)U_i^{pri} = \widehat{C}_{i,J}^{\mathsf{BFM}},$$

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

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

Remark 3.10

Connection between BHM, BFM and CLM

- BHM was independently developed by Benktander, see [14], and Hovinen, see [16].
- The BHM is a twice iterated BFM with Chain-Ladder link ratios.
- Iterating BFM further will finally lead to the CLM Best Estimate, see [17].

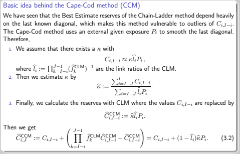
Basic idea behind the Cape-Cod method (CCM)

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 κ with
$$C_{i,I-i} \approx \kappa \widehat{l}_i P_i \,,$$
where $\widehat{l}_i := \prod_{k=I-i}^{J-1} (\widehat{f}_k^{\text{CLM}})^{-1}$ are the link ratios of the CLM.
2. Then we estimate κ by
$$\widehat{\kappa} := \frac{\sum_{i=I-J}^{I} C_{i,I-i}}{\sum_{i=I-J} \widehat{l}_i P_i}.$$
3. Finally, we calculate the reserves with CLM where the values $C_{i,I-i}$ are replaced by
$$\widehat{C}_{i,I-i}^{\text{CCM}} := \widehat{\kappa} \widehat{l}_i P_i.$$

Then we get

$$\widehat{C}_{i,J}^{\text{CCM}} := C_{i,I-i} + \left(\prod_{k=I-i}^{J-1} \widehat{f}_k^{\text{CLM}} \widehat{C}_{i,I-i}^{\text{CCM}} - \widehat{C}_{i,I-i}^{\text{CCM}} \right) = C_{i,I-i} + (1 - \widehat{l}_i) \widehat{\kappa} P_i. \tag{3.2}$$



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 κ with
$$C_{i,I-i} \approx \kappa \widehat{l}_i P_i \, ,$$
where $\widehat{l}_i := \prod_{k=I-i}^{J-1} (\widehat{f}_k^{\mathsf{CLM}})^{-1}$ are the link ratios of the CLM.
2. Then we estimate κ by
$$\widehat{\kappa} := \frac{\sum_{i=I-J}^{I} C_{i,I-i}}{\sum_{i=I-J} \widehat{l}_i P_i}.$$
3. Finally, we calculate the reserves with CLM where the values $C_{i,I-i}$ are replaced by
$$\widehat{C}_{i,I-i}^{\mathsf{CCM}} := \widehat{\kappa} \widehat{l}_i P_i.$$

Then we get
$$\widehat{C}_{i,J}^{\mathsf{CCM}} := C_{i,I-i} + \left(\prod_{k=I-i}^{J-1} \widehat{f}_k^{\mathsf{CLM}} \widehat{C}_{i,I-i}^{\mathsf{CCM}} - \widehat{C}_{i,I-i}^{\mathsf{CCM}} \right) = C_{i,I-i} + (1 - \widehat{l}_i) \widehat{\kappa} P_i. \tag{3.2}$$

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 $\widehat{\kappa}$) modified a priory ultimate $\widehat{\kappa} P_i$.

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 $\widehat{\kappa}$) modified a priory ultimate $\widehat{\kappa}P_i$.

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.

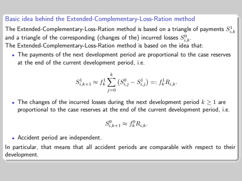
$$S_{i,k+1}^1 \approx f_k^1 \sum_{j=0}^{k} \left(S_{i,j}^0 - S_{i,j}^1\right) =: f_k^1 R_{i,k}.$$

- 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.

$$S_{i,k+1}^0 \approx f_k^0 R_{i,k}.$$

- 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 $S^1_{i,k}$ and a triangle of the corresponding (changes of the) incurred losses $S^0_{i,k}$.
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.

$$S^1_{i,k+1} \approx f^1_k \sum_{j=0}^{k} \left(S^0_{i,j} - S^1_{i,j}\right) =: f^1_k R_{i,k}.$$

- 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.

$$S^0_{i,k+1} \approx f^0_k R_{i,k}.$$

- 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 $S^0_{i,k}$

i\k	0	1	2	3
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
$\widehat{f}^0_k$	0.4	-0.4	0.0	

Payments $S^1_{i,k}$

i\k	0	1	2	3
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
$\widehat{f}^1_k$	0.5	0.5	1.0	

Case reserves $R_{i,k}$

i\k	0	1	2	3
0	400 $\xrightarrow{1.0}$	400 $\xrightarrow{0.1}$	40 $\xrightarrow{0.0}$	0
1	400 $\xrightarrow{1.0}$	400 $\xrightarrow{0.1}$	40 $\xrightarrow{0.0}$	0
2	400 $\xrightarrow{0.7}$	280 $\xrightarrow{0.1}$	28 $\xrightarrow{0.0}$	0
3	300 $\xrightarrow{0.9}$	270 $\xrightarrow{0.1}$	27 $\xrightarrow{0.0}$	0
$\widehat{f}_k$	0.9	0.1	0.0	

$$\widehat{f}^0_0 = \frac{200+160+120}{400+400+400} = 0.4 \qquad \widehat{f}^1_0 = \frac{200+160+240}{400+400+400} = 0.5 \qquad \widehat{f}_0 = 1 + 0.4 - 0.5 = 0.9$$

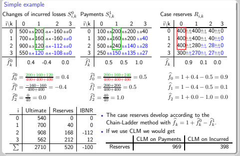
$$\widehat{f}^0_1 = \frac{-160-160}{400+400} = -0.4 \qquad \widehat{f}^1_1 = \frac{200+200}{400+400} = 0.5 \qquad \widehat{f}_1 = 1 - 0.4 - 0.5 = 0.1$$

$$\widehat{f}^0_2 = \frac{0}{40} = 0.0 \qquad \widehat{f}^1_2 = \frac{40}{40} = 1.0 \qquad \widehat{f}_2 = 1 + 0.0 - 1.0 = 0.0$$

i	Ultimate	Reserves	IBNR
0	540	0	0
1	700	40	0
2	908	168	-112
3	562	212	12
$\sum$	2710	520	-100

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

	CLM on Payments	CLM on Incurred
Reserves	969	398



Simple example

Changes of incurred losses $S^0_{i,k}$

i\k	0	1	2	3
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
$\widehat{f}^0_k$	0.4	-0.4	0.0	

Payments $S^1_{i,k}$

i\k	0	1	2	3
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
$\widehat{f}^1_k$	0.5	0.5	1.0	

Case reserves $R_{i,k}$

i\k	0	1	2	3
0	400 $\xrightarrow{1.0}$	400 $\xrightarrow{0.1}$	40 $\xrightarrow{0.0}$	0
1	400 $\xrightarrow{1.0}$	400 $\xrightarrow{0.1}$	40 $\xrightarrow{0.0}$	0
2	400 $\xrightarrow{0.7}$	280 $\xrightarrow{0.1}$	28 $\xrightarrow{0.0}$	0
3	300 $\xrightarrow{0.9}$	270 $\xrightarrow{0.1}$	27 $\xrightarrow{0.0}$	0
$\widehat{f}_k$	0.9	0.1	0.0	

$$\widehat{f}^0_0 = \frac{200+160+120}{400+400+400} = 0.4 \qquad \widehat{f}^1_0 = \frac{200+160+240}{400+400+400} = 0.5 \qquad \widehat{f}_0 = 1 + 0.4 - 0.5 = 0.9$$

$$\widehat{f}^0_1 = \frac{-160-160}{400+400} = -0.4 \qquad \widehat{f}^1_1 = \frac{200+200}{400+400} = 0.5 \qquad \widehat{f}_1 = 1 - 0.4 - 0.5 = 0.1$$

$$\widehat{f}^0_2 = \frac{0}{40} = 0.0 \qquad \widehat{f}^1_2 = \frac{40}{40} = 1.0 \qquad \widehat{f}_2 = 1 + 0.0 - 1.0 = 0.0$$

i	Ultimate	Reserves	IBNR
0	540	0	0
1	700	40	0
2	908	168	-112
3	562	212	12
$\sum$	2710	520	-100

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

	CLM on Payments	CLM on Incurred
Reserves	969	398

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*

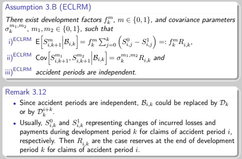
i)$^{\text{ECLRM}}$ $\mathsf{E}\left[S_{i,k+1}^m \middle| \mathcal{B}_{i,k}\right] = f_k^m \sum_{j=0}^k \left(S_{i,j}^0 - S_{i,j}^1\right) =: f_k^m R_{i,k}$,

ii)$^{\text{ECLRM}}$ $\mathsf{Cov}\left[S_{i,k+1}^{m_1}, S_{i,k+1}^{m_2} \middle| \mathcal{B}_{i,k}\right] = \sigma_k^{m_1,m_2} R_{i,k}$ *and*

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

Remark 3.12

- Since accident periods are independent, $\mathcal{B}_{i,k}$ could be replaced by $\mathcal{D}_k$ or by $\mathcal{D}_k^{i+k}$.
- 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 $R_{i,k}$ are the case reserves at the end of development period k for claims of accident period i.



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*

i)$^{\text{ECLRM}}$ $\mathsf{E}\left[S_{i,k+1}^m \middle| \mathcal{B}_{i,k}\right] = f_k^m \sum_{j=0}^k \left(S_{i,j}^0 - S_{i,j}^1\right) =: f_k^m R_{i,k}$,

ii)$^{\text{ECLRM}}$ $\mathsf{Cov}\left[S_{i,k+1}^{m_1}, S_{i,k+1}^{m_2} \middle| \mathcal{B}_{i,k}\right] = \sigma_k^{m_1,m_2} R_{i,k}$ *and*

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

Remark 3.12

- Since accident periods are independent, $\mathcal{B}_{i,k}$ could be replaced by $\mathcal{D}_k$ or by $\mathcal{D}_k^{i+k}$.
- 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 $R_{i,k}$ are the case reserves at the end of development period k for claims of accident period i.

Estimator 3.13 (Future development for ECLRM)

Assume Assumption 3.B is fulfilled. Then for every set of $\mathcal{D}_k$-conditionally unbiased estimators $\widehat{f}_k^m$ of f_k^m the estimators

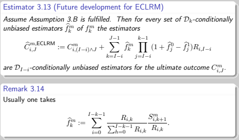
$$\widehat{C}_{i,J}^{m,\text{ECLRM}} := C_{i,(I-i)\wedge J}^m + \sum_{k=I-i}^{J-1} \widehat{f}_k^m \prod_{j=I-i}^{k-1} (1 + \widehat{f}_j^0 - \widehat{f}_j^1) R_{i,I-i}$$

are $\mathcal{D}_{I-i}$-conditionally unbiased estimators for the ultimate outcome $C_{i,J}^m$.

Remark 3.14

Usually one takes

$$\widehat{f}_k^m := \sum_{i=0}^{I-k-1} \frac{R_{i,k}}{\sum_{h=0}^{I-k-1} R_{i,k}} \frac{S_{i,k+1}^m}{R_{i,k}}.$$



Estimator 3.13 (Future development for ECLRM)

Assume Assumption 3.B is fulfilled. Then for every set of $\mathcal{D}_k$-conditionally unbiased estimators $\widehat{f}^m_k$ of f^m_k the estimators

$$\widehat{C}^{m,\textsf{ECLRM}}_{i,J} := C^m_{i,(I-i)\wedge J} + \sum_{k=I-i}^{J-1} \widehat{f}^m_k \prod_{j=I-i}^{k-1} (1 + \widehat{f}^0_j - \widehat{f}^1_j) R_{i,I-i}$$

are $\mathcal{D}_{I-i}$-conditionally unbiased estimators for the ultimate outcome $C^m_{i,J}$.

Remark 3.14

Usually one takes

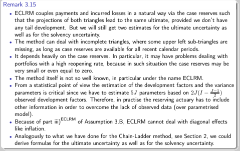
$$\widehat{f}^m_k := \sum_{i=0}^{I-k-1} \frac{R_{i,k}}{\sum_{h=0}^{I-k-1} R_{i,k}} \frac{S^m_{i,k+1}}{R_{i,k}}.$$

From Assumption 3.B.i)$^{\textsf{ECLRM}}$ it follows that $\mathsf{E}\left[R_{i,k+1} \middle| \mathcal{D}_k\right] = (1 + f^0_k - f^1_k) R_{i,k}$. Therefore, we get

$$\begin{aligned}
\mathsf{E}\left[\widehat{C}^{m,\textsf{ECLRM}}_{i,J} \middle| \mathcal{D}_{I-i}\right] &= C^m_{i,(I-i)\wedge J} + \sum_{k=I-i}^{J-1} \mathsf{E}\left[\widehat{f}^m_k \prod_{j=I-i}^{k-1} (1 + \widehat{f}^0_j - \widehat{f}^1_j) \middle| \mathcal{D}_{I-i}\right] R_{i,I-i} \\
&= C^m_{i,(I-i)\wedge J} + \sum_{k=I-i}^{J-1} \mathsf{E}\left[\mathsf{E}\left[\widehat{f}^m_k \middle| \mathcal{D}_k\right] \prod_{j=I-i}^{k-1} (1 + \widehat{f}^0_j - \widehat{f}^1_j) \middle| \mathcal{D}_{I-i}\right] R_{i,I-i} \\
&= C^m_{i,(I-i)\wedge J} + \sum_{k=I-i}^{J-1} f^m_k \mathsf{E}\left[\prod_{j=I-i}^{k-1} (1 + \widehat{f}^0_j - \widehat{f}^1_j) \middle| \mathcal{D}_{I-i}\right] R_{i,I-i} \\
= \ldots &= C^m_{i,(I-i)\wedge J} + \sum_{k=I-i}^{J-1} f^m_k \prod_{j=I-i}^{k-1} (1 + f^0_j - f^1_j) R_{i,I-i} \\
&= C^m_{i,(I-i)\wedge J} + \sum_{k=I-i}^{J-1} f^m_k \prod_{j=I-i+1}^{k-1} (1 + f^0_j - f^1_j) \mathsf{E}\left[R_{i,I-i+1} \middle| \mathcal{D}_{I-i}\right] \\
= \ldots &= C^m_{i,(I-i)\wedge J} + \sum_{k=I-i}^{J-1} f^m_k \mathsf{E}\left[R_{i,k} \middle| \mathcal{D}_{I-i}\right] \\
&= C^m_{i,(I-i)\wedge J} + \sum_{k=I-i}^{J-1} \mathsf{E}\left[S^m_{i,k+1} \middle| \mathcal{D}_{I-i}\right] = \mathsf{E}\left[C^m_{i,J} \middle| \mathcal{D}_{I-i}\right]
\end{aligned}$$

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)$^{\text{ECLRM}}$ 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

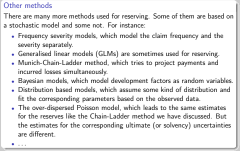
- ECLRM couples payments and incurred losses in a natural way via the case reserves such that the projections of both triangles lead to the same ultimate, provided we don't have any tail development. But we will still get two estimates for the ultimate uncertainty as well as for the solvency uncertainty.
- 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 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)$^{\text{ECLRM}}$ 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.

We will analyse uncertainties in the more general case of Linear Stochastic Reserving methods, see section 4.

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.
- . . .



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.
- . . .

Literature

[14] G. Benktander.
An approach to credibility in calculating IBNR for causalty excess reinsurance.
The Actuarial Review, 312:7, 1976.

[15] R.L. Bornhuetter and R.E. Ferguson.
The actuary and IBNR.
Proc. CAS, LIX:181–195, 1972.

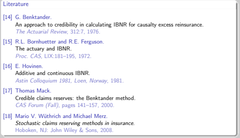
[16] E. Hovinen.
Additive and continuous IBNR.
Astin Colloquium 1981, Loen, Norway, 1981.

[17] Thomas Mack.
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.
Hoboken, NJ: John Wiley & Sons, 2008.

2021-04-26

Stochastic Reserving
Other classical reserving methods
Literature



Literature

[14] G. Benktander.
An approach to credibility in calculating IBNR for causalty excess reinsurance.
The Actuarial Review, 312:7, 1976.

[15] R.L. Bornhuetter and R.E. Ferguson.
The actuary and IBNR.
Proc. CAS, LIX:181–195, 1972.

[16] E. Hovinen.
Additive and continuous IBNR.
Astin Colloquium 1981, Loen, Norway, 1981.

[17] Thomas Mack.
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.
Hoboken, NJ: John Wiley & Sons, 2008.

Stochastic Reserving

Lecture 6

Linear-Stochastic-Reserving methods

René Dahms

ETH Zurich, Spring 2021

31 March 2021

(Last update: 26 April 2021)

Stochastic Reserving

Lecture 6

Linear-Stochastic-Reserving methods

René Dahms

ETH Zurich, Spring 2021

31 March 2021

(Last update: 26 April 2021)

4 Linear-Stochastic-Reserving methods

Stochastic Reserving

Lecture 6: Table of contents

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?

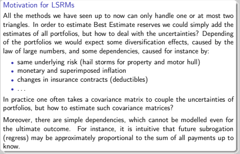
Moreover, there are simple dependencies, which cannot be modelled even for the ultimate outcome. For instance, it is intuitive that future subrogation (regress) may be approximately proportional to the sum of all payments up to know.

2021-04-26

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 (regress) 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

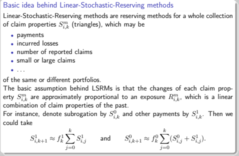
- payments
- incurred losses
- number of reported claims
- small or large claims
- ...

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

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



Basic idea behind Linear-Stochastic-Reserving methods

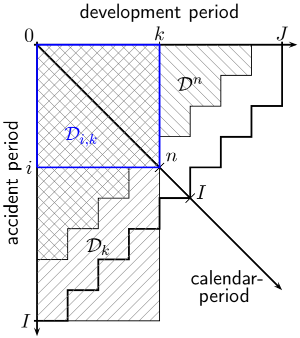
Linear-Stochastic-Reserving methods are reserving methods for a whole collection of claim properties $S_{i,k}^m$ (triangles), which may be

- 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

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

σ-algebras (repetition)



- $\mathcal{B}_{i,k}$ is the σ-algebra of all information of accident period i up to development period k:

$$\begin{aligned}\mathcal{B}_{i,k} :=& \sigma\left(S_{i,j}^m \colon 0 \leq j \leq k,\ 0 \leq m \leq M\right) \\ =& \sigma\left(C_{i,j}^m \colon 0 \leq j \leq k,\ 0 \leq m \leq M\right)\end{aligned}$$

- $\mathcal{D}_{i,k}$ is the σ-algebra containing all information up to accident period i and development period k:

$$\begin{aligned}\mathcal{D}_{i,k} :=& \sigma\left(S_{i,j} \colon 0 \leq h \leq i,\ 0 \leq j \leq k, 0 \leq m \leq M\right) \\ =& \sigma\left(\mathcal{B}_{h,k} \colon 0 \leq h \leq i\right)\end{aligned}$$

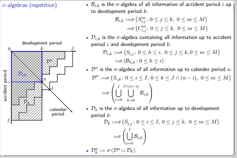
- $\mathcal{D}^n$ is the σ-algebra of all information up to calender period n:

$$\begin{aligned}\mathcal{D}^n :=& \sigma\left(S_{i,k} \colon 0 \leq i \leq I,\ 0 \leq k \leq J \wedge (n-i),\ 0 \leq m \leq M\right) \\ =& \sigma\left(\bigcup_{i=0}^{I} \bigcup_{k=0}^{J \wedge (n-i)} \mathcal{B}_{i,k}\right)\end{aligned}$$

- $\mathcal{D}_k$ is the σ-algebra of all information up to development period k:

$$\begin{aligned}\mathcal{D}_k :=& \sigma\left(S_{i,j} \colon 0 \leq i \leq I,\ 0 \leq j \leq k,\ 0 \leq m \leq M\right) \\ =& \sigma\left(\bigcup_{i=0}^{I} \mathcal{B}_{i,k}\right)\end{aligned}$$

- $\mathcal{D}_k^n := \sigma\left(\mathcal{D}^n \cup \mathcal{D}_k\right)$



σ-algebras (repetition)

- $\mathcal{B}_{i,k}$ is the σ-algebra of all information of accident period i up to development period k:
$$\begin{aligned}\mathcal{B}_{i,k} :=& \sigma\left(S_{i,j}^m \colon 0 \le j \le k,\ 0 \le m \le M\right)\\ =& \sigma\left(C_{i,j}^m \colon 0 \le j \le k,\ 0 \le m \le M\right)\end{aligned}$$
- $\mathcal{D}_{i,k}$ is the σ-algebra containing all information up to accident period i and development period k:
$$\begin{aligned}\mathcal{D}_{i,k} :=& \sigma\left(S_{i,j} \colon 0 \le h \le i,\ 0 \le j \le k, 0 \le m \le M\right)\\ =& \sigma\left(\mathcal{B}_{h,k} \colon 0 \le h \le i\right)\end{aligned}$$
- $\mathcal{D}^n$ is the σ-algebra of all information up to calender period n:
$$\begin{aligned}\mathcal{D}^n :=& \sigma\left(S_{i,k} \colon 0 \le i \le I, 0 \le k \le J \wedge (n-i),\ 0 \le m \le M\right)\\ =& \sigma\left(\bigcup_{i=0}^{I} \bigcup_{k=0}^{J\wedge(n-i)} \mathcal{B}_{i,k}\right)\end{aligned}$$
- $\mathcal{D}_k$ is the σ-algebra of all information up to development period k:
$$\begin{aligned}\mathcal{D}_k :=& \sigma\left(S_{i,j} \colon 0 \le i \le I, 0 \le j \le k,\ 0 \le m \le M\right)\\ =& \sigma\left(\bigcup_{i=0}^{I} \mathcal{B}_{i,k}\right)\end{aligned}$$
- $\mathcal{D}_k^n := \sigma\left(\mathcal{D}^n \cup \mathcal{D}_k\right)$

The σ-algebra $\mathcal{D}_k^n$ is used in order 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

$$\left(S_{i_1,k_1}^{m_1} \in \mathcal{D}_k^n \quad \text{and} \quad S_{i_2,k_2}^{m_2} \notin \mathcal{D}_k^n\right) \quad \text{or} \quad \left(S_{i_1,k_1}^{m_1} \notin \mathcal{D}_k^n \quad \text{and} \quad S_{i_2,k_2}^{m_2} \in \mathcal{D}_k^n\right).$$

Assumption 4.A (Linear-Stochastic-Reserving method)

We call the stochastic model of the increments $S_{i,k}^m$ *a Linear-Stochastic-Reserving method (*LSRM*) with*

- *development exposures* $R_{i,k}^m \in \mathcal{D}_{i,k}$*, which depend linearly on the claim properties, and*
- *covariance exposures* $R_{i,k}^{m_1,m_2} \in \mathcal{D}_{i,k}$*,*

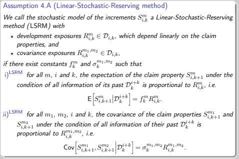
if there exist constants f_k^m *and* $\sigma_k^{m_1,m_2}$ *such that*

i)$^{\text{LSRM}}$ *for all* m*,* i *and* k*, the expectation of the claim property* $S_{i,k+1}^m$ *under the condition of all information of its past* $\mathcal{D}_k^{i+k}$ *is proportional to* $R_{i,k}^m$*, i.e.*

$$\mathsf{E}\left[S_{i,k+1}^m \middle| \mathcal{D}_k^{i+k}\right] = f_k^m R_{i,k}^m.$$

ii)$^{\text{LSRM}}$ *for all* m_1*,* m_2*,* i *and* k*, the covariance of the claim properties* $S_{i,k+1}^{m_1}$ *and* $S_{i,k+1}^{m_2}$ *under the condition of all information of their past* $\mathcal{D}_k^{i+k}$ *is proportional to* $R_{i,k}^{m_1,m_2}$*, i.e.*

$$\mathsf{Cov}\left[S_{i,k+1}^{m_1}, S_{i,k+1}^{m_2} \middle| \mathcal{D}_k^{i+k}\right] = \sigma_k^{m_1,m_2} R_{i,k}^{m_1,m_2}.$$



Assumption 4.A (Linear-Stochastic-Reserving method)

We call the stochastic model of the increments $S_{i,k}^m$ *a Linear-Stochastic-Reserving method (*LSRM*) with*

- *development exposures* $R_{i,k}^m \in \mathcal{D}_{i,k}$*, which depend linearly on the claim properties, and*
- *covariance exposures* $R_{i,k}^{m_1,m_2} \in \mathcal{D}_{i,k}$*,*

if there exist constants f_k^m *and* $\sigma_k^{m_1,m_2}$ *such that*

i)$^{\text{LSRM}}$ *for all* m, i *and* k*, the expectation of the claim property* $S_{i,k+1}^m$ *under the condition of all information of its past* $\mathcal{D}_k^{i+k}$ *is proportional to* $R_{i,k}^m$*, i.e.*

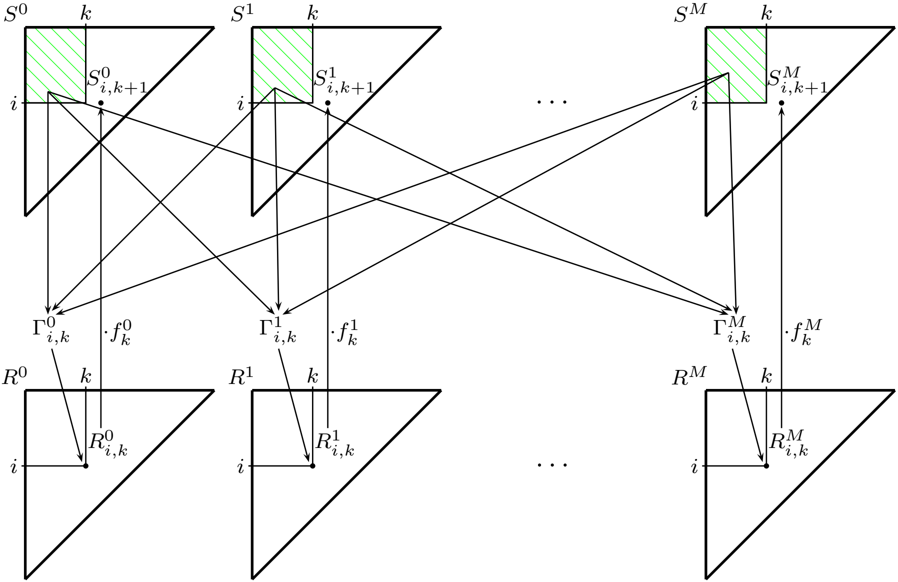
$$\mathsf{E}\left[S_{i,k+1}^m \middle| \mathcal{D}_k^{i+k}\right] = f_k^m R_{i,k}^m.$$

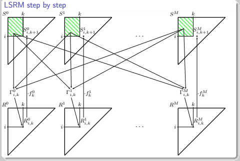
ii)$^{\text{LSRM}}$ *for all* m_1, m_2, i *and* k*, the covariance of the claim properties* $S_{i,k+1}^{m_1}$ *and* $S_{i,k+1}^{m_2}$ *under the condition of all information of their past* $\mathcal{D}_k^{i+k}$ *is proportional to* $R_{i,k}^{m_1,m_2}$*, i.e.*

$$\mathsf{Cov}\left[S_{i,k+1}^{m_1}, S_{i,k+1}^{m_2} \middle| \mathcal{D}_k^{i+k}\right] = \sigma_k^{m_1,m_2} R_{i,k}^{m_1,m_2}.$$

- We will call the parameters f_k^m and $\sigma_k^{m_1,m_2}$ development factors and covariance parameters, respectively.
- The stochastic model of LSRMs was introduced in [21]. Unfortunately, this article contains some typing errors, which make the implementation very hard. Therefore, a corrected version can be obtained by the lecturer. However, 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/.
- The choice of the exposures $R_{i,k}^m$ and $R_{i,k}^{m_1,m_2}$ 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.

LSRM step by step





LSRM step by step

$\Gamma^m_{i,k}$ denotes the linear operator that generates $R^m_{i,k}$.

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)$^{\mathsf{LSRM}}$ 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)$^{\mathsf{LSRM}}$, which means that all resulting corresponding covariance matrices have to be positive semi-definite. ☹

Lemma 4.2

Assume $S^m_{i,k}$ satisfy Assumption 4.A. Then

a)
$$\mathsf{E}\left[S^m_{i,k+1}\middle|\mathcal{D}^{i+k}\right] = \mathsf{E}\left[S^m_{i,k+1}\middle|\mathcal{D}_k\right] = \mathsf{E}\left[S^m_{i,k+1}\middle|\mathcal{D}_{i,k}\right]$$
$$= \mathsf{E}\left[S^m_{i,k+1}\middle|\mathcal{D}^{i+k}\cap\mathcal{D}_k\right] = f^m_k R^m_{i,k}.$$

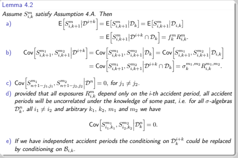
b)
$$\mathsf{Cov}\left[S^{m_1}_{i,k+1}, S^{m_2}_{i,k+1}\middle|\mathcal{D}^{i+k}\right] = \mathsf{Cov}\left[S^{m_1}_{i,k+1}, S^{m_2}_{i,k+1}\middle|\mathcal{D}_k\right] = \mathsf{Cov}\left[S^{m_1}_{i,k+1}, S^{m_2}_{i,k+1}\middle|\mathcal{D}_{i,k}\right]$$
$$= \mathsf{Cov}\left[S^{m_1}_{i,k+1}, S^{m_2}_{i,k+1}\middle|\mathcal{D}^{i+k}\cap\mathcal{D}_k\right] = \sigma^{m_1,m_2}_k R^{m_1,m_2}_{i,k}.$$

c) $\mathsf{Cov}\left[S^{m_1}_{n+1-j_1,j_1}, S^{m_2}_{n+1-j_2,j_2}\middle|\mathcal{D}^n\right] = 0$, *for* $j_1 \neq j_2$.

d) *provided that all exposures $R^m_{i,k}$ depend only on the i-th accident period, all accident periods will be uncorrelated under the knowledge of some past, i.e. for all σ-algebras $\mathcal{D}^n_k$, all $i_1 \neq i_2$ and arbitrary k_1, k_2, m_1 and m_2 we have*

$$\mathsf{Cov}\left[S^{m_1}_{i_1,k_1}, S^{m_2}_{i_2,k_2}\middle|\mathcal{D}^n_k\right] = 0.$$

e) *If we have independent accident periods the conditioning on $\mathcal{D}^{i+k}_k$ could be replaced by conditioning on $\mathcal{B}_{i,k}$.*



Lemma 4.2

Assume $S^m_{i,k}$ satisfy Assumption 4.A. Then

a)
$$\mathsf{E}\Big[S^m_{i,k+1}\Big|\mathcal{D}^{i+k}\Big] = \mathsf{E}\big[S^m_{i,k+1}\big|\mathcal{D}_k\big] = \mathsf{E}\big[S^m_{i,k+1}\big|\mathcal{D}_{i,k}\big] = \mathsf{E}\Big[S^m_{i,k+1}\Big|\mathcal{D}^{i+k}\cap\mathcal{D}_k\Big] = f^m_k R^m_{i,k}.$$

b)
$$\mathsf{Cov}\Big[S^{m_1}_{i,k+1}, S^{m_2}_{i,k+1}\Big|\mathcal{D}^{i+k}\Big] = \mathsf{Cov}\Big[S^{m_1}_{i,k+1}, S^{m_2}_{i,k+1}\Big|\mathcal{D}_k\Big] = \mathsf{Cov}\Big[S^{m_1}_{i,k+1}, S^{m_2}_{i,k+1}\Big|\mathcal{D}_{i,k}\Big] = \mathsf{Cov}\Big[S^{m_1}_{i,k+1}, S^{m_2}_{i,k+1}\Big|\mathcal{D}^{i+k}\cap\mathcal{D}_k\Big] = \sigma^{m_1,m_2}_k R^{m_1,m_2}_{i,k}.$$

c) $\mathsf{Cov}\Big[S^{m_1}_{n+1-j_1,j_1}, S^{m_2}_{n+1-j_2,j_2}\Big|\mathcal{D}^n\Big] = 0$, *for* $j_1 \neq j_2$.

d) *provided that all exposures $R^m_{i,k}$ depend only on the i-th accident period, all accident periods will be uncorrelated under the knowledge of some past, i.e. for all σ-algebras $\mathcal{D}^n_k$, all $i_1 \neq i_2$ and arbitrary k_1, k_2, m_1 and m_2 we have*

$$\mathsf{Cov}\Big[S^{m_1}_{i_1,k_1}, S^{m_2}_{i_2,k_2}\Big|\mathcal{D}^n_k\Big] = 0.$$

e) *If we have independent accident periods the conditioning on $\mathcal{D}^{i+k}_k$ could be replaced by conditioning on $\mathcal{B}_{i,k}$.*

a),b) Follows from the measurability of $R^m_{i,k}$ and $R^{m_1,m_2}_{i,k}$ with respect to $\mathcal{D}_{i,k}$.

c) Assume that $j_1 > j_2$. Then $S^{m_1}_{n+1-j_2,j_2}$ is $\mathcal{D}^n_{j_1-1}$-measurable and we get

$$\begin{aligned}
&\mathsf{Cov}\Big[S^{m_1}_{n+1-j_1,j_1}, S^{m_2}_{n+1-j_2,j_2}\Big|\mathcal{D}^n\Big]\\
&= \underbrace{\mathsf{Cov}\Big[\mathsf{E}\Big[S^{m_1}_{n+1-j_1,j_1}\Big|\mathcal{D}^n_{j_1-1}\Big], \mathsf{E}\Big[S^{m_2}_{n+1-j_2,j_2}\Big|\mathcal{D}^n_{j_1-1}\Big]\Big|\mathcal{D}^n\Big]}_{\mathsf{E}\Big[S^{m_1}_{n+1-j_1,j_1}\Big|\mathcal{D}^n_{j_1-1}\Big]=f^{m_1}_{j_1-1}R^{m_1}_{n+1-j_1,j_1-1}\ \textbf{is }\mathcal{D}^n\textbf{ measurable}} + \underbrace{\mathsf{E}\Big[\mathsf{Cov}\Big[S^{m_1}_{n+1-j_1,j_1}, S^{m_2}_{n+1-j_2,j_2}\Big|\mathcal{D}^n_{j_1-1}\Big]\Big|\mathcal{D}^n\Big]}_{S^{m_1}_{n+1-j_1,j_1}\textbf{ is }\mathcal{D}^n_{j_1-1}\textbf{-measurable}}\\
&= 0
\end{aligned}$$

d) If $S^{m_1}_{i_1,k_1}$ or $S^{m_2}_{i_2,k_2}$ is measurable with respect to $\mathcal{D}^n_k$ we are done. Otherwise, $\mathcal{D}^n_k$ is a subset of $\mathcal{D}^{i_1+k_1-1}_{k_1-1}$ and $\mathcal{D}^{i_2+k_2-1}_{k_2-1}$ and $S^{m_1}_{i_1,k_1}$ is measurable with respect to the past of $S^{m_2}_{i_2,k_2}$ or vice versa. Without loss of generality assume that $S^{m_1}_{i_1,k_1}$ is $\mathcal{D}^{i_2+k_2-1}_{k_2-1}$-measurable. Then we get

$$\begin{aligned}
\mathsf{Cov}\Big[S^{m_1}_{i_1,k_1}, S^{m_2}_{i_2,k_2}\Big|\mathcal{D}^n_k\Big] &= \mathsf{E}\Big[\mathsf{Cov}\Big[S^{m_1}_{i_1,k_1}, S^{m_2}_{i_2,k_2}\Big|\mathcal{D}^{i_2+k_2-1}_{k_2-1}\Big]\Big|\mathcal{D}^n_k\Big]\\
&\quad + \mathsf{Cov}\Big[\mathsf{E}\Big[S^{m_1}_{i_1,k_1}\Big|\mathcal{D}^{i_2+k_2-1}_{k_2-1}\Big], \mathsf{E}\Big[S^{m_2}_{i_2,k_2}\Big|\mathcal{D}^{i_2+k_2-1}_{k_2-1}\Big]\Big|\mathcal{D}^n_k\Big]\\
&= 0 + \mathsf{Cov}\Big[S^{m_1}_{i_1,k_1}, f^{m_2}_{k_2-1}R^{m_2}_{i_2,k_2-1}\Big|\mathcal{D}^n_k\Big].
\end{aligned}$$

Since $R^{m_2}_{i_2,k_2-1} \in \mathcal{B}_{i_2,k_2-1}$ and depends linearly on S it is enough to show that $S^{m_1}_{i_1,k_1}$ and $S^{m_2}_{i_2,k_2-1}$ are $\mathcal{D}^n_k$-conditional uncorrelated. Iteration until $S^{m_1}_{i_1,k_1-j}$ or $S^{m_2}_{i_2,k_2-j}$ is $\mathcal{D}^n_k$-measurable proves part d).

e) Because of independent accident periods.

Remark 4.3 (CLM as LSRM)

Because of Corollary 2.3, i.e.

$$\mathsf{E}\left[S^0_{i,k+1}\middle|\mathcal{D}^{i+k}_k\right] = (f_k - 1)\sum_{j=0}^{k} S^0_{i,j} = (f_k - 1)C_{i,k},$$

$$\mathsf{Cov}\left[S^0_{i,k+1}, S^0_{i,k+1}\middle|\mathcal{D}^{i+k}_k\right] = \sigma_k^2\sum_{j=0}^{k} S^0_{i,j} = \sigma_k^2 C_{i,k},$$

the Chain-Ladder method is a LSRM with exposures

$$R^0_{i,k} = R^{0,0}_{i,k} = C_{i,k}$$

and parameters

$$f^0_k = f_k - 1,$$
$$\sigma^{0,0}_k = \sigma^2_k.$$

Remark 4.3 (CLM as LSRM)

Because of Corollary 2.3, i.e.

$$\mathsf{E}\Big[S^0_{i,k+1}\big|\mathcal{D}^{i+k}_k\Big] = (f_k - 1)\sum_{j=0}^{k} S^0_{i,j} = (f_k - 1)C_{i,k},$$

$$\mathsf{Cov}\Big[S^0_{i,k+1}, S^0_{i,k+1}\big|\mathcal{D}^{i+k}_k\Big] = \sigma_k^2\sum_{j=0}^{k} S^0_{i,j} = \sigma_k^2 C_{i,k},$$

the Chain-Ladder method is a LSRM with exposures

$$R^0_{i,k} = R^{0,0}_{i,k} = C_{i,k}$$

and parameters

$$f^0_k = f_k - 1,$$
$$\sigma^{0,0}_k = \sigma^2_k.$$

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

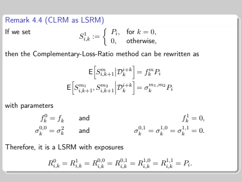
$$\mathsf{E}\left[S_{i,k+1}^{m}\middle|\mathcal{D}_k^{i+k}\right] = f_k^m P_i$$
$$\mathsf{E}\left[S_{i,k+1}^{m_1}, S_{i,k+1}^{m_2}\middle|\mathcal{D}_k^{i+k}\right] = \sigma_k^{m_1,m_2} P_i$$

with parameters

$$f_k^0 = f_k \qquad \text{and} \qquad f_k^1 = 0,$$
$$\sigma_k^{0,0} = \sigma_k^2 \qquad \text{and} \qquad \sigma_k^{0,1} = \sigma_k^{1,0} = \sigma_k^{1,1} = 0.$$

Therefore, it is a LSRM with exposures

$$R_{i,k}^0 = R_{i,k}^1 = R_{i,k}^{0,0} = R_{i,k}^{0,1} = R_{i,k}^{1,0} = R_{i,k}^{1,1} = P_i.$$



Remark 4.4 (CLRM as LSRM)

If we set
$$S^1_{i,k} := \begin{cases} P_i, & \text{for } k = 0, \\ 0, & \text{otherwise,} \end{cases}$$
then the Complementary-Loss-Ratio method can be rewritten as
$$\mathsf{E}\left[S^m_{i,k+1}\middle|\mathcal{D}^{i+k}_k\right] = f^m_k P_i$$
$$\mathsf{E}\left[S^{m_1}_{i,k+1}, S^{m_2}_{i,k+1}\middle|\mathcal{D}^{i+k}_k\right] = \sigma^{m_1,m_2}_k P_i$$
with parameters
$$f^0_k = f_k \quad \text{and} \quad f^1_k = 0,$$
$$\sigma^{0,0}_k = \sigma^2_k \quad \text{and} \quad \sigma^{0,1}_k = \sigma^{1,0}_k = \sigma^{1,1}_k = 0.$$
Therefore, it is a LSRM with exposures
$$R^0_{i,k} = R^1_{i,k} = R^{0,0}_{i,k} = R^{0,1}_{i,k} = R^{1,0}_{i,k} = R^{1,1}_{i,k} = P_i.$$

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 the Extended-Complementary-Loss-Ratio method is a LSRM with exposures

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

and parameters f_k^m and $\sigma_k^{m_1,m_2}$.

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 the Extended-Complementary-Loss-Ratio method is a LSRM with exposures

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

and parameters f_k^m and $\sigma_k^{m_1,m_2}$.

Estimator 4.7 (of the development parameter f_k^m)

Let $S_{i,k}^m$ satisfy Assumption 4.A. Then for each set of $\mathcal{D}^I \cap \mathcal{D}_k$-measurable weights $w_{i,k}^m$ with

- $w_{i,k}^m \geq 0$ *and* $R_{i,k}^m = 0$ *implies* $w_{i,k}^m = 0$,
- $\sum_{i=0}^{I-1-k} w_{i,k}^m = 1$ *if at least one* $R_{i,k}^m \neq 0$

we get that

- $$\widehat{f}_k^m := \sum_{i=0}^{I-1-k} w_{i,k}^m \frac{S_{i,k+1}^m}{R_{i,k}^m} \tag{4.1}$$

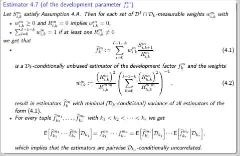
  is a $\mathcal{D}_k$-conditionally unbiased estimator of the development factor f_k^m and the weights

  $$w_{i,k}^m := \frac{\left(R_{i,k}^m\right)^2}{R_{i,k}^{m,m}} \left(\sum_{h=0}^{I-1-k} \frac{\left(R_{h,k}^m\right)^2}{R_{h,k}^{m,m}}\right)^{-1}, \tag{4.2}$$

  result in estimators $\widehat{f}_k^m$ with minimal ($\mathcal{D}_k$-conditional) variance of all estimators of the form (4.1).
- *For every tuple $\widehat{f}_{k_1}^{m_1}, \ldots, \widehat{f}_{k_r}^{m_r}$ with $k_1 < k_2 < \cdots < k_r$ we get*

  $$\mathsf{E}\left[\widehat{f}_{k_1}^{m_1} \cdots \widehat{f}_{k_r}^{m_r} \middle| \mathcal{D}_{k_1}\right] = f_{k_1}^{m_1} \cdots f_{k_r}^{m_r} = \mathsf{E}\left[\widehat{f}_{k_1}^{m_1} \middle| \mathcal{D}_{k_1}\right] \cdots \mathsf{E}\left[\widehat{f}_{k_r}^{m_r} \middle| \mathcal{D}_{k_r}\right],$$

  which implies that the estimators are pairwise $\mathcal{D}_{k_1}$-conditionally uncorrelated.



Estimator 4.7 (of the development parameter f_k^m)

Let $S_{i,k}^m$ satisfy Assumption 4.A. Then for each set of $\mathcal{D}^I \cap \mathcal{D}_k$-measurable weights $w_{i,k}^m$ with

- $w_{i,k}^m \geq 0$ *and* $R_{i,k}^m = 0$ *implies* $w_{i,k}^m = 0$,
- $\sum_{i=0}^{I-1-k} w_{i,k}^m = 1$ *if at least one* $R_{i,k}^m \neq 0$

we get that

- $$\widehat{f}_k^m := \sum_{i=0}^{I-1-k} w_{i,k}^m \frac{S_{i,k+1}^m}{R_{i,k}^m} \tag{4.1}$$

  is a $\mathcal{D}_k$-conditionally unbiased estimator of the development factor f_k^m and the weights

  $$w_{i,k}^m := \frac{\left(R_{i,k}^m\right)^2}{R_{i,k}^{m,m}} \left(\sum_{h=0}^{I-1-k} \frac{\left(R_{h,k}^m\right)^2}{R_{h,k}^{m,m}} \right)^{-1}, \tag{4.2}$$

  result in estimators $\widehat{f}_k^m$ with minimal ($\mathcal{D}_k$-conditional) variance of all estimators of the form (4.1).
- *For every tuple $\widehat{f}_{k_1}^{m_1}, \dots, \widehat{f}_{k_r}^{m_r}$ with $k_1 < k_2 < \dots < k_r$ we get*

  $$\mathsf{E}\left[\widehat{f}_{k_1}^{m_1} \cdots \widehat{f}_{k_r}^{m_r} \middle| \mathcal{D}_{k_1}\right] = f_{k_1}^{m_1} \cdots f_{k_r}^{m_r} = \mathsf{E}\left[\widehat{f}_{k_1}^{m_1} \middle| \mathcal{D}_{k_1}\right] \cdots \mathsf{E}\left[\widehat{f}_{k_r}^{m_r} \middle| \mathcal{D}_{k_r}\right],$$

  which implies that the estimators are pairwise $\mathcal{D}_{k_1}$-conditionally uncorrelated.

- **unbiased:** $$\mathsf{E}\left[\widehat{f}_k^m \middle| \mathcal{D}_k\right] = \sum_{i=0}^{I-1-k} w_{i,k}^m \frac{\mathsf{E}\left[\mathsf{E}\left[S_{i,k+1}^m \middle| \mathcal{D}_k^{i+k}\right] \middle| \mathcal{D}_k\right]}{R_{i,k}^m} = \sum_{i=0}^{I-1-k} w_{i,k}^m \frac{f_k^m R_{i,k}^m}{R_{i,k}^m} = f_k^m$$
- **minimal variance:** $$\mathsf{Var}\left[\widehat{f}_k^m\right] = \mathsf{E}\left[\mathsf{Var}\left[\widehat{f}_k^m \middle| \mathcal{D}_k\right]\right] + \mathsf{Var}\left[\mathsf{E}\left[\widehat{f}_k^m \middle| \mathcal{D}_k\right]\right] = \mathsf{E}\left[\mathsf{Var}\left[\widehat{f}_k^m \middle| \mathcal{D}_k\right]\right] + 0$$

$$\mathsf{Var}\left[\widehat{f}_k^m \middle| \mathcal{D}_k\right] = \mathsf{Var}\left[\sum_{i=0}^{I-1-k} w_{i,k}^m \frac{S_{i,k+1}^m}{R_{i,k}^m} \middle| \mathcal{D}_k\right] = \underbrace{\sum_{i=0}^{I-1-k} (w_{i,k}^m)^2 \frac{\mathsf{Var}\left[S_{i,k+1}^m \middle| \mathcal{D}_k\right]}{(R_{i,k}^m)^2}}_{\textbf{measurable with respect to } \mathcal{D}_k} = \underbrace{\sigma_k^{m,m} \sum_{i=0}^{I-1-k} (w_{i,k}^m)^2 \frac{R_{i,k}^{m,m}}{(R_{i,k}^m)^2}}_{\textbf{ii)}^{\textbf{LSRM}}}$$

Lagrange: minimize $$\textstyle\sum_{i=0}^{I-1-k} (w_{i,k}^m)^2 \frac{R_{i,k}^{m,m}}{(R_{i,k}^m)^2} + \lambda \left(1 - \sum_{i=0}^{I-1-k} w_{i,k}^m\right)$$

$$\frac{\partial}{\partial w_{i,k}^m} \bullet = 2 w_{i,k}^m \frac{R_{i,k}^{m,m}}{(R_{i,k}^m)^2} - \lambda \quad \Longrightarrow \quad w_{i,k}^m = \frac{\lambda}{2} \frac{(R_{i,k}^m)^2}{R_{i,k}^{m,m}} \quad \text{and} \quad \underbrace{\lambda = 2 \left(\sum_{i=0}^{I-1-k} \frac{(R_{i,k}^m)^2}{R_{i,k}^{m,m}} \right)^{-1}}_{\sum_{i=0}^{I-1-k} w_{i,k}^m = 1} \quad \Longrightarrow \quad \textbf{(4.2)}$$

- **uncorrelated:** $$\begin{aligned} \mathsf{E}\left[\widehat{f}_{k_1}^{m_1} \cdots \widehat{f}_{k_r}^{m_r} \middle| \mathcal{D}_{k_1}\right] &= \mathsf{E}\left[\mathsf{E}\left[\widehat{f}_{k_1}^{m_1} \cdots \widehat{f}_{k_r}^{m_n} \middle| \mathcal{D}_{k_r}\right] \middle| \mathcal{D}_{k_1}\right] \\ &= \mathsf{E}\left[\widehat{f}_{k_1}^{m_1} \cdots \widehat{f}_{k_{r-1}}^{m_{r-1}} \mathsf{E}\left[\widehat{f}_{k_r}^{m_r} \middle| \mathcal{D}_{k_r}\right] \middle| \mathcal{D}_{k_1}\right] \\ &= \mathsf{E}\left[\widehat{f}_{k_1}^{m_1} \cdots \widehat{f}_{k_{r-1}}^{m_{r-1}} \middle| \mathcal{D}_{k_1}\right] f_{k_r}^{m_r} = \dots = f_{k_1}^{m_1} \cdots f_{k_r}^{m_r} \end{aligned}$$

Definition 4.8 (Diagonal by diagonal projection)

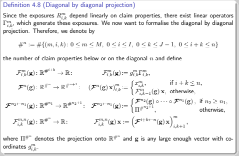
Since the exposures $R^m_{i,k}$ depend linearly on claim properties, there exist linear operators $\Gamma^m_{i,k}$, which generate these exposures. We now want to formalise the diagonal by diagonal projection. Therefore, we denote by

$$\#^n := \#\{(m,i,k)\colon 0 \le m \le M,\ 0 \le i \le I,\ 0 \le k \le J-1,\ 0 \le i+k \le n\}$$

the number of claim properties below or on the diagonal n and define

$$\begin{aligned}
\mathcal{F}^m_{i,k}(\mathbf{g})\colon \mathbb{R}^{\#^{i+k}} \to \mathbb{R}\colon \quad & \mathcal{F}^m_{i,k}(\mathbf{g}) := g^m_{i,k}\Gamma^m_{i,k},\\
\boldsymbol{\mathcal{F}}^n(\mathbf{g})\colon \mathbb{R}^{\#^n} \to \mathbb{R}^{\#^{n+1}}\colon \quad & (\boldsymbol{\mathcal{F}}^n(\mathbf{g})\,\mathbf{x})^m_{i,k} := \begin{cases} x^m_{i,k}, & \text{if } i+k \le n,\\ \mathcal{F}^m_{i,k-1}(\mathbf{g})\,\mathbf{x}, & \text{otherwise,} \end{cases}\\
\boldsymbol{\mathcal{F}}^{n_2\leftarrow n_1}(\mathbf{g})\colon \mathbb{R}^{\#^{n_1}} \to \mathbb{R}^{\#^{n_2+1}}\colon \quad & \boldsymbol{\mathcal{F}}^{n_2\leftarrow n_1}(\mathbf{g}) := \begin{cases} \boldsymbol{\mathcal{F}}^{n_2}(\mathbf{g}) \circ \cdots \circ \boldsymbol{\mathcal{F}}^{n_1}(\mathbf{g}), & \text{if } n_2 \ge n_1,\\ \Pi^{\#^{n_2+1}}, & \text{otherwise,} \end{cases}\\
\mathcal{F}^{m,n}_{i,k}(\mathbf{g})\colon \mathbb{R}^{\#^n} \to \mathbb{R}\colon \quad & \mathcal{F}^{m,n}_{i,k}(\mathbf{g})\,\mathbf{x} := \left(\boldsymbol{\mathcal{F}}^{i+k\leftarrow n}(\mathbf{g})\,\mathbf{x}\right)^m_{i,k+1},
\end{aligned}$$

where $\Pi^{\#^n}$ denotes the projection onto $\mathbb{R}^{\#^n}$ and $\mathbf{g}$ is any large enough vector with coordinates $g^m_{i,k}$.



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

$$\#^n := \#\{(m,i,k) \colon 0 \le m \le M,\ 0 \le i \le I,\ 0 \le k \le J-1,\ 0 \le i+k \le n\}$$

the number of claim properties below or on the diagonal n and define

$$\begin{aligned}
&\mathcal{F}_{i,k}^m(\mathbf{g}) \colon \mathbb{R}^{\#^{i+k}} \to \mathbb{R} \colon && \mathcal{F}_{i,k}^m(\mathbf{g}) := g_{i,k}^m \Gamma_{i,k}^m,\\
&\boldsymbol{\mathcal{F}}^n(\mathbf{g}) \colon \mathbb{R}^{\#^n} \to \mathbb{R}^{\#^{n+1}} \colon && (\boldsymbol{\mathcal{F}}^n(\mathbf{g})\,\mathbf{x})_{i,k}^m := \begin{cases} x_{i,k}^m, & \text{if } i+k \le n,\\ \mathcal{F}_{i,k-1}^m(\mathbf{g})\,\mathbf{x}, & \text{otherwise,} \end{cases}\\
&\boldsymbol{\mathcal{F}}^{n_2 \leftarrow n_1}(\mathbf{g}) \colon \mathbb{R}^{\#^{n_1}} \to \mathbb{R}^{\#^{n_2+1}} \colon && \boldsymbol{\mathcal{F}}^{n_2 \leftarrow n_1}(\mathbf{g}) := \begin{cases} \boldsymbol{\mathcal{F}}^{n_2}(\mathbf{g}) \circ \cdots \circ \boldsymbol{\mathcal{F}}^{n_1}(\mathbf{g}), & \text{if } n_2 \ge n_1,\\ \Pi^{\#^{n_2+1}}, & \text{otherwise,} \end{cases}\\
&\mathcal{F}_{i,k}^{m,n}(\mathbf{g}) \colon \mathbb{R}^{\#^n} \to \mathbb{R} \colon && \mathcal{F}_{i,k}^{m,n}(\mathbf{g})\,\mathbf{x} := \left(\boldsymbol{\mathcal{F}}^{i+k \leftarrow n}(\mathbf{g})\,\mathbf{x}\right)_{i,k+1}^m,
\end{aligned}$$

where $\Pi^{\#^n}$ denotes the projection onto $\mathbb{R}^{\#^n}$ and $\mathbf{g}$ is any large enough vector with coordinates $g_{i,k}^m$.

- Since the operators $\Gamma_{i,k}^m$ and $\mathcal{F}_{i,k}^m(\mathbf{g})$ only depend on coordinates (l,h,j) with

  $$0 \le m \le M,\ 0 \le h \le i,\ 0 \le j \le k$$

  they could be defined on a smaller domain, but than concatenation would not be possible.
- We added the parameter $\mathbf{g}$ in order to denote which development factors are used, for instance the real, but unknown, development vectors f_k^m or their estimates $\widehat{f}_k^m$.
- Often we will use parameters $\mathbf{g}$, which do not depend on the accident period i. Then we will skip the index i in $g_{i,k}^m$.

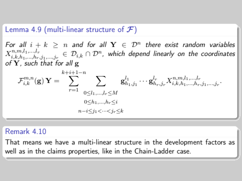
Lemma 4.9 (multi-linear structure of $\mathcal{F}$)

For all $i + k \geq n$ *and for all* $\mathbf{Y} \in \mathcal{D}^n$ *there exist random variables* $X^{n,m,l_1,\ldots,l_r}_{i,k,h_1,\ldots,h_r,j_1,\ldots,j_r} \in \mathcal{D}_{i,k} \cap \mathcal{D}^n$*, which depend linearly on the coordinates of* $\mathbf{Y}$*, such that for all* $\mathbf{g}$

$$\mathcal{F}^{m,n}_{i,k}(\mathbf{g})\,\mathbf{Y} = \sum_{r=1}^{k+i+1-n} \sum_{\substack{0\leq l_1,\ldots,l_r\leq M \\ 0\leq h_1,\ldots,h_r\leq i \\ n-i\leq j_1<\cdots<j_r\leq k}} \mathbf{g}^{l_1}_{h_1,j_1}\cdots\mathbf{g}^{l_r}_{h_r,j_r} X^{n,m,l_1,\ldots,l_r}_{i,k,h_1,\ldots,h_r,j_1,\ldots,j_r}.$$

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.



Lemma 4.9 (multi-linear structure of $\mathcal{F}$)

For all $i+k \geq n$ and for all $\mathbf{Y} \in \mathcal{D}^n$ there exist random variables $X^{n,m,l_1,\dots,l_r}_{i,k,h_1,\dots,h_r,j_1,\dots,j_r} \in \mathcal{D}_{i,k} \cap \mathcal{D}^n$, which depend linearly on the coordinates of $\mathbf{Y}$, such that for all $\mathbf{g}$

$$\mathcal{F}^{m,n}_{i,k}(\mathbf{g})\,\mathbf{Y} = \sum_{r=1}^{k+i+1-n} \sum_{\substack{0\leq l_1,\dots,l_r\leq M\\ 0\leq h_1,\dots,h_r\leq i\\ n-i\leq j_1<\dots<j_r\leq k}} \mathbf{g}^{l_1}_{h_1,j_1}\cdots\mathbf{g}^{l_r}_{h_r,j_r} X^{n,m,l_1,\dots,l_r}_{i,k,h_1,\dots,h_r,j_1,\dots,j_r}.$$

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.

If $i+k<n$ then

$$\mathcal{F}^{m,n}_{i,k}(\mathbf{g})\,\mathbf{Y} = Y_{i,k} \in \mathcal{D}_{i,k} \cap \mathcal{D}^n.$$

If $i+k=n$ we get

$$\mathcal{F}^{m,n}_{i,k}(\mathbf{g})\,\mathbf{Y} = \mathbf{g}^m_{i,k} \underbrace{\Gamma^m_{i,k}\,\mathbf{Y}}_{\in \mathcal{D}_{i,k}\cap\mathcal{D}^{i+k}},$$

because $\Gamma^m_{i,k}\mathbf{Y}$ depends only on coordinates of $\mathbf{Y}$ which are $\mathcal{D}_{i,k}\cap\mathcal{D}^{i+k}$ measurable.
Now assume that the statement is fulfilled for all n, h, j with $n \leq h+j < i+k$. Then we get

$$\mathcal{F}^{m,n}_{i,k}(\mathbf{g})\,\mathbf{Y} = \mathcal{F}^m_{i,k}(\mathbf{g}) \circ \boldsymbol{\mathcal{F}}^{i+k-1\leftarrow n}(\mathbf{g})\,\mathbf{Y} = \mathbf{g}^m_{i,k}\Gamma^m_{i,k} \circ \boldsymbol{\mathcal{F}}^{i+k-1\leftarrow n}(\mathbf{g})\,\mathbf{Y}.$$

By assumption the statement is fulfilled for each coordinate of $\left(\boldsymbol{\mathcal{F}}^{i+k-1\leftarrow n}(\mathbf{g})\,\mathbf{Y}\right)^l_{h,j}$ and since $\Gamma^m_{i,k}$ depends only on coordinates $h \leq i$ and $j \leq k$, only development factors $\mathbf{g}^l_{h,j}$ with $n-i \leq j < k$ are involved, which by induction proves our statement.

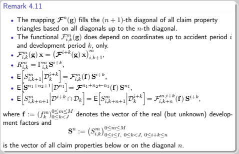
Remark 4.11

- The mapping $\boldsymbol{\mathcal{F}}^n(\mathbf{g})$ fills the $(n+1)$-th diagonal of all claim property triangles based on all diagonals up to the n-th diagonal.
- The functional $\mathcal{F}^m_{i,k}(\mathbf{g})$ does depend on coordinates up to accident period i and development period k, only.
- $\mathcal{F}^m_{i,k}(\mathbf{g})\,\mathbf{x} = \left(\boldsymbol{\mathcal{F}}^{i+k}(\mathbf{g})\,\mathbf{x}\right)^m_{i,k+1}$,
- $R^m_{i,k} = \Gamma^m_{i,k}\mathbf{S}^{i+k}$,
- $\mathsf{E}\left[S^m_{i,k+1}\middle|\mathcal{D}^{i+k}_k\right] = \mathcal{F}^m_{i,k}(\mathbf{f})\,\mathbf{S}^{i+k}$,
- $\mathsf{E}\left[\mathbf{S}^{n_1+n_2+1}\middle|\mathcal{D}^{n_1}\right] = \boldsymbol{\mathcal{F}}^{n_1+n_2\leftarrow n_1}(\mathbf{f})\,\mathbf{S}^{n_1}$,
- $\mathsf{E}\left[S^m_{i,k+n+1}\middle|\mathcal{D}^{i+k}\cap\mathcal{D}_k\right] = \mathsf{E}\left[S^m_{i,k+n+1}\middle|\mathcal{D}^{i+k}_k\right] = \mathcal{F}^{m,i+k}_{i,k+n}(\mathbf{f})\,\mathbf{S}^{i+k}$,

where $\mathbf{f} := (f^m_k)^{0\le m\le M}_{0\le k<J}$ denotes the vector of the real (but unknown) development factors and

$$\mathbf{S}^n := \left(S^m_{i,k}\right)^{0\le m\le M}_{0\le i\le I,\ 0\le k<J,\ 0\le i+k\le n}$$

is the vector of all claim properties below or on the diagonal n.



Remark 4.11

- The mapping $\boldsymbol{\mathcal{F}}^n(\mathbf{g})$ fills the $(n+1)$-th diagonal of all claim property triangles based on all diagonals up to the n-th diagonal.
- The functional $\mathcal{F}^m_{i,k}(\mathbf{g})$ does depend on coordinates up to accident period i and development period k, only.
- $\mathcal{F}^m_{i,k}(\mathbf{g})\,\mathbf{x} = \left(\boldsymbol{\mathcal{F}}^{i+k}(\mathbf{g})\,\mathbf{x}\right)^m_{i,k+1}$,
- $R^m_{i,k} = \Gamma^m_{i,k}\mathbf{S}^{i+k}$,
- $\mathsf{E}\left[S^m_{i,k+1}\middle|\mathcal{D}^{i+k}_k\right] = \mathcal{F}^m_{i,k}(\mathbf{f})\,\mathbf{S}^{i+k}$,
- $\mathsf{E}\left[\mathbf{S}^{n_1+n_2+1}\middle|\mathcal{D}^{n_1}\right] = \boldsymbol{\mathcal{F}}^{n_1+n_2\leftarrow n_1}(\mathbf{f})\,\mathbf{S}^{n_1}$,
- $\mathsf{E}\left[S^m_{i,k+n+1}\middle|\mathcal{D}^{i+k}\cap\mathcal{D}_k\right] = \mathsf{E}\left[S^m_{i,k+n+1}\middle|\mathcal{D}^{i+k}_k\right] = \mathcal{F}^{m,i+k}_{i,k+n}(\mathbf{f})\,\mathbf{S}^{i+k}$,

where $\mathbf{f} := (f^m_k)^{0\le m\le M}_{0\le k<J}$ denotes the vector of the real (but unknown) development factors and

$$\mathbf{S}^n := (S^m_{i,k})^{0\le m\le M}_{0\le i\le I,\ 0\le k<J,\ 0\le i+k\le n}$$

is the vector of all claim properties below or on the diagonal n.

$$\begin{aligned}\mathsf{E}\left[\mathbf{S}^{n_1+n_2+1}\middle|\mathcal{D}^{n_1}\right] &= \mathsf{E}\left[\mathsf{E}\left[\mathbf{S}^{n_1+n_2+1}\middle|\mathcal{D}^{n_1+n_2}\right]\middle|\mathcal{D}^{n_1}\right] = \mathsf{E}\left[\boldsymbol{\mathcal{F}}^{n_1+n_2}(\mathbf{f})\,\mathbf{S}^{n_1+n_2}\middle|\mathcal{D}^{n_1}\right]\\ &= \boldsymbol{\mathcal{F}}^{n_1+n_2}(\mathbf{f})\,\mathsf{E}\left[\mathbf{S}^{n_1+n_2}\middle|\mathcal{D}^{n_1}\right] = \ldots = \boldsymbol{\mathcal{F}}^{n_1+n_2\leftarrow n_1}(\mathbf{f})\,\mathbf{S}^{n_1}\end{aligned}$$

$$\mathsf{E}\left[S^m_{i,k+n+1}\middle|\mathcal{D}^{i+k}_k\right] = \left(\boldsymbol{\mathcal{F}}^{i+k+n\leftarrow i+k}(\mathbf{f})\,\mathbf{S}^{i+k}\right)^m_{i,k+n+1} = \mathcal{F}^{m,i+k}_{i,k+n}(\mathbf{f})\,\mathbf{S}^{i+k}$$

Estimator 4.12 (of the future development)

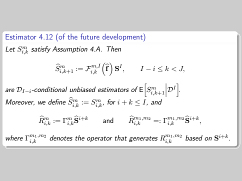
Let $S_{i,k}^{m}$ satisfy Assumption 4.A. Then

$$\widehat{S}_{i,k+1}^{m} := \mathcal{F}_{i,k}^{m,I}\left(\widehat{\mathbf{f}}\right)\mathbf{S}^{I}, \qquad I - i \leq k < J,$$

are $\mathcal{D}_{I-i}$-conditional unbiased estimators of $\mathsf{E}\left[S_{i,k+1}^{m}\middle|\mathcal{D}^{I}\right]$.
Moreover, we define $\widehat{S}_{i,k}^{m} := S_{i,k}^{m}$, for $i + k \leq I$, and

$$\widehat{R}_{i,k}^{m} := \Gamma_{i,k}^{m}\widehat{\mathbf{S}}^{i+k} \qquad \text{and} \qquad \widehat{R}_{i,k}^{m_1,m_2} =: \Gamma_{i,k}^{m_1,m_2}\widehat{\mathbf{S}}^{i+k},$$

where $\Gamma_{i,k}^{m_1,m_2}$ denotes the operator that generates $R_{i,k}^{m_1,m_2}$ based on $\mathbf{S}^{i+k}$.



Estimator 4.12 (of the future development)

Let $S_{i,k}^m$ *satisfy Assumption 4.A. Then*

$$\widehat{S}_{i,k+1}^m := \mathcal{F}_{i,k}^{m,I}\left(\widehat{\mathbf{f}}\right)\mathbf{S}^I, \qquad I-i \leq k < J,$$

are $\mathcal{D}_{I-i}$*-conditional unbiased estimators of* $\mathsf{E}\left[S_{i,k+1}^m\middle|\mathcal{D}^I\right]$.
Moreover, we define $\widehat{S}_{i,k}^m := S_{i,k}^m$*, for* $i+k \leq I$*, and*

$$\widehat{R}_{i,k}^m := \Gamma_{i,k}^m \widehat{\mathbf{S}}^{i+k} \qquad \text{and} \qquad \widehat{R}_{i,k}^{m_1,m_2} =: \Gamma_{i,k}^{m_1,m_2}\widehat{\mathbf{S}}^{i+k},$$

where $\Gamma_{i,k}^{m_1,m_2}$ *denotes the operator that generates* $R_{i,k}^{m_1,m_2}$ *based on* $\mathbf{S}^{i+k}$.

We will even prove that $\widehat{S}_{i,k+1}^m$ is an $\mathcal{D}_{I-h}$-conditionally unbiased estimator of $\mathsf{E}\left[S_{i,k+1}^m\middle|\mathcal{D}^I\right]$ for all $h \geq i$. We will do that by induction. If $i+k=I$ we get for all $h \geq i$

$$\begin{aligned}\mathsf{E}\left[\widehat{S}_{i,k+1}^m\middle|\mathcal{D}_{I-h}\right] &= \mathsf{E}\left[\mathsf{E}\left[\widehat{S}_{i,k+1}^m\middle|\mathcal{D}_k\right]\middle|\mathcal{D}_{I-h}\right] = \mathsf{E}\left[\mathsf{E}\left[\widehat{f}_k^m R_{i,k}^m\middle|\mathcal{D}_k\right]\middle|\mathcal{D}_{I-h}\right] = \mathsf{E}\left[\mathsf{E}\left[\widehat{f}_k^m\middle|\mathcal{D}_k\right]R_{i,k}^m\middle|\mathcal{D}_{I-h}\right]\\ &= \underbrace{\mathsf{E}\left[f_k^m R_{i,k}^m\middle|\mathcal{D}_{I-h}\right] = \mathsf{E}\left[\mathcal{F}_{i,k}^m(\mathbf{f})\,\mathbf{S}^I\middle|\mathcal{D}_{I-h}\right]}_{\textbf{Estimator 4.7}} = \underbrace{\mathsf{E}\left[\mathsf{E}\left[S_{i,k+1}^m\middle|\mathcal{D}^I\right]\middle|\mathcal{D}_{I-h}\right]}_{\textbf{Remark 4.11}}.\end{aligned}$$

Now assume that the statement is fulfilled for all $i+k<n$. Then we get for $i+k=n$ and all $h \geq i$

$$\begin{aligned}\mathsf{E}\left[\widehat{S}_{i,k+1}^m\middle|\mathcal{D}_{I-h}\right] &= \mathsf{E}\left[\mathcal{F}_{i,k}^m\left(\widehat{\mathbf{f}}\right)\widehat{\mathbf{S}}^{i+k}\middle|\mathcal{D}_{I-h}\right] = \mathsf{E}\left[\mathsf{E}\left[\mathcal{F}_{i,k}^m\left(\widehat{\mathbf{f}}\right)\middle|\mathcal{D}_k\right]\widehat{\mathbf{S}}^{i+k}\middle|\mathcal{D}_{I-h}\right]\\ &= \mathsf{E}\left[\mathcal{F}_{i,k}^m(\mathbf{f})\,\widehat{\mathbf{S}}^{i+k}\middle|\mathcal{D}_{I-h}\right] = \mathcal{F}_{i,k}^m(\mathbf{f})\,\mathsf{E}\left[\widehat{\mathbf{S}}^{i+k}\middle|\mathcal{D}_{I-h}\right].\end{aligned}$$

Since $\mathcal{F}_{i,k}^m(\mathbf{f})$ depends only on accident periods $h_1 \leq i$, all coordinates $\mathsf{E}\left[\widehat{S}_{h_1,j}^l\middle|\mathcal{D}_{I-h}\right]$ of $\mathsf{E}\left[\widehat{\mathbf{S}}^{i+k}\middle|\mathcal{D}_{I-h}\right]$ with $h_1>i$ will not be taken into account. For all others we can apply the induction hypotheses and proceed with

$$= \underbrace{\mathcal{F}_{i,k}^m(\mathbf{f})\,\mathsf{E}\left[\mathbf{S}^{i+k}\middle|\mathcal{D}_{I-h}\right]}_{\textbf{induction hypothesis}} = \underbrace{\mathsf{E}\left[\mathsf{E}\left[S_{i,k+1}^m\middle|\mathcal{D}_k^I\right]\middle|\mathcal{D}_{I-h}\right]}_{\textbf{Remark 4.11}} = \mathsf{E}\left[\mathsf{E}\left[S_{i,k+1}^m\middle|\mathcal{D}^I\right]\middle|\mathcal{D}_{I-h}\right].$$

□

Note, since $R_{i,k}^{m_1,m_2}$ is $\mathcal{D}^{i+k}$ measureable, there always extists an operator $\Gamma_{i,k}^{m_1,m_2}$ such that $R_{i,k}^{m_1,m_2} = \Gamma_{i,k}^{m_1,m_2}\mathbf{S}^{i+k}$.

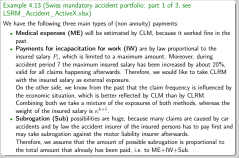
Example 4.13 (Swiss mandatory accident portfolio: part 1 of 3, see LSRM_Accident_ActiveX.xlsx)

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 κ^{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.



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 κ^{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.

Mathematical that means:

We have four claim properties with exposures

ME: $R^0_{i,k} = R^{0,0}_{i,k} = \sum_{j=0}^{k} S^0_{i,j}$

IW: $R^1_{i,k} = R^{1,1}_{i,k} = \sum_{j=0}^{k} \left(\kappa^{j+1} S^3_{i,j} + (1-\kappa^{j+1}) S^1_{i,j}\right)$

Sub: $R^2_{i,k} = R^{2,2}_{i,k} = \sum_{j=0}^{k} \left(S^0_{i,j} + S^1_{i,j} + S^2_{i,j}\right)$

Salary: $S^3_{i,0} = P_i$, $S^3_{i,j} = 0$, for $j > 0$, and

$$R^3_{i,k} = R^{3,0}_{i,k} = R^{0,3}_{i,k} = R^{3,1}_{i,k} = R^{1,3}_{i,k} = R^{3,2}_{i,k} = R^{2,3}_{i,k} = R^{3,3}_{i,k} = 0$$

For the not yet defined exposures we take the total payments up to now, i.e.

$$R^{0,1}_{i,k} = R^{1,0}_{i,k} = R^{0,2}_{i,k} = R^{2,0}_{i,k} = R^{1,2}_{i,k} = R^{2,1}_{i,k} = \sum_{j=0}^{k} \left(S^0_{i,j} + S^1_{i,j} + S^2_{i,j}\right).$$

Resulting Best Estimate reserves

- Depend almost linear on κ, 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 (small circle on the left), 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.

Mathematical that means:

We have four claim properties with exposures

ME: $R^0_{i,k} = R^{0,0}_{i,k} = \sum_{j=0}^k S^0_{i,j}$

IW: $R^1_{i,k} = R^{1,1}_{i,k} = \sum_{j=0}^k \left(\kappa^{j+1} S^3_{i,j} + (1-\kappa^{j+1}) S^1_{i,j}\right)$

Sub: $R^2_{i,k} = R^{2,2}_{i,k} = \sum_{j=0}^k \left(S^0_{i,j} + S^1_{i,j} + S^2_{i,j}\right)$

Salary: $S^3_{i,0} = P_i$, $S^3_{i,j} = 0$, for $j > 0$, and

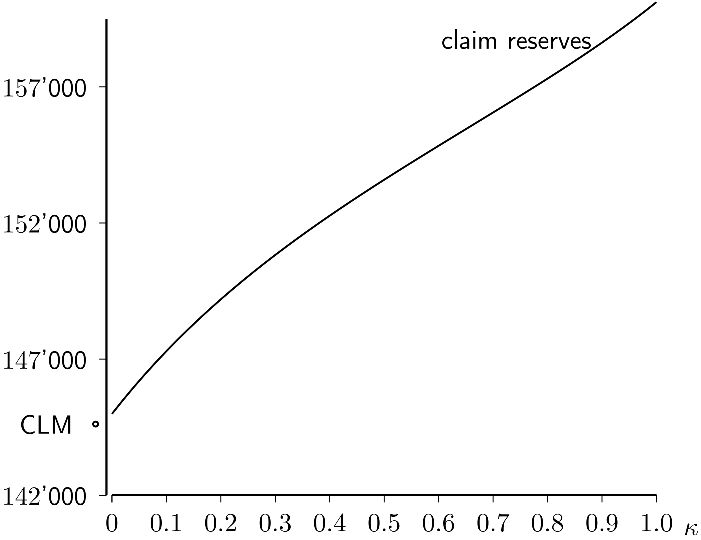
$$R^3_{i,k} = R^{3,0}_{i,k} = R^{0,3}_{i,k} = R^{3,1}_{i,k} = R^{1,3}_{i,k} = R^{3,2}_{i,k} = R^{2,3}_{i,k} = R^{3,3}_{i,k} = 0$$

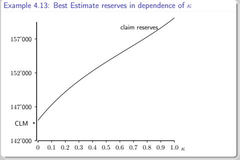
For the not yet defined exposures we take the total payments up to now, i.e. $R^{0,1}_{i,k} = R^{1,0}_{i,k} = R^{0,2}_{i,k} = R^{2,0}_{i,k} = R^{1,2}_{i,k} = R^{2,1}_{i,k} = \sum_{j=0}^k \left(S^0_{i,j} + S^1_{i,j} + S^2_{i,j}\right)$.

Resulting Best Estimate reserves

- Depend almost linear on κ, 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 (small circle on the left), 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.

Example 4.13: Best Estimate reserves in dependence of κ





Example 4.13: Best Estimate reserves in dependence of κ

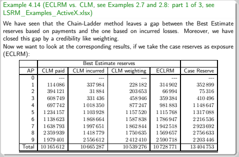
- The estimated covariance parameters $\widehat{\sigma}_k^{m_1,m_2}$ together with the estimated exposures $\widehat{R}_{i,k}^{m_1,m_2}$ lead to covariance matrices which are slightly non-positive definite for development periods $k \in \{5,6,7\}$. Since the corresponding negative eigenvalues are almost 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):

	Best Estimate reserves				
AP	CLM paid	CLM incurred	CLM weighting	ECLRM	Case Reserve
0	---	---	---	---	---
1	114 086	337 984	228 182	314 902	352 899
2	394 121	31 884	203 653	66 994	75 316
3	608 749	331 436	458 946	359 384	410 496
4	697 742	1 018 350	877 247	981 883	1 148 647
5	1 234 157	1 103 928	1 157 520	1 115 768	1 317 088
6	1 138 623	1 868 664	1 587 838	1 786 947	2 216 536
7	1 638 793	1 997 651	1 862 844	1 942 518	2 923 692
8	2 359 939	1 418 779	1 750 635	1 569 657	2 756 633
9	1 979 401	2 556 612	2 412 410	2 590 718	2 203 446
Total	10 165 612	10 665 287	10 539 276	10 728 771	13 404 753



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):

	Best Estimate reserves				
AP	CLM paid	CLM incurred	CLM weighting	ECLRM	Case Reserve
0	---	---	---	---	---
1	114 086	337 984	228 182	314 902	352 899
2	394 121	31 884	203 653	66 994	75 316
3	608 749	331 436	458 946	359 384	410 496
4	697 742	1 018 350	877 247	981 883	1 148 647
5	1 234 157	1 103 928	1 157 520	1 115 768	1 317 088
6	1 138 623	1 868 664	1 587 838	1 786 947	2 216 536
7	1 638 793	1 997 651	1 862 844	1 942 518	2 923 692
8	2 359 939	1 418 779	1 750 635	1 569 657	2 756 633
9	1 979 401	2 556 612	2 412 410	2 590 718	2 203 446
Total	10 165 612	10 665 287	10 539 276	10 728 771	13 404 753

- 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 quit different.

Stochastic Reserving

Lecture 7 (Continuation of Lecture 6)

Linear-Stochastic-Reserving methods

René Dahms

ETH Zurich, Spring 2021

14 April 2021

(Last update: 26 April 2021)

Stochastic Reserving
Lecture 7 (Continuation of Lecture 6)
Linear-Stochastic-Reserving methods

René Dahms

ETH Zurich, Spring 2021

14 April 2021
(Last update: 26 April 2021)

4 Linear-Stochastic-Reserving methods

2021-04-26

Stochastic Reserving
- Linear-Stochastic-Reserving methods
  - Future development
    - Lecture 7: Table of contents

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.

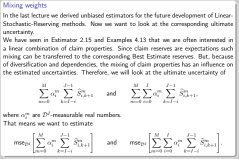
We have seen in Estimator 2.15 and Examples 4.13 that we are often interested in a linear combination of claim properties. Since claim reserves are expectations such mixing can be transferred to the corresponding Best Estimate reserves. But, because of diversification and dependencies, the mixing of claim properties has an influence on the estimated uncertainties. Therefore, we will look at the ultimate uncertainty of

$$\sum_{m=0}^{M}\alpha_i^m\sum_{k=I-i}^{J-1}\widehat{S}_{i,k+1}^m \qquad \text{and} \qquad \sum_{m=0}^{M}\sum_{i=0}^{I}\alpha_i^m\sum_{k=I-i}^{J-1}\widehat{S}_{i,k+1}^m,$$

where α_i^m are $\mathcal{D}^I$-measurable real numbers.

That means we want to estimate

$$\mathrm{mse}_{\mathcal{D}^I}\left[\sum_{m=0}^{M}\alpha_i^m\sum_{k=I-i}^{J-1}\widehat{S}_{i,k+1}^m\right] \qquad \text{and} \qquad \mathrm{mse}_{\mathcal{D}^I}\left[\sum_{m=0}^{M}\sum_{i=0}^{I}\alpha_i^m\sum_{k=I-i}^{J-1}\widehat{S}_{i,k+1}^m\right].$$



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

$$\sum_{m=0}^{M} \alpha_i^m \sum_{k=I-i}^{J-1} \widehat{S}_{i,k+1}^m \quad \text{and} \quad \sum_{m=0}^{M} \sum_{i=0}^{I} \alpha_i^m \sum_{k=I-i}^{J-1} \widehat{S}_{i,k+1}^m,$$

where α_i^m are $\mathcal{D}^I$-measurable real numbers.

That means we want to estimate

$$\mathsf{mse}_{\mathcal{D}^I}\left[\sum_{m=0}^{M} \alpha_i^m \sum_{k=I-i}^{J-1} \widehat{S}_{i,k+1}^m\right] \quad \text{and} \quad \mathsf{mse}_{\mathcal{D}^I}\left[\sum_{m=0}^{M} \sum_{i=0}^{I} \alpha_i^m \sum_{k=I-i}^{J-1} \widehat{S}_{i,k+1}^m\right].$$

Exapmles for mixing:

- Combination of two portfolios (weights are equal to one).
- Combination of two Chain-Ladder projections, one for payments and one for incurred losses (weights sum up to one for each i).
- Adding dependent payments, for instance subrogation and normal payments, which are projected separately (weights are equal to one).

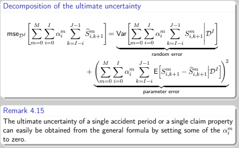
Decomposition of the ultimate uncertainty

$$\mathsf{mse}_{\mathcal{D}^I}\left[\sum_{m=0}^{M}\sum_{i=0}^{I}\alpha_i^m\sum_{k=I-i}^{J-1}\widehat{S}_{i,k+1}^m\right]=\underbrace{\mathsf{Var}\left[\sum_{m=0}^{M}\sum_{i=0}^{I}\alpha_i^m\sum_{k=I-i}^{J-1}S_{i,k+1}^m\middle|\mathcal{D}^I\right]}_{\text{random error}}$$

$$+\underbrace{\left(\sum_{m=0}^{M}\sum_{i=0}^{I}\alpha_i^m\sum_{k=I-i}^{J-1}\mathsf{E}\left[S_{i,k+1}^m-\widehat{S}_{i,k+1}^m\middle|\mathcal{D}^I\right]\right)^2}_{\text{parameter error}}$$

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 α_i^m to zero.



Decomposition of the ultimate uncertainty

$$\mathsf{mse}_{\mathcal{D}^I}\left[\sum_{m=0}^{M}\sum_{i=0}^{I}\alpha_i^m\sum_{k=I-i}^{J-1}\widehat{S}_{i,k+1}^m\right] = \underbrace{\mathsf{Var}\left[\sum_{m=0}^{M}\sum_{i=0}^{I}\alpha_i^m\sum_{k=I-i}^{J-1}S_{i,k+1}^m\middle|\mathcal{D}^I\right]}_{\text{random error}}$$
$$+\underbrace{\left(\sum_{m=0}^{M}\sum_{i=0}^{I}\alpha_i^m\sum_{k=I-i}^{J-1}\mathsf{E}\left[S_{i,k+1}^m-\widehat{S}_{i,k+1}^m\middle|\mathcal{D}^I\right]\right)^2}_{\text{parameter error}}$$

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 α_i^m to zero.

Taylor approximation

Like in the Chain-Ladder case we will look at the functional

$$U(\mathbf{g})\mathbf{x} := \sum_{m=0}^{M} \sum_{i=0}^{I} \alpha_i^m \left(\sum_{k=0}^{I-i} x_{i,k}^m + \sum_{k=I-i}^{J-1} \mathcal{F}_{i,k}^{m,I}(\mathbf{g})\,\mathbf{x} \right).$$

Then we get:

$$\partial_{h,j}^l U(\bar{\mathbf{g}})\mathbf{x} := \left.\frac{\partial U(\mathbf{g})\mathbf{x}}{\partial g_{h,j}^l}\right|_{\mathbf{g}=\bar{\mathbf{g}}} = \frac{U(\bar{\mathbf{g}})\mathbf{x} - U\Big(\bar{\mathbf{g}}_{h,j}^l|_0\Big)\mathbf{x}}{\bar{g}_{h,j}^l} = U\Big(\bar{\mathbf{g}}_{h,j}^l|_1\Big)\mathbf{x} - U\Big(\bar{\mathbf{g}}_{h,j}^l|_0\Big)\mathbf{x},$$

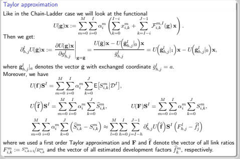
where $\bar{\mathbf{g}}_{h,j}^l|_a$ denotes the vector $\bar{\mathbf{g}}$ with exchanged coordinate $\bar{g}_{h,j}^l = a$.
Moreover, we have

$$U(\mathbf{f})\mathbf{S}^I = \sum_{m=0}^{M} \sum_{i=0}^{I} \alpha_i^m \sum_{k=0}^{J} \mathsf{E}\left[S_{i,k}^m \middle| \mathcal{D}^I\right],$$

$$U\Big(\widehat{\mathbf{f}}\Big)\mathbf{S}^I = \sum_{m=0}^{M} \sum_{i=0}^{I} \alpha_i^m \sum_{k=0}^{J} \widehat{S}_{i,k}^m, \qquad U(\mathbf{F})\mathbf{S}^I = \sum_{m=0}^{M} \sum_{i=0}^{I} \alpha_i^m \sum_{k=0}^{J} S_{i,k}^m,$$

$$\sum_{m=0}^{M} \sum_{i=0}^{I} \alpha_i^m \sum_{k=0}^{J} \left(\widehat{S}_{i,k}^m - S_{i,k}^m\right) \approx \sum_{l=0}^{M} \sum_{h=0}^{I} \sum_{j=I-h}^{J-1} \partial_{h,j}^l U\Big(\widehat{\mathbf{f}}\Big)\mathbf{S}^I \left(F_{h,j}^l - \widehat{f}_j^l\right)$$

where we used a first order Taylor approximation and $\mathbf{F}$ and $\widehat{\mathbf{f}}$ 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 $\widehat{f}_k^m$, respectively.



Taylor approximation

Like in the Chain-Ladder case we will look at the functional

$$U(\mathbf{g})\mathbf{x} := \sum_{m=0}^{M}\sum_{i=0}^{I}\alpha_i^m\left(\sum_{k=0}^{I-i}x_{i,k}^m + \sum_{k=I-i}^{J-1}\mathcal{F}_{i,k}^{m,I}(\mathbf{g})\,\mathbf{x}\right).$$

Then we get:

$$\partial_{h,j}^l U(\bar{\mathbf{g}})\mathbf{x} := \left.\frac{\partial U(\mathbf{g})\mathbf{x}}{\partial g_{h,j}^l}\right|_{\mathbf{g}=\bar{\mathbf{g}}} = \frac{U(\bar{\mathbf{g}})\mathbf{x} - U\left(\bar{\mathbf{g}}_{h,j}^l|_0\right)\mathbf{x}}{\bar{g}_{h,j}^l} = U\left(\bar{\mathbf{g}}_{h,j}^l|_1\right)\mathbf{x} - U\left(\bar{\mathbf{g}}_{h,j}^l|_0\right)\mathbf{x},$$

where $\bar{\mathbf{g}}_{h,j}^l|_a$ denotes the vector $\bar{\mathbf{g}}$ with exchanged coordinate $\bar{g}_{h,j}^l = a$.
Moreover, we have

$$U(\mathbf{f})\mathbf{S}^I = \sum_{m=0}^{M}\sum_{i=0}^{I}\alpha_i^m\sum_{k=0}^{J}\mathsf{E}\left[S_{i,k}^m|\mathcal{D}^I\right],$$

$$U\left(\widehat{\mathbf{f}}\right)\mathbf{S}^I = \sum_{m=0}^{M}\sum_{i=0}^{I}\alpha_i^m\sum_{k=0}^{J}\widehat{S}_{i,k}^m, \qquad U(\mathbf{F})\mathbf{S}^I = \sum_{m=0}^{M}\sum_{i=0}^{I}\alpha_i^m\sum_{k=0}^{J}S_{i,k}^m,$$

$$\sum_{m=0}^{M}\sum_{i=0}^{I}\alpha_i^m\sum_{k=0}^{J}\left(\widehat{S}_{i,k}^m - S_{i,k}^m\right) \approx \sum_{l=0}^{M}\sum_{h=0}^{I}\sum_{j=I-h}^{J-1}\partial_{h,j}^l U\left(\widehat{\mathbf{f}}\right)\mathbf{S}^I\left(F_{h,j}^l - \widehat{f}_j^l\right)$$

where we used a first order Taylor approximation and $\mathbf{F}$ and $\widehat{\mathbf{f}}$ 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 $\widehat{f}_k^m$, respectively.

Because of Lemma 4.9 $U(\mathbf{g})$ is an afine operator in each coordinate $g_{i,k}^m$ of $\mathbf{g}$. 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

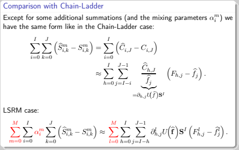
Except for some additional summations (and the mixing parameters α_i^m) we have the same form like in the Chain-Ladder case:

$$\sum_{i=0}^{I}\sum_{k=0}^{J}\left(\widehat{S}_{i,k}^{m}-S_{i,k}^{m}\right)=\sum_{i=0}^{I}\left(\widehat{C}_{i,J}-C_{i,J}\right)$$

$$\approx\sum_{h=0}^{I}\sum_{j=I-i}^{J-1}\underbrace{\frac{\widehat{C}_{h,J}}{\widehat{f}_j}}_{=\partial_{h,j}U\left(\widehat{f}\right)\mathbf{S}^I}\left(F_{h,j}-\widehat{f}_j\right).$$

LSRM case:

$$\sum_{m=0}^{M}\sum_{i=0}^{I}\alpha_i^m\sum_{k=0}^{J}\left(\widehat{S}_{i,k}^{m}-S_{i,k}^{m}\right)\approx\sum_{l=0}^{M}\sum_{h=0}^{I}\sum_{j=I-h}^{J-1}\partial_{h,j}^{l}U\left(\widehat{\mathbf{f}}\right)\mathbf{S}^I\left(F_{h,j}^{l}-\widehat{f}_j^{l}\right).$$



Comparison with Chain-Ladder

Except for some additional summations (and the mixing parameters α_i^m) we have the same form like in the Chain-Ladder case:

$$\begin{aligned}\sum_{i=0}^{I}\sum_{k=0}^{J}\left(\widehat{S}_{i,k}^{m}-S_{i,k}^{m}\right)&=\sum_{i=0}^{I}\left(\widehat{C}_{i,J}-C_{i,J}\right)\\&\approx\sum_{h=0}^{I}\sum_{j=I-i}^{J-1}\underbrace{\frac{\widehat{C}_{h,J}}{\widehat{f}_j}}_{=\partial_{h,j}U(\widehat{f})\mathbf{S}^I}\left(F_{h,j}-\widehat{f}_j\right).\end{aligned}$$

LSRM case:

$$\sum_{m=0}^{M}\sum_{i=0}^{I}\alpha_i^m\sum_{k=0}^{J}\left(\widehat{S}_{i,k}^{m}-S_{i,k}^{m}\right)\approx\sum_{l=0}^{M}\sum_{h=0}^{I}\sum_{j=I-h}^{J-1}\partial_{h,j}^{l}U\left(\widehat{\mathbf{f}}\right)\mathbf{S}^I\left(F_{h,j}^{l}-\widehat{f}_j^{l}\right).$$

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 $U\left(\widehat{\mathbf{f}}_j|_0\right)=0$.

Preparation for the derivation of the ultimate uncertainty

Like in the Chain-Ladder case we need some expectations and covariances of $\widehat{f}_k^m$ and $F_{i,k}^m$:

$$\mathsf{E}\left[F_{i,k}^m \middle| \mathcal{D}^I\right] = \mathsf{E}\left[\widehat{f}_k^m \middle| \mathcal{D}_k\right] = f_k^m \qquad i+k \geq I$$

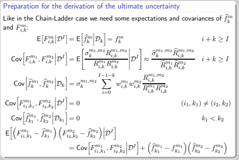
$$\mathsf{Cov}\left[F_{i,k}^{m_1}, F_{i,k}^{m_2} \middle| \mathcal{D}^I\right] = \mathsf{E}\left[\frac{\sigma_k^{m_1,m_2} R_{i,k}^{m_1,m_2}}{R_{i,k}^{m_1} R_{i,k}^{m_2}} \middle| \mathcal{D}^I\right] \approx \frac{\widehat{\sigma}_k^{m_1,m_2} \widehat{R}_{i,k}^{m_1,m_2}}{\widehat{R}_{i,k}^{m_1} \widehat{R}_{i,k}^{m_2}} \qquad i+k \geq I$$

$$\mathsf{Cov}\left[\widehat{f}_k^{m_1}, \widehat{f}_k^{m_2} \middle| \mathcal{D}_k\right] = \sigma_k^{m_1,m_2} \sum_{i=0}^{I-1-k} w_{i,k}^{m_1} w_{i,k}^{m_2} \frac{R_{i,k}^{m_1,m_2}}{R_{i,k}^{m_1} R_{i,k}^{m_2}}$$

$$\mathsf{Cov}\left[F_{i_1,k_1}^{m_1}, F_{i_2,k_2}^{m_2} \middle| \mathcal{D}^I\right] = 0 \qquad (i_1,k_1) \neq (i_2,k_2)$$

$$\mathsf{Cov}\left[\widehat{f}_{k_1}^{m_1}, \widehat{f}_{k_2}^{m_2} \middle| \mathcal{D}_{k_1}\right] = 0 \qquad k_1 < k_2$$

$$\mathsf{E}\left[\left(F_{i_1,k_1}^{m_1} - \widehat{f}_{k_1}^{m_1}\right)\left(F_{i_2,k_2}^{m_2} - \widehat{f}_{k_2}^{m_2}\right) \middle| \mathcal{D}^I\right]$$
$$= \mathsf{Cov}\left[F_{i_1,k_1}^{m_1}, F_{i_2,k_2}^{m_2} \middle| \mathcal{D}^I\right] + \left(\widehat{f}_{k_1}^{m_1} - f_{k_1}^{m_1}\right)\left(\widehat{f}_{k_2}^{m_2} - f_{k_2}^{m_2}\right)$$



Preparation for the derivation of the ultimate uncertainty

Like in the Chain-Ladder case we need some expectations and covariances of $\widehat{f}_k^m$ and $F_{i,k}^m$:

$$\mathsf{E}\left[F_{i,k}^m\middle|\mathcal{D}^I\right] = \mathsf{E}\left[\widehat{f}_k^m\middle|\mathcal{D}_k\right] = f_k^m \qquad i+k \geq I$$

$$\mathsf{Cov}\left[F_{i,k}^{m_1}, F_{i,k}^{m_2}\middle|\mathcal{D}^I\right] = \mathsf{E}\left[\frac{\sigma_k^{m_1,m_2} R_{i,k}^{m_1,m_2}}{R_{i,k}^{m_1} R_{i,k}^{m_2}}\middle|\mathcal{D}^I\right] \approx \frac{\widehat{\sigma}_k^{m_1,m_2} \widehat{R}_{i,k}^{m_1,m_2}}{\widehat{R}_{i,k}^{m_1} \widehat{R}_{i,k}^{m_2}} \qquad i+k \geq I$$

$$\mathsf{Cov}\left[\widehat{f}_k^{m_1}, \widehat{f}_k^{m_2}\middle|\mathcal{D}_k\right] = \sigma_k^{m_1,m_2} \sum_{i=0}^{I-1-k} w_{i,k}^{m_1} w_{i,k}^{m_2} \frac{R_{i,k}^{m_1,m_2}}{R_{i,k}^{m_1} R_{i,k}^{m_2}}$$

$$\mathsf{Cov}\left[F_{i_1,k_1}^{m_1}, F_{i_2,k_2}^{m_2}\middle|\mathcal{D}^I\right] = 0 \qquad (i_1,k_1) \neq (i_2,k_2)$$

$$\mathsf{Cov}\left[\widehat{f}_{k_1}^{m_1}, \widehat{f}_{k_2}^{m_2}\middle|\mathcal{D}_{k_1}\right] = 0 \qquad k_1 < k_2$$

$$\mathsf{E}\left[\left(F_{i_1,k_1}^{m_1} - \widehat{f}_{k_1}^{m_1}\right)\left(F_{i_2,k_2}^{m_2} - \widehat{f}_{k_2}^{m_2}\right)\middle|\mathcal{D}^I\right] = \mathsf{Cov}\left[F_{i_1,k_1}^{m_1}, F_{i_2,k_2}^{m_2}\middle|\mathcal{D}^I\right] + \left(\widehat{f}_{k_1}^{m_1} - f_{k_1}^{m_1}\right)\left(\widehat{f}_{k_2}^{m_2} - f_{k_2}^{m_2}\right)$$

$$\mathsf{E}\left[F_{i,k}^m\middle|\mathcal{D}^I\right] = \mathsf{E}\left[\mathsf{E}\left[F_{i,k}^m\middle|\mathcal{D}_k^{i+k}\right]\middle|\mathcal{D}^I\right] = \mathsf{E}\left[f_k^m\middle|\mathcal{D}^I\right] = f_k^m = \mathsf{E}\left[\widehat{f}_k^m\middle|\mathcal{D}_k\right] \qquad \text{because } \widehat{f}_k^m \text{ is } \mathcal{D}_k\text{-unbiased}$$

$$\mathsf{Cov}\left[F_{i,k}^{m_1}, F_{i,k}^{m_2}\middle|\mathcal{D}^I\right] = \mathsf{Cov}\left[\mathsf{E}\left[F_{i,k}^{m_1}\middle|\mathcal{D}_k^{i+k}\right], \mathsf{E}\left[F_{i,k}^{m_2}\middle|\mathcal{D}_k^{i+k}\right]\middle|\mathcal{D}^I\right] + \mathsf{E}\left[\mathsf{Cov}\left[F_{i,k}^{m_1}, F_{i,k}^{m_2}\middle|\mathcal{D}_k^{i+k}\right]\middle|\mathcal{D}^I\right]$$

$$= \mathsf{Cov}\left[f_k^{m_1}, f_k^{m_2}\middle|\mathcal{D}^I\right] + \mathsf{E}\left[\frac{\sigma_k^{m_1,m_2} R_{i,k}^{m_1,m_2}}{R_{i,k}^{m_1} R_{i,k}^{m_2}}\middle|\mathcal{D}^I\right] = 0 + \mathsf{E}\left[\frac{\sigma_k^{m_1,m_2} R_{i,k}^{m_1,m_2}}{R_{i,k}^{m_1} R_{i,k}^{m_2}}\middle|\mathcal{D}^I\right] \approx \frac{\widehat{\sigma}_k^{m_1,m_2} \widehat{R}_{i,k}^{m_1,m_2}}{\widehat{R}_{i,k}^{m_1} \widehat{R}_{i,k}^{m_2}}$$

$$\mathsf{Cov}\left[\widehat{f}_{k_1}^{m_1}, \widehat{f}_{k_2}^{m_2}\middle|\mathcal{D}_{k_1}\right] = \mathsf{Cov}\left[\mathsf{E}\left[\widehat{f}_{k_1}^{m_1}\middle|\mathcal{D}_{k_2}\right], \mathsf{E}\left[\widehat{f}_{k_2}^{m_2}\middle|\mathcal{D}_{k_2}\right]\middle|\mathcal{D}_{k_1}\right] + \mathsf{E}\left[\mathsf{Cov}\left[\widehat{f}_{k_1}^{m_1}, \widehat{f}_{k_2}^{m_2}\middle|\mathcal{D}_{k_2}\right]\middle|\mathcal{D}_{k_1}\right] = 0$$

$$\begin{aligned}
&\mathsf{E}\left[\left(F_{i,k_1}^{m_1} - \widehat{f}_{k_1}^{m_1}\right)\left(F_{i,k_2}^{m_2} - \widehat{f}_{k_2}^{m_2}\right)\middle|\mathcal{D}^I\right] \\
&\quad = \mathsf{E}\left[\left(\left(F_{i,k_1}^{m_1} - f_{k_1}^{m_1}\right) - \left(\widehat{f}_{k_1}^{m_1} - f_{k_1}^{m_1}\right)\right)\left(\left(F_{i,k_2}^{m_2} - f_{k_2}^{m_2}\right) - \left(\widehat{f}_{k_2}^{n_2} - f_{k_2}^{m_2}\right)\right)\middle|\mathcal{D}^I\right] \\
&\quad = \mathsf{E}\left[\left(F_{i,k_1}^{m_1} - f_{k_1}^{m_1}\right)\left(F_{i,k_2}^{m_2} - f_{k_2}^{m_2}\right)\middle|\mathcal{D}^I\right] - \mathsf{E}\left[\left(F_{i,k_1}^{m_1} - f_{k_1}^{m_1}\right)\left(\widehat{f}_{k_2}^{m_2} - f_{k_2}^{m_2}\right)\middle|\mathcal{D}^I\right] \\
&\qquad - \mathsf{E}\left[\left(\widehat{f}_{k_1}^{m_1} - f_{k_1}^{m_1}\right)\left(F_{i,k_2}^{m_2} - f_{k_2}^{m_2}\right)\middle|\mathcal{D}^I\right] + \mathsf{E}\left[\left(\widehat{f}_{k_1}^{m_1} - f_{k_1}^{m_1}\right)\left(\widehat{f}_{k_2}^{m_2} - f_{k_2}^{m_2}\right)\middle|\mathcal{D}^I\right] \\
&\quad = \mathsf{Cov}\left[F_{i,k_1}^{m_1}, F_{i,k_2}^{m_2}\middle|\mathcal{D}^I\right] - 0 - 0 + \left(\widehat{f}_{k_1}^{m_1} - f_{k_1}^{m_1}\right)\left(\widehat{f}_{k_2}^{m_2} - f_{k_2}^{m_2}\right)
\end{aligned}$$

If $i_1 + k_1 < I$ or $i_2 + k_2 < I$ then $F_{i_1,k_1}^{m_1} \in \mathcal{D}^I$ or $F_{i_2,k_2}^{m_2} \in \mathcal{D}^I$ and we are done. Otherwise, since $(i_1,k_1) \neq (i_2,k_2)$, either $F_{i_1,k_1}^{m_1} \in \mathcal{D}_{k_2}^{i_2+k_2}$ or $F_{i_2,k_2}^{m_2} \in \mathcal{D}_{k_1}^{i_1+k_1}$. Lets assume the first:

$$\begin{aligned}
\mathsf{Cov}\left[F_{i,k_1}^{m_1}, F_{i,k_2}^{m_2}\middle|\mathcal{D}^I\right] &= \mathsf{E}\left[\mathsf{Cov}\left[F_{i,k_1}^{m_1}, F_{i,k_2}^{m_2}\middle|\mathcal{D}_{k_2}^I\right]\middle|\mathcal{D}^I\right] + \mathsf{Cov}\left[\mathsf{E}\left[F_{i,k_1}^{m_1}\middle|\mathcal{D}_{k_2}^I\right], \mathsf{E}\left[F_{i,k_2}^{m_2}\middle|\mathcal{D}_{k_2}^I\right]\middle|\mathcal{D}^I\right] \\
&= 0 + \mathsf{Cov}\left[F_{i,k_1}^{m_1}, f_{k_2}^{m_2}\middle|\mathcal{D}^I\right] = 0
\end{aligned}$$

Estimator 4.16 (Linear approximation of the ultimate uncertainty)

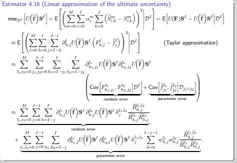
$$\mathsf{mse}_{\mathcal{D}^I}\left[U\left(\widehat{\mathbf{f}}\right)\mathbf{S}^I\right] = \mathsf{E}\left[\left(\sum_{m=0}^{M}\sum_{i=0}^{I}\alpha_i^m\sum_{k=0}^{J}\left(\widehat{S}_{i,k}^m - S_{i,k}^m\right)\right)^2\middle|\mathcal{D}^I\right] = \mathsf{E}\left[U(\mathbf{F})\mathbf{S}^I - U\left(\widehat{\mathbf{f}}\right)\mathbf{S}^I\middle|\mathcal{D}^I\right]$$

$$\approx \mathsf{E}\left[\left(\sum_{l=0}^{M}\sum_{h=0}^{I}\sum_{j=I-h}^{J-1}\partial_{h,j}^{l}U\left(\widehat{\mathbf{f}}\right)\mathbf{S}^I\left(F_{h,j}^{l} - \widehat{f}_j^{l}\right)\right)^2\middle|\mathcal{D}^I\right] \qquad \text{(Taylor approximation)}$$

$$\approx \sum_{l_1,l_2=0}^{M}\sum_{j_1,j_2=0}^{J-1}\sum_{h_1=I-j_1}^{I}\sum_{h_2=I-j_2}^{I}\partial_{h_1,j_1}^{l_1}U\left(\widehat{\mathbf{f}}\right)\mathbf{S}^I\partial_{h_2,j_2}^{l_2}U\left(\widehat{\mathbf{f}}\right)\mathbf{S}^I\left(\underbrace{\mathsf{Cov}\left[F_{h_1,j_1}^{l_1},F_{h_2,j_2}^{l_2}\middle|\mathcal{D}^I\right]}_{\text{random error}} + \underbrace{\mathsf{Cov}\left[\widehat{f}_{j_1}^{l_1},\widehat{f}_{j_2}^{l_2}\middle|\mathcal{D}_{j_1\wedge j_2}\right]}_{\text{parameter error}}\right)$$

$$\approx \underbrace{\sum_{l_1,l_2=0}^{M}\sum_{j=0}^{J-1}\sum_{h=I-j}^{I}\partial_{h,j}^{l_1}U\left(\widehat{\mathbf{f}}\right)\mathbf{S}^I\,\partial_{h,j}^{l_2}U\left(\widehat{\mathbf{f}}\right)\mathbf{S}^I\,\widehat{\sigma}_j^{l_1,l_2}\frac{\widehat{R}_{h,j}^{l_1,l_2}}{\widehat{R}_{h,j}^{l_1}\widehat{R}_{h,j}^{l_2}}}_{\text{random error}}$$

$$+\underbrace{\sum_{l_1,l_2=0}^{M}\sum_{j=0}^{J-1}\sum_{h_1,h_2=I-j}^{I}\partial_{h_1,j}^{l_1}U\left(\widehat{\mathbf{f}}\right)\mathbf{S}^I\,\partial_{h_2,j}^{l_2}U\left(\widehat{\mathbf{f}}\right)\mathbf{S}^I\,\widehat{\sigma}_j^{l_1,l_2}\sum_{h=0}^{I-j-1}w_{h,j}^{l_1}w_{h,j}^{l_2}\frac{\widehat{R}_{h,j}^{l_1,l_2}}{\widehat{R}_{h,j}^{l_1}\widehat{R}_{h,j}^{l_2}}}_{\text{parameter error}}.$$



Estimator 4.16 (Linear approximation of the ultimate uncertainty)

$$\begin{aligned}\mathsf{mse}_{\mathcal{D}^I}\Big[U\big(\widehat{\mathbf{f}}\big)\mathbf{S}^I\Big] &= \mathsf{E}\left[\left(\sum_{m=0}^{M}\sum_{i=0}^{I}\alpha_i^m\sum_{k=0}^{J}\left(\widehat{S}_{i,k}^m - S_{i,k}^m\right)\right)^2\middle|\mathcal{D}^I\right] = \mathsf{E}\Big[U(\mathbf{F})\mathbf{S}^I - U\big(\widehat{\mathbf{f}}\big)\mathbf{S}^I\Big|\mathcal{D}^I\Big]\\
&\approx \mathsf{E}\left[\left(\sum_{l=0}^{M}\sum_{h=0}^{I}\sum_{j=I-h}^{J-1}\partial_{h,j}^l U\big(\widehat{\mathbf{f}}\big)\mathbf{S}^I\left(F_{h,j}^l - \widehat{f}_j^l\right)\right)^2\middle|\mathcal{D}^I\right] \qquad \text{(Taylor approximation)}\\
&\approx \sum_{l_1,l_2=0}^{M}\sum_{j_1,j_2=0}^{J-1}\sum_{h_1=I-j_1}^{I}\sum_{h_2=I-j_2}^{I}\partial_{h_1,j_1}^{l_1}U\big(\widehat{\mathbf{f}}\big)\mathbf{S}^I\partial_{h_2,j_2}^{l_2}U\big(\widehat{\mathbf{f}}\big)\mathbf{S}^I\\
&\qquad\left(\underbrace{\mathsf{Cov}\Big[F_{h_1,j_1}^{l_1},F_{h_2,j_2}^{l_2}\Big|\mathcal{D}^I\Big]}_{\text{random error}} + \underbrace{\mathsf{Cov}\Big[\widehat{f}_{j_1}^{l_1},\widehat{f}_{j_2}^{l_2}\Big|\mathcal{D}_{j_1\wedge j_2}\Big]}_{\text{parameter error}}\right)\\
&\approx \underbrace{\sum_{l_1,l_2=0}^{M}\sum_{j=0}^{J-1}\sum_{h=I-j}^{I}\partial_{h,j}^{l_1}U\big(\widehat{\mathbf{f}}\big)\mathbf{S}^I\,\partial_{h,j}^{l_2}U\big(\widehat{\mathbf{f}}\big)\mathbf{S}^I\,\widehat{\sigma}_j^{l_1,l_2}\frac{\widehat{R}_{h,j}^{l_1,l_2}}{\widehat{R}_{h,j}^{l_1}\widehat{R}_{h,j}^{l_2}}}_{\text{random error}}\\
&\quad+\underbrace{\sum_{l_1,l_2=0}^{M}\sum_{j=0}^{J-1}\sum_{h_1,h_2=I-j}^{I}\partial_{h_1,j}^{l_1}U\big(\widehat{\mathbf{f}}\big)\mathbf{S}^I\,\partial_{h_2,j}^{l_2}U\big(\widehat{\mathbf{f}}\big)\mathbf{S}^I\,\widehat{\sigma}_j^{l_1,l_2}\sum_{h=0}^{I-j-1}w_{h,j}^{l_1}w_{h,j}^{l_2}\frac{\widehat{R}_{h,j}^{l_1,l_2}}{\widehat{R}_{h,j}^{l_1}\widehat{R}_{h,j}^{l_2}}}_{\text{parameter error}}.\end{aligned}$$

In the second approximation we used

$$\begin{aligned}\mathsf{E}\left[\left(F_{h_1,j_1}^{l_1} - \widehat{f}_{j_1}^{l_1}\right)\left(F_{h_2,j_2}^{l_2} - \widehat{f}_{j_2}^{l_2}\right)\middle|\mathcal{D}^I\right] &= \mathsf{Cov}\Big[F_{h_1,j_1}^{l_1},F_{h_2,j_2}^{l_2}\Big|\mathcal{D}^I\Big] + \left(\widehat{f}_{j_1}^{l_1} - f_{j_1}^{l_1}\right)\left(\widehat{f}_{j_2}^{l_1} - f_{j_2}^{l_2}\right)\\
&\approx \mathsf{Cov}\Big[F_{h_1,j_1}^{l_1},F_{h_2,j_2}^{l_2}\Big|\mathcal{D}^I\Big] + \mathsf{Cov}\Big[\widehat{f}_{j_1}^{l_1},\widehat{f}_{j_2}^{l_2}\Big|\mathcal{D}_{j_1\wedge j_2}\Big]\end{aligned}$$

and from the preparations above we know

$$\begin{aligned}\mathsf{Cov}\Big[F_{h_1,j_1}^{l_1},F_{h_2,j_2}^{l_2}\Big|\mathcal{D}^I\Big] &\approx \mathbf{1}_{j_1=j_2}\mathbf{1}_{h_1=h_2}\,\widehat{\sigma}_{j_1}^{l_1,l_2}\,\frac{\widehat{R}_{h_1,j_1}^{l_1,l_2}}{\widehat{R}_{h_1,j_1}^{l_1}\,\widehat{R}_{h_2,j_2}^{l_2}}\\
\mathsf{Cov}\Big[\widehat{f}_{j_1}^{l_1},\widehat{f}_{j_2}^{l_2}\Big|\mathcal{D}_{j_1\wedge j_2}\Big] &\approx \mathbf{1}_{j_1=j_2}\,\widehat{\sigma}_{j_1}^{l_1,l_2}\sum_{h=0}^{I-j_1-1}w_{h,j_1}^{l_1}\,w_{h,j_2}^{l_2}\,\frac{\widehat{R}_{h,j_1}^{l_1,l_2}}{\widehat{R}_{h,j_1}^{l_1}\,\widehat{R}_{h,j_2}^{l_2}}.\end{aligned}$$

This leads directly to the stated estimator.

Chain-Ladder estimator for the ultimate uncertainty

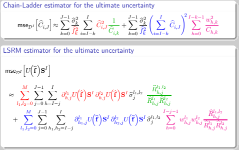
$$\mathsf{mse}_{\mathcal{D}^I}\left[\widehat{C}_{i,J}\right] \approx \sum_{k=0}^{J-1} \frac{\widehat{\sigma}_k^2}{\widehat{f}_k^2} \sum_{i=I-k}^{I} \widehat{C}_{i,J}^2 \frac{1}{\widehat{C}_{i,k}} + \sum_{k=0}^{J-1} \frac{\widehat{\sigma}_k^2}{\widehat{f}_k^2} \left(\sum_{i=I-k}^{I} \widehat{C}_{i,J}\right)^2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

LSRM estimator for the ultimate uncertainty

$$\mathsf{mse}_{\mathcal{D}^I}\left[U\left(\widehat{\mathbf{f}}\right)\mathbf{S}^I\right]$$

$$\approx \sum_{l_1,l_2=0}^{M} \sum_{j=0}^{J-1} \sum_{h=I-j}^{I} \partial_{h,j}^{l_1} U\left(\widehat{\mathbf{f}}\right)\mathbf{S}^I \, \partial_{h,j}^{l_2} U\left(\widehat{\mathbf{f}}\right)\mathbf{S}^I \, \widehat{\sigma}_j^{l_1,l_2} \, \frac{\widehat{R}_{h,j}^{l_1,l_2}}{\widehat{R}_{h,j}^{l_1}\widehat{R}_{h,j}^{l_2}}$$

$$+ \sum_{l_1,l_2=0}^{M} \sum_{j=0}^{J-1} \sum_{h_1,h_2=I-j}^{I} \partial_{h_1,j}^{l_1} U\left(\widehat{\mathbf{f}}\right)\mathbf{S}^I \, \partial_{h_2,j}^{l_2} U\left(\widehat{\mathbf{f}}\right)\mathbf{S}^I \, \widehat{\sigma}_j^{l_1,l_2} \sum_{h=0}^{I-j-1} w_{h,j}^{l_1} w_{h,j}^{l_2} \frac{\widehat{R}_{h,j}^{l_1,l_2}}{\widehat{R}_{h,j}^{l_1}\widehat{R}_{h,j}^{l_2}}$$



Chain-Ladder estimator for the ultimate uncertainty

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

LSRM estimator for the ultimate uncertainty

$$\mathsf{mse}_{\mathcal{D}^I}\left[U\left(\widehat{\mathbf{f}}\right)\mathbf{S}^I\right]$$

$$\approx \sum_{l_1,l_2=0}^{M} \sum_{j=0}^{J-1} \sum_{h=I-j}^{I} \partial_{h,j}^{l_1} U\left(\widehat{\mathbf{f}}\right)\mathbf{S}^I \, \partial_{h,j}^{l_2} U\left(\widehat{\mathbf{f}}\right)\mathbf{S}^I \, \widehat{\sigma}_j^{l_1,l_2} \, \frac{\widehat{R}_{h,j}^{l_1,l_2}}{\widehat{R}_{h,j}^{l_1} \widehat{R}_{h,j}^{l_2}}$$

$$+ \sum_{l_1,l_2=0}^{M} \sum_{j=0}^{J-1} \sum_{h_1,h_2=I-j}^{I} \partial_{h_1,j}^{l_1} U\left(\widehat{\mathbf{f}}\right)\mathbf{S}^I \, \partial_{h_2,j}^{l_2} U\left(\widehat{\mathbf{f}}\right)\mathbf{S}^I \, \widehat{\sigma}_j^{l_1,l_2} \sum_{h=0}^{I-j-1} w_{h,j}^{l_1} w_{h,j}^{l_2} \frac{\widehat{R}_{h,j}^{l_1,l_2}}{\widehat{R}_{h,j}^{l_1} \widehat{R}_{h,j}^{l_2}}$$

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

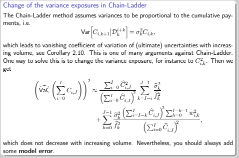
The Chain-Ladder method assumes variances to be proportional to the cumulative payments, i.e.

$$\mathsf{Var}\left[C_{i,k+1}\middle|\mathcal{D}_k^{i+k}\right] = \sigma_k^2 C_{i,k},$$

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

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

which does not decrease with increasing volume. Nevertheless, you should always add some **model error**.



Change of the variance exposures in Chain-Ladder

The Chain-Ladder method assumes variances to be proportional to the cumulative payments, i.e.

$$\mathsf{Var}\left[C_{i,k+1}\middle|\mathcal{D}_k^{i+k}\right] = \sigma_k^2 C_{i,k},$$

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

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

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 σ_k^2 will change, too, which compensates some effects.

Estimator 4.17 (of covariance parameter $\sigma_k^{m_1,m_2}$)

If the normalizing constant

$$Z_k^{m_1,m_2} := \sum_{i=0}^{I-1-k} \frac{w_{i,k}^{m_1} w_{i,k}^{m_2}}{R_{i,k}^{m_1} R_{i,k}^{m_2}} \left(1 - w_{i,k}^{m_1} - w_{i,k}^{m_2} + R_{i,k}^{m_1,m_2} \sum_{h=0}^{I-1-k} \frac{w_{h,k}^{m_1} w_{h,k}^{m_2}}{R_{h,k}^{m_1,m_2}}\right) > 0$$

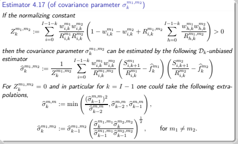
then the covariance parameter $\sigma_k^{m_1,m_2}$ can be estimated by the following $\mathcal{D}_k$-unbiased estimator

$$\widehat{\sigma}_k^{m_1,m_2} := \frac{1}{Z_k^{m_1,m_2}} \sum_{i=0}^{I-1-k} \frac{w_{i,k}^{m_1} w_{i,k}^{m_2}}{R_{i,k}^{m_1,m_2}} \left(\frac{S_{i,k+1}^{m_1}}{R_{i,k}^{m_1}} - \widehat{f}_k^{m_1}\right) \left(\frac{S_{i,k+1}^{m_2}}{R_{i,k}^{m_2}} - \widehat{f}_k^{m_2}\right)$$

For $Z_k^{m_1,m_2} = 0$ and in particular for $k = I - 1$ one could take the following extrapolations,

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

$$\widehat{\sigma}_k^{m_1,m_2} := \widehat{\sigma}_{k-1}^{m_1,m_2} \left(\frac{\widehat{\sigma}_k^{m_1,m_1} \widehat{\sigma}_k^{m_2,m_2}}{\widehat{\sigma}_{k-1}^{m_1,m_1} \widehat{\sigma}_{k-1}^{m_2,m_2}}\right)^{\frac{1}{2}}, \qquad \text{for } m_1 \neq m_2.$$



Estimator 4.17 (of covariance parameter $\sigma_k^{m_1,m_2}$)

If the normalizing constant

$$Z_k^{m_1,m_2} := \sum_{i=0}^{I-1-k} \frac{w_{i,k}^{m_1} w_{i,k}^{m_2}}{R_{i,k}^{m_1} R_{i,k}^{m_2}} \left(1 - w_{i,k}^{m_1} - w_{i,k}^{m_2} + R_{i,k}^{m_1,m_2} \sum_{h=0}^{I-1-k} \frac{w_{h,k}^{m_1} w_{h,k}^{m_2}}{R_{h,k}^{m_1,m_2}}\right) > 0$$

then the covariance parameter $\sigma_k^{m_1,m_2}$ can be estimated by the following $\mathcal{D}_k$-unbiased estimator

$$\widehat{\sigma}_k^{m_1,m_2} := \frac{1}{Z_k^{m_1,m_2}} \sum_{i=0}^{I-1-k} \frac{w_{i,k}^{m_1} w_{i,k}^{m_2}}{R_{i,k}^{m_1,m_2}} \left(\frac{S_{i,k+1}^{m_1}}{R_{i,k}^{m_1}} - \widehat{f}_k^{m_1}\right) \left(\frac{S_{i,k+1}^{m_2}}{R_{i,k}^{m_2}} - \widehat{f}_k^{m_2}\right)$$

For $Z_k^{m_1,m_2} = 0$ and in particular for $k = I-1$ one could take the following extrapolations,

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

$$\widehat{\sigma}_k^{m_1,m_2} := \widehat{\sigma}_{k-1}^{m_1,m_2} \left(\frac{\widehat{\sigma}_k^{m_1,m_1} \widehat{\sigma}_k^{m_2,m_2}}{\widehat{\sigma}_{k-1}^{m_1,m_1} \widehat{\sigma}_{k-1}^{m_2,m_2}}\right)^{\frac{1}{2}}, \qquad \text{for } m_1 \neq m_2.$$

$$\mathsf{E}\left[\left(\frac{S_{i,k+1}^{m_1}}{R_{i,k}^{m_1}} - \widehat{f}_k^{m_1}\right)\left(\frac{S_{i,k+1}^{m_2}}{R_{i,k}^{m_2}} - \widehat{f}_k^{m_2}\right)\middle|\mathcal{D}_k\right] = \underbrace{\mathsf{Cov}\left[\left(\frac{S_{i,k+1}^{m_1}}{R_{i,k}^{m_1}} - \widehat{f}_k^{m_1}\right), \left(\frac{S_{i,k+1}^{m_2}}{R_{i,k}^{m_2}} - \widehat{f}_k^{m_2}\right)\middle|\mathcal{D}_k\right]}_{\textbf{i)}^{\textbf{LSRM}} \textbf{ and Estimator 4.7}}$$

$$= \mathsf{Cov}\left[\frac{S_{i,k+1}^{m_1}}{R_{i,k}^{m_1}}, \frac{S_{i,k+1}^{m_2}}{R_{i,k}^{m_2}}\middle|\mathcal{D}_k\right] - \sum_{h=0}^{I-k-1} \mathsf{Cov}\left[w_{h,k}^{m_1}\frac{S_{h,k+1}^{m_1}}{R_{h,k}^{m_1}}, \frac{S_{i,k+1}^{m_2}}{R_{i,k}^{m_2}}\middle|\mathcal{D}_k\right]$$

$$- \sum_{h=0}^{I-k-1} \mathsf{Cov}\left[\frac{S_{i,k+1}^{m_1}}{R_{i,k}^{m_1}}, w_{h,k}^{m_2}\frac{S_{h,k+1}^{m_2}}{R_{h,k}^{m_2}}\middle|\mathcal{D}_k\right] + \sum_{h_1,h_2=0}^{I-k-1} \mathsf{Cov}\left[w_{h_1,k}^{m_1}\frac{S_{h_1,k+1}^{m_1}}{R_{h_1,k}^{m_1}}, w_{h_2,k}^{m_2}\frac{S_{h_2,k+1}^{m_2}}{R_{h_2,k}^{m_2}}\middle|\mathcal{D}_k\right]$$

$$\underbrace{= \sigma_k^{m_1,m_2}\frac{R_{i,k}^{m_1,m_2}}{R_{i,k}^{m_1}R_{i,k}^{m_2}}\left(1 - w_{i,k}^{m_1} - w_{i,k}^{m_2} + \frac{R_{i,k}^{m_1}R_{i,k}^{m_2}}{R_{i,k}^{m_1,m_2}}\sum_{h=0}^{I-k-1}\frac{w_{h,k}^{m_1}w_{h,k}^{m_2}R_{h,k}^{m_1,m_2}}{R_{h,k}^{m_1}R_{h,k}^{m_2}}\right)}_{\textbf{ii)}^{\textbf{LSRM}} \textbf{ and Lemma 4.2}}$$

$$\Longrightarrow \mathsf{E}\left[\widehat{\sigma}_k^{m_1,m_2}\middle|\mathcal{D}_k\right] = \frac{\sigma_k^{m_1,m_2}}{Z_k^{m_1,m_2}}\underbrace{\sum_{i=0}^{I-1-k}\frac{w_{i,k}^{m_1}w_{i,k}^{m_2}}{R_{i,k}^{m_1}R_{i,k}^{m_2}}\left(1 - w_{i,k}^{m_1} - w_{i,k}^{m_2} + \frac{R_{i,k}^{m_1}R_{i,k}^{m_2}}{R_{i,k}^{m_1,m_2}}\sum_{h=0}^{I-k-1}\frac{w_{h,k}^{m_1}w_{h,k}^{m_2}R_{h,k}^{m_1,m_2}}{R_{h,k}^{m_1}R_{h,k}^{m_2}}\right)}_{\textbf{change order of summation in the fourth term}}$$

$$= \sigma_k^{m_1,m_2}$$

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

$$\left(\sigma_k^{m_1,m_2} R_{i,k}^{m_1,m_2}\right)_{0\leq m_1,m_2\leq M}$$

  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 [22]. 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 $\sigma_k^{m_1,m_2}$)

- Even if the real covariance parameter $\sigma_k^{m_1,m_2}$ lead to positive semi-defined covariance matrices
$$\left(\sigma_k^{m_1,m_2} R_{i,k}^{m_1,m_2}\right)_{0 \leq m_1,m_2 \leq M}$$
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 [22]. 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.

Swiss mandatory accident portfolio: part 2 of 3, see Example 4.13

We have four claim properties with exposures

ME: $R^{0}_{i,k} = R^{0,0}_{i,k} = \sum_{j=0}^{k} S^{0}_{i,j}$

IW: $R^{1}_{i,k} = R^{1,1}_{i,k} = \sum_{j=0}^{k} \left(\kappa^{j+1} S^{3}_{i,j} + (1-\kappa^{j+1}) S^{1}_{i,j}\right)$

Sub: $R^{2}_{i,k} = R^{2,2}_{i,k} = \sum_{j=0}^{k} \left(S^{0}_{i,j} + S^{1}_{i,j} + S^{2}_{i,j}\right)$

Salary: $S^{3}_{i,0} = P_i$, $S^{3}_{i,j} = 0$, for $j > 0$, and

$$R^{3}_{i,k} = R^{3,0}_{i,k} = R^{0,3}_{i,k} = R^{3,1}_{i,k} = R^{1,3}_{i,k} = R^{3,2}_{i,k} = R^{2,3}_{i,k} = R^{3,3}_{i,k} = 0$$

For the not yet defined exposures we take the total payments up to now, i.e.

$$R^{0,1}_{i,k} = R^{1,0}_{i,k} = R^{0,2}_{i,k} = R^{2,0}_{i,k} = R^{1,2}_{i,k} = R^{2,1}_{i,k} = \sum_{j=0}^{k} \left(S^{0}_{i,j} + S^{1}_{i,j} + S^{2}_{i,j}\right).$$

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 (green circle on the left).

Swiss mandatory accident portfolio: part 2 of 3, see Example 4.13

We have four claim properties with exposures

ME: $R^0_{i,k} = R^{0,0}_{i,k} = \sum_{j=0}^k S^0_{i,j}$

IW: $R^1_{i,k} = R^{1,1}_{i,k} = \sum_{j=0}^k \left(\kappa^{j+1} S^3_{i,j} + (1-\kappa^{j+1}) S^1_{i,j}\right)$

Sub: $R^2_{i,k} = R^{2,2}_{i,k} = \sum_{j=0}^k \left(S^0_{i,j} + S^1_{i,j} + S^2_{i,j}\right)$

Salary: $S^3_{i,0} = P_i$, $S^3_{i,j} = 0$, for $j > 0$, and

$$R^3_{i,k} = R^{3,0}_{i,k} = R^{0,3}_{i,k} = R^{3,1}_{i,k} = R^{1,3}_{i,k} = R^{3,2}_{i,k} = R^{2,3}_{i,k} = R^{3,3}_{i,k} = 0$$

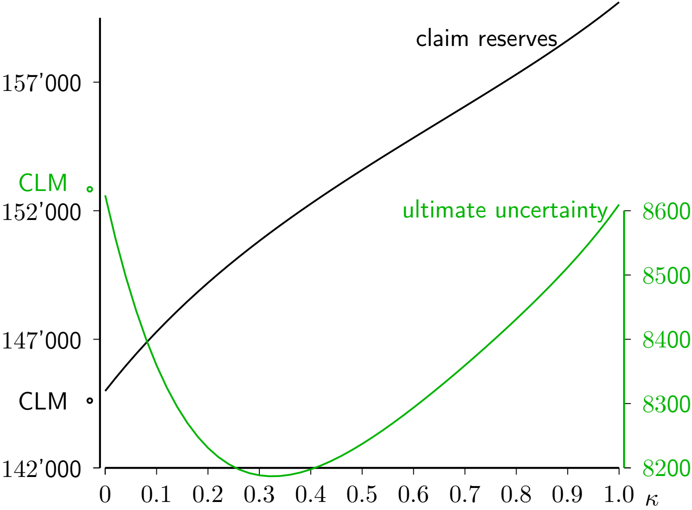
For the not yet defined exposures we take the total payments up to now, i.e.

$$R^{0,1}_{i,k} = R^{1,0}_{i,k} = R^{0,2}_{i,k} = R^{2,0}_{i,k} = R^{1,2}_{i,k} = R^{2,1}_{i,k} = \sum_{j=0}^k \left(S^0_{i,j} + S^1_{i,j} + S^2_{i,j}\right).$$

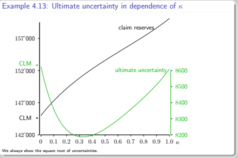
Resulting ultimate uncertainty

- 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 (green circle on the left).

Example 4.13: Ultimate uncertainty in dependence of κ



We always show the square root of uncertainties.



Example 4.13: Ultimate uncertainty in dependence of κ

We always show the square root of uncertainties.

Even if it looks tempting you must not use the estimates of the ultimate uncertainty to evaluate which model is the best!

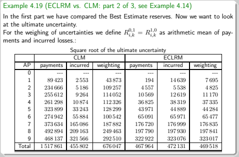
Example 4.19 (ECLRM vs. CLM: part 2 of 3, see Example 4.14)

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 define $R_{i,k}^{0,1} = R_{i,k}^{1,0}$ as arithmetic mean of payments and incurred losses.:

Square root of the ultimate uncertainty

	CLM			ECLRM		
AP	payments	incurred	weighting	payments	incurred	weighting
0	---	---	---	---	---	---
1	89 423	2 553	43 873	194	14 639	7 695
2	234 666	5 186	109 257	4 557	5 538	4 825
3	255 612	9 264	114 052	10 569	12 619	11 170
4	261 298	10 874	112 326	36 825	38 319	37 335
5	323 899	33 243	128 299	43 971	44 889	44 284
6	274 942	55 884	100 542	65 091	65 971	65 477
7	373 634	165 086	187 882	176 720	176 999	176 835
8	492 894	209 163	249 463	197 790	197 930	197 841
9	468 137	321 566	292 510	322 922	323 076	323 017
Total	1 517 861	455 802	676 047	467 964	472 131	469 518



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 define $R^{0,1}_{i,k} = R^{1,0}_{i,k}$ as arithmetic mean of payments and incurred losses.:

Square root of the ultimate uncertainty

	CLM			ECLRM		
AP	payments	incurred	weighting	payments	incurred	weighting
0	---	---	---	---	---	---
1	89 423	2 553	43 873	194	14 639	7 695
2	234 666	5 186	109 257	4 557	5 538	4 825
3	255 612	9 264	114 052	10 569	12 619	11 170
4	261 298	10 874	112 326	36 825	38 319	37 335
5	323 899	33 243	128 299	43 971	44 889	44 284
6	274 942	55 884	100 542	65 091	65 971	65 477
7	373 634	165 086	187 882	176 720	176 999	176 835
8	492 894	209 163	249 463	197 790	197 930	197 841
9	468 137	321 566	292 510	322 922	323 076	323 017
Total	1 517 861	455 802	676 047	467 964	472 131	469 518

- Taking the arithmetic mean

$$R^{m_1,m_2}_{i,k} := \frac{1}{2}\left(R^{m_1,m_1}_{i,k} + R^{m_2,m_2}_{i,k}\right)$$

  for the coupling exposures works fine if $R^{m_1,m_1}_{i,k}$ and $R^{m_2,m_2}_{i,k}$ are similar. In general the geometric mean

$$R^{m_1,m_2}_{i,k} := \sqrt{R^{m_1,m_1}_{i,k}\, R^{m_2,m_2}_{i,k}}$$

  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 6)

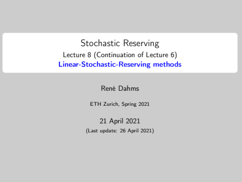
Linear-Stochastic-Reserving methods

René Dahms

ETH Zurich, Spring 2021

21 April 2021

(Last update: 26 April 2021)



Stochastic Reserving

Lecture 8 (Continuation of Lecture 6)

Linear-Stochastic-Reserving methods

René Dahms

ETH Zurich, Spring 2021

21 April 2021

(Last update: 26 April 2021)

4 Linear-Stochastic-Reserving methods

Lecture 8: Table of contents

Consistent estimation over time

In this section we want to look at the solvency uncertainty, i.e. the uncertainty of the claims development result

$$\mathrm{CDR}_i^{I+1} := \sum_{m=0}^{M} \alpha_i^m \sum_{k=I-i}^{J-1} \left(\widehat{S}_{i,k+1}^{m,I} - \widehat{S}_{i,k+1}^{m,I+1} \right) \quad \text{and} \quad \mathrm{CDR}^{I+1} := \sum_{i=0}^{I} \mathrm{CDR}_i^{I+1},$$

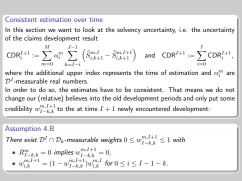
where the additional upper index represents the time of estimation and α_i^m are $\mathcal{D}^I$-measurable real numbers.

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 $\mathcal{D}^I \cap \mathcal{D}_k$-measurable weights $0 \leq w_{I-k,k}^{m,I+1} \leq 1$ with

- $R_{I-k,k}^m = 0$ *implies* $w_{I-k,k}^{m,I+1} = 0$,
- $w_{i,k}^{m,I+1} = (1 - w_{I-k,k}^{m,I+1}) w_{i,k}^{m,I}$ *for* $0 \leq i \leq I-1-k$.



Consistent estimation over time

In this section we want to look at the solvency uncertainty, i.e. the uncertainty of the claims development result

$$\mathsf{CDR}_i^{I+1} := \sum_{m=0}^{M} \alpha_i^m \sum_{k=I-i}^{J-1} \left(\widehat{S}_{i,k+1}^{m,I} - \widehat{S}_{i,k+1}^{m,I+1} \right) \quad \text{and} \quad \mathsf{CDR}^{I+1} := \sum_{i=0}^{I} \mathsf{CDR}_i^{I+1},$$

where the additional upper index represents the time of estimation and α_i^m are $\mathcal{D}^I$-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 $\mathcal{D}^I \cap \mathcal{D}_k$-measurable weights $0 \leq w_{I-k,k}^{m,I+1} \leq 1$ with

- $R_{I-k,k}^m = 0$ *implies* $w_{I-k,k}^{m,I+1} = 0$,
- $w_{i,k}^{m,I+1} = (1 - w_{I-k,k}^{m,I+1}) w_{i,k}^{m,I}$ *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

$$\widehat{f}_k^{m,I+1} := \sum_{i=0}^{I-k} w_{i,k}^{m,I+1} \frac{S_{i,k+1}^m}{R_{i,k}^m} = (1 - w_{I-k,k}^{m,I+1})\widehat{f}_k^{m,I} + w_{I-k,k}^{m,I+1} \frac{S_{I-k,k+1}^m}{R_{I-k,k}^m}$$

satisfy:

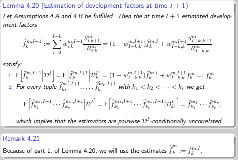
1. $\mathsf{E}\left[\widehat{f}_k^{m,I+1}\middle|\mathcal{D}^I\right] = \mathsf{E}\left[\widehat{f}_k^{m,I+1}\middle|\mathcal{D}_k^I\right] = (1 - w_{I-k,k}^{m,I+1})\widehat{f}_k^{m,I} + w_{I-k,k}^{m,I+1} f_k^m =: \bar{f}_k^m$
2. *For every tuple* $\widehat{f}_{k_1}^{m_1,I+1}, \dots, \widehat{f}_{k_r}^{m_r,I+1}$ *with* $k_1 < k_2 < \cdots < k_r$ *we get*

$$\mathsf{E}\left[\widehat{f}_{k_1}^{m_1,I+1} \cdots \widehat{f}_{k_r}^{m_r,I+1}\middle|\mathcal{D}^I\right] = \mathsf{E}\left[\widehat{f}_{k_1}^{m_1,I+1} \cdots \widehat{f}_{k_r}^{m_r,I+1}\middle|\mathcal{D}_{k_1}^I\right] = \bar{f}_{k_1}^{m_1} \cdots \bar{f}_{k_r}^{m_r},$$

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 $\widehat{\bar{f}}_k^m := \widehat{f}_k^{m,I}$.



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

$$\widehat{f}_k^{m,I+1} := \sum_{i=0}^{I-k} w_{i,k}^{m,I+1} \frac{S_{i,k+1}^m}{R_{i,k}^m} = (1 - w_{I-k,k}^{m,I+1})\widehat{f}_k^{m,I} + w_{I-k,k}^{m,I+1} \frac{S_{I-k,k+1}^m}{R_{I-k,k}^m}$$

satisfy:

1. $\mathsf{E}\left[\widehat{f}_k^{m,I+1}\middle|\mathcal{D}^I\right] = \mathsf{E}\left[\widehat{f}_k^{m,I+1}\middle|\mathcal{D}_k^I\right] = (1 - w_{I-k,k}^{m,I+1})\widehat{f}_k^{m,I} + w_{I-k,k}^{m,I+1} f_k^m =: \bar{f}_k^m$
2. *For every tuple $\widehat{f}_{k_1}^{m_1,I+1}, \ldots, \widehat{f}_{k_r}^{m_r,I+1}$ with $k_1 < k_2 < \cdots < k_r$ we get*

$$\mathsf{E}\left[\widehat{f}_{k_1}^{m_1,I+1} \cdots \widehat{f}_{k_r}^{m_r,I+1}\middle|\mathcal{D}^I\right] = \mathsf{E}\left[\widehat{f}_{k_1}^{m_1,I+1} \cdots \widehat{f}_{k_r}^{m_r,I+1}\middle|\mathcal{D}_{k_1}^I\right] = \bar{f}_{k_1}^{m_1} \cdots \bar{f}_{k_r}^{m_r},$$

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 $\widehat{\widehat{f}}_k^{m} := \widehat{f}_k^{m,I}$.

$$\mathsf{E}\left[\widehat{f}_k^{m,I+1}\middle|\mathcal{D}^I\right] = \mathsf{E}\left[\mathsf{E}\left[\widehat{f}_k^{m,I+1}\middle|\mathcal{D}_k^I\right]\middle|\mathcal{D}^I\right] = (1 - w_{I-k,k}^{m,I+1})\widehat{f}_k^{m,I} + w_{I-k,k}^{m,I+1} \frac{\mathsf{E}\left[\mathsf{E}\left[S_{I-k,k+1}^m\middle|\mathcal{D}_k^I\right]\middle|\mathcal{D}^I\right]}{R_{I-k,k}^m}$$

$$= (1 - w_{I-k,k}^{m,I+1})\widehat{f}_k^{m,I} + w_{I-k,k}^{m,I+1} f_k^m = \bar{f}_k^m$$

$$\begin{aligned}\mathsf{E}\left[\widehat{f}_{k_1}^{m_1,I+1} \cdots \widehat{f}_{k_r}^{m_r,I+1}\middle|\mathcal{D}^I\right] &= \mathsf{E}\left[\widehat{f}_{k_1}^{m_1,I+1} \cdots \widehat{f}_{k}^{m_{r-1},I+1} \mathsf{E}\left[\widehat{f}_{k_r}^{m_r,I+1}\middle|\mathcal{D}_{k_r}^I\right]\middle|\mathcal{D}^I\right] \\ &= \mathsf{E}\left[\widehat{f}_{k_1}^{m_1,I+1} \cdots \widehat{f}_{k}^{m_{r-1},I+1} \bar{f}_{k_r}^{m_r}\middle|\mathcal{D}^I\right] \\ = \cdots &= \bar{f}_{k_1}^{m_1} \cdots \bar{f}_{k_r}^{m_r}\end{aligned}$$

and similar for $\mathcal{D}_{k_1}^I$ instead of $\mathcal{D}^I$.

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*

$$\bar{S}_{i,k+1}^{m} := \mathsf{E}\left[\widehat{S}_{i,k+1}^{m,I+1}\middle|\mathcal{D}^{I}\right] = \mathcal{F}_{i,k}^{m,I+1}(\bar{\mathbf{f}})\,\boldsymbol{\mathcal{F}}^{I}(\mathbf{f})\,\mathbf{S}^{I}.$$

Hence, we will use the estimates

$$\widehat{\mathsf{E}}\left[\widehat{S}_{i,k+1}^{m,I+1}\middle|\mathcal{D}^{I}\right] = \widehat{\bar{S}}_{i,k+1}^{m} := \mathcal{F}_{i,k}^{m,I+1}\left(\widehat{\bar{\mathbf{f}}}\right)\boldsymbol{\mathcal{F}}^{I}\left(\widehat{\mathbf{f}}^{I}\right)\mathbf{S}^{I} = \mathcal{F}_{i,k}^{m,I}\left(\widehat{\mathbf{f}}^{I}\right)\mathbf{S}^{I},$$

which implies $\widehat{\mathsf{E}}\left[\mathsf{CDR}^{I+1}\middle|\mathcal{D}^{I}\right] := 0$. *That means, we have Best Estimate reserves.*

Notation

As always we will use $\widehat{S}_{i,k}^{m,I+1} := S_{i,k}^{m}$ for $i+k \le I+1$ and

$$\widehat{R}_{i,k}^{m,I+1} := \Gamma_{i,k}^{m}\widehat{\mathbf{S}}^{i+k,I+1} \qquad \text{and} \qquad \widehat{R}_{i,k}^{m_1,m_2,I+1} := \Gamma_{i,k}^{m_1,m_2}\widehat{\mathbf{S}}^{i+k,I+1}.$$

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

$$\bar{S}^{m}_{i,k+1} := \mathsf{E}\left[\widehat{S}^{m,I+1}_{i,k+1}\middle|\mathcal{D}^I\right] = \mathcal{F}^{m,I+1}_{i,k}\left(\bar{\mathbf{f}}\right)\boldsymbol{\mathcal{F}}^I\left(\mathbf{f}\right)\mathbf{S}^I.$$

Hence, we will use the estimates

$$\widehat{\mathsf{E}}\left[\widehat{S}^{m,I+1}_{i,k+1}\middle|\mathcal{D}^I\right] = \widehat{\bar{S}}^{m}_{i,k+1} := \mathcal{F}^{m,I+1}_{i,k}\left(\widehat{\bar{\mathbf{f}}}\right)\boldsymbol{\mathcal{F}}^I\left(\widehat{\mathbf{f}}^I\right)\mathbf{S}^I = \mathcal{F}^{m,I}_{i,k}\left(\widehat{\mathbf{f}}^I\right)\mathbf{S}^I,$$

which implies $\widehat{\mathsf{E}}\left[\mathsf{CDR}^{I+1}\middle|\mathcal{D}^I\right] := 0$. *That means, we have Best Estimate reserves.*

Notation

As always we will use $\widehat{S}^{m,I+1}_{i,k} := S^m_{i,k}$ for $i+k \le I+1$ and

$$\widehat{R}^{m,I+1}_{i,k} := \Gamma^m_{i,k}\widehat{\mathbf{S}}^{i+k,I+1} \quad \text{and} \quad \widehat{R}^{m_1,m_2,I+1}_{i,k} := \Gamma^{m_1,m_2}_{i,k}\widehat{\mathbf{S}}^{i+k,I+1}.$$

At estimation time $I+1$ we have

$$\widehat{S}^{m,I+1}_{i,k+1} = \mathcal{F}^{m,I+1}_{i,k}\left(\widehat{\mathbf{f}}^{I+1}\right)\mathbf{S}^{I+1}.$$

Induction: If $i+k \le I$ then $\widehat{S}^{m,I+1}_{i,k+1} = S^m_{i,k+1}$ and therefore

$$\mathsf{E}\left[\widehat{S}^{m,I+1}_{i,k+1}\middle|\mathcal{D}^I\right] = \mathsf{E}\left[S^m_{i,k+1}\middle|\mathcal{D}^I\right] = \underbrace{f^m_k R^m_{i,k}}_{\text{i)LSRM}} = \mathcal{F}^{m,I}_{i,k}\left(\mathbf{f}\right)\mathbf{S}^I.$$

Now assume that Lemma 4.22 is fulfilled for all $i+k<n$. Then we get for $i+k=n$

$$\begin{aligned}
\mathsf{E}\left[\widehat{S}^{m,I+1}_{i,k+1}\middle|\mathcal{D}^I\right] &= \mathsf{E}\left[\widehat{f}^{I+1}_k\widehat{R}^{m,I+1}_{i,k}\middle|\mathcal{D}^I\right]\\
&= \mathsf{E}\left[\mathsf{E}\left[\widehat{f}^{I+1}_k\middle|\mathcal{D}^I_k\right]\widehat{R}^{I+1}_{i,k}\middle|\mathcal{D}^I\right]\\
&= \mathsf{E}\left[\bar{f}_k\widehat{R}^m_{i,k}\middle|\mathcal{D}^I\right]\\
&= \mathcal{F}^{m,n}_{i,k}\left(\bar{\mathbf{f}}\right)\mathsf{E}\left[\widehat{\mathbf{S}}^n\middle|\mathcal{D}^I\right]\\
&= \underbrace{\mathcal{F}^{m,n}_{i,k}\left(\bar{\mathbf{f}}\right)\boldsymbol{\mathcal{F}}^{n\leftarrow I+1}\left(\bar{\mathbf{f}}\right)\boldsymbol{\mathcal{F}}^I\left(\mathbf{f}\right)\mathbf{S}^I}_{\textbf{induction hypothesis}}\\
&= \mathcal{F}^{m,I+1}_{i,k}\left(\bar{\mathbf{f}}\right)\boldsymbol{\mathcal{F}}^I\left(\mathbf{f}\right)\mathbf{S}^I.
\end{aligned}$$

$\square$

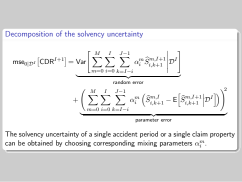
Note, a proof without induction can be done by a combination of the tower property, the multilinearity of $\boldsymbol{\mathcal{F}}^n\left(\mathbf{f}\right)$, see Lemma 4.9, and the product formula of Lemma 4.20.

Decomposition of the solvency uncertainty

$$\mathsf{mse}_{0|\mathcal{D}^I}\left[\mathsf{CDR}^{I+1}\right] = \underbrace{\mathsf{Var}\left[\sum_{m=0}^{M}\sum_{i=0}^{I}\sum_{k=I-i}^{J-1}\alpha_i^m\,\widehat{S}_{i,k+1}^{m,I+1}\middle|\mathcal{D}^I\right]}_{\text{random error}}$$

$$+\underbrace{\left(\sum_{m=0}^{M}\sum_{i=0}^{I}\sum_{k=I-i}^{J-1}\alpha_i^m\left(\widehat{S}_{i,k+1}^{m,I}-\mathsf{E}\left[\widehat{S}_{i,k+1}^{m,I+1}\middle|\mathcal{D}^I\right]\right)\right)^2}_{\text{parameter error}}$$

The solvency uncertainty of a single accident period or a single claim property can be obtained by choosing corresponding mixing parameters α_i^m.



Decomposition of the solvency uncertainty

$$\mathsf{mse}_{0|\mathcal{D}^I}\left[\mathsf{CDR}^{I+1}\right] = \underbrace{\mathsf{Var}\left[\sum_{m=0}^{M}\sum_{i=0}^{I}\sum_{k=I-i}^{J-1}\alpha_i^m \widehat{S}_{i,k+1}^{m,I+1}\middle|\mathcal{D}^I\right]}_{\text{random error}} + \underbrace{\left(\sum_{m=0}^{M}\sum_{i=0}^{I}\sum_{k=I-i}^{J-1}\alpha_i^m\left(\widehat{S}_{i,k+1}^{m,I} - \mathsf{E}\left[\widehat{S}_{i,k+1}^{m,I+1}\middle|\mathcal{D}^I\right]\right)\right)^2}_{\text{parameter error}}$$

The solvency uncertainty of a single accident period or a single claim property can be obtained by choosing corresponding mixing parameters α_i^m.

Taylor approximation of next years estimates

Recall the (multi-linear) functional:

$$U(\mathbf{g})\mathbf{x} := \sum_{m=0}^{M}\sum_{i=0}^{I}\alpha_i^m\left(\sum_{k=0}^{I-i}x_{i,k}^m + \sum_{k=I-i}^{J-1}\mathcal{F}_{i,k}^{m,I}(\mathbf{g})\,\mathbf{x}\right).$$

Then we have

$$U(\mathbf{f})\mathbf{S}^I = \sum_{m=0}^{M}\sum_{i=0}^{I}\alpha_i^m\sum_{k=0}^{J}\mathsf{E}\left[S_{i,k}^m\middle|\mathcal{D}^I\right],$$

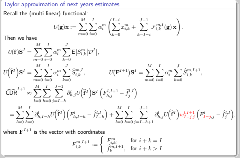
$$U\left(\widehat{\mathbf{f}}^I\right)\mathbf{S}^I = \sum_{m=0}^{M}\sum_{i=0}^{I}\alpha_i^m\sum_{k=0}^{J}\widehat{S}_{i,k}^{m,I}, \qquad U(\mathbf{F}^{I+1})\mathbf{S}^I = \sum_{m=0}^{M}\sum_{i=0}^{I}\alpha_i^m\sum_{k=0}^{J}\widehat{S}_{i,k}^{m,I+1},$$

$$\widehat{\mathsf{CDR}}^{I+1} \approx \sum_{l=0}^{M}\sum_{h=0}^{I}\sum_{j=I-h}^{J-1}\partial_{h,j}^l U\left(\widehat{\mathbf{f}}^I\right)\mathbf{S}^I\left(F_{h,j}^{l,I+1}-\widehat{f}_j^{l,I}\right)$$

$$= \sum_{l=0}^{M}\sum_{h=0}^{I}\partial_{h,I-h}^l U\left(\widehat{\mathbf{f}}^I\right)\left(F_{h,I-h}^l-\widehat{f}_{I-h}^{l,I}\right) + \sum_{l=0}^{M}\sum_{h=0}^{I}\sum_{j=I-h+1}^{J-1}\partial_{h,j}^l U\left(\widehat{\mathbf{f}}^I\right)w_{I-j,j}^{l,I+1}\left(F_{I-j,j}^l-\widehat{f}_j^{l,I}\right),$$

where $\mathbf{F}^{I+1}$ is the vector with coordinates

$$F_{i,k}^{m,I+1} := \begin{cases} F_{i,k}^m, & \text{for } i+k=I \\ \widehat{f}_k^{m,I+1}, & \text{for } i+k>I \end{cases}$$



Taylor approximation of next years estimates

Recall the (multi-linear) functional:

$$U(\mathbf{g})\mathbf{x} := \sum_{m=0}^{M}\sum_{i=0}^{I}\alpha_i^m\left(\sum_{k=0}^{I-i}x_{i,k}^m + \sum_{k=I-i}^{J-1}\mathcal{F}_{i,k}^{m,I}(\mathbf{g})\,\mathbf{x}\right).$$

Then we have

$$U(\mathbf{f})\mathbf{S}^I = \sum_{m=0}^{M}\sum_{i=0}^{I}\alpha_i^m\sum_{k=0}^{J}\mathsf{E}\left[S_{i,k}^m\middle|\mathcal{D}^I\right],$$

$$U\left(\widehat{\mathbf{f}}^I\right)\mathbf{S}^I = \sum_{m=0}^{M}\sum_{i=0}^{I}\alpha_i^m\sum_{k=0}^{J}\widehat{S}_{i,k}^{m,I}, \qquad U(\mathbf{F}^{I+1})\mathbf{S}^I = \sum_{m=0}^{M}\sum_{i=0}^{I}\alpha_i^m\sum_{k=0}^{J}\widehat{S}_{i,k}^{m,I+1},$$

$$\widehat{\mathsf{CDR}}^{I+1} \approx \sum_{l=0}^{M}\sum_{h=0}^{I}\sum_{j=I-h}^{J-1}\partial_{h,j}^l U\left(\widehat{\mathbf{f}}^I\right)\mathbf{S}^I\left(F_{h,j}^{l,I+1} - \widehat{f}_j^{l,I}\right)$$

$$= \sum_{l=0}^{M}\sum_{h=0}^{I}\partial_{h,I-h}^l U\left(\widehat{\mathbf{f}}^I\right)\left(F_{h,I-h}^l - \widehat{f}_{I-h}^{l,I}\right) + \sum_{l=0}^{M}\sum_{h=0}^{I}\sum_{j=I-h+1}^{J-1}\partial_{h,j}^l U\left(\widehat{\mathbf{f}}^I\right)w_{I-j,j}^{l,I+1}\left(F_{I-j,j}^l - \widehat{f}_j^{l,I}\right),$$

where $\mathbf{F}^{I+1}$ is the vector with coordinates

$$F_{i,k}^{m,I+1} := \begin{cases} F_{i,k}^m, & \text{for } i+k=I \\ \widehat{f}_k^{m,I+1}, & \text{for } i+k>I \end{cases}$$

For $k = I - i$ we get, see Lemma 4.20,

$$F_{i,I-i}^{m,I+1} - \widehat{f}_{I-i}^{m,I} = F_{i,I-i}^m - \widehat{f}_{I-i}^{m,I}$$

and for $k > I - i$ it is

$$\begin{aligned} F_{i,k}^{m,I+1} - \widehat{f}_k^{m,I} &= \widehat{f}_k^{m,I+1} - \widehat{f}_k^{m,I} = (1 - w_{I-k,k}^{m,I+1})\widehat{f}_k^{m,I} + w_{I-k,k}^{m,I+1}F_{I-k,k}^m - \widehat{f}_k^{m,I} \\ &= w_{I-k,k}^{m,I+1}\left(F_{I-k,k}^m - \widehat{f}_k^{m,I}\right). \end{aligned}$$

Note, since

$$F_{i,I-i}^{m,I+1} = F_{i,I-i}^m = \frac{S_{i,I-i+1}^m}{R_{i,I-i}^m}$$

we get

$$\mathcal{F}_{i,I-i}^{m,I}\left(\mathbf{F}^{I+1}\right)\mathbf{S}^I = F_{i,I-i}^{m,I+1}R_{i,I-i}^m = S_{i,I-i+1}^m.$$

That means, the operator $U(\mathbf{F}^{I+1})$ (re)contructs 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

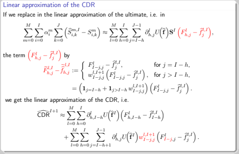
$$\sum_{m=0}^{M}\sum_{i=0}^{I}\alpha_i^m\sum_{k=0}^{J}\left(\widehat{S}_{i,k}^{m,I}-S_{i,k}^m\right)\approx\sum_{l=0}^{M}\sum_{h=0}^{I}\sum_{j=I-h}^{J-1}\partial_{h,j}^l U\Big(\widehat{\mathbf{f}}\Big)\mathbf{S}^I\left(F_{h,j}^l-\widehat{f}_j^{l,I}\right),$$

the term $\left(F_{h,j}^l-\widehat{f}_j^{l,I}\right)$ by

$$\begin{aligned}\widetilde{F}_{h,j}^{l,I}-\widetilde{\widehat{f}}_{h,j}^{l,I}&:=\begin{cases}F_{I-j,j}^l-\widehat{f}_j^{l,I}, & \text{for } j=I-h,\\ w_{I-j,j}^{l,I+1}\left(F_{I-j,j}^l-\widehat{f}_j^{l,I}\right), & \text{for } j>I-h,\end{cases}\\ &=\left(\mathbf{1}_{j=I-h}+\mathbf{1}_{j>I-h}\,w_{I-j,j}^{l,I+1}\right)\left(F_{I-j,j}^l-\widehat{f}_j^{l,I}\right).\end{aligned}$$

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

$$\begin{aligned}\widehat{\mathsf{CDR}}^{I+1}\approx&\sum_{l=0}^{M}\sum_{h=0}^{I}\partial_{h,I-h}^l U\Big(\widehat{\mathbf{f}}^I\Big)\left(F_{h,I-h}^l-\widehat{f}_{I-h}^{l,I}\right)\\ &+\sum_{l=0}^{M}\sum_{h=0}^{I}\sum_{j=I-h+1}^{J-1}\partial_{h,j}^l U\Big(\widehat{\mathbf{f}}^I\Big)w_{I-j,j}^{l,I+1}\left(F_{I-j,j}^l-\widehat{f}_j^{l,I}\right).\end{aligned}$$



Linear approximation of the CDR

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

$$\sum_{m=0}^{M}\sum_{i=0}^{I}\alpha_i^m\sum_{k=0}^{J}\left(\widehat{S}_{i,k}^{m,I}-S_{i,k}^m\right)\approx\sum_{l=0}^{M}\sum_{h=0}^{I}\sum_{j=I-h}^{J-1}\partial_{h,j}^l U\left(\widehat{\mathbf{f}}\right)\mathbf{S}^I\left(F_{h,j}^l-\widehat{f}_j^{l,I}\right),$$

the term $\left(F_{h,j}^l-\widehat{f}_j^{l,I}\right)$ by

$$\begin{aligned}\widetilde{F}_{h,j}^{l,I}-\widetilde{\widehat{f}}_{h,j}^{l,I}&:=\begin{cases}F_{I-j,j}^l-\widehat{f}_j^{l,I}, & \text{for } j=I-h,\\ w_{I-j,j}^{l,I+1}\left(F_{I-j,j}^l-\widehat{f}_j^{l,I}\right), & \text{for } j>I-h,\end{cases}\\ &=\left(\mathbf{1}_{j=I-h}+\mathbf{1}_{j>I-h}\,w_{I-j,j}^{l,I+1}\right)\left(F_{I-j,j}^l-\widehat{f}_j^{l,I}\right).\end{aligned}$$

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

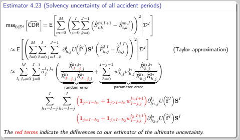
$$\begin{aligned}\widehat{\mathsf{CDR}}^{I+1}\approx&\sum_{l=0}^{M}\sum_{h=0}^{I}\partial_{h,I-h}^l U\left(\widehat{\mathbf{f}}^I\right)\left(F_{h,I-h}^l-\widehat{f}_{I-h}^{l,I}\right)\\ &+\sum_{l=0}^{M}\sum_{h=0}^{I}\sum_{j=I-h+1}^{J-1}\partial_{h,j}^l U\left(\widehat{\mathbf{f}}^I\right)w_{I-j,j}^{l,I+1}\left(F_{I-j,j}^l-\widehat{f}_j^{l,I}\right).\end{aligned}$$

The term $\widetilde{F}_{h,j}^{l,I}-\widetilde{\widehat{f}}_{h,j}^{l,I}$ depends on the accident period h only via the indicator functions $\mathbf{1}_{j=I-h}$ and $\mathbf{1}_{j>I-h}$.

Estimator 4.23 (Solvency uncertainty of all accident periods)

$$
\begin{aligned}
&\mathrm{mse}_{0|\mathcal{D}^I}\left[\widehat{\mathrm{CDR}}\right] = \mathrm{E}\left[\sum_{m=0}^{M}\left(\sum_{i=0}^{I}\sum_{k=0}^{J-1}\left(\widehat{S}^{m,I+1}_{i,k}-\widehat{S}^{m,I}_{i,k}\right)\right)^2\middle|\mathcal{D}^I\right]\\
&\approx \mathrm{E}\left[\left(\sum_{l=0}^{M}\sum_{h=0}^{I}\sum_{j=I-h}^{J-1}\partial^{l}_{h,j}U\!\left(\widehat{\mathbf{f}}^I\right)\mathbf{S}^I\left(\widetilde{F}^{l,I}_{h,j}-\widetilde{\widehat{f}}^{l,I}_{h,j}\right)\right)^2\middle|\mathcal{D}^I\right] \qquad \text{(Taylor approximation)}\\
&\approx \sum_{l_1,l_2=0}^{M}\sum_{j=0}^{J-1}\widehat{\sigma}^{l_1,l_2}_j\left(\underbrace{\frac{\widehat{R}^{l_1,l_2}_{I-j,j}}{\widehat{R}^{l_1}_{I-j,j}\widehat{R}^{l_2}_{I-j,j}}}_{\text{random error}}+\underbrace{\sum_{h=0}^{I-j-1}w^{l_1}_{h,j}w^{l_2}_{h,j}\frac{\widehat{R}^{l_1,l_2}_{h,j}}{\widehat{R}^{l_1}_{h,j}\widehat{R}^{l_2}_{h,j}}}_{\text{parameter error}}\right)\\
&\qquad\sum_{h_1=I-j}^{I}\sum_{h_2=I-j}^{I}\left(\mathbf{1}_{j=I-h_1}+\mathbf{1}_{j>I-h_1}w^{l_1,I+1}_{I-j,j}\right)\partial^{l_1}_{h_1,j}U\!\left(\widehat{\mathbf{f}}^I\right)\mathbf{S}^I\\
&\qquad\qquad\left(\mathbf{1}_{j=I-h_2}+\mathbf{1}_{j>I-h_2}w^{l_2,I+1}_{I-j,j}\right)\partial^{l_2}_{h_2,j}U\!\left(\widehat{\mathbf{f}}^I\right)\mathbf{S}^I
\end{aligned}
$$

The red terms indicate the differences to our estimator of the ultimate uncertainty.



Estimator 4.23 (Solvency uncertainty of all accident periods)

$$\begin{aligned}\mathsf{mse}_{0|\mathcal{D}^I}\left[\widehat{\mathsf{CDR}}\right] &= \mathsf{E}\left[\sum_{m=0}^{M}\left(\sum_{i=0}^{I}\sum_{k=0}^{J-1}\left(\widehat{S}_{i,k}^{m,I+1}-\widehat{S}_{i,k}^{m,I}\right)\right)^2\middle|\mathcal{D}^I\right]\\ &\approx \mathsf{E}\left[\left(\sum_{l=0}^{M}\sum_{h=0}^{I}\sum_{j=I-h}^{J-1}\partial_{h,j}^{l}U\left(\widehat{\mathbf{f}}^I\right)\mathbf{S}^I\left(\widetilde{F}_{h,j}^{l,I}-\widetilde{\widehat{f}}_{h,j}^{l,I}\right)\right)^2\middle|\mathcal{D}^I\right] \qquad \text{(Taylor approximation)}\\ &\approx \sum_{l_1,l_2=0}^{M}\sum_{j=0}^{J-1}\widehat{\sigma}_j^{l_1,l_2}\left(\underbrace{\frac{\widehat{R}_{I-j,j}^{l_1,l_2}}{\widehat{R}_{I-j,j}^{l_1}\widehat{R}_{I-j,j}^{l_2}}}_{\text{random error}}+\underbrace{\sum_{h=0}^{I-j-1}w_{h,j}^{l_1}w_{h,j}^{l_2}\frac{\widehat{R}_{h,j}^{l_1,l_2}}{\widehat{R}_{h,j}^{l_1}\widehat{R}_{h,j}^{l_2}}}_{\text{parameter error}}\right)\\ &\qquad \sum_{h_1=I-j}^{I}\sum_{h_2=I-j}^{I}\left(\mathbf{1}_{j=I-h_1}+\mathbf{1}_{j>I-h_1}w_{I-j,j}^{l_1,I+1}\right)\partial_{h_1,j}^{l_1}U\left(\widehat{\mathbf{f}}^I\right)\mathbf{S}^I\\ &\qquad \left(\mathbf{1}_{j=I-h_2}+\mathbf{1}_{j>I-h_2}w_{I-j,j}^{l_2,I+1}\right)\partial_{h_2,j}^{l_2}U\left(\widehat{\mathbf{f}}^I\right)\mathbf{S}^I\end{aligned}$$

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

$$\begin{aligned}&\mathsf{mse}_{0|\mathcal{D}^I}\left[\widehat{\mathsf{CDR}}\right]\\ &= \sum_{l_1,l_2=0}^{M}\sum_{h_1,h_2=0}^{I}\sum_{j_1=I-h_1}^{J-1}\sum_{j_2=I-h_2}^{J-1}\left(\mathbf{1}_{j_1=I-h_1}+\mathbf{1}_{j_1>I-h_1}w_{I-j_1,j_1}^{l_1,I+1}\right)\left(\mathbf{1}_{j_2=I-h_2}+\mathbf{1}_{j_2>I-h_2}w_{I-j_2,j_2}^{l_2,I+1}\right)\\ &\qquad \partial_{h_1,j_1}^{l_1}U\left(\widehat{\mathbf{f}}^I\right)\mathbf{S}^I\,\partial_{h_2,j_2}^{l_2}U\left(\widehat{\mathbf{f}}^I\right)\mathbf{S}^I\,\mathsf{E}\left[\left(F_{I-j_1,j_1}^{l_1}-\widehat{f}_{j_1}^{l_1,I}\right)\left(F_{I-j_2,j_2}^{l_2}-\widehat{f}_{j_2}^{l_2,I}\right)\middle|\mathcal{D}^I\right]\end{aligned}$$

and from the estimation of the ultimate uncertainty we know

$$\mathsf{E}\left[\left(F_{I-j_1,j_1}^{l_1}-\widehat{f}_{j_1}^{l_1,I}\right)\left(F_{I-j_2,j_2}^{l_2}-\widehat{f}_{j_2}^{l_2,I}\right)\middle|\mathcal{D}^I\right]\approx \mathbf{1}_{j_1=j_2}\widehat{\sigma}_{j_1}^{l_1,l_2}\left(\frac{\widehat{R}_{I-j,j}^{l_1,l_2}}{\widehat{R}_{I-j,j}^{l_1}\widehat{R}_{I-j,j}^{l_2}}+\sum_{h=0}^{I-j-1}w_{h,j}^{l_1}w_{h,j}^{l_2}\frac{\widehat{R}_{h,j}^{l_1,l_2}}{\widehat{R}_{h,j}^{l_1}\widehat{R}_{h,j}^{l_2}}\right).$$

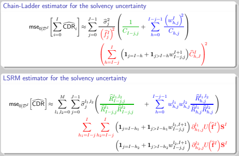
Both together lead to the stated estimator.
Note, if it wasn't for different claim properties (indeces l_1 and l_2) the last two lines of the estimator would have been a square of a sum over accident periods.
Moreover, for the random error part we had in the ultimate uncertainty case only one sum over accident periods h from $I-j$ to I, i.e. we had $h_1 = h_2$.

Chain-Ladder estimator for the solvency uncertainty

$$\mathrm{mse}_{0|\mathcal{D}^I}\left[\sum_{i=0}^{I}\widehat{\mathrm{CDR}}_i\right] \approx \sum_{j=0}^{J-1}\frac{\widehat{\sigma}_j^2}{\left(\widehat{f}_j^I\right)^2}\left(\frac{1}{C_{I-j,j}}+\sum_{h=0}^{I-j-1}\frac{\left(w_{h,j}^I\right)^2}{C_{h,j}}\right)\left(\sum_{h=I-j}^{I}\left(\mathbf{1}_{j=I-h}+\mathbf{1}_{j>I-h}w_{I-j,j}^{I+1}\right)\widehat{C}_{h,J}^I\right)^2$$

LSRM estimator for the solvency uncertainty

$$\mathrm{mse}_{0|\mathcal{D}^I}\left[\widehat{\mathrm{CDR}}\right]\approx\sum_{l_1,l_2=0}^{M}\sum_{j=0}^{J-1}\widehat{\sigma}_j^{l_1,l_2}\left(\frac{\widehat{R}^{l_1,l_2}_{I-j,j}}{\widehat{R}^{l_1}_{I-j,j}\widehat{R}^{l_2}_{I-j,j}}+\sum_{h=0}^{I-j-1}w^{l_1}_{h,j}w^{l_2}_{h,j}\frac{\widehat{R}^{l_1,l_2}_{h,j}}{\widehat{R}^{l_1}_{h,j}\widehat{R}^{l_2}_{h,j}}\right)\sum_{h_1=I-j}^{I}\sum_{h_2=I-j}^{I}\left(\mathbf{1}_{j=I-h_1}+\mathbf{1}_{j>I-h_1}w^{l_1,I+1}_{I-j,j}\right)\partial^{l_1}_{h_1,j}U\left(\widehat{\mathbf{f}}^I\right)\mathbf{S}^I\left(\mathbf{1}_{j=I-h_2}+\mathbf{1}_{j>I-h_2}w^{l_2,I+1}_{I-j,j}\right)\partial^{l_2}_{h_2,j}U\left(\widehat{\mathbf{f}}^I\right)\mathbf{S}^I$$



Chain-Ladder estimator for the solvency uncertainty

$$\mathsf{mse}_{0|\mathcal{D}^I}\left[\sum_{i=0}^{I}\widehat{\mathsf{CDR}}_i\right] \approx \sum_{j=0}^{J-1}\frac{\widehat{\sigma}_j^2}{\left(\widehat{f}_j^I\right)^2}\left(\frac{1}{C_{I-j,j}} + \sum_{h=0}^{I-j-1}\frac{\left(w_{h,j}^I\right)^2}{C_{h,j}}\right)$$
$$\left(\sum_{h=I-j}^{I}\left(\mathbf{1}_{j=I-h} + \mathbf{1}_{j>I-h}w_{I-j,j}^{I+1}\right)\widehat{C}_{h,J}^I\right)^2$$

LSRM estimator for the solvency uncertainty

$$\mathsf{mse}_{0|\mathcal{D}^I}\left[\widehat{\mathsf{CDR}}\right] \approx \sum_{l_1,l_2=0}^{M}\sum_{j=0}^{J-1}\widehat{\sigma}_j^{l_1,l_2}\left(\frac{\widehat{R}_{I-j,j}^{l_1,l_2}}{\widehat{R}_{I-j,j}^{l_1}\widehat{R}_{I-j,j}^{l_2}} + \sum_{h=0}^{I-j-1}w_{h,j}^{l_1}w_{h,j}^{l_2}\frac{\widehat{R}_{h,j}^{l_1,l_2}}{\widehat{R}_{h,j}^{l_1}\widehat{R}_{h,j}^{l_2}}\right)$$
$$\sum_{h_1=I-j}^{I}\sum_{h_2=I-j}^{I}\left(\mathbf{1}_{j=I-h_1} + \mathbf{1}_{j>I-h_1}w_{I-j,j}^{l_1,I+1}\right)\partial_{h_1,j}^{l_1}U\left(\widehat{\mathbf{f}}^I\right)\mathbf{S}^I$$
$$\left(\mathbf{1}_{j=I-h_2} + \mathbf{1}_{j>I-h_2}w_{I-j,j}^{l_2,I+1}\right)\partial_{h_2,j}^{l_2}U\left(\widehat{\mathbf{f}}^I\right)\mathbf{S}^I$$

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

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

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

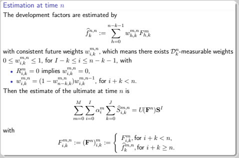
- $R_{i,k}^m = 0$ implies $w_{i,k}^{m,n} = 0$,
- $w_{i,k}^{m,n} = (1 - w_{n-k,k}^{m,n}) w_{i,k}^{m,n-1}$, for $i + k < n$.

Then the estimate of the ultimate at time n is

$$\sum_{m=0}^{M} \sum_{i=0}^{I} \alpha_i^m \sum_{k=0}^{J} \widehat{S}_{i,k}^{m,n} = U(\mathbf{F}^n)\mathbf{S}^I$$

with

$$F_{i,k}^{m,n} := (\mathbf{F}^n)_{i,k}^m := \begin{cases} F_{i,k}^m, \text{for } i + k < n, \\ \widehat{f}_k^{m,n}, \text{for } i + k \geq n. \end{cases}$$



Estimation at time n

The development factors are estimated by

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

with consistent future weights $w_{i,k}^{m,n}$, which means there exists $\mathcal{D}_k^n$-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

$$\sum_{m=0}^{M} \sum_{i=0}^{I} \alpha_i^m \sum_{k=0}^{J} \widehat{S}_{i,k}^{m,n} = U(\mathbf{F}^n)\mathbf{S}^I$$

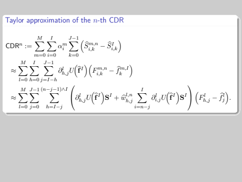
with

$$F_{i,k}^{m,n} := (\mathbf{F}^n)_{i,k}^m := \begin{cases} F_{i,k}^m, \text{for } i + k < n, \\ \widehat{f}_k^{m,n}, \text{for } i + k \geq n. \end{cases}$$

- 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 $\mathcal{D}^I$ measurable, we will consider them as constant in our estimations.

Taylor approximation of the n-th CDR

$$\begin{aligned}
\mathsf{CDR}^n &:= \sum_{m=0}^{M} \sum_{i=0}^{I} \alpha_i^m \sum_{k=0}^{J-1} \left(\widehat{S}_{i,k}^{m,n} - \widehat{S}_{i,k}^{I} \right) \\
&\approx \sum_{l=0}^{M} \sum_{h=0}^{I} \sum_{j=I-h}^{J-1} \partial_{h,j}^l U\Big(\widehat{\mathbf{f}}^I\Big) \Big(F_{i,k}^{m,n} - \widehat{f}_k^{m,I} \Big) \\
&\approx \sum_{l=0}^{M} \sum_{j=0}^{J-1} \sum_{h=I-j}^{(n-j-1)\wedge I} \left(\partial_{h,j}^l U\Big(\widehat{\mathbf{f}}^I\Big) \mathbf{S}^I + \widehat{w}_{h,j}^{l,n} \sum_{i=n-j}^{I} \partial_{i,j}^l U\Big(\widehat{\mathbf{f}}^I\Big) \mathbf{S}^I \right) \Big(F_{h,j}^l - \widehat{f}_j^l \Big).
\end{aligned}$$



Taylor approximation of the n-th CDR

$$\mathsf{CDR}^n := \sum_{m=0}^{M}\sum_{i=0}^{I}\alpha_i^m\sum_{k=0}^{J-1}\left(\widehat{S}_{i,k}^{m,n}-\widehat{S}_{i,k}^{I}\right)$$
$$\approx\sum_{l=0}^{M}\sum_{h=0}^{I}\sum_{j=I-h}^{J-1}\partial_{h,j}^{l}U\left(\widehat{\mathbf{f}}^{I}\right)\left(F_{i,k}^{m,n}-\widehat{f}_k^{m,I}\right)$$
$$\approx\sum_{l=0}^{M}\sum_{j=0}^{J-1}\sum_{h=I-j}^{(n-j-1)\wedge I}\left(\partial_{h,j}^{l}U\left(\widehat{\mathbf{f}}^{I}\right)\mathbf{S}^{I}+\widehat{w}_{h,j}^{l,n}\sum_{i=n-j}^{I}\partial_{i,j}^{l}U\left(\widehat{\mathbf{f}}^{I}\right)\mathbf{S}^{I}\right)\left(F_{h,j}^{l}-\widehat{f}_j^{l}\right).$$

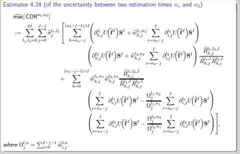
Here we used that

$$\widehat{f}_k^{m,n} := \sum_{i=0}^{n-k-1} w_{i,k}^{m,n} F_{i,k}^{m} = \sum_{i=I-k}^{n-k-1} w_{i,k}^{m,n} F_{i,k}^{m} + \left(1-\sum_{i=I-k}^{n-k-1} w_{i,k}^{m,n}\right)\widehat{f}_k^{m,I}.$$

Estimator 4.24 (of the uncertainty between two estimation times n_1 and n_2)

$$\begin{aligned}
&\widehat{\mathsf{mse}}\left[\,\mathsf{CDR}^{n_1,n_2}\right] \\
&:= \sum_{l_1,l_2=0}^{M} \sum_{j=0}^{J-1} \widehat{\sigma}_j^{l_1,l_2} \Bigg[\sum_{h=n_1-j}^{(n_2-j-1)\wedge I} \left(\partial_{h,j}^{l_1} U\left(\widehat{\mathbf{f}}^I\right)\mathbf{S}^I + \widehat{w}_{h,j}^{l_1,n_2} \sum_{i=n_2-j}^{I} \partial_{i,j}^{l_1} U\left(\widehat{\mathbf{f}}^I\right)\mathbf{S}^I \right) \\
&\qquad\qquad \left(\partial_{h,j}^{l_2} U\left(\widehat{\mathbf{f}}^I\right)\mathbf{S}^I + \widehat{w}_{h,j}^{l_2,n_2} \sum_{i=n_2-j}^{I} \partial_{i,j}^{l_2} U\left(\widehat{\mathbf{f}}^I\right)\mathbf{S}^I \right) \frac{\widehat{R}_{h,j}^{l_1,l_2,I}}{\widehat{R}_{h,j}^{l_1,I}\,\widehat{R}_{h,j}^{l_2,I}} \\
&\qquad + \sum_{h=0}^{(n_1-j-1)\wedge I} \widehat{w}_{h,j}^{l_1,n_1}\,\widehat{w}_{h,j}^{l_2,n_1} \frac{\widehat{R}_{h,j}^{l_1,l_2,I}}{\widehat{R}_{h,j}^{l_1,I}\,\widehat{R}_{h,j}^{l_2,I}} \\
&\qquad\qquad \left(\sum_{i=n_1-j}^{I} \partial_{i,j}^{l_1} U\left(\widehat{\mathbf{f}}^I\right)\mathbf{S}^I - \frac{\Omega_j^{l_1,n_2}}{\Omega_j^{l_1,n_1}} \sum_{i=n_2-j}^{I} \partial_{i,j}^{l_1} U\left(\widehat{\mathbf{f}}^I\right)\mathbf{S}^I \right) \\
&\qquad\qquad \left(\sum_{i=n_1-j}^{I} \partial_{i,j}^{l_2} U\left(\widehat{\mathbf{f}}^I\right)\mathbf{S}^I - \frac{\Omega_j^{l_2,n_2}}{\Omega_j^{l_2,n_1}} \sum_{i=n_2-j}^{I} \partial_{i,j}^{l_2} U\left(\widehat{\mathbf{f}}^I\right)\mathbf{S}^I \right) \Bigg],
\end{aligned}$$

where $\Omega_j^{l,n} = \sum_{i=0}^{I-j-1} \widehat{w}_{i,j}^{l,n}$.



Estimator 4.24 (of the uncertainty between two estimation times n_1 and n_2)

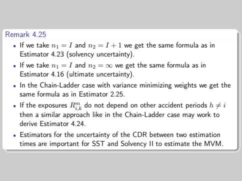
$$
\begin{aligned}
&\widehat{\mathsf{mse}}[\,\mathsf{CDR}^{n_1,n_2}]\\
&:= \sum_{l_1,l_2=0}^{M}\sum_{j=0}^{J-1}\widehat{\sigma}_j^{l_1,l_2}\Bigg[\sum_{h=n_1-j}^{(n_2-j-1)\wedge I}\Bigg(\partial_{h,j}^{l_1}U\Big(\widehat{\mathbf{f}}^I\Big)\mathbf{S}^I+\widehat{w}_{h,j}^{l_1,n_2}\sum_{i=n_2-j}^{I}\partial_{i,j}^{l_1}U\Big(\widehat{\mathbf{f}}^I\Big)\mathbf{S}^I\Bigg)\\
&\qquad\Bigg(\partial_{h,j}^{l_2}U\Big(\widehat{\mathbf{f}}^I\Big)\mathbf{S}^I+\widehat{w}_{h,j}^{l_2,n_2}\sum_{i=n_2-j}^{I}\partial_{i,j}^{l_2}U\Big(\widehat{\mathbf{f}}^I\Big)\mathbf{S}^I\Bigg)\frac{\widehat{R}_{h,j}^{l_1,l_2,I}}{\widehat{R}_{h,j}^{l_1,I}\widehat{R}_{h,j}^{l_2,I}}\\
&\quad+\sum_{h=0}^{(n_1-j-1)\wedge I}\widehat{w}_{h,j}^{l_1,n_1}\widehat{w}_{h,j}^{l_2,n_1}\frac{\widehat{R}_{h,j}^{l_1,l_2,I}}{\widehat{R}_{h,j}^{l_1,I}\widehat{R}_{h,j}^{l_2,I}}\\
&\qquad\Bigg(\sum_{i=n_1-j}^{I}\partial_{i,j}^{l_1}U\Big(\widehat{\mathbf{f}}^I\Big)\mathbf{S}^I-\frac{\Omega_j^{l_1,n_2}}{\Omega_j^{l_1,n_1}}\sum_{i=n_2-j}^{I}\partial_{i,j}^{l_1}U\Big(\widehat{\mathbf{f}}^I\Big)\mathbf{S}^I\Bigg)\\
&\qquad\Bigg(\sum_{i=n_1-j}^{I}\partial_{i,j}^{l_2}U\Big(\widehat{\mathbf{f}}^I\Big)\mathbf{S}^I-\frac{\Omega_j^{l_2,n_2}}{\Omega_j^{l_2,n_1}}\sum_{i=n_2-j}^{I}\partial_{i,j}^{l_2}U\Big(\widehat{\mathbf{f}}^I\Big)\mathbf{S}^I\Bigg)\Bigg],
\end{aligned}
$$

where $\Omega_j^{l,n}=\sum_{i=0}^{I-j-1}\widehat{w}_{i,j}^{l,n}$.

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: $R^0_{i,k} = R^{0,0}_{i,k} = \sum_{j=0}^k S^0_{i,j}$

IW: $R^1_{i,k} = R^{1,1}_{i,k} = \sum_{j=0}^k \left(\kappa^{j+1} S^3_{i,j} + (1-\kappa^{j+1}) S^1_{i,j}\right)$

Sub: $R^2_{i,k} = R^{2,2}_{i,k} = \sum_{j=0}^k \left(S^0_{i,j} + S^1_{i,j} + S^2_{i,j}\right)$

Salary: $S^3_{i,0} = P_i$, $S^3_{i,j} = 0$, for $j > 0$, and

$$R^3_{i,k} = R^{3,0}_{i,k} = R^{0,3}_{i,k} = R^{3,1}_{i,k} = R^{1,3}_{i,k} = R^{3,2}_{i,k} = R^{2,3}_{i,k} = R^{3,3}_{i,k} = 0$$

For the not yet defined exposures we take the total payments up to now, i.e.

$$R^{0,1}_{i,k} = R^{1,0}_{i,k} = R^{0,2}_{i,k} = R^{2,0}_{i,k} = R^{1,2}_{i,k} = R^{2,1}_{i,k} = \sum_{j=0}^k \left(S^0_{i,j} + S^1_{i,j} + S^2_{i,j}\right).$$

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 (small blue circle on the left).

Swiss mandatory accident portfolio: part 3 of 3, see Example 4.13

We have four claim properties with exposures

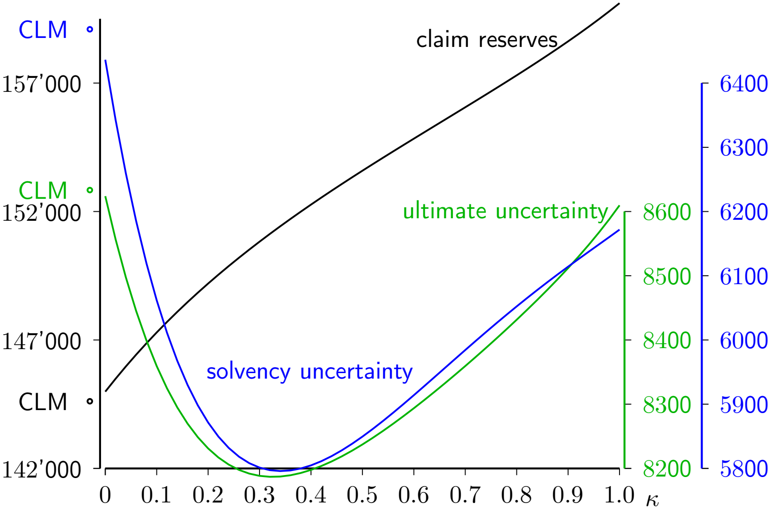
ME: $R^0_{i,k} = R^{0,0}_{i,k} = \sum_{j=0}^k S^0_{i,j}$
IW: $R^1_{i,k} = R^{1,1}_{i,k} = \sum_{j=0}^k \left(\kappa^{j+1} S^3_{i,j} + (1-\kappa^{j+1}) S^1_{i,j}\right)$
Sub: $R^2_{i,k} = R^{2,2}_{i,k} = \sum_{j=0}^k \left(S^0_{i,j} + S^1_{i,j} + S^2_{i,j}\right)$
Salary: $S^3_{i,0} = P_i$, $S^3_{i,j} = 0$, for $j > 0$, and
$R^3_{i,k} = R^{3,0}_{i,k} = R^{0,3}_{i,k} = R^{3,1}_{i,k} = R^{1,3}_{i,k} = R^{3,2}_{i,k} = R^{2,3}_{i,k} = R^{3,3}_{i,k} = 0$

For the not yet defined exposures we take the total payments up to now, i.e.
$R^{0,1}_{i,k} = R^{1,0}_{i,k} = R^{0,2}_{i,k} = R^{2,0}_{i,k} = R^{1,2}_{i,k} = R^{2,1}_{i,k} = \sum_{j=0}^k \left(S^0_{i,j} + S^1_{i,j} + S^2_{i,j}\right)$.

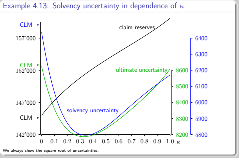
Resulting solvency uncertainty

- 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 (small blue circle on the left).

Example 4.13: Solvency uncertainty in dependence of κ



We always show the square root of uncertainties.



Example 4.13: Solvency uncertainty in dependence of κ

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.

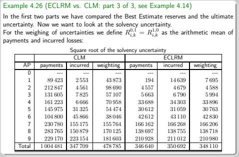
In our example the solvency uncertainty is about 70% of the ultimate uncertainty. In general this ratio usually lies between 50% (general liability) and 90% (NatCat). One minus this ratio represents the gain of information over one year in comparission to all unknown information about the reserves.

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 define $R_{i,k}^{0,1} = R_{i,k}^{1,0}$ as the arithmetic mean of payments and incurred losses:

Square root of the solvency uncertainty

	CLM			ECLRM		
AP	payments	incurred	weighting	payments	incurred	weighting
0	---	---	---	---	---	---
1	89 423	2 553	43 873	194	14 639	7 695
2	212 847	4 561	98 690	4 557	4 679	4 588
3	131 605	7 825	57 107	5 663	6 790	5 994
4	161 223	6 666	70 958	33 688	34 303	33 896
5	145 975	31 325	54 474	30 612	31 059	30 763
6	104 800	45 866	38 046	42 612	43 110	42 830
7	230 780	155 175	155 764	166 162	166 268	166 206
8	283 765	150 879	170 125	138 697	138 755	138 718
9	229 170	223 154	181 603	210 928	211 012	210 980
Total	1 004 481	347 709	478 785	346 640	350 692	348 110



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 define $R_{i,k}^{0,1} = R_{i,k}^{1,0}$ as the arithmetic mean of payments and incurred losses:

Square root of the solvency uncertainty

	CLM			ECLRM		
AP	payments	incurred	weighting	payments	incurred	weighting
0	---	---	---	---	---	---
1	89 423	2 553	43 873	194	14 639	7 695
2	212 847	4 561	98 690	4 557	4 679	4 588
3	131 605	7 825	57 107	5 663	6 790	5 994
4	161 223	6 666	70 958	33 688	34 303	33 896
5	145 975	31 325	54 474	30 612	31 059	30 763
6	104 800	45 866	38 046	42 612	43 110	42 830
7	230 780	155 175	155 764	166 162	166 268	166 206
8	283 765	150 879	170 125	138 697	138 755	138 718
9	229 170	223 154	181 603	210 928	211 012	210 980
Total	1 004 481	347 709	478 785	346 640	350 692	348 110

- 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

$$R_{i,k}^{m_1,m_2} := \sqrt{R_{i,k}^{m_1,m_1}\,R_{i,k}^{m_2,m_2}}$$

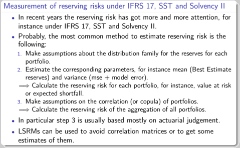
  usually works better.
- Although the Best Estimate reserves are similar, the solvency uncertainties are not, in particular CLM on payments leads to a much higher solvency uncertainty than the others.
- Again, you must not use estimates of the ultimate uncertainty to evaluate which model is the best.

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).

  $\Longrightarrow$ 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.

  $\Longrightarrow$ 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.



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).

  $\Longrightarrow$ 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.

  $\Longrightarrow$ 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.

Part of the correlation matrix of the SST-Standardmodel 2014

	MFH	MFK	Sach	ES-Pool	Haft	UVG	
MFH	1.00	0.15	0.15	0.15	0.25	0.50	
MFK	0.15	1.00	0.15	0.15	0.15	0.15	
Sach	0.15	0.15	1.00	0.15	0.15	0.15	⋯
ES-Pool	0.15	0.15	0.15	1.00	0.15	0.15	
Haft	0.25	0.15	0.15	0.15	1.00	0.25	
UVG	0.50	0.15	0.15	0.15	0.25	1.00	
			⋮				⋱

- The entries are based on actuarial judgement.
- The correlation matrix under Solvency II contains similar entries.

Part of the correlation matrix of the SST-Standardmodel 2014

	MFH	MFK	Sach	ES-Pool	Haft	UVG	
MFH	1.00	0.15	0.15	0.15	0.25	0.50	
MFK	0.15	1.00	0.15	0.15	0.15	0.15	
Sach	0.15	0.15	1.00	0.15	0.15	0.15	⋯
ES-Pool	0.15	0.15	0.15	1.00	0.15	0.15	
Haft	0.25	0.15	0.15	0.15	1.00	0.25	
UVG	0.50	0.15	0.15	0.15	0.25	1.00	
			⋮				⋱

- 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. Chose a distribution family for the total reserve of all portfolios.
3. Estimate the corresponding parameter, for instance

$$\begin{aligned} \text{mean} &= \text{Best Estimate reserves and} \\ \text{variance} &= \text{ultimate or solvency uncertainty} + \text{model error.} \end{aligned}$$

   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.

$\Longrightarrow$ 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 $R_{i,k}^{m_1,m_2}$ 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. Chose a distribution family for the total reserve of all portfolios.
3. Estimate the corresponding parameter, for instance

   $$\begin{aligned} \text{mean} &= \text{Best Estimate reserves and} \\ \text{variance} &= \text{ultimate or solvency uncertainty} + \text{model error.} \end{aligned}$$

   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.

$\Longrightarrow$ Calculate the reserving risk.

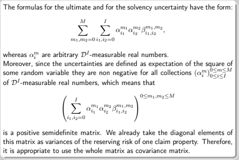
The formulas for the ultimate and for the solvency uncertainty have the form:

$$\sum_{m_1,m_2=0}^{M} \sum_{i_1,i_2=0}^{I} \alpha_{i_1}^{m_1} \alpha_{i_2}^{m_2} \beta_{i_1,i_2}^{m_1,m_2},$$

whereas α_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 $(\alpha_i^m)_{0\leq i\leq I}^{0\leq m\leq M}$ of $\mathcal{D}^I$-measurable real numbers, which means that

$$\left(\sum_{i_1,i_2=0}^{I} \alpha_{i_1}^{m_1} \alpha_{i_2}^{m_2} \beta_{i_1,i_2}^{m_1,m_2} \right)^{0\leq m_1,m_2\leq M}$$

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:

$$\sum_{m_1,m_2=0}^{M}\sum_{i_1,i_2=0}^{I}\alpha_{i_1}^{m_1}\alpha_{i_2}^{m_2}\beta_{i_1,i_2}^{m_1,m_2},$$

whereas α_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 $(\alpha_i^m)_{0\leq i\leq I}^{0\leq m\leq M}$ of $\mathcal{D}^I$-measurable real numbers, which means that

$$\left(\sum_{i_1,i_2=0}^{I}\alpha_{i_1}^{m_1}\alpha_{i_2}^{m_2}\beta_{i_1,i_2}^{m_1,m_2}\right)^{0\leq m_1,m_2\leq M}$$

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 $\boldsymbol{x} = (x_m)_{0\leq m\leq M}$ we get

$$\begin{aligned}\boldsymbol{x}'\left(\sum_{i_1,i_2=0}^{I}\alpha_{i_1}^{m_1}\alpha_{i_2}^{m_2}\beta_{i_1,i_2}^{m_1,m_2}\right)^{0\leq m_1,m_2\leq M}\boldsymbol{x} &= \sum_{m_1,m_2=0}^{M} x_{m_1}\sum_{i_1,i_2=0}^{I}\alpha_{i_1}^{m_1}\alpha_{i_2}^{m_2}\beta_{i_1,i_2}^{m_1,m_2}x_{m_2}\\ &= \sum_{m_1,m_2=0}^{M}\sum_{i_1,i_2=0}^{I}\alpha_{i_1}^{m_1}x_{m_1}\alpha_{i_2}^{m_2}x_{m_2}\beta_{i_1,i_2}^{m_1,m_2}\geq 0.\end{aligned}$$

Estimating correlation of reserve risk, see LSRM_Cor_ActiveX.xlsx

Based on the example of article [28] by A. Gisler and M. Wüthrich with

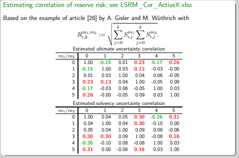
$$R_{i,k}^{m_1,m_2} := \sqrt{\sum_{j=0}^{k} S_{i,j}^{m_1} \sum_{j=0}^{k} S_{i,j}^{m_2}}.$$

Estimated ultimate uncertainty correlation

m_1/m_2	0	1	2	3	4	5
0	1.00	**-0.15**	0.01	**0.23**	**-0.17**	**0.26**
1	**-0.15**	1.00	0.03	**0.13**	-0.03	-0.00
2	0.01	0.03	1.00	0.04	0.06	-0.05
3	**0.23**	**0.13**	0.04	1.00	-0.05	0.09
4	**-0.17**	-0.03	0.06	-0.05	1.00	0.03
5	**0.26**	-0.00	-0.05	0.09	0.03	1.00

Estimated solvency uncertainty correlation

m_1/m_2	0	1	2	3	4	5
0	1.00	0.04	0.05	**0.30**	**-0.26**	**0.31**
1	0.04	1.00	0.04	**0.30**	-0.10	0.00
2	0.05	0.04	1.00	0.09	0.08	-0.06
3	**0.30**	**0.30**	0.09	1.00	-0.08	**0.16**
4	**-0.26**	-0.10	0.08	-0.08	1.00	0.03
5	**0.31**	0.00	-0.06	**0.16**	0.03	1.00



Estimating correlation of reserve risk, see LSRM_Cor_ActiveX.xlsx

Based on the example of article [28] by A. Gisler and M. Wüthrich with

$$R_{i,k}^{m_1,m_2} := \sqrt{\sum_{j=0}^{k} S_{i,j}^{m_1} \sum_{j=0}^{k} S_{i,j}^{m_2}}.$$

Estimated ultimate uncertainty correlation

m_1/m_2	0	1	2	3	4	5
0	1.00	-0.15	0.01	0.23	-0.17	0.26
1	-0.15	1.00	0.03	0.13	-0.03	-0.00
2	0.01	0.03	1.00	0.04	0.06	-0.05
3	0.23	0.13	0.04	1.00	-0.05	0.09
4	-0.17	-0.03	0.06	-0.05	1.00	0.03
5	0.26	-0.00	-0.05	0.09	0.03	1.00

Estimated solvency uncertainty correlation

m_1/m_2	0	1	2	3	4	5
0	1.00	0.04	0.05	0.30	-0.26	0.31
1	0.04	1.00	0.04	0.30	-0.10	0.00
2	0.05	0.04	1.00	0.09	0.08	-0.06
3	0.30	0.30	0.09	1.00	-0.08	0.16
4	-0.26	-0.10	0.08	-0.08	1.00	0.03
5	0.31	0.00	-0.06	0.16	0.03	1.00

- 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

$$S_{i,k}^{5} \text{ vs. } S_{i,k}^{0} \qquad \text{and} \qquad S_{i,k}^{3} \text{ vs. } S_{i,k}^{0},\ S_{i,k}^{1} \text{ and } S_{i,k}^{5}$$

  and some diversification related to

$$S_{i,k}^{0} \text{ vs. } S_{i,k}^{4} \qquad \text{and maybe} \qquad S_{i,k}^{0} \text{ vs. } S_{i,k}^{1} \text{ and } S_{i,k}^{4}.$$

- Strictly taken, the model is not valid, because of some negative eigenvalues of the covariance matrices $(\widehat{\sigma}_k^{m_1,m_2} R_{i,k}^{m_1,m_2})^{0 \leq m_1,m_2 \leq M}$ 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.

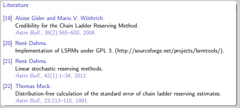
Literature

[19] Aloise Gisler and Mario V. Wüthrich.
Credibility for the Chain Ladder Reserving Method.
Astin Bull., 38(2):565–600, 2008.

[20] René Dahms.
Implementation of LSRMs under GPL 3. (http://sourceforge.net/projects/lsrmtools/).

[21] René Dahms.
Linear stochastic reserving methods.
Astin Bull., 42(1):1–34, 2012.

[22] Thomas Mack.
Distribution-free calculation of the standard error of chain ladder reserving estimates.
Astin Bull., 23:213–115, 1991.



Literature

[19] Aloise Gisler and Mario V. Wüthrich.
Credibility for the Chain Ladder Reserving Method.
Astin Bull., 38(2):565–600, 2008.

[20] René Dahms.
Implementation of LSRMs under GPL 3. (http://sourceforge.net/projects/lsrmtools/).

[21] René Dahms.
Linear stochastic reserving methods.
Astin Bull., 42(1):1–34, 2012.

[22] Thomas Mack.
Distribution-free calculation of the standard error of chain ladder reserving estimates.
Astin Bull., 23:213–115, 1991.

Stochastic Reserving

Lecture 9

Poission-Model

René Dahms

ETH Zurich, Spring 2021

28 April 2021

(Last update: 26 April 2021)

Stochastic Reserving

Lecture 9

Poission-Model

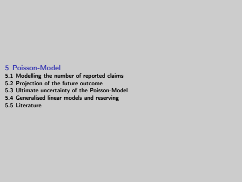
René Dahms

ETH Zurich, Spring 2021

28 April 2021
(Last update: 26 April 2021)

5 Poisson-Model

Lecture 9: Table of contents

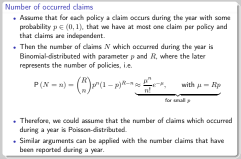


Number of occurred claims

- 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 claims N which occurred during the year is Binomial-distributed with parameter p and R, where the later represents the number of policies, i.e.

$$\mathsf{P}\left(N = n\right) = \binom{R}{n} p^n (1-p)^{R-n} \underbrace{\approx \frac{\mu^n}{n!} e^{-\mu}, \qquad \text{with } \mu = Rp}_{\text{for small } p}$$

- Therefore, we could assume that the number of claims which occurred during a year is Poisson-distributed.
- Similar arguments can be applied with the number claims that have been reported during a year.



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 claims N which occurred during the year is Binomial-distributed with parameter p and R, where the later represents the number of policies, i.e.

$$\mathsf{P}\left(N=n\right)=\binom{R}{n}p^n(1-p)^{R-n}\underbrace{\approx\frac{\mu^n}{n!}e^{-\mu},\qquad\text{with } \mu=Rp}_{\text{for small } p}$$

- Therefore, we could assume that the number of claims which occurred during a year is Poisson-distributed.
- Similar arguments can be applied with the number claims that have been reported during a year.

Assumption 5.A (Poisson-Model)

Assume that there are parameters $\mu_0, \ldots, \mu_I > 0$ *and* $\gamma_0, \ldots, \gamma_J > 0$ *such that*

i)$^{\text{Poi}}$ $S_{i,k}$ *are independent Poisson-distributed random variables with*

$$\mathsf{E}\left[S_{i,k}\right] = \gamma_k \mu_i.$$

ii)$^{\text{Poi}}$ $\sum_{k=0}^{J} \gamma_k = 1.$

Remark 5.1

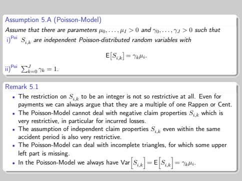
- The restriction on $S_{i,k}$ to be an integer is not so restrictive at all. Even for payments we can always argue that they are a multiple of one Rappen or Cent.
- The Poisson-Model cannot deal with negative claim properties $S_{i,k}$ which is very restrictive, in particular for incurred losses.
- 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 $\mathsf{Var}\left[S_{i,k}\right] = \mathsf{E}\left[S_{i,k}\right] = \gamma_k \mu_i$.

2021-04-26

Stochastic Reserving

Poisson-Model

Modelling the number of reported claims



Assumption 5.A (Poisson-Model)

Assume that there are parameters $\mu_0, \dots, \mu_I > 0$ *and* $\gamma_0, \dots, \gamma_J > 0$ *such that*

i)$^{\text{Poi}}$ $S_{i,k}$ *are independent Poisson-distributed random variables with*

$$\mathsf{E}[S_{i,k}] = \gamma_k \mu_i.$$

ii)$^{\text{Poi}}$ $\sum_{k=0}^{J} \gamma_k = 1.$

Remark 5.1

- 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 Rappen or 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 $\mathsf{Var}\left[S_{i,k}\right] = \mathsf{E}\left[S_{i,k}\right] = \gamma_k \mu_i$.

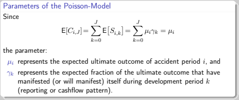
Parameters of the Poisson-Model

Since

$$\mathsf{E}[C_{i,J}] = \sum_{k=0}^{J} \mathsf{E}\left[S_{i,k}\right] = \sum_{k=0}^{J} \mu_i \gamma_k = \mu_i$$

the parameter:

- μ_i represents the expected ultimate outcome of accident period i, and
- γ_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

$$\mathsf{E}[C_{i,J}] = \sum_{k=0}^{J} \mathsf{E}\left[S_{i,k}\right] = \sum_{k=0}^{J} \mu_i \gamma_k = \mu_i$$

the parameter:

- μ_i represents the expected ultimate outcome of accident period i, and
- γ_k represents the expected fraction of the ultimate outcome that have manifested (or will manifest) itself during development period k (reporting or cashflow pattern).

Probability of the observed triangle

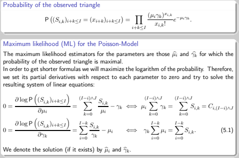
$$\mathsf{P}\left((S_{i,k})_{i+k\leq I} = (x_{i+k})_{i+k\leq I}\right) = \prod_{i+k\leq I} \frac{(\mu_i\gamma_k)^{x_{i,k}}}{x_{i,k}!} e^{-\mu_i\gamma_k}.$$

Maximum likelihood (ML) for the Poisson-Model

The maximum likelihood estimators for the parameters are those $\widehat{\mu_i}$ and $\widehat{\gamma_k}$ 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:

$$\begin{aligned}
0 &= \frac{\partial \log \mathsf{P}\left((S_{i,k})_{i+k\leq I}\right)}{\partial \mu_i} = \sum_{k=0}^{(I-i)\wedge J} \frac{S_{i,k}}{\mu_i} - \gamma_k \iff \mu_i \sum_{k=0}^{(I-i)\wedge J} \gamma_k = \sum_{k=0}^{(I-i)\wedge J} S_{i,k} = C_{i,(I-i)\wedge J} \\
0 &= \frac{\partial \log \mathsf{P}\left((S_{i,k})_{i+k\leq I}\right)}{\partial \gamma_k} = \sum_{i=0}^{I-k} \frac{S_{i,k}}{\gamma_k} - \mu_i \iff \gamma_k \sum_{i=0}^{I-k} \mu_i = \sum_{i=0}^{I-k} S_{i,k}. \qquad (5.1)
\end{aligned}$$

We denote the solution (if it exists) by $\widehat{\mu_i}$ and $\widehat{\gamma_k}$.



Probability of the observed triangle

$$\mathsf{P}\left((S_{i,k})_{i+k\le I} = (x_{i+k})_{i+k\le I}\right) = \prod_{i+k\le I} \frac{(\mu_i\gamma_k)^{x_{i,k}}}{x_{i,k}!} e^{-\mu_i\gamma_k}.$$

Maximum likelihood (ML) for the Poisson-Model

The maximum likelihood estimators for the parameters are those $\widehat{\mu}_i$ and $\widehat{\gamma}_k$ 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:

$$\begin{aligned} 0 &= \frac{\partial \log \mathsf{P}\left((S_{i,k})_{i+k\le I}\right)}{\partial \mu_i} = \sum_{k=0}^{(I-i)\wedge J} \frac{S_{i,k}}{\mu_i} - \gamma_k \iff \mu_i \sum_{k=0}^{(I-i)\wedge J} \gamma_k = \sum_{k=0}^{(I-i)\wedge J} S_{i,k} = C_{i,(I-i)\wedge J} \\ 0 &= \frac{\partial \log \mathsf{P}\left((S_{i,k})_{i+k\le I}\right)}{\partial \gamma_k} = \sum_{i=0}^{I-k} \frac{S_{i,k}}{\gamma_k} - \mu_i \iff \gamma_k \sum_{i=0}^{I-k} \mu_i = \sum_{i=0}^{I-k} S_{i,k}. \end{aligned} \tag{5.1}$$

We denote the solution (if it exists) by $\widehat{\mu}_i$ and $\widehat{\gamma}_k$.

One can proof that if the observed data are not too strange then there exists an unique solution of (5.1), which represents a maximum.
An example for ‘too strange’ is $S_{i,k} = 0$ for all observed accident and development periods.

Estimator 5.2 (for the future outcome within the Poisson-Model)

$$\widehat{S}_{i,k}^{\mathsf{Poi}} := \widehat{\mathsf{E}}\big[S_{i,k}\big] := \widehat{\mu}_i\widehat{\gamma}_k$$

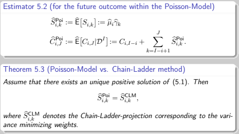
$$\widehat{C}_{i,J}^{\mathsf{Poi}} := \widehat{\mathsf{E}}\big[C_{i,J}|\mathcal{D}^I\big] := C_{i,I-i} + \sum_{k=I-i+1}^{J} \widehat{S}_{i,k}^{\mathsf{Poi}}.$$

Theorem 5.3 (Poisson-Model vs. Chain-Ladder method)

Assume that there exists an unique positive solution of (5.1). *Then*

$$\widehat{S}_{i,k}^{\mathsf{Poi}} = \widehat{S}_{i,k}^{\mathsf{CLM}},$$

where $\widehat{S}_{i,k}^{\mathsf{CLM}}$ *denotes the Chain-Ladder-projection corresponding to the variance minimizing weights.*



Estimator 5.2 (for the future outcome within the Poisson-Model)

$$\widehat{S}^{\mathsf{Poi}}_{i,k} := \widehat{\mathsf{E}}[S_{i,k}] := \widehat{\mu}_i\widehat{\gamma}_k$$

$$\widehat{C}^{\mathsf{Poi}}_{i,J} := \widehat{\mathsf{E}}[C_{i,J}|\mathcal{D}^I] := C_{i,I-i} + \sum_{k=I-i+1}^{J} \widehat{S}^{\mathsf{Poi}}_{i,k}.$$

Theorem 5.3 (Poisson-Model vs. Chain-Ladder method)

Assume that there exists an unique positive solution of (5.1). *Then*

$$\widehat{S}^{\mathsf{Poi}}_{i,k} = \widehat{S}^{\mathsf{CLM}}_{i,k},$$

where $\widehat{S}^{\mathsf{CLM}}_{i,k}$ denotes the Chain-Ladder-projection corresponding to the variance minimizing weights.

Lemma

$$\sum_{i=0}^{I-k} C_{i,k} = \sum_{i=0}^{I-k} \widehat{\mu}_i \sum_{j=0}^{k} \widehat{\gamma}_j \quad \textbf{and} \quad \sum_{i=0}^{I-k} C_{i,k-1} = \sum_{i=0}^{I-k} (C_{i,k} - S_{i,k}) = \sum_{i=0}^{I-k} \widehat{\mu}_i \sum_{j=0}^{k-1} \widehat{\gamma}_j, \qquad \textbf{for all } i+k \le I$$

Proof of the above lemma (by induction): Start with $k = J$

$$\sum_{i=0}^{I-J} C_{i,J} = \sum_{i=0}^{I-J} \sum_{j=0}^{J} S_{i,j} = \sum_{i=0}^{I-J} \widehat{\mu}_i \underbrace{\sum_{j=0}^{J} \widehat{\gamma}_j}_{\textbf{(5.1)}}.$$

Now assume that the lemma is true for some $k > 0$ then we get

$$\begin{aligned}\sum_{i=0}^{I-(k-1)} C_{i,k-1} &= \sum_{i=0}^{I-k} C_{i,k-1} + C_{I-(k-1),k-1} = \sum_{i=0}^{I-k} C_{i,k} - \sum_{i=0}^{I-k} S_{i,k} + \sum_{j=0}^{k-1} S_{I-(k-1),j} \\ &= \sum_{i=0}^{I-k} \widehat{\mu}_i \sum_{j=0}^{k} \widehat{\gamma}_j - \widehat{\gamma}_k \underbrace{\sum_{i=0}^{I-k} \widehat{\mu}_i}_{\textbf{(5.1)}} + \underbrace{\widehat{\mu}_{I-(k-1)} \sum_{j=0}^{k-1} \widehat{\gamma}_k}_{\textbf{(5.1)}} = \sum_{i=0}^{I-(k-1)} \sum_{j=0}^{k-1} \widehat{\mu}_i \widehat{\gamma}_j.\end{aligned}$$

☐

Proof of Theorem 5.3:

$$\begin{aligned}\widehat{C}^{\textbf{Poi}}_{i,J} &= \underbrace{C_{i,I-i} + \widehat{\mu}_i \sum_{k=I-i+1}^{J} \widehat{\gamma}_k}_{\textbf{Estimator 5.2}} = C_{i,I-i} + \underbrace{\frac{C_{i,I-i}}{\sum_{k=0}^{I-i} \widehat{\gamma}_k}}_{\textbf{(5.1)}} \sum_{k=I-i+1}^{J} \widehat{\gamma}_k = C_{i,I-i} \left(1 + \frac{\sum_{k=I-i+1}^{J} \widehat{\gamma}_k}{\sum_{k=0}^{I-i} \widehat{\gamma}_k}\right) \\ &= C_{i,I-i} \frac{\sum_{k=0}^{J} \widehat{\gamma}_k}{\sum_{k=0}^{I-i} \widehat{\gamma}_k} = C_{i,I-i} \frac{\sum_{k=0}^{I-i+1} \widehat{\gamma}_k}{\sum_{k=0}^{I-i} \widehat{\gamma}_k} \cdots \frac{\sum_{k=0}^{J} \widehat{\gamma}_k}{\sum_{k=0}^{J-1} \widehat{\gamma}_k} = C_{i,I-i} \underbrace{\frac{\sum_{h=0}^{I-(I-i+1)} C_{i,I-i+1}}{\sum_{h=0}^{I-(I-i+1)} C_{i,I-i}}}_{\textbf{above lemma}} \cdots \underbrace{\frac{\sum_{h=0}^{I-J} C_{i,J}}{\sum_{h=0}^{I-J} C_{i,J-1}}}_{\textbf{above lemma}} \\ &= C_{i,I-i}(1 + \widehat{f}_{I-i}) \cdots (1 + \widehat{f}_{J-1}) = \widehat{C}^{\textbf{CLM}}_{i,J}\end{aligned}$$

☐

Corollary 5.4 (Poisson-Model vs. Chain-Ladder method)

Taking CLM *as* LSRM *with the variance minimizing weights we have*

$$\widehat{S}_{i,k}^{\mathsf{CLM}} = \widehat{\mathsf{E}}\big[S_{i,k}\big] = \widehat{f}_{k-1}(1+\widehat{f}_{k-2})\cdots(1+\widehat{f}_{I-i})C_{i,I-i}.$$

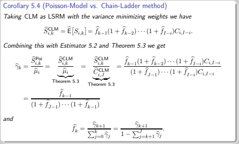
Combining this with Estimator 5.2 and Theorem 5.3 we get

$$\widehat{\gamma}_k = \frac{\widehat{S}_{i,k}^{\mathsf{Poi}}}{\widehat{\mu}_i} = \underbrace{\frac{\widehat{S}_{i,k}^{\mathsf{CLM}}}{\widehat{\mu}_i}}_{\text{Theorem 5.3}} = \underbrace{\frac{\widehat{S}_{i,k}^{\mathsf{CLM}}}{\widehat{C}_{i,J}^{\mathsf{CLM}}}}_{\text{Theorem 5.3}} = \frac{\widehat{f}_{k-1}(1+\widehat{f}_{k-2})\cdots(1+\widehat{f}_{I-i})C_{i,I-i}}{(1+\widehat{f}_{J-1})\cdots(1+\widehat{f}_{I-i})C_{i,I-i}}$$

$$= \frac{\widehat{f}_{k-1}}{(1+\widehat{f}_{J-1})\cdots(1+\widehat{f}_{k-1})}$$

and

$$\widehat{f}_k = \frac{\widehat{\gamma}_{k+1}}{\sum_{j=0}^{k}\widehat{\gamma}_j} = \frac{\widehat{\gamma}_{k+1}}{1-\sum_{j=k+1}^{J}\widehat{\gamma}_j}.$$



Corollary 5.4 (Poisson-Model vs. Chain-Ladder method)

Taking CLM *as* LSRM *with the variance minimizing weights we have*

$$\widehat{S}_{i,k}^{\mathsf{CLM}} = \widehat{\mathsf{E}}[S_{i,k}] = \widehat{f}_{k-1}(1+\widehat{f}_{k-2})\cdots(1+\widehat{f}_{I-i})C_{i,I-i}.$$

Combining this with Estimator 5.2 and Theorem 5.3 we get

$$\widehat{\gamma}_k = \frac{\widehat{S}_{i,k}^{\mathsf{Poi}}}{\widehat{\mu}_i} = \underbrace{\frac{\widehat{S}_{i,k}^{\mathsf{CLM}}}{\widehat{\mu}_i}}_{\text{Theorem 5.3}} = \underbrace{\frac{\widehat{S}_{i,k}^{\mathsf{CLM}}}{\widehat{C}_{i,J}^{\mathsf{CLM}}}}_{\text{Theorem 5.3}} = \frac{\widehat{f}_{k-1}(1+\widehat{f}_{k-2})\cdots(1+\widehat{f}_{I-i})C_{i,I-i}}{(1+\widehat{f}_{J-1})\cdots(1+\widehat{f}_{I-i})C_{i,I-i}}$$

$$= \frac{\widehat{f}_{k-1}}{(1+\widehat{f}_{J-1})\cdots(1+\widehat{f}_{k-1})}$$

and

$$\widehat{f}_k = \frac{\widehat{\gamma}_{k+1}}{\sum_{j=0}^{k}\widehat{\gamma}_j} = \frac{\widehat{\gamma}_{k+1}}{1-\sum_{j=k+1}^{J}\widehat{\gamma}_j}.$$

Proof of the last statement: From the proof of Theorem 5.3 we know that

$$1+\widehat{f}_k = \frac{\sum_{j=0}^{k+1}\widehat{\gamma}_j}{\sum_{j=0}^{k}\widehat{\gamma}_j}.$$

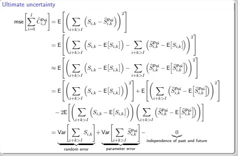
From this we compute

$$\widehat{f}_k = \frac{\sum_{j=0}^{k+1}\widehat{\gamma}_j}{\sum_{j=0}^{k}\widehat{\gamma}_j} - 1 = \frac{\widehat{\gamma}_{k+1}}{\sum_{j=0}^{k}\widehat{\gamma}_j} = \underbrace{\frac{\widehat{\gamma}_{k+1}}{1-\sum_{j=k+1}^{J}\widehat{\gamma}_j}}_{\sum_{j=0}^{J}\widehat{\gamma}_j=1}.$$

□

Ultimate uncertainty

$$
\begin{aligned}
\mathsf{mse}\left[\sum_{i=0}^{I}\widehat{C}_{i,J}^{\mathsf{Poi}}\right] &= \mathsf{E}\left[\left(\sum_{i+k>I}\left(S_{i,k}-\widehat{S}_{i,k}^{\mathsf{Poi}}\right)\right)^2\right]\\
&= \mathsf{E}\left[\left(\sum_{i+k>I}\left(S_{i,k}-\mathsf{E}\left[S_{i,k}\right]\right)-\sum_{i+k>I}\left(\widehat{S}_{i,k}^{\mathsf{Poi}}-\mathsf{E}\left[S_{i,k}\right]\right)\right)^2\right]\\
&\approx \mathsf{E}\left[\left(\sum_{i+k>I}\left(S_{i,k}-\mathsf{E}\left[S_{i,k}\right]\right)-\sum_{i+k>I}\left(\widehat{S}_{i,k}^{\mathsf{Poi}}-\mathsf{E}\left[\widehat{S}_{i,k}^{\mathsf{Poi}}\right]\right)\right)^2\right]\\
&= \mathsf{E}\left[\left(\sum_{i+k>I}\left(S_{i,k}-\mathsf{E}\left[S_{i,k}\right]\right)\right)^2\right]+\mathsf{E}\left[\left(\sum_{i+k>I}\left(\widehat{S}_{i,k}^{\mathsf{Poi}}-\mathsf{E}\left[\widehat{S}_{i,k}^{\mathsf{Poi}}\right]\right)\right)^2\right]\\
&\quad -2\mathsf{E}\left[\left(\sum_{i+k>I}\left(S_{i,k}-\mathsf{E}\left[S_{i,k}\right]\right)\right)\left(\sum_{i+k>I}\left(\widehat{S}_{i,k}^{\mathsf{Poi}}-\mathsf{E}\left[\widehat{S}_{i,k}^{\mathsf{Poi}}\right]\right)\right)\right]\\
&= \underbrace{\mathsf{Var}\left[\sum_{i+k>I}S_{i,k}\right]}_{\text{random error}}+\underbrace{\mathsf{Var}\left[\sum_{i+k>I}\widehat{S}_{i,k}^{\mathsf{Poi}}\right]}_{\text{parameter error}}-\underbrace{0}_{\text{independence of past and future}}
\end{aligned}
$$



Ultimate uncertainty

$$\begin{aligned}
\mathsf{mse}\left[\sum_{i=0}^{I} \widehat{C}_{i,J}^{\mathsf{Poi}}\right] &= \mathsf{E}\left[\left(\sum_{i+k>I}\left(S_{i,k}-\widehat{S}_{i,k}^{\mathsf{Poi}}\right)\right)^2\right] \\
&= \mathsf{E}\left[\left(\sum_{i+k>I}\left(S_{i,k}-\mathsf{E}[S_{i,k}]\right)-\sum_{i+k>I}\left(\widehat{S}_{i,k}^{\mathsf{Poi}}-\mathsf{E}[S_{i,k}]\right)\right)^2\right] \\
&\approx \mathsf{E}\left[\left(\sum_{i+k>I}\left(S_{i,k}-\mathsf{E}[S_{i,k}]\right)-\sum_{i+k>I}\left(\widehat{S}_{i,k}^{\mathsf{Poi}}-\mathsf{E}\left[\widehat{S}_{i,k}^{\mathsf{Poi}}\right]\right)\right)^2\right] \\
&= \mathsf{E}\left[\left(\sum_{i+k>I}\left(S_{i,k}-\mathsf{E}[S_{i,k}]\right)\right)^2\right]+\mathsf{E}\left[\left(\sum_{i+k>I}\left(\widehat{S}_{i,k}^{\mathsf{Poi}}-\mathsf{E}\left[\widehat{S}_{i,k}^{\mathsf{Poi}}\right]\right)\right)^2\right] \\
&\quad - 2\mathsf{E}\left[\left(\sum_{i+k>I}\left(S_{i,k}-\mathsf{E}[S_{i,k}]\right)\right)\left(\sum_{i+k>I}\left(\widehat{S}_{i,k}^{\mathsf{Poi}}-\mathsf{E}\left[\widehat{S}_{i,k}^{\mathsf{Poi}}\right]\right)\right)\right] \\
&= \underbrace{\mathsf{Var}\left[\sum_{i+k>I} S_{i,k}\right]}_{\text{random error}}+\underbrace{\mathsf{Var}\left[\sum_{i+k>I} \widehat{S}_{i,k}^{\mathsf{Poi}}\right]}_{\text{parameter error}}-\underbrace{0}_{\text{independence of past and future}}
\end{aligned}$$

Random error

Since all $S_{i,k}$ are independent we get

$$\mathrm{Var}\left[\sum_{i+k>I} S_{i,k}\right] = \sum_{i+k>I} \mathrm{Var}\left[S_{i,k}\right] = \sum_{i+k>I} \gamma_k \mu_i \approx \sum_{i+k>I} \widehat{\gamma}_k \widehat{\mu}_i.$$

Parameter error

In order to analyse the parameter error we use the following Tylor expansion:

$$\ln(z) \approx \ln(z_0) + \frac{1}{z_0}(z - z_0) \qquad \text{for } z_0 = 1 \text{ and } z = \frac{\widehat{\gamma}_k \widehat{\mu}_i}{\gamma_k \mu_i}.$$

Therefore, we get

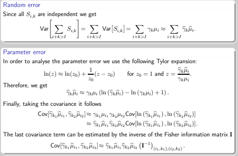
$$\widehat{\gamma}_k \widehat{\mu}_i \approx \gamma_k \mu_i \left(\ln\left(\widehat{\gamma}_k \widehat{\mu}_i\right) - \ln\left(\gamma_k \mu_i\right) + 1\right).$$

Finally, taking the covariance it follows

$$\begin{aligned}\mathrm{Cov}[\widehat{\gamma}_{k_1} \widehat{\mu}_{i_1}, \widehat{\gamma}_{k_2} \widehat{\mu}_{i_2}] &\approx \gamma_{k_1} \mu_{i_1} \gamma_{k_2} \mu_{i_2} \mathrm{Cov}[\ln\left(\widehat{\gamma}_{k_1} \widehat{\mu}_{i_1}\right), \ln\left(\widehat{\gamma}_{k_2} \widehat{\mu}_{i_2}\right)] \\ &\approx \widehat{\gamma}_{k_1} \widehat{\mu}_{i_1} \widehat{\gamma}_{k_2} \widehat{\mu}_{i_2} \mathrm{Cov}[\ln\left(\widehat{\gamma}_{k_1} \widehat{\mu}_{i_1}\right), \ln\left(\widehat{\gamma}_{k_2} \widehat{\mu}_{i_2}\right)].\end{aligned}$$

The last covariance term can be estimated by the inverse of the Fisher information matrix $\mathbf{I}$

$$\mathrm{Cov}[\widehat{\gamma}_{k_1} \widehat{\mu}_{i_1}, \widehat{\gamma}_{k_2} \widehat{\mu}_{i_2}] \approx \widehat{\gamma}_{k_1} \widehat{\mu}_{i_1} \widehat{\gamma}_{k_2} \widehat{\mu}_{i_2} \left(\mathbf{I}^{-1}\right)_{(i_1,k_1),(i_2,k_2)}.$$



Random error

Since all $S_{i,k}$ are independent we get

$$\mathsf{Var}\left[\sum_{i+k>I} S_{i,k}\right] = \sum_{i+k>I} \mathsf{Var}[S_{i,k}] = \sum_{i+k>I} \gamma_k\mu_i \approx \sum_{i+k>I} \widehat{\gamma}_k\widehat{\mu}_i.$$

Parameter error

In order to analyse the parameter error we use the following Tylor expansion:

$$\ln(z) \approx \ln(z_0) + \frac{1}{z_0}(z - z_0) \qquad \text{for } z_0 = 1 \text{ and } z = \frac{\widehat{\gamma}_k\widehat{\mu}_i}{\gamma_k\mu_i}.$$

Therefore, we get

$$\widehat{\gamma}_k\widehat{\mu}_i \approx \gamma_k\mu_i \left(\ln\left(\widehat{\gamma}_k\widehat{\mu}_i\right) - \ln\left(\gamma_k\mu_i\right) + 1\right).$$

Finally, taking the covariance it follows

$$\begin{aligned}\mathsf{Cov}[\widehat{\gamma}_{k_1}\widehat{\mu}_{i_1}, \widehat{\gamma}_{k_2}\widehat{\mu}_{i_2}] &\approx \gamma_{k_1}\mu_{i_1}\gamma_{k_2}\mu_{i_2}\mathsf{Cov}[\ln\left(\widehat{\gamma}_{k_1}\widehat{\mu}_{i_1}\right), \ln\left(\widehat{\gamma}_{k_2}\widehat{\mu}_{i_2}\right)]\\ &\approx \widehat{\gamma}_{k_1}\widehat{\mu}_{i_1}\widehat{\gamma}_{k_2}\widehat{\mu}_{i_2}\mathsf{Cov}[\ln\left(\widehat{\gamma}_{k_1}\widehat{\mu}_{i_1}\right), \ln\left(\widehat{\gamma}_{k_2}\widehat{\mu}_{i_2}\right)].\end{aligned}$$

The last covariance term can be estimated by the inverse of the Fisher information matrix $\mathbf{I}$

$$\mathsf{Cov}[\widehat{\gamma}_{k_1}\widehat{\mu}_{i_1}, \widehat{\gamma}_{k_2}\widehat{\mu}_{i_2}] \approx \widehat{\gamma}_{k_1}\widehat{\mu}_{i_1}\widehat{\gamma}_{k_2}\widehat{\mu}_{i_2}\left(\mathbf{I}^{-1}\right)_{(i_1,k_1),(i_2,k_2)}.$$

Estimator 5.5 (of the ultimate uncertainty)

$$\widehat{\mathsf{mse}}\left[\sum_{i=0}^{I}\widehat{C}_{i,J}^{\mathsf{Poi}}\right] \approx \underbrace{\sum_{i+k>I}\widehat{\gamma}_k\widehat{\mu}_i}_{\text{random error}} + \underbrace{\sum_{i_1+k_1,i_2+k_2>I}\widehat{\gamma}_{k_1}\widehat{\mu}_{i_1}\widehat{\gamma}_{k_2}\widehat{\mu}_{i_2}\left(\mathbf{I}^{-1}\right)_{(i_1,k_1),(i_2,k_2)}}_{\text{parameter error}}.$$

Remark 5.6

- The ultimate uncertainty of a single accident period i can be estimated by

$$\widehat{\mathsf{mse}}\left[\widehat{C}_{i,J}^{\mathsf{Poi}}\right] \approx \underbrace{\sum_{k=I-i+1}^{J}\widehat{\gamma}_k\widehat{\mu}_i}_{\text{random error}} + \underbrace{\sum_{k_1,k_2=I-i+1}^{J}\widehat{\gamma}_{k_1}\widehat{\mu}_i\widehat{\gamma}_{k_2}\widehat{\mu}_i\left(\mathbf{I}^{-1}\right)_{(i,k_1),(i,k_2)}}_{\text{parameter error}}.$$

- 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.

Estimator 5.5 (of the ultimate uncertainty)

$$\widehat{\mathsf{mse}}\left[\sum_{i=0}^{I}\widehat{C}_{i,J}^{\mathsf{Poi}}\right] \approx \underbrace{\sum_{i+k>I}\widehat{\gamma}_k\widehat{\mu}_i}_{\text{random error}} + \underbrace{\sum_{i_1+k_1,i_2+k_2>I}\widehat{\gamma}_{k_1}\widehat{\mu}_{i_1}\widehat{\gamma}_{k_2}\widehat{\mu}_{i_2}\left(\mathbf{I}^{-1}\right)_{(i_1,k_1),(i_2,k_2)}}_{\text{parameter error}}.$$

Remark 5.6

- The ultimate uncertainty of a single accident period i can be estimated by

$$\widehat{\mathsf{mse}}\left[\widehat{C}_{i,J}^{\mathsf{Poi}}\right] \approx \underbrace{\sum_{k=I-i+1}^{J}\widehat{\gamma}_k\widehat{\mu}_i}_{\text{random error}} + \underbrace{\sum_{k_1,k_2=I-i+1}^{J}\widehat{\gamma}_{k_1}\widehat{\mu}_i\widehat{\gamma}_{k_2}\widehat{\mu}_i\left(\mathbf{I}^{-1}\right)_{(i,k_1),(i,k_2)}}_{\text{parameter error}}.$$

- 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

$$\mathsf{E}\left[S_{i,k}\right] = \mathsf{Var}\left[S_{i,k}\right] = \gamma_k \mu_i.$$

Overdispersed Poisson-Model

The restriction on the variance to be equal to the expectation can be softened by taking

$$\mathsf{E}\left[S_{i,k}\right] = \gamma_k \mu_i \qquad \text{and} \qquad \mathsf{Var}\left[S_{i,k}\right] = \varphi_k \gamma_k \mu_i,$$

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

$$\mathsf{E}[S_{i,k}] = \mathsf{Var}[S_{i,k}] = \gamma_k \mu_i.$$

Overdispersed Poisson-Model

The restriction on the variance to be equal to the expectation can be softened by taking

$$\mathsf{E}[S_{i,k}] = \gamma_k \mu_i \qquad \text{and} \qquad \mathsf{Var}[S_{i,k}] = \varphi_k \gamma_k \mu_i,$$

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.

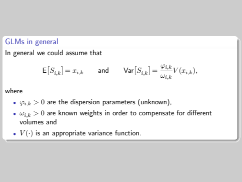
GLMs in general

In general we could assume that

$$\mathsf{E}\left[S_{i,k}\right] = x_{i,k} \qquad \text{and} \qquad \mathsf{Var}\left[S_{i,k}\right] = \frac{\varphi_{i,k}}{\omega_{i,k}} V(x_{i,k}),$$

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

$$\mathsf{E}[S_{i,k}] = x_{i,k} \qquad \text{and} \qquad \mathsf{Var}[S_{i,k}] = \frac{\varphi_{i,k}}{\omega_{i,k}} V(x_{i,k}),$$

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.

Literature

[23] McCullagh, P. and Nelder, J.A.
Generalised Linear Models.
Chapman & Hall, London, 1989.

Literature

[23] McCullagh, P. and Nelder, J.A.
Generalised Linear Models.
Chapman & Hall, London, 1989.

Stochastic Reserving

Lecture 10

Bootstrapping

René Dahms

ETH Zurich, Spring 2021

5 May 2021

Jürg Schelldorfer is standing in

(Last update: 26 April 2021)

Stochastic Reserving

Lecture 10

Bootstrapping

René Dahms

ETH Zurich, Spring 2021

5 May 2021

Jürg Schelldorfer is standing in

(Last update: 26 April 2021)

6 Bootstrap for CLM

2021-04-26

Stochastic Reserving

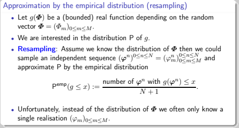
Lecture 10: Table of contents

Approximation by the empirical distribution (resampling)

- Let $g(\boldsymbol{\Phi})$ be a (bounded) real function depending on the random vector $\boldsymbol{\Phi} = (\Phi_m)_{0 \le m \le M}$.
- We are interested in the distribution P of g.
- **Resampling**: Assume we know the distribution of $\boldsymbol{\Phi}$ then we could sample an independent sequence $(\boldsymbol{\varphi}^n)^{0 \le n \le N} = (\varphi_m^n)_{0 \le m \le M}^{0 \le n \le N}$ and approximate P by the empirical distribution

$$\mathsf{P}^{\mathsf{emp}}(g \le x) := \frac{\text{number of } \boldsymbol{\varphi}^n \text{ with } g(\boldsymbol{\varphi}^n) \le x}{N+1}.$$

- Unfortunately, instead of the distribution of $\boldsymbol{\Phi}$ we often only know a single realisation $(\varphi_m)_{0 \le m \le M}$.



Approximation by the empirical distribution (resampling)

- Let $g(\boldsymbol{\Phi})$ be a (bounded) real function depending on the random vector $\boldsymbol{\Phi} = (\Phi_m)_{0 \le m \le M}$.
- We are interested in the distribution P of g.
- **Resampling**: Assume we know the distribution of $\boldsymbol{\Phi}$ then we could sample an independent sequence $(\boldsymbol{\varphi}^n)^{0 \le n \le N} = (\varphi^n_m)^{0 \le n \le N}_{0 \le m \le M}$ and approximate P by the empirical distribution

$$\mathsf{P}^{\mathsf{emp}}(g \le x) := \frac{\text{number of } \boldsymbol{\varphi}^n \text{ with } g(\boldsymbol{\varphi}^n) \le x}{N+1}.$$

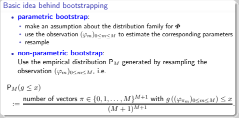
- Unfortunately, instead of the distribution of $\boldsymbol{\Phi}$ we often only know a single realisation $(\varphi_m)_{0 \le m \le M}$.

- If the function g is 'nice enough' it is well known that the empirical distribution converges to P in some sense.
  If φ_m are realisations of i.i.d. random variables one could take $(\varphi_{\pi_m})_{0 \le m \le M}$, with some $\pi \in \{0, 1, \ldots, M\}^{M+1}$, instead of independent realisations of $\boldsymbol{\Phi}$, which leads us to non-parametric bootstrap.

Basic idea behind bootstrapping

- **parametric bootstrap**:
  - make an assumption about the distribution family for $\boldsymbol{\Phi}$
  - use the observation $(\varphi_m)_{0\le m\le M}$ to estimate the corresponding parameters
  - resample
- **non-parametric bootstrap**:
  Use the empirical distribution P_M generated by resampling the observation $(\varphi_m)_{0\le m\le M}$, i.e.

$$\mathsf{P}_M(g\le x) := \frac{\text{number of vectors } \pi\in\{0,1,\dots,M\}^{M+1} \text{ with } g\left((\varphi_{\pi_m})_{0\le m\le M}\right)\le x}{(M+1)^{M+1}}.$$



Basic idea behind bootstrapping

- **parametric bootstrap**:
  - make an assumption about the distribution family for $\boldsymbol{\Phi}$
  - 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.

$$\mathsf{P}_M(g \leq x) := \frac{\text{number of vectors } \pi \in \{0,1,\ldots,M\}^{M+1} \text{ with } g\left((\varphi_{\pi_m})_{0\leq m\leq M}\right) \leq x}{(M+1)^{M+1}}.$$

Parametric bootstrap:

- Which distribution family should we take?
- The estimation of the parameters of the underlying distribution of $\boldsymbol{\Phi}$ based on a single observation is very uncertain.

Non-parametric bootstrap:

- At least we have to assume that the components of $\boldsymbol{\Phi}$ are independent and identical distributed.
- In most cases the number $(M+1)^{M+1}$ is too large. So instead of calculating all combinations, we use resampling replacement to approximate the emirical distribution.
- 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 $\boldsymbol{\Phi}$ 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

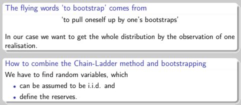
‘to pull oneself up by one’s bootstraps’

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. and
- define the reserves.



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. and
- define the reserves.

It goes back to the fairy tale by Baron Munchausen who claimed to saved himself from being drowned in a swamp by pulling on his own hair.

Recapitulation

Let $C_{i,k} := \sum_{j=0}^{k} S_{i,j}$. If we have

i')$^{\text{CLM}}$ $\mathsf{E}\left[S_{i,k+1}\middle|\mathcal{B}_{i,k}\right] = f_k C_{i,k}$,

ii')$^{\text{CLM}}$ $\mathsf{Var}\left[S_{i,k+1}\middle|\mathcal{B}_{i,k}\right] = \sigma_k^2 C_{i,k}$ and

iii')$^{\text{CLM}}$ accident periods are independent.

Then

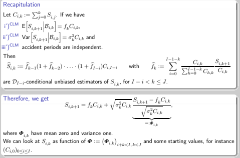
$$\widehat{S}_{i,k} := \widehat{f}_{k-1}(1+\widehat{f}_{k-2})\cdot\ldots\cdot(1+\widehat{f}_{I-i})C_{i,I-i} \qquad \text{with} \qquad \widehat{f}_k := \sum_{i=0}^{I-1-k} \frac{C_{i,k}}{\sum_{h=0}^{I-1-k} C_{h,k}} \frac{S_{i,k+1}}{C_{i,k}}$$

are $\mathcal{D}_{I-i}$-conditional unbiased estimators of $S_{i,k}$, for $I-i<k\leq J$.

Therefore, we get

$$S_{i,k+1} = f_k C_{i,k} + \sqrt{\sigma_k^2 C_{i,k}}\,\underbrace{\frac{S_{i,k+1}-f_k C_{i,k}}{\sqrt{\sigma_k^2 C_{i,k}}}}_{=:\boldsymbol{\Phi}_{i,k}},$$

where $\boldsymbol{\Phi}_{i,k}$ have mean zero and variance one.
We can look at $S_{i,k}$ as function of $\boldsymbol{\Phi} := \left(\boldsymbol{\Phi}_{i,k}\right)_{i+k<I,\,k<J}$ 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

i')$^{\text{CLM}}$ $\mathsf{E}\left[S_{i,k+1}\middle|\mathcal{B}_{i,k}\right] = f_k C_{i,k}$,

ii')$^{\text{CLM}}$ $\mathsf{Var}\left[S_{i,k+1}\middle|\mathcal{B}_{i,k}\right] = \sigma_k^2 C_{i,k}$ and

iii')$^{\text{CLM}}$ accident periods are independent.

Then

$$\widehat{S}_{i,k} := \widehat{f}_{k-1}(1+\widehat{f}_{k-2}) \cdot \ldots \cdot (1+\widehat{f}_{I-i})C_{i,I-i} \qquad \text{with} \qquad \widehat{f}_k := \sum_{i=0}^{I-1-k} \frac{C_{i,k}}{\sum_{h=0}^{I-1-k} C_{h,k}} \frac{S_{i,k+1}}{C_{i,k}}$$

are $\mathcal{D}_{I-i}$-conditional unbiased estimators of $S_{i,k}$, for $I-i < k \leq J$.

Therefore, we get

$$S_{i,k+1} = f_k C_{i,k} + \sqrt{\sigma_k^2 C_{i,k}}\underbrace{\frac{S_{i,k+1} - f_k C_{i,k}}{\sqrt{\sigma_k^2 C_{i,k}}}}_{=:\boldsymbol{\Phi}_{i,k}},$$

where $\boldsymbol{\Phi}_{i,k}$ have mean zero and variance one.
We can look at $S_{i,k}$ as function of $\boldsymbol{\Phi} := \left(\boldsymbol{\Phi}_{i,k}\right)_{i+k<I,\,k<J}$ 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 σ_k^2.
- $\boldsymbol{\Phi}_{i,k}$ 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:
  - $\boldsymbol{\Phi}_{I,J-1}$ in the case where $I = J$, because it is equal to zero (deterministically).
  - all $\boldsymbol{\Phi}_{i,k}$ 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.

6.3 Bootstrapping Chain-Ladder step by step, variant 1 (1/3)

Step 1: Chain-Ladder method

claim property: $S_{i,k}$

estimated development factors: $\widehat{f}_k$

estimated variance parameters: $\widehat{\sigma}_k^2$

$S_{i,k}$	0	1	2	3	ultimate	reserve
0	100	100	50	0	250	0
1	300	190	88	0	578	0
2	100	85	37	0	222	37
3	200	150	70	0	420	220
$\widehat{f}_k$	0.75	0.2	0		1470	257
$\widehat{\sigma}_k^2$	6.67	0.70	0.09			

Step 2: Residuals

Pearson residuals inclusive variance adjustment:

$$\varphi_{i,k} := \frac{S_{i,k+1} - \widehat{f}_k C_{i,k}}{\sqrt{\widehat{\sigma}_k^2 C_{i,k}}} \sqrt{\frac{I-k}{I-k-1}}$$

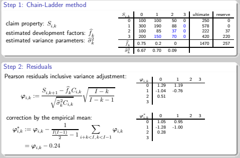
$\varphi_{i,k}$	0	1	2	3
0	1.29	1.19		
1	-1.04	-0.76		
2	0.51			
3				

correction by the empirical mean:

$$\varphi_{i,k}^* := \varphi_{i,k} - \frac{1}{\frac{I(I-1)}{2} - 1} \sum_{i+k<I,\, k<I-1} \varphi_{i,k}$$

$$= \varphi_{i,k} - 0.24$$

$\varphi_{i,k}^*$	0	1	2	3
0	1.05	0.95		
1	-1.28	-1.00		
2	0.28			
3				



Step 1: Chain-Ladder method

claim property: $S_{i,k}$
estimated development factors: $\widehat{f}_k$
estimated variance parameters: $\widehat{\sigma}_k^2$

$S_{i,k}$	0	1	2	3	ultimate	reserve
0	100	100	50	0	250	0
1	300	190	88	0	578	0
2	100	85	37	0	222	37
3	200	150	70	0	420	220
$\widehat{f}_k$	0.75	0.2	0		1470	257
$\widehat{\sigma}_k^2$	6.67	0.70	0.09			

Step 2: Residuals

Pearson residuals inclusive variance adjustment:

$$\varphi_{i,k} := \frac{S_{i,k+1} - \widehat{f}_k C_{i,k}}{\sqrt{\widehat{\sigma}_k^2 C_{i,k}}} \sqrt{\frac{I-k}{I-k-1}}$$

$\varphi_{i,k}$	0	1	2	3
0	1.29	1.19		
1	-1.04	-0.76		
2	0.51			
3				

correction by the empirical mean:

$$\varphi^*_{i,k} := \varphi_{i,k} - \frac{1}{\frac{I(I-1)}{2} - 1} \sum_{i+k<I, k<I-1} \varphi_{i,k} = \varphi_{i,k} - 0.24$$

$\varphi^*_{i,k}$	0	1	2	3
0	1.05	0.95		
1	-1.28	-1.00		
2	0.28			
3				

- No residuals for development periods k, where $\widehat{\sigma}_k^2$ has to be approximated.
- Although $\boldsymbol{\Phi}_{i,k}$ has zero mean and variance equal to one, its estimate $\boldsymbol{\varphi}_{i,k}$ doesn't. The reason for this is that we do not know the parameters f_k and σ_k and use some estimators instead. Lets assume we know σ_k^2 and take the variance minimizing weights $w_{i,k}$:

$$\mathsf{Var}\left[\frac{S_{i,k+1} - \widehat{f}_k C_{i,k}}{\sqrt{\sigma_k^2 C_{i,k}}}\middle|\mathcal{D}_k\right] = \frac{1}{\sigma_k^2 C_{i,k}} \left(\mathsf{Var}\Big[S_{i,k+1}\Big|\mathcal{D}_k\Big] - 2C_{i,k}\mathsf{Cov}\Big[S_{i,k+1}, \widehat{f}_k\Big|\mathcal{D}_k\Big] + C_{i,k}^2\mathsf{Var}\Big[\widehat{f}_k\Big|\mathcal{D}_k\Big]\right)$$

$$= \frac{1}{\sigma_k^2 C_{i,k}} \left(\sigma_k^2 C_{i,k} - 2C_{i,k}\mathsf{Cov}\left[S_{i,k+1}, w_{i,k}\frac{S_{i,k+1}}{C_{i,k}}\middle|\mathcal{D}_k\right] + C_{i,k}^2 \sum_{h=0}^{I-k-1} \frac{w_{h,k}^2}{C_{h,k}^2}\mathsf{Var}\Big[S_{h,k+1}\Big|\mathcal{D}_k\Big]\right)$$

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

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 σ_k^2 we take $\sqrt{\frac{I-k}{I-k-1}}$ instead. This insures that

$$\frac{1}{I-k} \sum_{i=0}^{I-k-1} \mathsf{Var}\Big[\boldsymbol{\varphi}_{i,k}\Big|\mathcal{D}_k\Big] = 1 \qquad \text{provided we know } \sigma_k^2.$$

Step 3: Resampled residuals (non-parametric bootstrap)

set of residuals:

$$\{-1.28,\ -1.00,\ 0.28,\ 0.95,\ 1.05\}$$

$\varphi^*_{i,k}$	0	1	2	3
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	

Step 4a: Resampled triangle and Chain-Ladder method without process variance

$$S^*_{i,0} := S_{i,0}$$

$$S^*_{i,k+1} := \widehat{f}_k C^*_{i,k} + \sqrt{\widehat{\sigma}^2_k C^*_{i,k}}\,\varphi^*_{i,k}, \qquad i+k \le I$$

$$S^*_{i,k+1} := \widehat{f}^*_k C^*_{i,k}, \qquad i+k > I$$

$S^*_{i,k}$	0	1	2	3	ultimate	reserve
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
$\widehat{f}^*_k$	0.73	0.21	0.02		1497	288

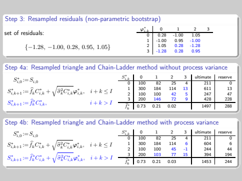
Step 4b: Resampled triangle and Chain-Ladder method with process variance

$$S^*_{i,0} := S_{i,0}$$

$$S^*_{i,k+1} := \widehat{f}_k C^*_{i,k} + \sqrt{\widehat{\sigma}^2_k C^*_{i,k}}\,\varphi^*_{i,k}, \qquad i+k \le I$$

$$S^*_{i,k+1} := \widehat{f}^*_k C^*_{i,k} + \sqrt{\widehat{\sigma}^2_k C^*_{i,k}}\,\varphi^*_{i,k}, \qquad i+k > I$$

$S^*_{i,k}$	0	1	2	3	ultimate	reserve
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
$\widehat{f}^*_k$	0.73	0.21	0.03		1453	244



Step 3: Resampled residuals (non-parametric bootstrap)

set of residuals:

$$\{-1.28,\ -1.00,\ 0.28,\ 0.95,\ 1.05\}$$

$\varphi^*_{i,k}$	0	1	2	3
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	

Step 4a: Resampled triangle and Chain-Ladder method without process variance

$$S^*_{i,0} := S_{i,0}$$

$$S^*_{i,k+1} := \widehat{f}_k C^*_{i,k} + \sqrt{\widehat{\sigma}^2_k C^*_{i,k}}\varphi^*_{i,k}, \quad i+k \le I$$

$$S^*_{i,k+1} := \widehat{f}^*_k C^*_{i,k}, \qquad i+k > I$$

$S^*_{i,k}$	0	1	2	3	ultimate	reserve
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
$\widehat{f}^*_k$	0.73	0.21	0.02		1497	288

Step 4b: Resampled triangle and Chain-Ladder method with process variance

$$S^*_{i,0} := S_{i,0}$$

$$S^*_{i,k+1} := \widehat{f}_k C^*_{i,k} + \sqrt{\widehat{\sigma}^2_k C^*_{i,k}}\varphi^*_{i,k}, \quad i+k \le I$$

$$S^*_{i,k+1} := \widehat{f}^*_k C^*_{i,k} + \sqrt{\widehat{\sigma}^2_k C^*_{i,k}}\varphi^*_{i,k}, \quad i+k > I$$

$S^*_{i,k}$	0	1	2	3	ultimate	reserve
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
$\widehat{f}^*_k$	0.73	0.21	0.03		1453	244

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.

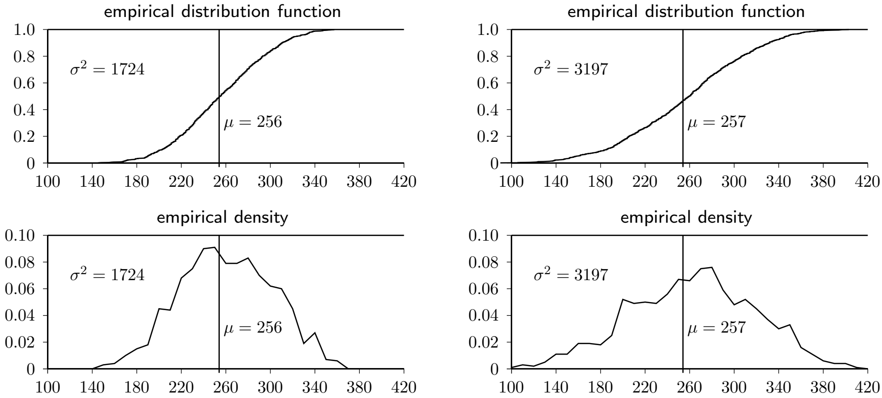
Step 5: Repeat steps 3 and 4 and collect the resulting reserves

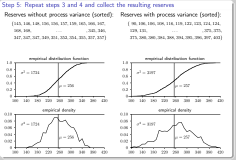
Reserves without process variance (sorted):

$$\{145, 146, 148, 156, 156, 157, 159, 165, 166, 167,\ 168, 168, \quad \ldots \quad, 345, 346,\ 347, 347, 347, 349, 351, 352, 354, 355, 357, 357\}$$

Reserves with process variance (sorted):

$$\{\ 90, 106, 106, 108, 116, 119, 122, 123, 124, 124,\ 129, 131, \quad \ldots \quad, 375, 375,\ 375, 380, 380, 384, 388, 394, 395, 396, 397, 403\}$$





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):

$\{90, 106, 106, 108, 116, 119, 122, 123, 124, 124, 129, 131, \ldots, 375, 375, 375, 380, 380, 384, 388, 394, 395, 396, 397, 403\}$

- without process variance: σ^2 represents the squared parameter estimation error
- with process variance: σ^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,
  - $S^*_{i,k+1} := \widehat{f}^*_k C_{i,k} + \sqrt{\widehat{\sigma}^2_k C_{i,k}} \boldsymbol{\varphi}^*_{i,k}$

Recapitulation: (overdispersed) Poisson-Model

If we have

i)$^{\text{Poi}}$ $S_{i,k}$ are independent random variables,

ii)$^{\text{Poi}}$ the distribution of $S_{i,k}$ belongs to the exponential dispersion family and

iii)$^{\text{Poi}}$ $\mathrm{Var}\left[S_{i,k}\right] = \vartheta_k \mathrm{E}\left[S_{i,k}\right] = \vartheta_k \gamma_k \mu_i$.

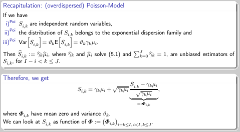
Then $\widehat{S}_{i,k} := \widehat{\gamma}_k \widehat{\mu}_i$, where $\widehat{\gamma}_k$ and $\widehat{\mu}_i$ solve (5.1) and $\sum_{k=0}^{J} \widehat{\gamma}_k = 1$, are unbiased estimators of $S_{i,k}$, for $I - i < k \leq J$.

Therefore, we get

$$S_{i,k} = \gamma_k \mu_i + \sqrt{\gamma_k \mu_i}\, \underbrace{\frac{S_{i,k} - \gamma_k \mu_i}{\sqrt{\gamma_k \mu_i}}}_{=:\boldsymbol{\Phi}_{i,k}},$$

where $\boldsymbol{\Phi}_{i,k}$ have mean zero and variance ϑ_k.

We can look at $S_{i,k}$ as function of $\boldsymbol{\Phi} := \left(\boldsymbol{\Phi}_{i,k}\right)_{i+k \leq I,\, i<I,\, k \leq J}$.



Recapitulation: (overdispersed) Poisson-Model

If we have

- i)$^{\text{Poi}}$ $S_{i,k}$ are independent random variables,
- ii)$^{\text{Poi}}$ the distribution of $S_{i,k}$ 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 $\widehat{S}_{i,k} := \widehat{\gamma}_k \widehat{\mu}_i$, where $\widehat{\gamma}_k$ and $\widehat{\mu}_i$ solve (5.1) and $\sum_{k=0}^{J} \widehat{\gamma}_k = 1$, are unbiased estimators of $S_{i,k}$, for $I - i < k \leq J$.

Therefore, we get

$$S_{i,k} = \gamma_k \mu_i + \sqrt{\gamma_k \mu_i} \underbrace{\frac{S_{i,k} - \gamma_k \mu_i}{\sqrt{\gamma_k \mu_i}}}_{=:\boldsymbol{\Phi}_{i,k}},$$

where $\boldsymbol{\Phi}_{i,k}$ have mean zero and variance ϑ_k.
We can look at $S_{i,k}$ as function of $\boldsymbol{\Phi} := \left(\boldsymbol{\Phi}_{i,k}\right)_{i+k \leq I, i<I, k \leq J}$.

- In the last formula we still have some unknown parameters, i.e. γ_k, μ_i and ϑ_k.

Step 1: Chain-Ladder method (Poisson-Model)

claim property: $S_{i,k}$
estimated payment pattern: $\widehat{\gamma}_k$
estimated ultimate: $\widehat{\mu}_i$

$S_{i,k}$	0	1	2	3	ultimate ($\widehat{\mu}_i$)	reserve
0	100	100	50	0	250	0
1	300	190	88	0	578	0
2	100	85	37	0	222	37
3	200	150	70	0	420	220
$\widehat{\gamma}_k$	0.48	0.36	0.17	0.00	1470	257

Step 2: Residuals

Residuals inclusive variance adjustment:

$$\varphi_{i,k} := \frac{S_{i,k} - \widehat{\gamma}_k \widehat{\mu}_i}{\sqrt{\widehat{\gamma}_k \widehat{\mu}_i}} \underbrace{\sqrt{\frac{I-k}{\sum_{h=0}^{I-k} \frac{\left(S_{h,k} - \widehat{\gamma}_k \widehat{\mu}_h\right)^2}{\widehat{\gamma}_k \widehat{\mu}_h}}}}_{=:\sqrt{\widehat{\vartheta}_k}}$$

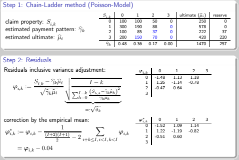
$\varphi_{i,k}$	0	1	2	3
0	-1.48	1.13	1.18	
1	1.26	-1.14	-0.78	
2	-0.47	0.64		
3				

correction by the empirical mean:

$$\varphi^*_{i,k} := \varphi_{i,k} - \frac{1}{\frac{(I+2)(I+1)}{2} - 2} \sum_{i+k \leq I,\, i<I,\, k<I} \varphi_{i,k}$$

$$= \varphi_{i,k} - 0.04$$

$\varphi^*_{i,k}$	0	1	2	3
0	-1.52	1.09	1.14	
1	1.22	-1.19	-0.82	
2	-0.51	0.60		
3				



Step 1: Chain-Ladder method (Poisson-Model)

claim property: $S_{i,k}$
estimated payment pattern: $\widehat{\gamma}_k$
estimated ultimate: $\widehat{\mu}_i$

$S_{i,k}$	0	1	2	3	ultimate ($\widehat{\mu}_i$)	reserve
0	100	100	50	0	250	0
1	300	190	88	0	578	0
2	100	85	37	0	222	37
3	200	150	70	0	420	220
$\widehat{\gamma}_k$	0.48	0.36	0.17	0.00	1470	257

Step 2: Residuals

Residuals inclusive variance adjustment:

$$\varphi_{i,k} := \frac{S_{i,k} - \widehat{\gamma}_k \widehat{\mu}_i}{\sqrt{\widehat{\gamma}_k \widehat{\mu}_i}} \underbrace{\sqrt{\frac{I-k}{\sum_{h=0}^{I-k} \frac{(S_{h,k} - \widehat{\gamma}_k \widehat{\mu}_h)^2}{\widehat{\gamma}_k \widehat{\mu}_h}}}}_{=:\sqrt{\widehat{\vartheta}_k}}$$

$\varphi_{i,k}$	0	1	2	3
0	-1.48	1.13	1.18	
1	1.26	-1.14	-0.78	
2	-0.47	0.64		
3				

correction by the empirical mean:

$$\varphi^*_{i,k} := \varphi_{i,k} - \frac{1}{\frac{(I+2)(I+1)}{2} - 2} \sum_{i+k \leq I, i<I, k<I} \varphi_{i,k}$$
$$= \varphi_{i,k} - 0.04$$

$\varphi^*_{i,k}$	0	1	2	3
0	-1.52	1.09	1.14	
1	1.22	-1.19	-0.82	
2	-0.51	0.60		
3				

- No residual for $(i, k) = (I, 0)$, because it is equal to zero (deterministically).
- Although $\boldsymbol{\Phi}_{i,k}$ has zero mean and variance equal to ϑ_k, its estimate $\boldsymbol{\varphi}_{i,k}$ doesn't. The reason for this is that we do not know the parameters f_k and σ_k and use some estimators instead. The variance adjustment

$$\sqrt{\widehat{\vartheta}_k} = \sqrt{\frac{I-k}{\sum_{h=0}^{I-k} \frac{\left(S_{h,k} - \widehat{\gamma}_k \widehat{\mu}_h\right)^2}{\widehat{\gamma}_k \widehat{\mu}_h}}}$$

ensures that the empirical variance equals one, i.e. that

$$\frac{1}{I-k} \sum_{i=0}^{I-k} (\boldsymbol{\varphi}^*_{i,k} - 0)^2 = 1.$$

6.5 Bootstrapping Chain-Ladder step by step, variant 2 (2/3)

Step 3: Resampled residuals (non-parametric bootstrap)

set of residuals:

$$\{-1.52, -1.19, -0.82, -0.51, 0.60, 1.09, 1.14, 1.22\}$$

$\boldsymbol{\varphi}^*_{i,k}$	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

Step 4a: Resampled triangle and Chain-Ladder method without process variance

$$S^*_{i,k} := \widehat{\gamma}_k \widehat{\mu}_i + \sqrt{\widehat{\vartheta}_k \widehat{\gamma}_k \widehat{\mu}_i}\,\boldsymbol{\varphi}^*_{i,k}, \quad i+k \le I,$$

$$S^*_{i,k} := \widehat{\gamma}^*_k \widehat{\mu}^*_i, \quad i+k > I$$

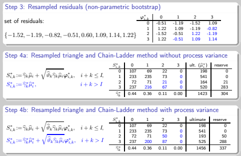
$S^*_{i,k}$	0	1	2	3	ult. $(\widehat{\mu}^*_i)$	reserve
0	107	69	22	0	198	0
1	233	235	73	0	541	0
2	72	71	21	0	164	21
3	237	216	67	0	520	283
$\widehat{\gamma}^*_k$	0.44	0.36	0.11	0.00	1423	304

Step 4b: Resampled triangle and Chain-Ladder method with process variance

$$S^*_{i,k} := \widehat{\gamma}_k \widehat{\mu}_i + \sqrt{\widehat{\vartheta}_k \widehat{\gamma}_k \widehat{\mu}_i}\,\boldsymbol{\varphi}^*_{i,k}, \quad i+k \le I,$$

$$S^*_{i,k} := \widehat{\gamma}^*_k \widehat{\mu}^*_i + \sqrt{\widehat{\vartheta}_k \widehat{\gamma}_k \widehat{\mu}_i}\,\boldsymbol{\varphi}^*_{i,k}, \quad i+k > I$$

$S^*_{i,k}$	0	1	2	3	ultimate	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
$\widehat{\gamma}^*_k$	0.44	0.36	0.11	0.00	1456	337



Step 3: Resampled residuals (non-parametric bootstrap)

set of residuals:

$\{-1.52, -1.19, -0.82, -0.51, 0.60, 1.09, 1.14, 1.22\}$

$\varphi^*_{i,k}$	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

Step 4a: Resampled triangle and Chain-Ladder method without process variance

$S^*_{i,k} := \widehat{\gamma}_k \widehat{\mu}_i + \sqrt{\widehat{\vartheta}_k \widehat{\gamma}_k \widehat{\mu}_i} \varphi^*_{i,k}, \quad i+k \le I,$

$S^*_{i,k} := \widehat{\gamma}^*_k \widehat{\mu}^*_i, \quad i+k > I$

$S^*_{i,k}$	0	1	2	3	ult. $(\widehat{\mu}^*_i)$	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
$\widehat{\gamma}^*_k$	0.44	0.36	0.11	0.00	1423	304

Step 4b: Resampled triangle and Chain-Ladder method with process variance

$S^*_{i,k} := \widehat{\gamma}_k \widehat{\mu}_i + \sqrt{\widehat{\vartheta}_k \widehat{\gamma}_k \widehat{\mu}_i} \varphi^*_{i,k}, \quad i+k \le I,$

$S^*_{i,k} := \widehat{\gamma}^*_k \widehat{\mu}^*_i + \sqrt{\widehat{\vartheta}_k \widehat{\gamma}_k \widehat{\mu}_i} \varphi^*_{i,k}, \quad i+k > I$

$S^*_{i,k}$	0	1	2	3	ultimate	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
$\widehat{\gamma}^*_k$	0.44	0.36	0.11	0.00	1456	337

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.

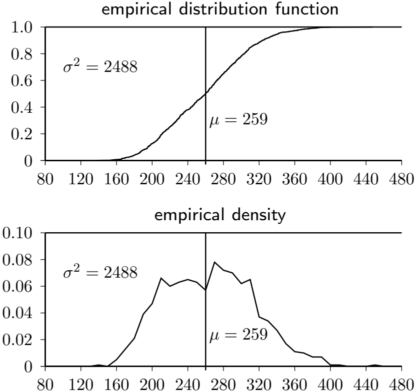
Step 5: Repeat steps 3 and 4 and collect the resulting reserves

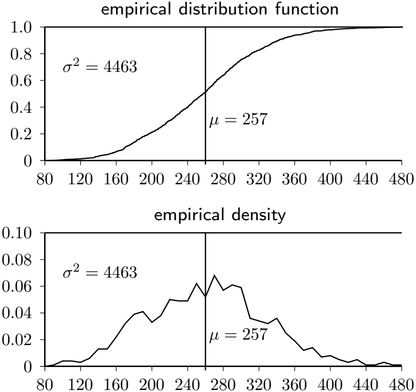
Reserves without process variance (sorted):

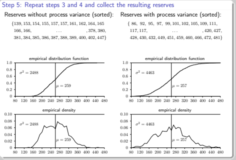
$$\{139, 153, 154, 155, 157, 157, 161, 162, 164, 165$$
$$166, 166, \quad \ldots \quad , 378, 380,$$
$$381, 384, 385, 386, 387, 388, 389, 400, 402, 447\}$$

Reserves with process variance (sorted):

$$\{\ 86,\ 92,\ 95,\ 97,\ 99, 101, 102, 105, 109, 111,$$
$$117, 117, \quad \ldots \quad , 420, 427,$$
$$428, 430, 432, 449, 451, 459, 460, 466, 472, 481\}$$







Step 5: Repeat steps 3 and 4 and collect the resulting reserves

Reserves without process variance (sorted):

$\{139, 153, 154, 155, 157, 157, 161, 162, 164, 165$
$166, 166, \dots, 378, 380,$
$381, 384, 385, 386, 387, 388, 389, 400, 402, 447\}$

Reserves with process variance (sorted):

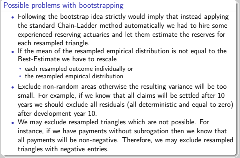
$\{86, 92, 95, 97, 99, 101, 102, 105, 109, 111,$
$117, 117, \dots, 420, 427,$
$428, 430, 432, 449, 451, 459, 460, 466, 472, 481\}$

- without process variance: σ^2 represents the squared parameter estimation error
- with process variance: σ^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

$$S^*_{i,k} := \widehat{\gamma}^*_k \widehat{\mu}^*_i + \sqrt{\widehat{\vartheta}_k \widehat{\gamma}^*_k \widehat{\mu}^*_i}\,\boldsymbol{\varphi}^*_{i,k}, \quad i + k > I.$$

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.



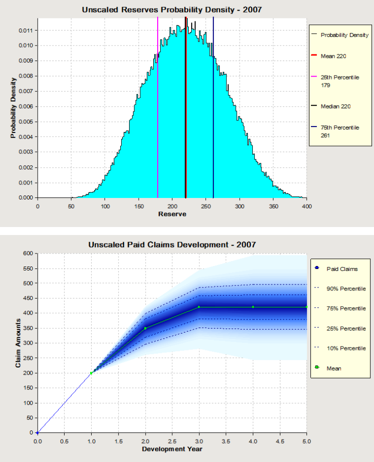
Possible problems with bootstrapping

- 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.

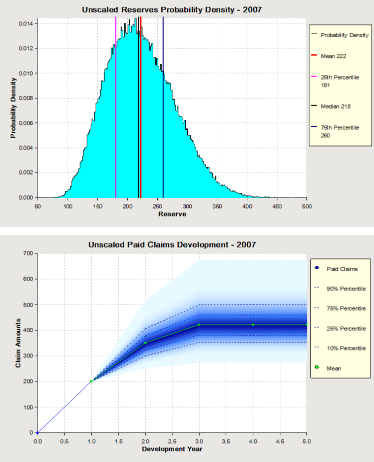
In the case of the last bullet point it could even happen that the cumulative payments get negative.

Bootstrapped probabilities (inclusive process variance) of both variants

Variant 1 (ResQ output)

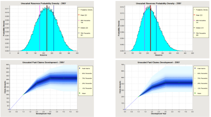


Variant 2 (ResQ output)



The bootstrapped distribution using variant 1 looks a bit too symmetric. Therefore, I would prefer variant 2 in this case.

Bootstrapped probabilities (inclusive process variance) of both variants



Variant 1 (ResQ output)

Variant 2 (ResQ output)

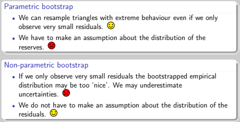
The bootstrapped distribution using variant 1 looks a bit too symmetric. Therefore, I would prefer variant 2 in this case.

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. ☺



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

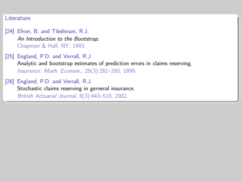
The results indicate, that the uncertainty may be underestimated by bootstrapping.

Literature

[24] Efron, B. and Tibshirani, R.J.
An Introduction to the Bootstrap.
Chapman & Hall, NY, 1993.

[25] England, P.D. and Verrall, R.J.
Analytic and bootstrap estimates of prediction errors in claims reserving.
Insurance: Math. Econom., 25(3):281–293, 1999.

[26] England, P.D. and Verrall, R.J.
Stochastic claims reserving in gerneral insurance.
British Actuarial Journal, 8(3):443–518, 2002.



Literature

[24] Efron, B. and Tibshirani, R.J.
An Introduction to the Bootstrap.
Chapman & Hall, NY, 1993.

[25] England, P.D. and Verrall, R.J.
Analytic and bootstrap estimates of prediction errors in claims reserving.
Insurance: Math. Econom., 25(3):281–293, 1999.

[26] England, P.D. and Verrall, R.J.
Stochastic claims reserving in gerneral insurance.
British Actuarial Journal, 8(3):443–518, 2002.

Stochastic Reserving

Lecture 11

Mid-Year Reserving

René Dahms

ETH Zurich, Spring 2021

12 May 2021

(Last update: 26 April 2021)

Stochastic Reserving

Lecture 11

Mid-Year Reserving

René Dahms

ETH Zurich, Spring 2021

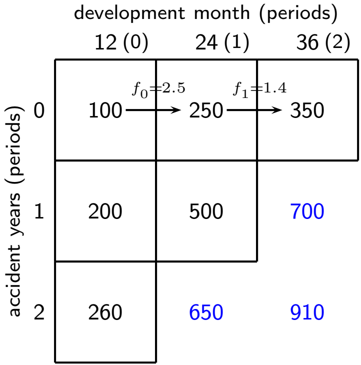
12 May 2021

(Last update: 26 April 2021)

7 Mid year reserving

Lecture 11: Table of contents

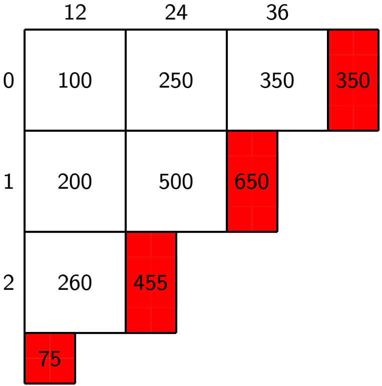
Chain-Ladder method at year end



development month (periods)

accident years (periods)	12 (0)	24 (1)	36 (2)
0	100	250	350
1	200	500	700
2	260	650	910

Chain-Ladder method at mid-year



	12	24	36	
0	100	250	350	350
1	200	500	650	
2	260	455		
	75			

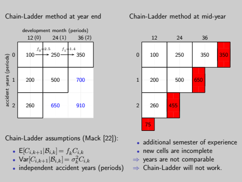
Chain-Ladder assumptions (Mack [22]):

- $\mathrm{E}[C_{i,k+1}|\mathcal{B}_{i,k}] = f_k C_{i,k}$
- $\mathrm{Var}[C_{i,k+1}|\mathcal{B}_{i,k}] = \sigma_k^2 C_{i,k}$
- independent accident years (periods)

- additional semester of experience
- new cells are incomplete

$\Rightarrow$ years are not comparable

$\Rightarrow$ Chain-Ladder will not work.



Chain-Ladder method at year end

development month (periods)

accident years (periods)	12 (0)	24 (1)	36 (2)
0	100	$\xrightarrow{f_0=2.5}$ 250	$\xrightarrow{f_1=1.4}$ 350
1	200	500	700
2	260	650	910

Chain-Ladder method at mid-year

	12	24	36	
0	100	250	350	350
1	200	500	650	
2	260	455		
	75			

Chain-Ladder assumptions (Mack [22]):

- $\mathrm{E}[C_{i,k+1}|\mathcal{B}_{i,k}] = f_k C_{i,k}$
- $\mathrm{Var}[C_{i,k+1}|\mathcal{B}_{i,k}] = \sigma_k^2 C_{i,k}$
- independent accident years (periods)

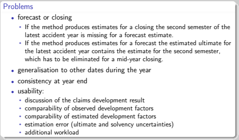
- additional semester of experience
- new cells are incomplete

⇒ 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



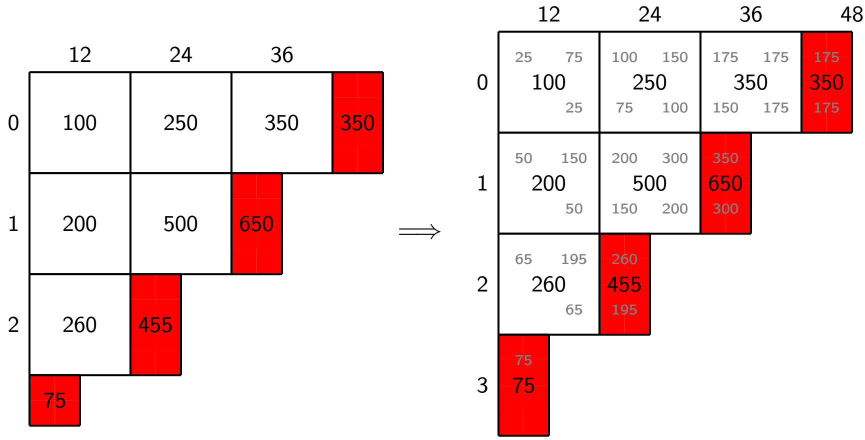
Problems

- 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 applyed 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 posibility 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?

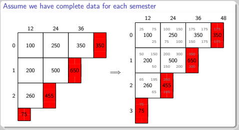
Assume we have complete data for each semester



	12	24	36	
0	100	250	350	350
1	200	500	650	
2	260	455		
3	75			

$\Longrightarrow$

	12	24	36	48
0	25 75 100 25	100 150 250 75 100	175 175 350 150 175	175 350 175
1	50 150 200 50	200 300 500 150 200	350 650 300	
2	65 195 260 65	260 455 195		
3	75 75			



Assume we have complete data for each semester

	12	24	36	
0	100	250	350	350
1	200	500	650	
2	260	455		
	75			

⟹

	12	24	36	48
0	25 75 100 25	100 150 250 75 100	175 175 350 150 175	175 350 175
1	50 150 200 50	200 300 500 150 200	350 650 300	
2	65 195 260 65	260 455 195		
3	75 75			

For the numerical example we took for each accident semester the following non-random development pattern

development month	6	12	18	24	30	36	42
cumulative	25	75	100	150	175	175	175
incremental	25	50	25	50	25	0	0

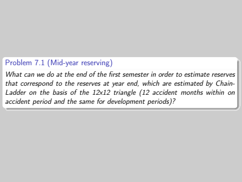
and accident semester volumes

accident semester	1H 0	2H 0	1H 1	2H 1	1H 2	2H 2	1H 3
volume	1	1	2	2	2.6	2.6	3

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 7.1 (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 on accident period and the same for development periods)?



Problem 7.1 (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 on accident period and the same for development periods)?

Step by step

	12	24	36	48
0	25 75 100 25	100 150 250 75 100	175 175 350 150 175	175 350 175
1	50 150 200 50	200 300 500 150 200	350 650 300	
2	65 195 260 65	260 455 195		
3	75 75			

$\Longrightarrow$

	6	12	18	24	30	36	42
0	25 25	75 100 25	100 175 75	150 250 100	175 325 150	175 350 175	175 350 175
1	50 50	150 200 50	200 350 150	300 500 200	350 650 300	700	700
2	65 65	195 260 65	260 455 195	650	845	910	910
3	75 75	300	525	750	975	1050	1050
f_k	4	7/4	10/7	1.3	14/13	1	

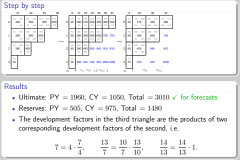
$\Longrightarrow$

	6	18	30	42
0	25 25	75 100 175 25 75	150 175 325 100 150	175 175 350 175 175
1	50 50	150 200 350 50 150	300 350 650 200 300	700
2	65 65	195 260 455 65 195	845	910
3	75 75	525	975	1050
f_k	7	13/7	14/13	

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.

$$7 = 4 \cdot \frac{7}{4}, \qquad \frac{13}{7} = \frac{10}{7} \cdot \frac{13}{10}, \qquad \frac{14}{13} = \frac{14}{13} \cdot 1.$$



Step by step

	12	24	36	48
0	100	250	350	350
1	200	500	650	
2	260	455		
3	75			

→

	6	12	18	24	30	36	42
0	25	100	175	250	325	350	350
1	50	200	350	500	650	700	700
2	65	260	455	650	845	910	910
3	75	300	525	750	975	1050	1050
f_k	4	$7/4$	$10/7$	1.3	$14/13$	1	

→

	6	18	30	42
0	25	175	325	350
1	50	350	650	700
2	65	455	845	910
3	75	525	975	1050
f_k	7	$13/7$	$14/13$	

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.

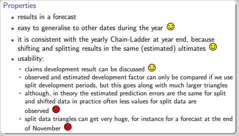
$$7 = 4 \cdot \frac{7}{4}, \qquad \frac{13}{7} = \frac{10}{7} \cdot \frac{13}{10}, \qquad \frac{14}{13} = \frac{14}{13} \cdot 1.$$

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 ☹



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

$$\mathcal{B}_{i,k} := \sigma\left(C_{i,j},\ 0 \le j \le k\right) \qquad \text{and} \qquad \mathcal{B}^*_{i,k} := \sigma\left(C^*_{i,j},\ 0 \le j \le k\right)$$

the corresponding information of the past. Then we have

$$C^*_{i,k} = C_{i,2k+1} \qquad \text{and} \qquad \mathcal{B}^*_{i,k} \subseteq \mathcal{B}_{i,2k+1}.$$

Assume that the semester data $C_{i,k}$ satisfies the Chain-Ladder assumptions, i.e.

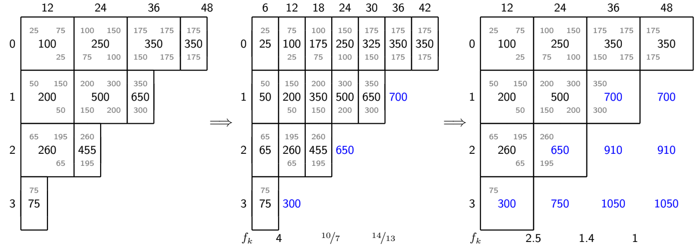
- $\mathsf{E}[C_{i,k+1}|\mathcal{B}_{i,k}] = f_k C_{i,k}$
- $\mathsf{Var}[C_{i,k+1}|\mathcal{B}_{i,k}] = \sigma_k^2 C_{i,k}$
- accident years are independent.

Then $C^*_{i,k}$ satisfies the Chain-Ladder assumptions, too:

- $\mathsf{E}\left[C^*_{i,k+1}\middle|\mathcal{B}^*_{i,k}\right] = \mathsf{E}\left[\mathsf{E}\left[C_{i,2(k+1)+1}\middle|\mathcal{B}_{i,2k+1}\right]\middle|\mathcal{B}^*_{i,k}\right] = \mathsf{E}\left[f_{2k+2}f_{2k+1}C_{i,2k+1}\middle|\mathcal{B}^*_{i,k}\right] = f_{2k+2}f_{2k+1}C^*_{i,k}$
- $$\begin{aligned}\mathsf{Var}\left[C^*_{i,k+1}\middle|\mathcal{B}^*_{i,k}\right] &= \mathsf{E}\left[\mathsf{Var}\left[C_{i,2(k+1)+1}\middle|\mathcal{B}_{i,2k+2}\right]\middle|\mathcal{B}^*_{i,k}\right] + \mathsf{Var}\left[\mathsf{E}\left[C_{i,2(k+1)+1}\middle|\mathcal{B}_{i,2k+2}\right]\middle|\mathcal{B}^*_{i,k}\right]\\ &= \mathsf{E}\left[\sigma^2_{2k+2}C_{i,2k+2}\middle|\mathcal{B}^*_{i,k}\right] + \mathsf{Var}\left[f_{2k+2}C_{i,2k+2}\middle|\mathcal{B}^*_{i,k}\right]\\ &= \sigma^2_{2k+2}\mathsf{E}\left[\mathsf{E}[C_{i,2k+2}|\mathcal{B}_{i,2k+1}]\middle|\mathcal{B}^*_{i,k}\right]\\ &\quad + f^2_{2k+2}\left(\mathsf{Var}\left[\mathsf{E}[C_{i,2k+2}|\mathcal{B}_{i,2k+1}]\middle|\mathcal{B}^*_{i,k}\right] + \mathsf{E}\left[\mathsf{Var}[C_{i,2k+2}|\mathcal{B}_{i,k+1}]\middle|\mathcal{B}^*_{i,k}\right]\right)\\ &= \sigma^2_{2k+2}f_{2k+1}C^*_{i,k} + f^2_{2k+2}(0 + \sigma^2_{2k+1}C^*_{i,k6}) = \underbrace{\left(\sigma^2_{2k+2}f_{2k+1} + f^2_{2k+2}\sigma^2_{2k+1}\right)}_{=\sigma_k^{*2}}C^*_{i,k}.\end{aligned}$$
- accident years are independent.

But in practice one often observes $\sigma_k^{*2} > \sigma^2_{2k+2}f_{2k+1} + f^2_{2k+2}\sigma^2_{2k+1}$

Step by step



	12	24	36	48
0	25 75 100 25	100 150 250 75 100	175 175 350 150 175	175 350 175
1	50 150 200 50	200 300 500 150 200	350 650 300	
2	65 195 260 65	260 455 195		
3	75 75			

⟹

	6	12	18	24	30	36	42
0	25 25	75 100 25	100 175 75	150 250 100	175 325 150	175 350 175	175 350 175
1	50 50	150 200 50	200 350 150	300 500 200	350 650 300	700	
2	65 65	195 260 65	260 455 195	650			
3	75 75	300					
f_k	4		$^{10}/_{7}$		$^{14}/_{13}$		

⟹

	12	24	36	48
0	25 75 100 25	100 150 250 75 100	175 175 350 150 175	175 350 175
1	50 150 200 50	200 300 500 150 200	350 700 300	700
2	65 195 260 65	260 650 195	910	910
3	75 300	750	1050	1050
f_k	2.5	1.4	1	

Results

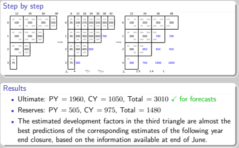
- Ultimate: PY $= 1960$, CY $= 1050$, Total $= 3010$ ✓ for forecasts
- Reserves: PY $= 505$, CY $= 975$, Total $= 1480$
- The estimated development factors in the third triangle are almost the best predictions of the corresponding estimates of the following year end closure, based on the information available at end of June.

2021-04-26

Stochastic Reserving

Mid year reserving

Methods for mid-year reserving



Step by step

	12	24	36	48
0	100	250	350	350
1	200	500	650	
2	260	455		
3	75			

→

	6	12	18	24	30	36	42
0	25	100	175	250	325	350	350
1	50	200	350	500	650	700	
2	65	260	455	650			
3	75	300					
f_k		4		$^{10}/_{7}$		$^{14}/_{13}$	

→

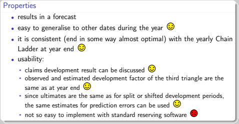
	12	24	36	48
0	100	250	350	350
1	200	500	700	700
2	260	650	910	910
3	300	750	1050	1050
f_k		2.5	1.4	1

Results

- Ultimate: PY $= 1960$, CY $= 1050$, Total $= 3010$ ✓ for forecasts
- Reserves: PY $= 505$, CY $= 975$, Total $= 1480$
- The estimated development factors in the third triangle are almost the best predictions of the corresponding estimates of the following year end closure, based on the information available at end of June.

Properties

- results in a forecast
- easy to generalise to other dates during the year ☺
- it is consistent (end in some way almost optimal) with the yearly Chain Ladder at year end ☺
- usability:
    - claims development result can be discussed ☺
    - observed and estimated development factor of the third triangle 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 ☺
    - not so easy to implement with standard reserving software ☹



Properties

- results in a forecast
- easy to generalise to other dates during the year ☺
- it is consistent (end in some way almost optimal) with the yearly Chain Ladder at year end ☺
- usability:
  - claims development result can be discussed ☺
  - observed and estimated development factor of the third triangle 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 ☺
  - not so easy to implement with standard reserving software ☹

Using the same notation like in the case of split development periods we get

$$\begin{aligned}\widehat{f}_k &= \frac{\sum_{i=0}^{I-k} \widehat{C}^*_{i,k+1}}{\sum_{i=0}^{I-k} C^*_{i,k}} = \frac{\sum_{i=0}^{I-k-1} C^*_{i,k+1} + \widehat{f}^{hy}_{2k+2} C_{I-k,2k+2}}{\sum_{i=0}^{I-k} C^*_{i,k}}\\ &= \frac{\sum_{i=0}^{I-k-1} C^*_{i,k}}{\sum_{i=0}^{I-k} C^*_{i,k}} \widehat{f}^{ye}_k + \left(1 - \frac{\sum_{i=0}^{I-k-1} C^*_{i,k}}{\sum_{i=0}^{I-k} C^*_{i,k}}\right) \frac{C_{I-k,2k+2}}{C_{I-k,2k+1}} \widehat{f}^{hy}_{2k+2}.\end{aligned}$$

That means the estimated development factors $\widehat{f}_k$ are a weighted mean of the estimated development factors $\widehat{f}^{ye}_k$ 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 $\widehat{f}^{hy}_{2k+2}$.

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 α_k in order to forecast the estimated development factors $\widehat{f}^{ye+1}_k$ of the next year end closing, i. e. α_k that minimize (see [27] for details)

$$\mathsf{E}\left[\left((1-\alpha_k)\widehat{f}^{ye}_k + \alpha_k \frac{C_{I-k,2k+2}}{C_{I-k,2k+1}} \widehat{f}^{hy}_{2k+2} - \widehat{f}^{ye+1}_k\right)^2 \middle| C_{i,j} \text{ known at end of June}\right].$$

Step by step

	12	24	36	48
0	25 75 100 25	100 150 250 75 100	175 175 350 150 175	175 350 175
1	50 150 200 50	200 300 500 150 200	350 650 300	
2	65 195 260 65	260 455 195		
3	75 75			

$\Longrightarrow$

	12	24	36	
	25 75	100 150	175 175	175
0	25 75 125 50	100 150 350 150 200	175 175 525 300 350	
1	50 150 215 65	200 300 560 195 260	840 805	
2	65 195 270 75	723 690	1084 980	
f_k		2.68	1.5	

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'

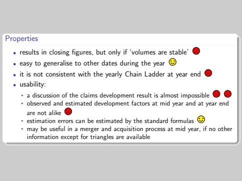
Step by step

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

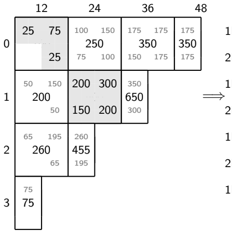


Properties

- 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 next method (subsection 11.2.4).

Step by step



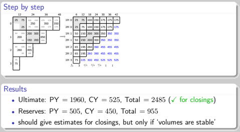
	6	12	18	24	30	36	42
1H 0	25	75	100	150	175	175	175
2H 0	25	75	100	150	175	175	175
1H 1	50	150	200	300	350	350	350
2H 1	50	150	200	300	350	350	350
1H 2	65	195	260	390	455	455	455
2H 2	65	195	260	390	455	455	455
1H 3	75	225	300	450	525	525	525
f_k	3	$4/3$	$3/2$	$7/6$	1	1	

Results

- Ultimate: PY $= 1960$, CY $= 525$, Total $= 2485$ (✓ for closings)
- Reserves: PY $= 505$, CY $= 450$, Total $= 955$
- should give estimates for closings, but only if 'volumes are stable'

2021-04-26

Stochastic Reserving
└ Mid year reserving
  └ Methods for mid-year reserving



Step by step

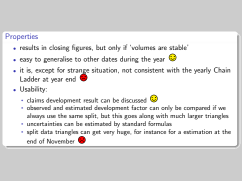
	6	12	18	24	30	36	42
1H 0	25	75	100	150	175	175	175
2H 0	25	75	100	150	175	175	175
1H 1	50	150	200	300	350	350	350
2H 1	50	150	200	300	350	350	350
1H 2	65	195	260	390	455	455	455
2H 2	65	195	260	390	455	455	455
1H 3	75	225	300	450	525	525	525
f_k	3	$4/3$	$3/2$	$7/4$	1	1	

Results

- Ultimate: PY $= 1960$, CY $= 525$, Total $= 2485$ (✓ for closings)
- Reserves: PY $= 505$, CY $= 450$, Total $= 955$
- 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, 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 goes along with much larger triangles
    - uncertainties can be estimated by standard formulas
    - split data triangles can get very huge, for instance for a estimation at the end of November ☹



Properties

- results in closing figures, but only if ‘volumes are stable’
- 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 goes along with much larger triangles
  - uncertainties can be estimated by standard formulas
  - split data triangles can get very huge, for instance for a estimation 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

$$\mathcal{B}_{i,k} := \sigma\left(C_{i,j},\ 0 \leq j \leq k\right) \qquad \text{and} \qquad \mathcal{B}^*_{i,k} := \sigma\left(C^*_{i,j},\ 0 \leq j \leq k\right)$$

the corresponding information of the past. Then we have

$$C^*_{i,k} = C_{2i,2k+1} + C_{2i+1,2k} \qquad \text{and} \qquad \mathcal{B}^*_{i,k} \subseteq \sigma\left(\mathcal{B}_{2i,2k+1} \cup \mathcal{B}_{2i+1,2k}\right).$$

Assume that $C_{i,k}$ and $C^*_{i,k}$ satisfy the Chain-Ladder assumptions, i.e.

- $\mathsf{E}\left[C_{i,k+1}\middle|\mathcal{B}_{i,k}\right] = f_k C_{i,k}$
- $\mathsf{Var}\left[C_{i,k+1}\middle|\mathcal{B}_{i,k}\right] = \sigma_k^2 C_{i,k}$
- accident semester are independent.

- $\mathsf{E}\left[C^*_{i,k+1}\middle|\mathcal{B}^*_{i,k}\right] = g_k C^*_{i,k}$
- $\mathsf{Var}\left[C^*_{i,k+1}\middle|\mathcal{B}^*_{i,k}\right] = \tau_k^2 C^*_{i,k}$
- accident years are independent.

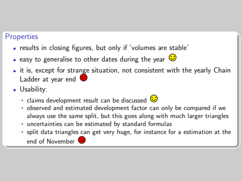
Then we get:

$$\begin{aligned} g_k(C_{2i,2k+1} + C_{2i+1,2k}) = g_k C^*_{i,k} &= \mathsf{E}\left[C_{2i,2(k+1)+1} + C_{2i+1,2(k+1)}\middle|\mathcal{B}^*_{i,k}\right] \\ &= \mathsf{E}\left[\mathsf{E}\left[C_{2i,2(k+1)+1} + C_{2i+1,2(k+1)}\middle|\mathcal{B}_{2i,2k+1}, \mathcal{B}_{2i+1,2k}\right]\middle|\mathcal{B}^*_{i,k}\right] \\ &= \mathsf{E}\left[f_{2k+2}f_{2k+1}C_{2i,2k+1} + f_{2k+1}f_{2k}C_{2i+1,2k}\middle|\mathcal{B}^*_{i,k}\right]. \qquad (7.1) \end{aligned}$$

Therefore, it follows

$$0 = (g_k - f_{2k+2}f_{2k+1})\mathsf{E}\left[C_{2i,2k+1}\middle|\mathcal{D}\right] + (g_k - f_{2k+1}f_{2k})\mathsf{E}\left[C_{2i+1,2k}\middle|\mathcal{D}\right], \qquad (7.2)$$

for each σ-algebra $\mathcal{D} \subseteq \mathcal{B}^*_{i,k}$.



Properties

- results in closing figures, but only if ‘volumes are stable’
- 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 goes along with much larger triangles
  - uncertainties can be estimated by standard formulas
  - split data triangles can get very huge, for instance for a estimation at the end of November ☹

Moreover, multiplying (7.1) by $(C_{2i,2k+1} + C_{2i+1,2k})$, resorting the terms and using (7.2) we get and

$$0 = (g_k - f_{2k+2}f_{2k+1})\mathsf{Var}\big[C_{2i,2k+1}\big|\mathcal{D}\big] + (g_k - f_{2k+1}f_{2k})\mathsf{Var}\big[C_{2i+1,2k}\big|\mathcal{D}\big],$$

which is only possible if

- $\mathsf{Var}\big[C_{i,k}\big] = 0$, 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
- $$\frac{\mathsf{Var}\big[C_{2i,2k+1}\big|\mathcal{D}\big]}{f_{2k}\mathsf{Var}\big[C_{2i+1,2k}\big|\mathcal{D}\big]} = -\frac{g_k - f_{2k+1}f_{2k}}{f_{2k}(g_k - f_{2k+2}f_{2k+1})} = \frac{\mathsf{E}\big[C_{2i,2k+1}\big|\mathcal{D}\big]}{f_{2k}\mathsf{E}\big[C_{2i+1,2k}\big|\mathcal{D}\big]} = \frac{\mathsf{E}\big[C_{2i,0}\big]}{\mathsf{E}\big[C_{2i+1,0}\big]},$$

  where the last equation is true, because the second term is independent of the σ-algebra $\mathcal{D}$ and we can take the trivial σ-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 theses cases are very strange circumstances. □

Properties

- results in closing figures, but only if ‘volumes are stable’
- 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 goes along with much larger triangles
  - uncertainties can be estimated by standard formulas
  - split data triangles can get very huge, for instance for a estimation at the end of November ☹

Step by step

	12	24	36	48
0	25 75 100 25	100 150 250 75 100	175 175 350 150 175	175 350 175
1	50 150 200 50	200 300 500 150 200	350 650 300	
2	65 195 260 65	260 455 195		
3	75 75			

$\Longrightarrow$

	6	18		30		42	
1H 0	25	75	100	150	175	175	175
1H 1	50	150	200	300	350		
1H 2	65	195	260				
1H 3	75						

$\Longrightarrow$

	6	18	30	42
1H 0	25	100	175	175
1H 1	50	200	350	350
1H 2	65	260	455	455
1H 3	75	300	525	525
f_k	4	7/4	1	

	12		24		36	
2H 0	25	75	100	150	175	175
2H 1	50	150	200	300		
2H 2	65	195				

	12	24	36
2H 0	75	150	175
2H 1	150	300	350
2H 2	195	390	455
f_k	2	7/6	

Results

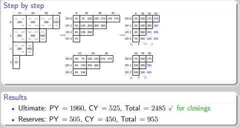
- Ultimate: PY $= 1960$, CY $= 525$, Total $= 2485$ ✓ for closings
- Reserves: PY $= 505$, CY $= 450$, Total $= 955$

2021-04-26

Stochastic Reserving

Mid year reserving

Methods for mid-year reserving



Step by step

	12	24	36	48
0	100	250	350	350
1	200	500	650	
2	260	455		
3	75			

$\longrightarrow$

	6		18		30		42
1H 0	25	75	100	150	175	175	175
1H 1	50	150	200	300	350		
1H 2	65	195	260				
1H 3	75						

	12		24		36	
2H 0	25	75	100	150	175	175
2H 1	50	150	200	300		
2H 2	65	195				

$\longrightarrow$

	6	18	30	42
1H 0	25	100	175	175
1H 1	50	200	350	350
1H 2	65	260	455	455
1H 3	75	300	525	525
f_Δ	4	$7/4$	1	

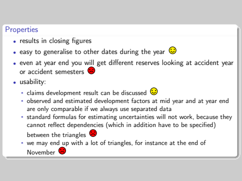
	12	24	36
2H 0	75	150	175
2H 1	150	300	350
2H 2	195	390	455
f_Δ	2	$7/6$	

Results

- Ultimate: PY $= 1960$, CY $= 525$, Total $= 2485$ ✓ for closings
- Reserves: PY $= 505$, CY $= 450$, Total $= 955$

Properties

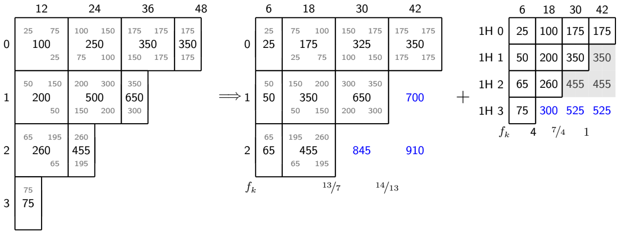
- results in closing figures
- easy to generalise to other dates during the year ☺
- even at year end you will get different reserves looking at accident year or accident semesters ☹
- usability:
  - claims development result can be discussed ☺
  - observed and estimated development factors at mid year 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) between the triangles ☹
  - we may end up with a lot of triangles, for instance at the end of November ☹

Step by step



	12	24	36	48
0	25 75 100 25	100 150 250 75 100	175 175 350 150 175	175 350 175
1	50 150 200 50	200 300 500 150 200	350 650 300	
2	65 195 260 65	260 455 195		
3	75 75			

$\Longrightarrow$

	6	18	30	42
0	25 25	75 100 175 25 75	150 175 325 100 150	175 175 350 175 175
1	50 50	150 200 350 50 150	300 350 650 200 300	700
2	65 65	195 260 455 65 195	845	910
f_k		$^{13}/_{7}$	$^{14}/_{13}$	

$+$

	6	18	30	42
1H 0	25	100	175	175
1H 1	50	200	350	350
1H 2	65	260	455	455
1H 3	75	300	525	525
f_k	4	$^{7}/_{4}$	1	

Results

- Ultimate: PY $= 1960$, CY $= 525$, Total $= 2485$ ✓ for closings
- Reserves: PY $= 505$, CY $= 450$, Total $= 955$

Step by step

Results

- Ultimate: PY $= 1960$, CY $= 525$, Total $= 2485$ ✓ for closings
- Reserves: PY $= 505$, CY $= 450$, Total $= 955$

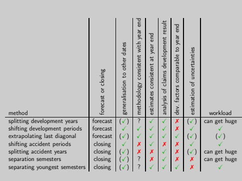
Properties

- 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

method	forecast or closing	generalisation to other dates	methodology consistent with year end	estimates consistent at year end	analysis of claims development result	dev. factors comparable to year end	estimation of uncertainties	workload
splitting development years	forecast	(✓)	?	✓	✓	✗	(✓)	can get huge
shifting development periods	forecast	✓	✓	✓	✓	✗	✓	✓
extrapolating last diagonal	forecast	(✓)	✓	✓	✓	✓	(✓)	(✓)
shifting accident periods	closing	✓	✗	✓	✗	✗	✓	✓
splitting accident years	closing	(✓)	✗	✗	✓	✗	(✓)	can get huge
separation semesters	closing	(✓)	?	✗	✓	✗	✗	can get huge
separating youngest semesters	closing	(✓)	?	✓	✓	✓	✗	✓

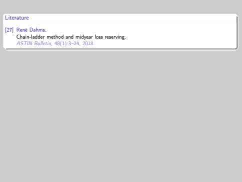


method	forecast or closing	generalisation to other dates	methodology consistent with year end	estimates consistent at year end	analysis of claims development result	dev. factors comparable to year end	estimation of uncertainties	workload
splitting development years	forecast	(✓)	?	✓	✓	✗	(✓)	can get huge
shifting development periods	forecast	✓	✓	✓	✓	✗	✓	✓
extrapolating last diagonal	forecast	(✓)	✓	✓	✓	✓	(✓)	(✓)
shifting accident periods	closing	✓	✗	✓	✗	✗	✓	✓
splitting accident years	closing	(✓)	✗	✗	✓	✗	(✓)	can get huge
separation semesters	closing	(✓)	?	✗	✓	✗	✗	can get huge
separating youngest semesters	closing	(✓)	?	✓	✓	✓	✗	✓

- (✓) 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 (and the 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, which I would prefer if I had to implement a software

Literature

[27] René Dahms.
Chain-ladder method and midyear loss reserving.
ASTIN Bulletin, 48(1):3–24, 2018.



Literature

[27] René Dahms.
Chain-ladder method and midyear loss reserving.
ASTIN Bulletin, 48(1):3–24, 2018.

Stochastic Reserving

Lecture 12

CLM: **Bayesian & credibility approach**

René Dahms

ETH Zurich, Spring 2021

19 May 2021

(Last update: 26 April 2021)

Stochastic Reserving

Stochastic Reserving

Lecture 12

CLM: **Bayesian & credibility approach**

René Dahms

ETH Zurich, Spring 2021

19 May 2021

(Last update: 26 April 2021)

8 CLM: Bayesian & credibility approach

Lecture 12: Table of contents

Recapitulation of the Chain-Ladder method

Let $C_{i,k} := \sum_{j=0}^{k} S_{i,j}$. If we have

i)$^{\text{CLM}}$ $\mathsf{E}[C_{i,k+1}|\mathcal{B}_{i,k}] = f_k C_{i,k}$,
ii)$^{\text{CLM}}$ $\mathsf{Var}[C_{i,k+1}|\mathcal{B}_{i,k}] = \sigma_k^2 C_{i,k}$ and
iii)$^{\text{CLM}}$ accident periods are independent.

Then $\widehat{C}_{i,k+1} := \widehat{f}_k \cdot \ldots \cdot \widehat{f}_{I-i} C_{i,I-i}$ with

$$\widehat{f}_k := \sum_{i=0}^{I-1-k} \frac{C_{i,k}}{\sum_{h=0}^{I-1-k} C_{h,k}} \frac{C_{i,k+1}}{C_{i,k}}$$

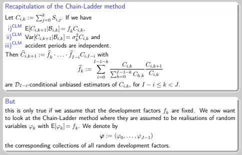
are $\mathcal{D}_{I-i}$-conditional unbiased estimators of $C_{i,k}$, for $I - i \leq k < J$.

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 φ_k with $\mathsf{E}[\varphi_k] = f_k$. We denote by

$$\boldsymbol{\varphi} := (\varphi_0, \ldots, \varphi_{J-1})$$

the corresponding collections of all random development factors.



Recapitulation of the Chain-Ladder method

Let $C_{i,k} := \sum_{j=0}^{k} S_{i,j}$. If we have

i)$^{\mathsf{CLM}}$ $\mathsf{E}[C_{i,k+1}|\mathcal{B}_{i,k}] = f_k C_{i,k}$,
ii)$^{\mathsf{CLM}}$ $\mathsf{Var}[C_{i,k+1}|\mathcal{B}_{i,k}] = \sigma_k^2 C_{i,k}$ and
iii)$^{\mathsf{CLM}}$ accident periods are independent.

Then $\widehat{C}_{i,k+1} := \widehat{f}_k \cdot \ldots \cdot \widehat{f}_{I-i} C_{i,I-i}$ with

$$\widehat{f}_k := \sum_{i=0}^{I-1-k} \frac{C_{i,k}}{\sum_{h=0}^{I-1-k} C_{h,k}} \frac{C_{i,k+1}}{C_{i,k}}$$

are $\mathcal{D}_{I-i}$-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 φ_k with $\mathsf{E}[\varphi_k] = f_k$. We denote by

$$\boldsymbol{\varphi} := (\varphi_0, \ldots, \varphi_{J-1})$$

the corresponding collections of all random development factors.

Note, everything will stay correct if we replace $\mathcal{B}_{i,k}$ with $\mathcal{D}_k^{i+k}$ and skip the independence assumption.

Assumption 8.A (Bayesian Chain-Ladder method)

We assume that

i)^{Bay} $\mathsf{E}[C_{i,k+1}|\boldsymbol{\varphi},\mathcal{B}_{i,k}] = \varphi_k C_{i,k}$,

ii)^{Bay} $\mathsf{Var}[C_{i,k+1}|\boldsymbol{\varphi},\mathcal{B}_{i,k}] = \sigma_k^2(\boldsymbol{\varphi}) C_{i,k}$,

iii)^{Bay} *conditional given* $\boldsymbol{\varphi}$ *the accident periods are independent and*

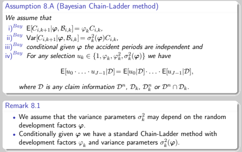
iv)^{Bay} *For any selection* $u_k \in \{1, \varphi_k, \varphi_k^2, \sigma_k^2(\boldsymbol{\varphi})\}$ *we have*

$$\mathsf{E}[u_0 \cdot \ldots \cdot u_{J-1}|\mathcal{D}] = \mathsf{E}[u_0|\mathcal{D}] \cdot \ldots \cdot \mathsf{E}[u_{J-1}|\mathcal{D}],$$

where $\mathcal{D}$ *is any claim information* $\mathcal{D}^n$, $\mathcal{D}_k$, $\mathcal{D}_k^n$ *or* $\mathcal{D}^n \cap \mathcal{D}_k$.

Remark 8.1

- We assume that the variance parameters σ_k^2 may depend on the random development factors $\boldsymbol{\varphi}$.
- Conditionally given $\boldsymbol{\varphi}$ we have a standard Chain-Ladder method with development factors φ_k and variance parameters $\sigma_k^2(\boldsymbol{\varphi})$.



Assumption 8.A (Bayesian Chain-Ladder method)

We assume that

- i)^{Bay} $\mathsf{E}[C_{i,k+1}|\boldsymbol{\varphi},\mathcal{B}_{i,k}]=\varphi_k C_{i,k}$,
- ii)^{Bay} $\mathsf{Var}[C_{i,k+1}|\boldsymbol{\varphi},\mathcal{B}_{i,k}]=\sigma_k^2(\boldsymbol{\varphi}) C_{i,k}$,
- iii)^{Bay} *conditional given $\boldsymbol{\varphi}$ the accident periods are independent and*
- iv)^{Bay} *For any selection $u_k \in \{1,\varphi_k,\varphi_k^2,\sigma_k^2(\boldsymbol{\varphi})\}$ we have*

$$\mathsf{E}[u_0\cdot\ldots\cdot u_{J-1}|\mathcal{D}]=\mathsf{E}[u_0|\mathcal{D}]\cdot\ldots\cdot\mathsf{E}[u_{J-1}|\mathcal{D}],$$

where $\mathcal{D}$ is any claim information $\mathcal{D}^n$, $\mathcal{D}_k$, $\mathcal{D}_k^n$ or $\mathcal{D}^n\cap\mathcal{D}_k$.

Remark 8.1

- We assume that the variance parameters σ_k^2 may depend on the random development factors $\boldsymbol{\varphi}$.
- Conditionally given $\boldsymbol{\varphi}$ we have a standard Chain-Ladder method with development factors φ_k and variance parameters $\sigma_k^2(\boldsymbol{\varphi})$.

Note, everything will stay correct if we replace $\mathcal{B}_{i,k}$ with $\mathcal{D}_k^{i+k}$ and skip the independence assumption.

Definition 8.2 (Bayes estimators)

Let Z be a random variable and $\mathcal{D}$ some σ-algebra (for instance the information contained in some observations). The Bayes estimator Z^{Bay} of Z given $\mathcal{D}$ is defined by

$$Z^{Bay} := \mathsf{E}[Z|\mathcal{D}].$$

Corollary 8.3

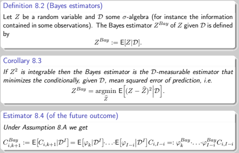
If Z^2 is integrable then the Bayes estimator is the $\mathcal{D}$-measurable estimator that minimizes the conditionally, given $\mathcal{D}$, mean squared error of prediction, i.e.

$$Z^{Bay} = \underset{\widehat{Z}}{\operatorname{argmin}} \, \mathsf{E}\left[(Z - \widehat{Z})^2 \middle| \mathcal{D}\right].$$

Estimator 8.4 (of the future outcome)

Under Assumption 8.A we get

$$C_{i,k+1}^{Bay} := \mathsf{E}\left[C_{i,k+1}|\mathcal{D}^I\right] = \mathsf{E}\left[\varphi_k|\mathcal{D}^I\right] \cdot \ldots \cdot \mathsf{E}\left[\varphi_{I-i}\middle|\mathcal{D}^I\right] C_{i,I-i} =: \varphi_k^{Bay} \cdot \ldots \cdot \varphi_{I-i}^{Bay} C_{i,I-i}$$



Definition 8.2 (Bayes estimators)

Let Z be a random variable and $\mathcal{D}$ some σ-algebra (for instance the information contained in some observations). The Bayes estimator Z^{Bay} of Z given $\mathcal{D}$ is defined by

$$Z^{Bay} := \mathsf{E}[Z|\mathcal{D}].$$

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.

$$Z^{Bay} = \underset{\widehat{Z}}{\operatorname{argmin}}\, \mathsf{E}\Big[(Z-\widehat{Z})^2\Big|\mathcal{D}\Big].$$

Estimator 8.4 (of the future outcome)

Under Assumption 8.A we get

$$C_{i,k+1}^{Bay} := \mathsf{E}[C_{i,k+1}|\mathcal{D}^I] = \mathsf{E}[\varphi_k|\mathcal{D}^I]\cdot\ldots\cdot\mathsf{E}[\varphi_{I-i}|\mathcal{D}^I]C_{i,I-i} =: \varphi_k^{Bay}\cdot\ldots\cdot\varphi_{I-i}^{Bay}C_{i,I-i}$$

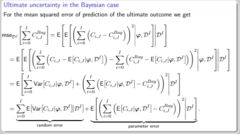
- The corollary is true, because the conditional expectation is the orthogonal projection onto the subspace of all $\mathcal{D}$ measurable functions (within the space of all square integrable functions).
- **Proof of Estimator 8.4**:

$$\begin{aligned} C_{i,k+1}^{Bay} &= \mathsf{E}\Big[C_{i,k+1}\Big|\mathcal{D}^I\Big] = \mathsf{E}\Big[\mathsf{E}\Big[C_{i,k+1}\Big|\boldsymbol{\varphi},\mathcal{D}^I\Big]\Big|\mathcal{D}^I\Big] \\ &= \qquad \underbrace{\mathsf{E}\Big[\varphi_k\cdot\ldots\cdot\varphi_{I-i}C_{i,I-i}\Big|\mathcal{D}^I\Big]}_{\text{standard CLM for fixed development factors}} \\ &= \underbrace{\mathsf{E}\Big[\varphi_k|\mathcal{D}^I\Big]\cdot\ldots\cdot\mathsf{E}\Big[\varphi_{I-i}\Big|\mathcal{D}^I\Big]C_{i,I-i}\,.}_{\text{iv)}^{Bay}} \end{aligned}$$

Ultimate uncertainty in the Bayesian case

For the mean squared error of prediction of the ultimate outcome we get

$$
\begin{aligned}
\mathrm{mse}_{\mathcal{D}^I}\left[\sum_{i=0}^{I} C_{i,J}^{Bay}\right] &= \mathsf{E}\left[\mathsf{E}\left[\left(\sum_{i=0}^{I}\left(C_{i,J} - C_{i,J}^{Bay}\right)\right)^2 \middle| \boldsymbol{\varphi}, \mathcal{D}^I\right] \middle| \mathcal{D}^I\right] \\
&= \mathsf{E}\left[\mathsf{E}\left[\left(\sum_{i=0}^{I}\left(C_{i,J} - \mathsf{E}\left[C_{i,J}|\boldsymbol{\varphi}, \mathcal{D}^I\right]\right) - \sum_{i=0}^{I}\left(C_{i,J}^{Bay} - \mathsf{E}\left[C_{i,J}|\boldsymbol{\varphi}, \mathcal{D}^I\right]\right)\right)^2 \middle| \boldsymbol{\varphi}, \mathcal{D}^I\right] \middle| \mathcal{D}^I\right] \\
&= \mathsf{E}\left[\sum_{i=0}^{I} \mathrm{Var}\left[C_{i,J}|\boldsymbol{\varphi}, \mathcal{D}^I\right] + \left(\sum_{i=0}^{I}\left(\mathsf{E}\left[C_{i,J}|\boldsymbol{\varphi}, \mathcal{D}^I\right] - C_{i,J}^{Bay}\right)\right)^2 \middle| \mathcal{D}^I\right] \\
&= \underbrace{\sum_{i=0}^{I} \mathsf{E}\left[\mathrm{Var}\left[C_{i,J}|\boldsymbol{\varphi}, \mathcal{D}^I\right] \middle| \mathcal{D}^I\right]}_{\text{random error}} + \underbrace{\mathsf{E}\left[\left(\sum_{i=0}^{I}\left(\mathsf{E}\left[C_{i,J}|\boldsymbol{\varphi}, \mathcal{D}^I\right] - C_{i,J}^{Bay}\right)\right)^2 \middle| \mathcal{D}^I\right]}_{\text{parameter error}}.
\end{aligned}
$$



Ultimate uncertainty in the Bayesian case

For the mean squared error of prediction of the ultimate outcome we get

$$\begin{aligned}
\mathsf{mse}_{\mathcal{D}^I}\left[\sum_{i=0}^{I} C_{i,J}^{Bay}\right] &= \mathsf{E}\left[\mathsf{E}\left[\left(\sum_{i=0}^{I}\left(C_{i,J}-C_{i,J}^{Bay}\right)\right)^2\middle|\boldsymbol{\varphi},\mathcal{D}^I\right]\middle|\mathcal{D}^I\right]\\
&= \mathsf{E}\left[\mathsf{E}\left[\left(\sum_{i=0}^{I}\left(C_{i,J}-\mathsf{E}[C_{i,J}|\boldsymbol{\varphi},\mathcal{D}^I]\right)-\sum_{i=0}^{I}\left(C_{i,J}^{Bay}-\mathsf{E}[C_{i,J}|\boldsymbol{\varphi},\mathcal{D}^I]\right)\right)^2\middle|\boldsymbol{\varphi},\mathcal{D}^I\right]\middle|\mathcal{D}^I\right]\\
&= \mathsf{E}\left[\sum_{i=0}^{I}\mathsf{Var}[C_{i,J}|\boldsymbol{\varphi},\mathcal{D}^I]+\left(\sum_{i=0}^{I}\left(\mathsf{E}[C_{i,J}|\boldsymbol{\varphi},\mathcal{D}^I]-C_{i,J}^{Bay}\right)\right)^2\middle|\mathcal{D}^I\right]\\
&= \underbrace{\sum_{i=0}^{I}\mathsf{E}\left[\mathsf{Var}[C_{i,J}|\boldsymbol{\varphi},\mathcal{D}^I]\middle|\mathcal{D}^I\right]}_{\text{random error}}+\underbrace{\mathsf{E}\left[\left(\sum_{i=0}^{I}\left(\mathsf{E}[C_{i,J}|\boldsymbol{\varphi},\mathcal{D}^I]-C_{i,J}^{Bay}\right)\right)^2\middle|\mathcal{D}^I\right]}_{\text{parameter error}}.
\end{aligned}$$

Derivation of the random error

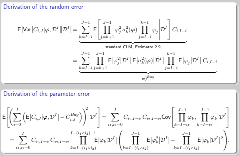
$$\mathsf{E}\left[\mathsf{Var}\left[C_{i,J}|\boldsymbol{\varphi},\mathcal{D}^I\right]\middle|\mathcal{D}^I\right] = \underbrace{\sum_{k=I-i}^{J-1} \mathsf{E}\left[\prod_{j=k+1}^{J-1} \varphi_j^2\,\sigma_k^2(\boldsymbol{\varphi}) \prod_{j=I-i}^{k-1} \varphi_j \middle| \mathcal{D}^I\right] C_{i,I-i}}_{\text{standard CLM, Estimator 2.9}}$$

$$= \underbrace{\sum_{k=I-i}^{J-1} \prod_{j=k+1}^{J-1} \mathsf{E}\left[\varphi_j^2\middle|\mathcal{D}^I\right] \mathsf{E}\left[\sigma_k^2(\boldsymbol{\varphi})\middle|\mathcal{D}^I\right] \prod_{j=I-i}^{k-1} \mathsf{E}\left[\varphi_j\middle|\mathcal{D}^I\right] C_{i,I-i}\,.}_{\text{iv})^{Bay}}$$

Derivation of the parameter error

$$\mathsf{E}\left[\left(\sum_{i=0}^{I}\left(\mathsf{E}\left[C_{i,J}|\boldsymbol{\varphi},\mathcal{D}^I\right] - C_{i,J}^{Bay}\right)\right)^2\middle|\mathcal{D}^I\right] = \sum_{i_1,i_2=0}^{I} C_{i_1,I-i_1} C_{i_2,I-i_2} \mathsf{Cov}\left[\prod_{k=I-i_1}^{J-1}\varphi_k, \prod_{k=I-i_2}^{J-1}\varphi_k\middle|\mathcal{D}^I\right]$$

$$= \sum_{i_1,i_2=0}^{I} C_{i_1,I-i_1} C_{i_2,I-i_2} \prod_{k=I-(i_1\vee i_2)}^{I-(i_1\wedge i_2)-1} \mathsf{E}\left[\varphi_k|\mathcal{D}^I\right] \left(\prod_{k=I-(i_1\wedge i_2)}^{J-1} \mathsf{E}\left[\varphi_k^2|\mathcal{D}^I\right] - \prod_{k=I-(i_1\wedge i_2)}^{J-1} \mathsf{E}\left[\varphi_k|\mathcal{D}^I\right]^2\right).$$



Derivation of the random error

$$\mathsf{E}\Big[\mathsf{Var}\big[C_{i,J}|\boldsymbol{\varphi},\mathcal{D}^I\big]\Big|\mathcal{D}^I\Big] = \underbrace{\sum_{k=I-i}^{J-1} \mathsf{E}\left[\prod_{j=k+1}^{J-1}\varphi_j^2\,\sigma_k^2(\boldsymbol{\varphi})\prod_{j=I-i}^{k-1}\varphi_j\middle|\mathcal{D}^I\right]C_{i,I-i}}_{\text{standard CLM, Estimator 2.9}}$$
$$= \underbrace{\sum_{k=I-i}^{J-1}\prod_{j=k+1}^{J-1}\mathsf{E}\big[\varphi_j^2|\mathcal{D}^I\big]\,\mathsf{E}\big[\sigma_k^2(\boldsymbol{\varphi})|\mathcal{D}^I\big]\prod_{j=I-i}^{k-1}\mathsf{E}\big[\varphi_j|\mathcal{D}^I\big]\,C_{i,I-i}\,.}_{\text{iv)}^{Bay}}$$

Derivation of the parameter error

$$\mathsf{E}\left[\left(\sum_{i=0}^{I}\left(\mathsf{E}\big[C_{i,J}|\boldsymbol{\varphi},\mathcal{D}^I\big]-C_{i,J}^{Bay}\right)\right)^2\middle|\mathcal{D}^I\right] = \sum_{i_1,i_2=0}^{I} C_{i_1,I-i_1}C_{i_2,I-i_2}\mathsf{Cov}\left[\prod_{k=I-i_1}^{J-1}\varphi_k,\prod_{k=I-i_2}^{J-1}\varphi_k\middle|\mathcal{D}^I\right]$$
$$= \sum_{i_1,i_2=0}^{I} C_{i_1,I-i_1}C_{i_2,I-i_2}\prod_{k=I-(i_1\vee i_2)}^{I-(i_1\wedge i_2)-1}\mathsf{E}\big[\varphi_k|\mathcal{D}^I\big]\left(\prod_{k=I-(i_1\wedge i_2)}^{J-1}\mathsf{E}\big[\varphi_k^2|\mathcal{D}^I\big]-\prod_{k=I-(i_1\wedge i_2)}^{J-1}\mathsf{E}\big[\varphi_k|\mathcal{D}^I\big]^2\right).$$

parameter error:

$$\mathsf{E}\left[\left(\sum_{i=0}^{I}\left(\mathsf{E}\Big[C_{i,J}\Big|\boldsymbol{\varphi},\mathcal{D}^I\Big]-C_{i,J}^{Bay}\right)\right)^2\middle|\mathcal{D}^I\right] = \underbrace{\mathsf{Var}\left[\sum_{i=0}^{I}\mathsf{E}\Big[C_{i,J}\Big|\boldsymbol{\varphi},\mathcal{D}^I\Big]\middle|\mathcal{D}^I\right]}_{\mathsf{E}\left[\mathsf{E}\left[C_{i,J}\middle|\boldsymbol{\varphi},\mathcal{D}^I\right]\middle|\mathcal{D}^I\right]=C_{i,J}^{Bay}}$$

$$= \sum_{i_1,i_2=0}^{I}\mathsf{Cov}\left[C_{i_1,I-i_1}\prod_{k=I-i_1}^{J-1}\varphi_k,\,C_{i_2,I-i_2}\prod_{k=I-i_2}^{J-1}\varphi_k\middle|\mathcal{D}^I\right]$$

$$= \sum_{i_1,i_2=0}^{I}C_{i_1,I-i_1}C_{i_2,I-i_2}\mathsf{Cov}\left[\prod_{k=I-i_1}^{J-1}\varphi_k,\,\prod_{k=I-i_2}^{J-1}\varphi_k\middle|\mathcal{D}^I\right]$$

$$= \sum_{i_1,i_2=0}^{I}C_{i_1,I-i_1}C_{i_2,I-i_2}\left(\mathsf{E}\left[\prod_{k=I-i_1}^{J-1}\varphi_k\prod_{k=I-i_2}^{J-1}\varphi_k\middle|\mathcal{D}^I\right]-\mathsf{E}\left[\prod_{k=I-i_1}^{J-1}\varphi_k\middle|\mathcal{D}^I\right]\mathsf{E}\left[\prod_{k=I-i_2}^{J-1}\varphi_k\middle|\mathcal{D}^I\right]\right)$$

$$= \sum_{i_1,i_2=0}^{I}C_{i_1,I-i_1}C_{i_2,I-i_2}\underbrace{\prod_{k=I-(i_1\vee i_2)}^{I-(i_1\wedge i_2)-1}\mathsf{E}\Big[\varphi_k\Big|\mathcal{D}^I\Big]\left(\prod_{k=I-(i_1\wedge i_2)}^{J-1}\mathsf{E}\Big[\varphi_k^2\Big|\mathcal{D}^I\Big]-\prod_{k=I-(i_1\wedge i_2)}^{J-1}\mathsf{E}\Big[\varphi_k\Big|\mathcal{D}^I\Big]^2\right)}_{\text{iv)}^{Bay}}.$$

Note: Although accident periods are independent given $\mathcal{D}^I$ **and** $\boldsymbol{\varphi}$ they are usually not independent given $\mathcal{D}^I$.

Problem 8.5

We still have to estimate

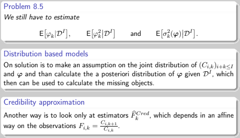
$$\mathsf{E}\left[\varphi_k \middle| \mathcal{D}^I\right], \qquad \mathsf{E}\left[\varphi_k^2 \middle| \mathcal{D}^I\right] \qquad \text{and} \qquad \mathsf{E}\left[\sigma_k^2(\boldsymbol{\varphi}) \middle| \mathcal{D}^I\right].$$

Distribution based models

On solution is to make an assumption on the joint distribution of $(C_{i,k})_{i+k \le I}$ and $\boldsymbol{\varphi}$ and than calculate the a posteriori distribution of $\boldsymbol{\varphi}$ given $\mathcal{D}^I$, which then can be used to calculate the missing objects.

Credibility approximation

Another way is to look only at estimators $\widehat{F}_k^{Cred}$, 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

$$\mathsf{E}[\varphi_k|\mathcal{D}^I], \quad \mathsf{E}[\varphi_k^2|\mathcal{D}^I] \quad \text{and} \quad \mathsf{E}[\sigma_k^2(\boldsymbol{\varphi})|\mathcal{D}^I].$$

Distribution based models

On solution is to make an assumption on the joint distribution of $(C_{i,k})_{i+k\leq I}$ and $\boldsymbol{\varphi}$ and than calculate the a posteriori distribution of $\boldsymbol{\varphi}$ given $\mathcal{D}^I$, which then can be used to calculate the missing objects.

Credibility approximation

Another way is to look only at estimators $\widehat{F}_k^{Cred}$, which depends in an affine way on the observations $F_{i,k} = \frac{C_{i,k+1}}{C_{i,k}}$.

- Note, we know $\mathsf{E}[\varphi_k] = f_k$, but usually $\mathsf{E}[\varphi_k|\mathcal{D}^I] \neq f_k$.
- Even if we have a good model for the joint distribution of $(C_{i,k})_{i+k\leq I}$ and $\boldsymbol{\varphi}$, 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 onto the affine subspace of $\mathcal{D}^I$ generated by the link ratios $F_{i,k}$ instead of the projection onto $\mathcal{D}^I$ itself.

Definition 8.6 (Credibility estimators of the development factors)

$$\widehat{F}_k^{Cred} := \underset{\widehat{\varphi}=a_k+\sum_{i=0}^{I-k-1} a_{i,k}\frac{C_{i,k+1}}{C_{i,k}}}{\operatorname{argmin}} \mathsf{E}\left[\left(\varphi_k - \widehat{\varphi}\right)^2 \middle| \mathcal{D}^I\right].$$

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*

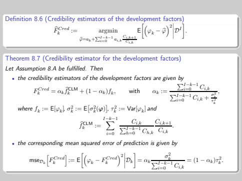
$$F_k^{Cred} = \alpha_k \widehat{f}_k^{\mathsf{CLM}} + (1-\alpha_k) f_k, \qquad \text{with} \qquad \alpha_k := \frac{\sum_{i=0}^{I-k-1} C_{i,k}}{\sum_{i=0}^{I-k-1} C_{i,k} + \frac{\sigma_k^2}{\tau_k^2}},$$

where $f_k := \mathsf{E}[\varphi_k]$, $\sigma_k^2 := \mathsf{E}\left[\sigma_k^2(\boldsymbol{\varphi})\right]$, $\tau_k^2 := \mathsf{Var}[\varphi_k]$ *and*

$$\widehat{f}_k^{\mathsf{CLM}} := \sum_{i=0}^{I-k-1} \frac{C_{i,k}}{\sum_{h=0}^{I-k-1} C_{h,k}} \frac{C_{i,k+1}}{C_{i,k}}.$$

- *the corresponding mean squared error of prediction is given by*

$$\mathsf{mse}_{\mathcal{D}_k}\left[F_k^{Cred}\right] := \mathsf{E}\left[\left(\varphi_k - F_k^{Cred}\right)^2 \middle| \mathcal{D}_k\right] = \alpha_k \frac{\sigma_k^2}{\sum_{i=0}^{I-k-1} C_{i,k}} = (1-\alpha_k)\tau_k^2.$$



Definition 8.6 (Credibility estimators of the development factors)

$$\widehat{F}_k^{Cred} := \underset{\widehat{\varphi}=a_k+\sum_{i=0}^{I-k-1} a_{i,k}\frac{C_{i,k+1}}{C_{i,k}}}{\operatorname{argmin}} \mathsf{E}\left[\left(\varphi_k - \widehat{\varphi}\right)^2 \middle| \mathcal{D}^I\right].$$

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*

$$F_k^{Cred} = \alpha_k \widehat{f}_k^{\mathsf{CLM}} + (1-\alpha_k) f_k, \qquad \text{with} \qquad \alpha_k := \frac{\sum_{i=0}^{I-k-1} C_{i,k}}{\sum_{i=0}^{I-k-1} C_{i,k} + \frac{\sigma_k^2}{\tau_k^2}},$$

  where $f_k := \mathsf{E}[\varphi_k]$, $\sigma_k^2 := \mathsf{E}[\sigma_k^2(\varphi)]$, $\tau_k^2 := \mathsf{Var}[\varphi_k]$ *and*

$$\widehat{f}_k^{\mathsf{CLM}} := \sum_{i=0}^{I-k-1} \frac{C_{i,k}}{\sum_{h=0}^{I-k-1} C_{h,k}} \frac{C_{i,k+1}}{C_{i,k}}.$$

- *the corresponding mean squared error of prediction is given by*

$$\mathsf{mse}_{\mathcal{D}_k}\left[F_k^{Cred}\right] := \mathsf{E}\left[\left(\varphi_k - F_k^{Cred}\right)^2 \middle| \mathcal{D}_k\right] = \alpha_k \frac{\sigma_k^2}{\sum_{i=0}^{I-k-1} C_{i,k}} = (1-\alpha_k)\tau_k^2.$$

- Conditionally given $\mathcal{D}_k$, the random variables $F_{i,k} = \frac{C_{i,k+1}}{C_{i,k}}$, $i = 0, \dots, I-k-1$, fulfil the assumptions of the Bühlmann and Straub model (see [29, 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 [29, 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 $f_k = \mathsf{E}[\varphi_k]$ we don't mark it with a hat like other estimatior.

Estimator 8.8 (Credibility estimator of the future development)

$$\widehat{C}^{Cred}_{i,k} := \widehat{F}^{Cred}_{k-1} \cdot \ldots \cdot \widehat{F}^{Cred}_{I-i} C_{i,I-i}, \qquad \text{for } i+k > I.$$

Estimation of the structural parameters f_k, σ_k^2, and τ_k^2, see [29, 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

$$\widehat{F}^{m,Cred}_k := \widehat{\alpha}^m_k \widehat{f}^{m,\mathsf{CLM}}_k + (1-\widehat{\alpha}^m_k)\widehat{f}_k \quad \text{and} \quad \widehat{C}^{m,Cred}_{i,k} := \widehat{F}^{m,Cred}_{k-1} \cdot \ldots \cdot \widehat{F}^{m,Cred}_{I-i} C^m_{i,I-i} \quad \text{with}$$

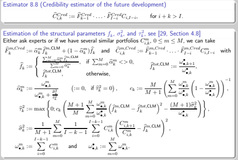
$$\widehat{f}_k := \begin{cases} \frac{\sum_{m=0}^M \widehat{\alpha}^m_k \widehat{f}^{m,\mathsf{CLM}}_k}{\sum_{m=0}^M \widehat{\alpha}^m_k}, & \text{if } \sum_{m=0}^M \widehat{\alpha}^m_k <> 0, \\ \widehat{f}^{tot,\mathsf{CLM}}_k, & \text{otherwise}, \end{cases} \qquad \widehat{f}^{tot,\mathsf{CLM}}_k := \frac{\omega^\bullet_{\bullet,k+1}}{\omega^\bullet_{\bullet,k}},$$

$$\widehat{\alpha}^m_k := \frac{\omega^m_{\bullet,k}}{\omega^m_{\bullet,k} + \frac{\widehat{\sigma}^2_k}{\widehat{\tau}^2_k}} \qquad (:= 0, \ \text{if } \widehat{\tau}^2_k = 0), \qquad c_k := \frac{M}{M+1}\left(\sum_{m=0}^M \frac{\omega^m_{\bullet,k}}{\omega^\bullet_{\bullet,k}}\left(1 - \frac{\omega^m_{\bullet,k}}{\omega^\bullet_{\bullet,k}}\right)\right)^{-1},$$

$$\widehat{\tau}^2_k := \max\left\{0; c_k\left(\frac{M+1}{M}\sum_{m=0}^M \frac{\omega^m_{\bullet,k}}{\omega^\bullet_{\bullet,k}}\left(\widehat{f}^{m,\mathsf{CLM}}_k - \widehat{f}^{tot,\mathsf{CLM}}_k\right)^2 - \frac{(M+1)\widehat{\sigma}^2_k}{\omega^\bullet_{\bullet,k}}\right)\right\},$$

$$\widehat{\sigma}^2_k := \frac{1}{M+1}\sum_{m=0}^M \frac{1}{I-k-1}\sum_{i=0}^{I-k-1} C^m_{i,k}\left(\frac{C^m_{i,k+1}}{C^m_{i,k}} - \widehat{f}^{m,\mathsf{CLM}}_k\right)^2,$$

$$\omega^m_{\bullet,k} := \sum_{i=0}^{I-k-1} C^m_{i,k} \quad \text{and} \quad \omega^\bullet_{\bullet,k} := \sum_{m=0}^M \omega^m_{\bullet,k}.$$



Estimator 8.8 (Credibility estimator of the future development)

$$\widehat{C}^{Cred}_{i,k} := \widehat{F}^{Cred}_{k-1} \cdot \ldots \cdot \widehat{F}^{Cred}_{I-i} C_{i,I-i}, \qquad \text{for } i+k>I.$$

Estimation of the structural parameters f_k, σ_k^2, and τ_k^2, see [29, Section 4.8]

Either ask experts or if we have several similar portfolios $C^m_{i,k}$, $0 \le m \le M$, we can take

$$\widehat{F}^{m,Cred}_k := \widehat{\alpha}^m_k \widehat{f}^{m,\mathsf{CLM}}_k + (1-\widehat{\alpha}^m_k)\widehat{f}_k \quad \text{and} \quad \widehat{C}^{m,Cred}_{i,k} := \widehat{F}^{m,Cred}_{k-1} \cdot \ldots \cdot \widehat{F}^{m,Cred}_{I-i} C^m_{i,I-i} \quad \text{with}$$

$$\widehat{f}_k := \begin{cases} \frac{\sum_{m=0}^M \widehat{\alpha}^m_k \widehat{f}^{m,\mathsf{CLM}}_k}{\sum_{m=0}^M \widehat{\alpha}^m_k}, & \text{if } \sum_{m=0}^M \widehat{\alpha}^m_k <> 0, \\ \widehat{f}^{tot,\mathsf{CLM}}_k, & \text{otherwise}, \end{cases} \qquad \widehat{f}^{tot,\mathsf{CLM}}_k := \frac{\omega^{\bullet}_{\bullet,k+1}}{\omega^{\bullet}_{\bullet,k}},$$

$$\widehat{\alpha}^m_k := \frac{\omega^m_{\bullet,k}}{\omega^m_{\bullet,k} + \frac{\widehat{\sigma}^2_k}{\widehat{\tau}^2_k}} \qquad (:= 0, \ \text{if } \widehat{\tau}^2_k = 0), \qquad c_k := \frac{M}{M+1}\left(\sum_{m=0}^M \frac{\omega^m_{\bullet,k}}{\omega^{\bullet}_{\bullet,k}}\left(1 - \frac{\omega^m_{\bullet,k}}{\omega^{\bullet}_{\bullet,k}}\right)\right)^{-1},$$

$$\widehat{\tau}^2_k := \max\left\{0; c_k\left(\frac{M+1}{M}\sum_{m=0}^M \frac{\omega^m_{\bullet,k}}{\omega^{\bullet}_{\bullet,k}}\left(\widehat{f}^{m,\mathsf{CLM}}_k - \widehat{f}^{tot,\mathsf{CLM}}_k\right)^2 - \frac{(M+1)\widehat{\sigma}^2_k}{\omega^{\bullet}_{\bullet,k}}\right)\right\},$$

$$\widehat{\sigma}^2_k := \frac{1}{M+1}\sum_{m=0}^M \frac{1}{I-k-1}\sum_{i=0}^{I-k-1} C^m_{i,k}\left(\frac{C^m_{i,k+1}}{C^m_{i,k}} - \widehat{f}^{m,\mathsf{CLM}}_k\right)^2,$$

$$\omega^m_{\bullet,k} := \sum_{i=0}^{I-k-1} C^m_{i,k} \quad \text{and} \quad \omega^{\bullet}_{\bullet,k} := \sum_{m=0}^M \omega^m_{\bullet,k}.$$

- In the case of non-informative priors, i.e. $\tau_k^2 \to \infty$, the estimators of the future development are the same as for the standard Chain-Ladder method introduced in Section 2.
- $\widehat{f}_k^{tot,\mathsf{CLM}} = \frac{\omega^{\bullet}_{\bullet,k+1}}{\omega^{\bullet}_{\bullet,k}}$ are the standard estimates of the development factors of the combined portfolio $\sum_{m=0}^M C^m_{i,k}$.
- The factors c_k are normalizing factors that makes the estimators $\widehat{\tau}_k$ unbiased (conditioned $\widehat{\tau}_k > 0$).

Estimator 8.9 (of the ultimate uncertainty)

Let Assumption 8.A be fulfilled. Then the ultimate uncertainty is given by

$$\mathrm{mse}_{\mathcal{D}^I}\left[\sum_{i=0}^{I}\widehat{C}_{i,J}^{Cred}\right] = \underbrace{\sum_{i=0}^{I}\sum_{k=I-i}^{J-1}\prod_{j=k+1}^{J-1}\mathrm{E}\left[\varphi_j^2\middle|\mathcal{D}^I\right]\mathrm{E}\left[\sigma_k^2(\boldsymbol{\varphi})\middle|\mathcal{D}^I\right]\prod_{j=I-i}^{k-1}\mathrm{E}\left[\varphi_j\middle|\mathcal{D}^I\right]C_{i,I-i}}_{\text{random error}}$$

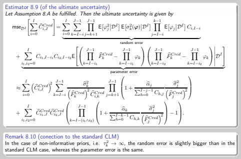
$$+\underbrace{\sum_{i_1,i_2=0}^{I}C_{i_1,I-i_1}C_{i_2,I-i_2}\mathrm{E}\left[\left(\prod_{k=I-i_1}^{J-1}\widehat{F}_k^{Cred}-\prod_{k=I-i_1}^{J-1}\varphi_k\right)\left(\prod_{k=I-i_2}^{J-1}\widehat{F}_k^{Cred}-\prod_{k=I-i_2}^{J-1}\varphi_k\right)\middle|\mathcal{D}^I\right]}_{\text{parameter error}}$$

$$\approx\sum_{i=0}^{I}\left(\widehat{C}_{i,J}^{Cred}\right)^2\sum_{k=I-i}^{J-1}\frac{\widehat{\sigma}_k^2}{\left(\widehat{F}_k^{Cred}\right)^2\widehat{C}_{i,k}^{Cred}}\prod_{j=k+1}^{J-1}\left(1+\frac{\widehat{\alpha}_j}{\sum_{h=0}^{I-j-1}C_{h,j}}\frac{\widehat{\sigma}_j^2}{\left(\widehat{F}_j^{Cred}\right)^2}\right)$$

$$+\sum_{i_1,i_2=0}^{I}\widehat{C}_{i_1,J}^{Cred}\widehat{C}_{i_2,J}^{Cred}\left(\prod_{k=I-(i_1\wedge i_2)}^{J-1}\left(1+\frac{\widehat{\alpha}_k}{\sum_{h=0}^{I-k-1}C_{h,k}}\frac{\widehat{\sigma}_j^2}{\left(\widehat{F}_j^{Cred}\right)^2}\right)-1\right).$$

Remark 8.10 (conection 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.



Estimator 8.9 (of the ultimate uncertainty)

Let Assumption 8.A be fulfilled. Then the ultimate uncertainty is given by

$$\mathsf{mse}_{\mathcal{D}^I}\left[\sum_{i=0}^{I}\widehat{C}_{i,J}^{Cred}\right]=\underbrace{\sum_{i=0}^{I}\sum_{k=I-i}^{J-1}\prod_{j=k+1}^{J-1}\mathsf{E}\left[\varphi_j^2\middle|\mathcal{D}^I\right]\mathsf{E}\left[\sigma_k^2(\boldsymbol{\varphi})\middle|\mathcal{D}^I\right]\prod_{j=I-i}^{k-1}\mathsf{E}\left[\varphi_j\middle|\mathcal{D}^I\right]C_{i,I-i}}_{\text{random error}}$$

$$+\underbrace{\sum_{i_1,i_2=0}^{I}C_{i_1,I-i_1}C_{i_2,I-i_2}\mathsf{E}\left[\left(\prod_{k=I-i_1}^{J-1}\widehat{F}_k^{Cred}-\prod_{k=I-i_1}^{J-1}\varphi_k\right)\left(\prod_{k=I-i_2}^{J-1}\widehat{F}_k^{Cred}-\prod_{k=I-i_2}^{J-1}\varphi_k\right)\middle|\mathcal{D}^I\right]}_{\text{parameter error}}$$

$$\approx\sum_{i=0}^{I}\left(\widehat{C}_{i,J}^{Cred}\right)^2\sum_{k=I-i}^{J-1}\frac{\widehat{\sigma}_k^2}{\left(\widehat{F}_k^{Cred}\right)^2\widehat{C}_{i,k}^{Cred}}\prod_{j=k+1}^{J-1}\left(1+\frac{\widehat{\alpha}_j}{\sum_{h=0}^{I-j-1}C_{h,j}}\frac{\widehat{\sigma}_j^2}{\left(\widehat{F}_j^{Cred}\right)^2}\right)$$

$$+\sum_{i_1,i_2=0}^{I}\widehat{C}_{i_1,J}^{Cred}\widehat{C}_{i_2,J}^{Cred}\left(\prod_{k=I-(i_1\wedge i_2)}^{J-1}\left(1+\frac{\widehat{\alpha}_k}{\sum_{h=0}^{I-k-1}C_{h,k}}\frac{\widehat{\sigma}_j^2}{\left(\widehat{F}_j^{Cred}\right)^2}\right)-1\right).$$

Remark 8.10 (conection 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

$$\mathsf{mse}_{\mathcal{D}^I}\left[\sum_{i=0}^{I}C_{i,J}^{Cred}\right]=\underbrace{\sum_{i=0}^{I}\mathsf{E}\left[\mathsf{Var}\left[C_{i,J}|\boldsymbol{\varphi},\mathcal{D}^I\right]\middle|\mathcal{D}^I\right]}_{\textbf{random error}}+\underbrace{\mathsf{E}\left[\left(\sum_{i=0}^{I}\left(\mathsf{E}\left[C_{i,J}|\boldsymbol{\varphi},\mathcal{D}^I\right]-C_{i,J}^{Cred}\right)\right)^2\middle|\mathcal{D}^I\right]}_{\textbf{parameter error}}.$$

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 $\mathsf{E}\left[\sigma_k^2(\boldsymbol{\varphi})\middle|\mathcal{D}^I\right]\approx\sigma_k^2$ and $\mathsf{E}\left[\varphi_k|\mathcal{D}^I\right]\approx\widehat{F}_k^{Cred}$.

Moreover, we estimate

$$\mathsf{E}\left[\varphi_j^2\middle|\mathcal{D}^I\right]=\mathsf{E}\left[\left(\varphi_j-\mathsf{E}\left[\varphi_j\middle|\mathcal{D}^I\right]\right)^2\middle|\mathcal{D}^I\right]+\left(\mathsf{E}\left[\varphi_k|\mathcal{D}^I\right]\right)^2\approx\underbrace{\widehat{\alpha}_j\frac{\widehat{\sigma}_j^2}{\sum_{i=0}^{I-j-1}C_{i,j}}}_{\textbf{Theorem 8.7}}+\left(\widehat{F}_j^{Cred}\right)^2.$$

Finally, we compute

$$\mathsf{E}\left[\left(\prod_{k=I-i_1}^{J-1}F_k^{Cred}-\prod_{k=I-i_1}^{J-1}\varphi_k\right)\left(\prod_{k=I-i_2}^{J-1}F_k^{Cred}-\prod_{k=I-i_2}^{J-1}\varphi_k\right)\middle|\mathcal{D}^I\right]$$

$$\approx\mathsf{E}\left[\left(\prod_{k=I-i_1}^{J-1}F_k^{Bay}-\prod_{k=I-i_1}^{J-1}\varphi_k\right)\left(\prod_{k=I-i_2}^{J-1}F_k^{Bay}-\prod_{k=I-i_2}^{J-1}\varphi_k\right)\middle|\mathcal{D}^I\right]=\mathsf{Cov}\left[\prod_{k=I-i_1}^{J-1}\varphi_k,\prod_{k=I-i_2}^{J-1}\varphi_k\middle|\mathcal{D}^I\right]$$

$$=\prod_{k=I-(i_1\vee i_2)}^{I-(i_1\wedge i_2)-1}\mathsf{E}\left[\varphi_k|\mathcal{D}^I\right]\left(\prod_{k=I-(i_1\wedge i_2)}^{J-1}\mathsf{E}\left[\varphi_k^2\middle|\mathcal{D}^I\right]-\prod_{k=I-(i_1\wedge i_2)}^{J-1}\mathsf{E}\left[\varphi_k|\mathcal{D}^I\right]^2\right)$$

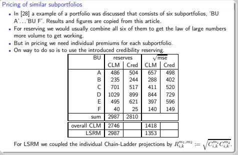
and replace all unknown parameters by they estimates and take the factors $\left(\widehat{C}_{i,J}^{Cred}\right)^2$ and $\widehat{C}_{i_1,J}^{Cred}\widehat{C}_{i_2,J}^{Cred}$ out.

Pricing of similar subportfolios

- In [28] a 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.
- On way to do so is to use the introduced credibility reserving.

BU	reserves		$\sqrt{\text{mse}}$	
	CLM	Cred	CLM	Cred
A	486	504	657	498
B	235	244	288	402
C	701	517	411	520
D	1029	899	844	729
E	495	621	397	596
F	40	25	140	149
sum	2987	2810		
overall CLM	2746		1418	
LSRM	2987		1353	

For LSRM we coupled the individual Chain-Ladder projections by $R^{m_1,m_2}_{i,k} := \sqrt{C^{m_1}_{i,k} C^{m_2}_{i,k}}$.



Pricing of similar subportfolios

- In [28] a 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.
- On way to do so is to use the introduced credibility reserving.

BU	reserves		$\sqrt{\text{mse}}$	
	CLM	Cred	CLM	Cred
A	486	504	657	498
B	235	244	288	402
C	701	517	411	520
D	1029	899	844	729
E	495	621	397	596
F	40	25	140	149
sum	2987	2810		
overall CLM	2746		1418	
LSRM	2987		1353	

For LSRM we coupled the individual Chain-Ladder projections by $R_{i,k}^{m_1,m_2} := \sqrt{C_{i,k}^{m_1} C_{i,k}^{m_2}}$.

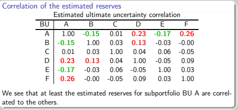
- The total reserves differ only by 6 %, but per subportfolio the differences are much larger (up to 46 %).
- The CLM reserves for the combined portfolio are even smaller.
- The mse of the combined portfolios is about 25 % larger than the sum of the individual ones. This may be a hint that the estimated 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 on slide 147, the CLM and the LSRM estimates are (re)calculated. The presented figures for CLM, which are taken from the original article [28], differ slightly from the recalculated once, because of rounding effects.

Correlation of the estimated reserves

Estimated ultimate uncertainty correlation

BU	A	B	C	D	E	F
A	1.00	**-0.15**	0.01	**0.23**	**-0.17**	**0.26**
B	**-0.15**	1.00	0.03	**0.13**	-0.03	-0.00
C	0.01	0.03	1.00	0.04	0.06	-0.05
D	**0.23**	**0.13**	0.04	1.00	-0.05	0.09
E	**-0.17**	-0.03	0.06	-0.05	1.00	0.03
F	**0.26**	-0.00	-0.05	0.09	0.03	1.00

We see that at least the estimated reserves for subportfolio BU A are correlated to the others.



Correlation of the estimated reserves

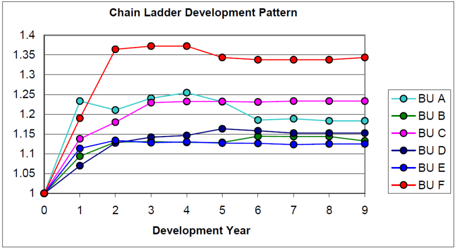
Estimated ultimate uncertainty correlation

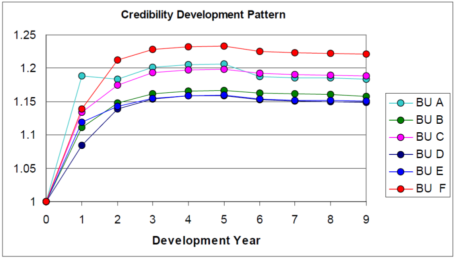
BU	A	B	C	D	E	F
A	1.00	**-0.15**	0.01	**0.23**	**-0.17**	**0.26**
B	**-0.15**	1.00	0.03	**0.13**	-0.03	-0.00
C	0.01	0.03	1.00	0.04	0.06	-0.05
D	**0.23**	**0.13**	0.04	1.00	-0.05	0.09
E	**-0.17**	-0.03	0.06	-0.05	1.00	0.03
F	**0.26**	-0.00	-0.05	0.09	0.03	1.00

We see that at least the estimated reserves for subportfolio BU A are correlated to the others.

Comparison of the estimated development pattern (1/2)

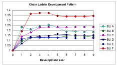
The individual CLM development pattern are smoothed by the credibility approach:





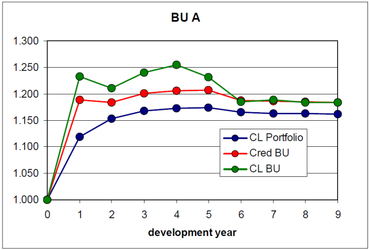
Comparison of the estimated development pattern (1/2)

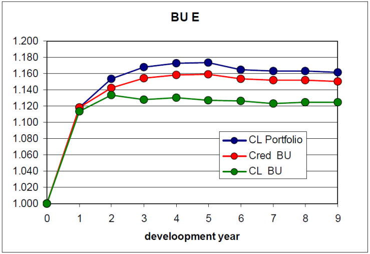
The individual CLM development pattern are smoothed by the credibility approach:

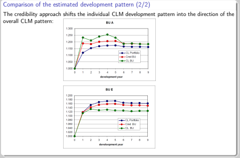


Comparison of the estimated development pattern (2/2)

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.
Astin Bull., 38(2):565–600, 2008.

[29] Bühlmann, H. and Gisler, A.
A Course in Credibility Theory and its Applications.
Universitext, Springer Verlag, 2005.

Literature

[28] Aloise Gisler and Mario V. Wüthrich.
Credibility for the Chain Ladder Reserving Method.
Astin Bull., 38(2):565–600, 2008.

[29] Bühlmann, H. and Gisler, A.
A Course in Credibility Theory and its Applications.
Universitext, Springer Verlag, 2005.

Stochastic Reserving

Lecture 13

Separation of Large and Small Claims

René Dahms

ETH Zurich, Spring 2021

26 April 2021

(Last update: 26 April 2021)

Stochastic Reserving

Lecture 13

Separation of Large and Small Claims

René Dahms

ETH Zurich, Spring 2021

26 April 2021

(Last update: 26 April 2021)

9 Separation of small and large claims

Lecture 13: Table of contents

Increments of incurred losses with individual Chain-Ladder development factors

included losses of all claims

i\k	0	1	2	3	4
0	296 $\xrightarrow{7.5\%}$	22 $\xrightarrow{-3.0\%}$	-10 $\xrightarrow{-4.4\%}$	-14 $\xrightarrow{-6.1\%}$	-18
1	285 $\xrightarrow{1.6\%}$	5 $\xrightarrow{-5.9\%}$	-17 $\xrightarrow{-7.2\%}$	-20	
2	259 $\xrightarrow{4.4\%}$	11 $\xrightarrow{-8.0\%}$	-22		
3	277 $\xrightarrow{5.9\%}$	16			
4	268				

(ever) large claims excluded

i\k	0	1	2	3	4
0	269 $\xrightarrow{4.4\%}$	12 $\xrightarrow{-3.8\%}$	-11 $\xrightarrow{-5.7\%}$	-15 $\xrightarrow{-7.9\%}$	-20
1	274 $\xrightarrow{1.2\%}$	3 $\xrightarrow{-6.4\%}$	-18 $\xrightarrow{-7.5\%}$	-19	
2	250 $\xrightarrow{4.5\%}$	11 $\xrightarrow{-8.0\%}$	-21		
3	254 $\xrightarrow{3.3\%}$	8			
4	263				

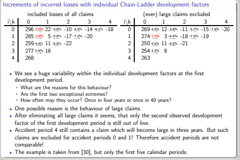
- We see a huge variability within the individual development factors at the first development period.
  - What are the reasons for this behaviour?
  - Are the first two exceptional extremes?
  - How often may they occur? Once in four years or once in 40 years?
- One possible reason is the behaviour 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 [30], but only the first five calendar periods.

2021-04-26

Stochastic Reserving

Separation of small and large claims

What is the problem with large claims



Increments of incurred losses with individual Chain-Ladder development factors

included losses of all claims

i\k	0	1	2	3	4
0	296 $\xrightarrow{7.5\%}$	22 $\xrightarrow{-3.0\%}$	-10 $\xrightarrow{-4.4\%}$	-14 $\xrightarrow{-6.1\%}$	-18
1	285 $\xrightarrow{1.6\%}$	5 $\xrightarrow{-5.9\%}$	-17 $\xrightarrow{-7.2\%}$	-20	
2	259 $\xrightarrow{4.4\%}$	11 $\xrightarrow{-8.0\%}$	-22		
3	277 $\xrightarrow{5.9\%}$	16			
4	268				

(ever) large claims excluded

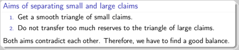
i\k	0	1	2	3	4
0	269 $\xrightarrow{4.4\%}$	12 $\xrightarrow{-3.8\%}$	-11 $\xrightarrow{-5.7\%}$	-15 $\xrightarrow{-7.9\%}$	-20
1	274 $\xrightarrow{1.2\%}$	3 $\xrightarrow{-6.4\%}$	-18 $\xrightarrow{-7.5\%}$	-19	
2	250 $\xrightarrow{4.5\%}$	11 $\xrightarrow{-8.0\%}$	-21		
3	254 $\xrightarrow{3.3\%}$	8			
4	263				

- 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 [30], 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

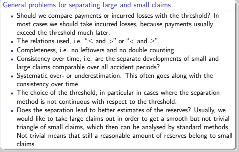
1. Get a smooth triangle of small claims.
2. Do not transfer too much reserves to the triangle of large claims.

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. "$\leq$ 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.



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. "$\leq$ 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.

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)
- ...

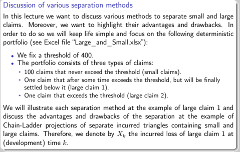
small claims:
- small losses
- normal claims (or losses)
- attritional claims (or losses)
- ...

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.



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

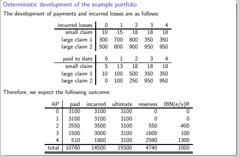
The development of payments and incurred losses are as follows:

incurred losses	0	1	2	3	4
small claim	10	15	18	18	18
large claim 1	300	700	800	350	350
large claim 2	500	800	900	950	950

paid to date	0	1	2	3	4
small claim	5	13	18	18	18
large claim 1	10	100	500	350	350
large claim 2	0	100	250	950	950

Therefore, we expect the following outcome:

AP	paid	incurred	ultimate	reserves	IBN(e/y)R
0	3100	3100	3100	0	0
1	3100	3100	3100	0	0
2	2550	3500	3100	550	-400
3	1500	3000	3100	1600	100
4	510	1800	3100	2590	1300
total	10760	14500	15500	4740	1000



Deterministic development of the example portfolio

The development of payments and incurred losses are as follows:

incurred losses	0	1	2	3	4
small claim	10	15	18	18	18
large claim 1	300	700	800	350	350
large claim 2	500	800	900	950	950

paid to date	0	1	2	3	4
small claim	5	13	18	18	18
large claim 1	10	100	500	350	350
large claim 2	0	100	250	950	950

Therefore, we expect the following outcome:

AP	paid	incurred	ultimate	reserves	IBN(e/y)R
0	3100	3100	3100	0	0
1	3100	3100	3100	0	0
2	2550	3500	3100	550	-400
3	1500	3000	3100	1600	100
4	510	1800	3100	2590	1300
total	10760	14500	15500	4740	1000

cumulative incurred losses of 100 small claims

	0	1	2	3	4
0	1000	1500	1800	1800	1800
1	1000	1500	1800	1800	
2	1000	1500	1800		
3	1000	1500			
4	1000				

cumulative incurred losses of large claim 1

	0	1	2	3	4
0	300	700	800	350	350
1	300	700	800	350	
2	300	700	800		
3	300	700			
4	300				

cumulative incurred losses of large claim 2

	0	1	2	3	4
0	500	800	900	950	950
1	500	800	900	950	
2	500	800	900		
3	500	800			
4	500				

cumulative incurred losses of all claims

	0	1	2	3	4
0	1800	3000	3500	3100	3100
1	1800	3000	3500	3100	
2	1800	3000	3500		
3	1800	3000			
4	1800				

cumulative payments for 100 small claims

	0	1	2	3	4
0	500	1300	1800	1800	1800
1	500	1300	1800	1800	
2	500	1300	1800		
3	500	1300			
4	500				

cumulative payments for large claim 1

	0	1	2	3	4
0	10	100	500	350	350
1	10	100	500	350	
2	10	100	500		
3	10	100			
4	10				

cumulative payments for large claim 2

	0	1	2	3	4
0	0	100	250	950	950
1	0	100	250	950	
2	0	100	250		
3	0	100			
4	0				

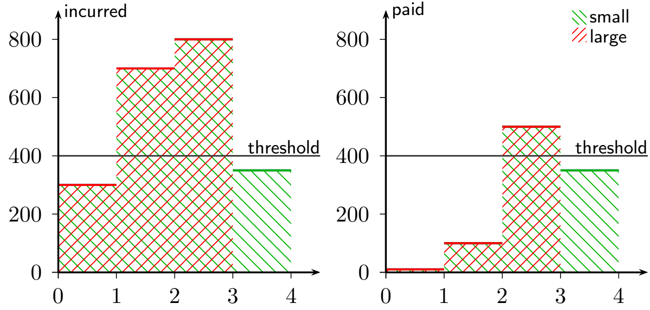
cumulative payments for all claims

	0	1	2	3	4
0	510	1500	2550	3100	3100
1	510	1500	2550	3100	
2	510	1500	2550		
3	510	1500			
4	510				

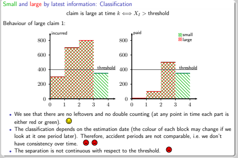
Small and large by latest information: Classification

$$\text{claim is large at time } k \Longleftrightarrow X_I > \text{threshold}$$

Behaviour of large claim 1:



- 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. ☹



Small and large by latest information: Classification

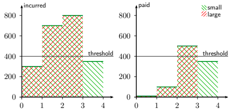
claim is large at time $k \iff X_I >$ threshold

Behaviour of large claim 1:

- 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.

Small and large by latest information: Projection



small	0	1	2	3	4
0	1300	900	400	-450	0
1	1300	900	400	-450	0
2	1000	500	300	-312	0
3	1000	500	280	-308	0
4	1300	791	390	-429	0
$\widehat{f}_k$	0.61	0.19	-0.17	0.00	

large	0	1	2	3	4
0	500	300	100	50	0
1	500	300	100	50	0
2	800	700	200	94	0
3	800	700	194	94	0
4	500	385	114	55	0
$\widehat{f}_k$	0.77	0.13	0.06	0.00	

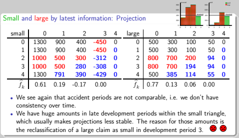
- We see again that accident periods are not comparable, i.e. we don't have consistency over time.
- We have huge amounts in late development periods within the small triangle, which usually makes projections less stable. The reason for those amounts is the reclassification of a large claim as small in development period 3. ☹ ☹

2021-04-26

Stochastic Reserving

└─Separation of small and large claims

└─How to separate small from large claims

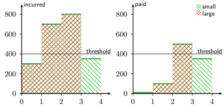


Small and large by latest information: Projection

small	0	1	2	3	4
0	1300	900	400	**-450**	0
1	1300	900	400	**-450**	**0**
2	**1000**	**500**	**300**	**-312**	**0**
3	**1000**	**500**	**280**	**-308**	**0**
4	1300	**791**	**390**	**-429**	**0**
$\widehat{f}_k$	0.61	0.19	-0.17	0.00	

large	0	1	2	3	4
0	500	300	100	50	0
1	500	300	100	50	**0**
2	**800**	**700**	**200**	**94**	**0**
3	**800**	**700**	**194**	**94**	**0**
4	500	**385**	**114**	**55**	**0**
$\widehat{f}_k$	0.77	0.13	0.06	0.00	

- 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. ☹ ☹



Small and large by latest information: Results

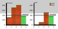
	expected results		estimated results			
AP	ultimate	reserves	ultimate	reserves	small res.	large res.
0	3100	0	3100	0	0	0
1	3100	0	3100	0	0	0
2	3100	550	**3283**	**733**	-312	1044
3	3100	1600	**3259**	**1759**	172	1588
4	3100	2590	**3106**	**2596**	1542	1054
total	15500	4740	**15848**	**5088**	1402	**3686**

- 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.

Small and large by latest information: Results



	expected results		estimated results			
AP	ultimate	reserves	ultimate	reserves	small res.	large res.
0	3100	0	3100	0	0	0
1	3100	0	3100	0	0	0
2	3100	550	3283	733	-312	1044
3	3100	1600	3259	1759	172	1588
4	3100	2590	3106	2596	1542	1054
total	15500	4740	15848	5088	1402	3686

- Under- and overestimation. ☹ ☹
- More than 75% of the reserves belong to the large triangle, which is usually less stable. ☹ ☹

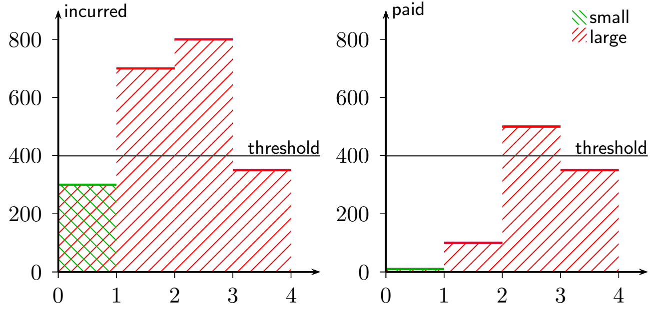
Conclusion (pros: 1 ☺ versus cons: 1 ☹ and 4 ☹ ☹)

Do not use the separation method 'small and large by latest information' for the estimation of reserves.

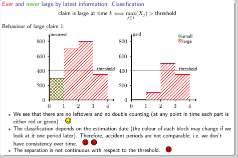
Ever and never large by latest information: Classification

$$\text{claim is large at time } k \iff \max_{j\le I}(X_j) > \text{threshold}$$

Behaviour of large claim 1:



- 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. ☹



Ever and never large by latest information: Classification

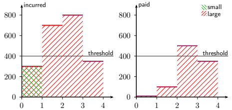
$$\text{claim is large at time } k \iff \max_{j \leq I}(X_j) > \text{threshold}$$

Behaviour of large claim 1:

- 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.

Ever and never large by latest information: Projection



small	0	1	2	3	4
0	1000	500	300	0	0
1	1000	500	300	0	**0**
2	1000	500	300	**0**	**0**
3	1000	500	**300**	**0**	**0**
4	1300	**650**	**390**	**0**	**0**
$\widehat{f}_k$	0.5	0.2	0.00	0.00	

large	0	1	2	3	4
0	800	700	200	-400	0
1	800	700	200	-400	**0**
2	800	700	200	**-400**	**0**
3	800	700	**200**	**-400**	**0**
4	500	**438**	**125**	**-250**	**0**
$\widehat{f}_k$	0.88	0.13	-0.24	0.00	

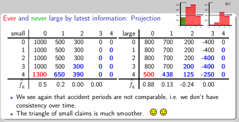
- We see again that accident periods are not comparable, i.e. we don't have consistency over time.
- The triangle of small claims is much smoother. ☺ ☺

2021-04-26

Stochastic Reserving

Separation of small and large claims

How to separate small from large claims

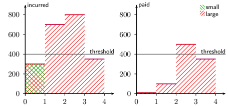


Ever and never large by latest information: Projection

small	0	1	2	3	4
0	1000	500	300	0	0
1	1000	500	300	0	**0**
2	1000	500	300	**0**	**0**
3	1000	500	**300**	**0**	**0**
4	**1300**	**650**	**390**	**0**	**0**
$\hat{f}_k$	0.5	0.2	0.00	0.00	

large	0	1	2	3	4
0	800	700	200	-400	0
1	800	700	200	-400	**0**
2	800	700	200	**-400**	**0**
3	800	700	**200**	**-400**	**0**
4	**500**	**438**	**125**	**-250**	**0**
$\hat{f}_k$	0.88	0.13	-0.24	0.00	

- 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. ☺ ☺



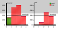
Ever and never large by latest information: Results

	expected results		estimated results			
AP	ultimate	reserves	ultimate	reserves	small res.	large res.
0	3100	0	3100	0	0	0
1	3100	0	3100	0	0	0
2	3100	550	3100	550	0	550
3	3100	1600	3100	1600	500	1100
4	3100	2590	**3153**	**2642**	1830	813
total	15500	4740	**15553**	**4792**	2330	**2463**

- 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.



Ever and never large by latest information: Results

	expected results		estimated results			
AP	ultimate	reserves	ultimate	reserves	small res.	large res.
0	3100	0	3100	0	0	0
1	3100	0	3100	0	0	0
2	3100	550	3100	550	0	550
3	3100	1600	3100	1600	500	1100
4	3100	2590	3153	2642	1830	813
total	15500	4740	15553	4792	2330	2463

- Under- and overestimation. ☹ ☹
- More than 50% of the reserves belong to the large triangle, which is usually less stable. ☹ ☹

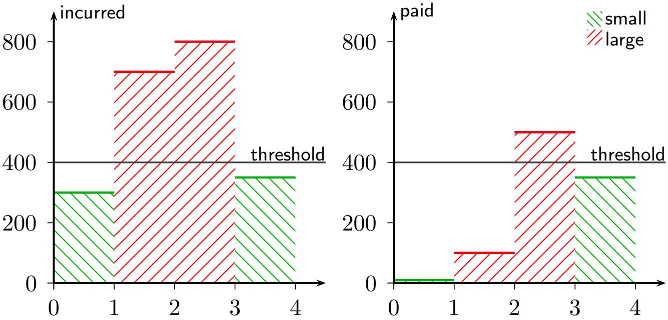
Conclusion (pros: 1 ☺ ☺ and 1 ☺ versus cons: 1 ☹ and 3 ☹ ☹)

Do not use the separation method 'ever and never large by latest information' for the estimation of reserves.

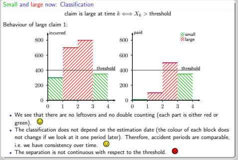
Small and large now: Classification

$$\text{claim is large at time } k \iff X_k > \text{threshold}$$

Behaviour of large claim 1:



- 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. ☹



Small and large now: Classification

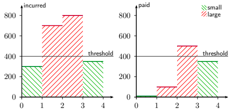
claim is large at time $k \iff X_k >$ threshold

Behaviour of large claim 1:

- 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. ☹

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



small	0	1	2	3	4
0	1300	200	300	**350**	0
1	1300	200	300	**350**	**0**
2	1300	200	300	**350**	**0**
3	1300	200	**300**	**350**	**0**
4	1300	**200**	**300**	**350**	**0**
$\widehat{f}_k$	0.15	0.20	0.19	0.00	

large	0	1	2	3	4
0	500	1000	200	**-750**	0
1	500	1000	200	**-750**	**0**
2	500	1000	200	**-750**	**0**
3	500	1000	**200**	**-750**	**0**
4	500	**1000**	**200**	**-750**	**0**
$\widehat{f}_k$	2.00	0.13	-0.44	0.00	

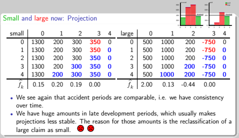
- We see again that accident periods are comparable, i.e. we have consistency over time.
- We have huge amounts in late development periods, which usually makes projections less stable. The reason for those amounts is the reclassification of a large claim as small. ☹ ☹

2021-04-26

Stochastic Reserving

Separation of small and large claims

How to separate small from large claims

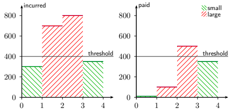


Small and large now: Projection

small	0	1	2	3	4
0	1300	200	300	**350**	0
1	1300	200	300	**350**	**0**
2	1300	200	300	**350**	**0**
3	1300	200	**300**	**350**	**0**
4	1300	**200**	**300**	**350**	**0**
$\widehat{f}_k$	0.15	0.20	0.19	0.00	

large	0	1	2	3	4
0	500	1000	200	**-750**	0
1	500	1000	200	**-750**	**0**
2	500	1000	200	**-750**	**0**
3	500	1000	**200**	**-750**	**0**
4	500	**1000**	**200**	**-750**	**0**
$\widehat{f}_k$	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. ☹ ☹



Small and large now: Results

	expected results		estimated results			
AP	ultimate	reserves	ultimate	reserves	small res.	large res.
0	3100	0	3100	0	0	0
1	3100	0	3100	0	0	0
2	3100	550	3100	550	350	200
3	3100	1600	3100	1600	850	750
4	3100	2590	3100	2590	1640	950
total	15500	4740	15500	4740	2840	**1900**

- 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 ☹ ☹)

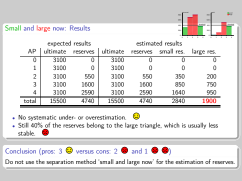
Do not use the separation method ‘small and large now’ for the estimation of reserves.

2021-04-26

Stochastic Reserving

Separation of small and large claims

How to separate small from large claims



Small and large now: Results

	expected results		estimated results			
AP	ultimate	reserves	ultimate	reserves	small res.	large res.
0	3100	0	3100	0	0	0
1	3100	0	3100	0	0	0
2	3100	550	3100	550	350	200
3	3100	1600	3100	1600	850	750
4	3100	2590	3100	2590	1640	950
total	15500	4740	15500	4740	2840	**1900**

- No systematic under- or overestimation. ☺
- Still 40% of the reserves belong to the large triangle, which is usually less stable. ☹

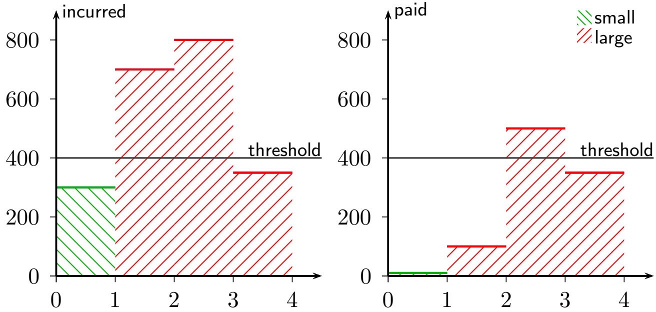
Conclusion (pros: 3 ☺ versus cons: 2 ☹ and 1 ☹ ☹)

Do not use the separation method 'small and large now' for the estimation of reserves.

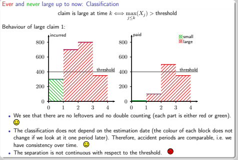
Ever and never large up to now: Classification

$$\text{claim is large at time } k \iff \max_{j\le k}(X_j) > \text{threshold}$$

Behaviour of large claim 1:



- 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. ☹



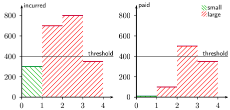
Ever and never large up to now: Classification

$$\text{claim is large at time } k \Longleftrightarrow \max_{j\le k}(X_j) > \text{threshold}$$

Behaviour of large claim 1:

- 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.

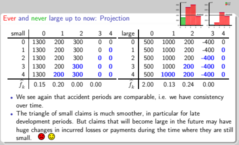


Ever and never large up to now: Projection

small	0	1	2	3	4
0	1300	200	300	0	0
1	1300	200	300	0	**0**
2	1300	200	300	**0**	**0**
3	1300	200	**300**	**0**	**0**
4	1300	**200**	**300**	**0**	**0**
$\widehat{f}_k$	0.15	0.20	0.00	0.00	

large	0	1	2	3	4
0	500	1000	200	-400	0
1	500	1000	200	-400	**0**
2	500	1000	200	**-400**	**0**
3	500	1000	**200**	**-400**	**0**
4	500	**1000**	**200**	**-400**	**0**
$\widehat{f}_k$	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. ☹ ☺

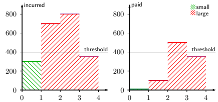


Ever and never large up to now: Projection

small	0	1	2	3	4
0	1300	200	300	0	0
1	1300	200	300	0	**0**
2	1300	200	300	**0**	**0**
3	1300	200	**300**	**0**	**0**
4	1300	**200**	**300**	**0**	**0**
$\widehat{f}_k$	0.15	0.20	0.00	0.00	

large	0	1	2	3	4
0	500	1000	200	-400	0
1	500	1000	200	-400	**0**
2	500	1000	200	**-400**	**0**
3	500	1000	**200**	**-400**	**0**
4	500	**1000**	**200**	**-400**	**0**
$\widehat{f}_k$	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. ☹ ☺



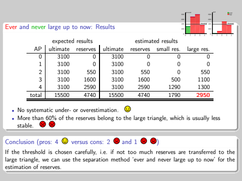
Ever and never large up to now: Results

	expected results		estimated results			
AP	ultimate	reserves	ultimate	reserves	small res.	large res.
0	3100	0	3100	0	0	0
1	3100	0	3100	0	0	0
2	3100	550	3100	550	0	550
3	3100	1600	3100	1600	500	1100
4	3100	2590	3100	2590	1290	1300
total	15500	4740	15500	4740	1790	**2950**

- 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.



Ever and never large up to now: Results

	expected results		estimated results			
AP	ultimate	reserves	ultimate	reserves	small res.	large res.
0	3100	0	3100	0	0	0
1	3100	0	3100	0	0	0
2	3100	550	3100	550	0	550
3	3100	1600	3100	1600	500	1100
4	3100	2590	3100	2590	1290	1300
total	15500	4740	15500	4740	1790	2950

- 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 ☹ ☹)

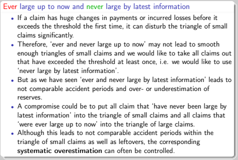
If the threshold is chosen carefully, i.e. if not too much reserves are transferred to the large triangle, we can use the separation method 'ever and never large up to now' for the estimation of reserves.

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 not be 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.

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 exceeded 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.



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 exceeded 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

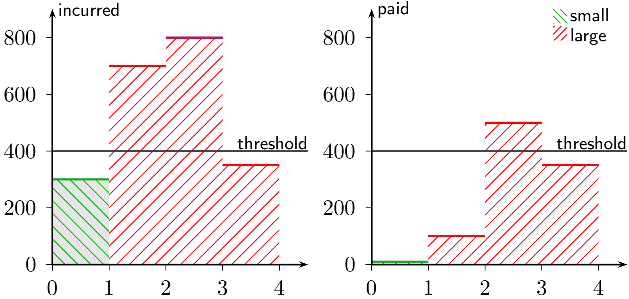
has good properties but may still leave a lot of reserves within the triangle of large claims. On way to get around this is to take the following method.

Ever large up to now and never large by latest information: Classification

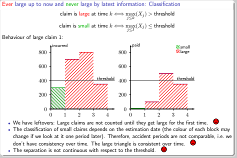
$$\text{claim is large at time } k \iff \max_{j\le k}(X_j) > \text{threshold}$$

$$\text{claim is small at time } k \iff \max_{j\le I}(X_j) \le \text{threshold}$$

Behaviour of large claim 1:



- 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. ☹



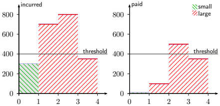
Ever large up to now and never large by latest information: Classification

$$\text{claim is large at time } k \iff \max_{j\le k}(X_j) > \text{threshold}$$

$$\text{claim is small at time } k \iff \max_{j\le I}(X_j) \le \text{threshold}$$

Behaviour of large claim 1:

- 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: Projection

small	0	1	2	3	4
0	1000	500	300	0	0
1	1000	500	300	0	**0**
2	1000	500	300	**0**	**0**
3	1000	500	**300**	**0**	**0**
4	**1300**	**650**	**390**	**0**	**0**
$\widehat{f}_k$	0.5	0.2	0.00	0.00	

large	0	1	2	3	4
0	500	1000	200	-400	0
1	500	1000	200	-400	**0**
2	500	1000	200	**-400**	**0**
3	500	1000	**200**	**-400**	**0**
4	500	**1000**	**200**	**-400**	**0**
$\widehat{f}_k$	2.00	0.13	0.24	0.00	

- 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

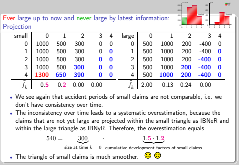
$$540 = \underbrace{300}_{\text{size at time } k=0} \cdot \underbrace{1.5 \cdot 1.2}_{\text{cumulative development factors of small claims}}$$

- The triangle of small claims is much smoother. ☺ ☺

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

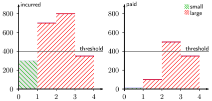
small	0	1	2	3	4
0	1000	500	300	0	0
1	1000	500	300	0	0
2	1000	500	300	0	0
3	1000	500	300	0	0
4	1300	650	390	0	0
$\widehat{f}_k$	0.5	0.2	0.00	0.00	

large	0	1	2	3	4
0	500	1000	200	-400	0
1	500	1000	200	-400	0
2	500	1000	200	-400	0
3	500	1000	200	-400	0
4	500	1000	200	-400	0
$\widehat{f}_k$	2.00	0.13	0.24	0.00	

- 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

$$540 = \underbrace{300}_{\text{size at time } k=0} \cdot \underbrace{1.5 \cdot 1.2}_{\text{cumulative development factors of small claims}}$$

- The triangle of small claims is much smoother. ☺ ☺



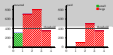
Ever large up to now and never large by latest information: Results

	expected results		estimated results			
AP	ultimate	reserves	ultimate	reserves	small res.	large res.
0	3100	0	3100	0	0	0
1	3100	0	3100	0	0	0
2	3100	550	3100	550	0	550
3	3100	1600	3100	1600	500	1100
4	3100	2590	**3640**	**3130**	1830	1300
total	15500	4740	**16040**	**5280**	2330	**2950**

- 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 control the systematic overestimation the separation method ‘ever large up to now and never large by latest information’ can be used.



Ever large up to now and never large by latest information: Results

	expected results		estimated results			
AP	ultimate	reserves	ultimate	reserves	small res.	large res.
0	3100	0	3100	0	0	0
1	3100	0	3100	0	0	0
2	3100	550	3100	550	0	550
3	3100	1600	3100	1600	500	1100
4	3100	2590	**3640**	**3130**	1830	1300
total	15500	4740	**16040**	**5280**	2330	**2950**

- 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 ☹)

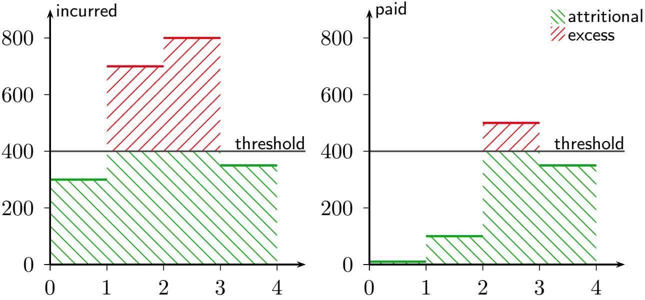
If we control the systematic overestimation the separation method 'ever large up to now and never large by latest information' can be used.

Attritional and excess: Classification

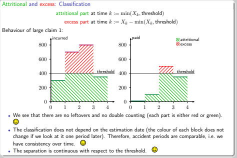
$$\text{attritional part at time } k := \min(X_k, \text{threshold})$$

$$\text{excess part at time } k := X_k - \min(X_k, \text{threshold})$$

Behaviour of large claim 1:



- 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. ☺



Attritional and excess: Classification

$$\text{attritional part at time } k := \min(X_k, \text{threshold})$$

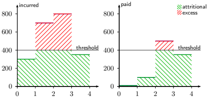
$$\text{excess part at time } k := X_k - \min(X_k, \text{threshold})$$

Behaviour of large claim 1:

- 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.

Attritional and excess: Projection

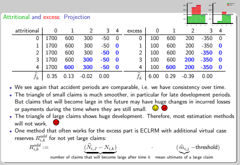


attritional	0	1	2	3	4
0	1700	600	300	-50	0
1	1700	600	300	-50	**0**
2	1700	600	300	**-50**	**0**
3	1700	600	**300**	**-50**	**0**
4	1700	**600**	**300**	**-50**	**0**
$\widehat{f}_k$	0.35	0.13	-0.02	0.00	

excess	0	1	2	3	4
0	100	600	200	-350	0
1	100	600	200	-350	**0**
2	100	600	200	**-350**	**0**
3	100	600	**200**	**-350**	**0**
4	100	**600**	**200**	**-350**	**0**
$\widehat{f}_k$	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. ☺ ☹
- 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 $R^{add}_{i,k}$ for not yet large claims:

$$R^{add}_{i,k} := \underbrace{(\widehat{N}_{i,J} - N_{i,k})}_{\text{number of claims that will become large after time } k} \cdot \underbrace{(\widehat{m}_{i,j}}_{\text{mean ultimate of a large claim}} - \text{threshold})$$



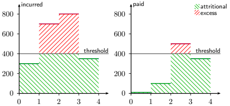
Attritional and excess: Projection

attritional	0	1	2	3	4
0	1700	600	300	-50	0
1	1700	600	300	-50	**0**
2	1700	600	300	**-50**	**0**
3	1700	600	**300**	**-50**	**0**
4	1700	**600**	**300**	**-50**	**0**
$\widehat{f}_k$	0.35	0.13	-0.02	0.00	

excess	0	1	2	3	4
0	100	600	200	-350	0
1	100	600	200	-350	**0**
2	100	600	200	**-350**	**0**
3	100	600	**200**	**-350**	**0**
4	100	**600**	**200**	**-350**	**0**
$\widehat{f}_k$	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. ☺ ☹
- 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 $R_{i,k}^{add}$ for not yet large claims:

$$R_{i,k}^{add} := \underbrace{(\widehat{N}_{i,J} - N_{i,k})}_{\text{number of claims that will become large after time } k} \cdot \underbrace{(\widehat{m}_{i,j}}_{\text{mean ultimate of a large claim}} - \text{threshold})$$



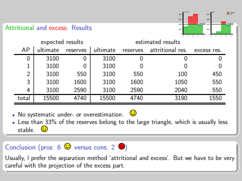
Attritional and excess: Results

	expected results		estimated results			
AP	ultimate	reserves	ultimate	reserves	attritional res.	excess res.
0	3100	0	3100	0	0	0
1	3100	0	3100	0	0	0
2	3100	550	3100	550	100	450
3	3100	1600	3100	1600	1050	550
4	3100	2590	3100	2590	2040	550
total	15500	4740	15500	4740	3190	1550

- 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.



Attritional and excess: Results

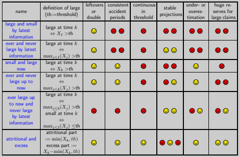
	expected results		estimated results			
AP	ultimate	reserves	ultimate	reserves	attritional res.	excess res.
0	3100	0	3100	0	0	0
1	3100	0	3100	0	0	0
2	3100	550	3100	550	100	450
3	3100	1600	3100	1600	1050	550
4	3100	2590	3100	2590	2040	550
total	15500	4740	15500	4740	3190	1550

- 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.

name	definition of large (th:=threshold)	leftovers or double	consistent accident periods	continuous in threshold	stable projections	under- or overestimation	huge reserves for large claims
large and small by latest information	large at time k $\Leftrightarrow X_I >$th	☺	☹ ☹	☹	☹ ☹	☹ ☹	☹ ☹
ever and never large by latest information	large at time k $\Leftrightarrow$ $\max_{j\leq I}(X_j) >$th	☺	☹ ☹	☹	☺ ☺	☹ ☹	☹ ☹
small and large now	large at time k $\Leftrightarrow X_k >$th	☺	☺	☹	☹ ☹	☺	☹
ever and never large up to now	large at time k $\Leftrightarrow$ $\max_{j\leq k}(X_j) >$th	☺	☺	☹	☹ ☺	☺	☹ ☹
ever large up to now and never large by latest information	large at time k $\Leftrightarrow$ $\max_{j\leq k}(X_j) >$th small at time k $\Leftrightarrow$ $\max_{j\leq I}(X_j) \leq$th	☹	☹	☹	☺ ☺	☹ ☺	☹ ☺
attritional and excess	attritional part $:= \min(X_k, th)$ excess part $:= X_k - \min(X_k, th)$	☺	☺	☺	☹ ☺ ☹	☺	☺ ☺



name	definition of large (th:=threshold)	leftovers or double	consistent accident periods	continuous in threshold	stable projections	under- or overestimation	huge reserves for large claims
large and small by latest information	large at time k $\Leftrightarrow X_I >$th	☺	☹ ☹	☹	☹ ☹	☹ ☹	☹ ☹
ever and never large by latest information	large at time k $\Leftrightarrow$ $\max_{j<I}(X_j) >$th	☺	☹ ☹	☹	☺ ☺	☹ ☹	☹ ☹
small and large now	large at time k $\Leftrightarrow X_k >$th	☺	☺	☹	☹ ☹	☺	☹
ever and never large up to now	large at time k $\Leftrightarrow$ $\max_{j<k}(X_j) >$th	☺	☺	☹	☹ ☺	☺	☹ ☹
ever large up to now and never large by latest information	large at time k $\Leftrightarrow$ $\max_{j\leq k}(X_j) >$th small at time k $\Leftrightarrow$ $\max_{j<I}(X_j) \leq$th	☹	☹	☹	☺ ☺	☹ ☺	☹ ☺
attritional and excess	attritional part $:= \min(X_k, th)$ excess part $:=$ $X_k - \min(X_k, th)$	☺	☺	☺	☹ ☺ ☹	☺	☺ ☺

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.

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 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.

Bifurcation of large and small losses: Basic idea

The separation methods we have seen up to now do not look at stochastic 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 [30]. 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.

2021-04-26

Stochastic Reserving

Separation of small and large claims

Modelling the transition from small to large

Bifurcation of large and small losses: Basic idea

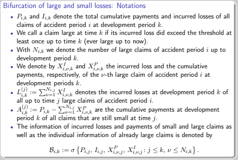
The separation methods we have seen up to now do not look at stochastic transition of claims from the triangle of small claims to the one of large claims. We now want to try to model these transitions and will follow the notation of U. Riegel [30]. The basic idea is to look separately at:

- 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 $X^{I}_{i,\nu,k}$ and $X^{P}_{i,\nu,k}$ the incurred loss and the cumulative payments, respectively, of the ν-th large claim of accident period i at development periods k.
- $L^{(j)}_{i,k} := \sum_{\nu=1}^{N_{i,j}} X^{I}_{i,\nu,k}$ denotes the incurred losses at development period k of all up to time j large claims of accident period i.
- $A^{(j)}_{i,k} := P_{i,k} - \sum_{\nu=1}^{N_{i,j}} X^{P}_{i,\nu,k}$ 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

$$\mathcal{B}_{i,k} := \sigma\left\{P_{i,j},\, I_{i,j},\, X^{P}_{i,\nu,j},\, X^{I}_{i,\nu,j} : j \leq k,\, \nu \leq N_{i,k}\right\}.$$



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 $X^I_{i,\nu,k}$ and $X^P_{i,\nu,k}$ the incurred loss and the cumulative payments, respectively, of the ν-th large claim of accident period i at development periods k.
- $L^{(j)}_{i,k} := \sum_{\nu=1}^{N_{i,j}} X^I_{i,\nu,k}$ denotes the incurred losses at development period k of all up to time j large claims of accident period i.
- $A^{(j)}_{i,k} := P_{i,k} - \sum_{\nu=1}^{N_{i,j}} X^P_{i,\nu,k}$ 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

$$\mathcal{B}_{i,k} := \sigma\left\{P_{i,j},\, I_{i,j},\, X^P_{i,\nu,j},\, X^I_{i,\nu,j} \colon j \le k,\, \nu \le N_{i,k}\right\}.$$

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.

$$\mathsf{E}[N_{i,k+1}|\mathcal{B}_{i,k}] = n_k N_{i,k}.$$

3. The cumulative payments of small claims as long as they stay small develop according to CLM, i.e.

$$\mathsf{E}\left[A_{i,k+1}^{(k+1)}\middle|\mathcal{B}_{i,k},\ A_{i,k}^{(k+1)}\right] = a_k A_{i,k}^{(k+1)}.$$

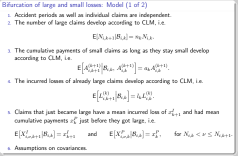
4. The incurred losses of already large claims develop according to CLM, i.e.

$$\mathsf{E}\left[L_{i,k+1}^{(k)}\middle|\mathcal{B}_{i,k}\right] = l_k L_{i,k}^{(k)}.$$

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.

$$\mathsf{E}\left[X_{i,\nu,k+1}^I\middle|\mathcal{B}_{i,k}\right] = x_{k+1}^I \qquad \text{and} \qquad \mathsf{E}\left[X_{i,\nu,k}^P\middle|\mathcal{B}_{i,k}\right] = x_k^P, \qquad \text{for } N_{i,k} < \nu \leq N_{i,k+1}.$$

6. Assumptions on covariances.



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.
$$\mathsf{E}[N_{i,k+1}|\mathcal{B}_{i,k}] = n_k N_{i,k}.$$
3. The cumulative payments of small claims as long as they stay small develop according to CLM, i.e.
$$\mathsf{E}\left[A_{i,k+1}^{(k+1)}\middle|\mathcal{B}_{i,k},\, A_{i,k}^{(k+1)}\right] = a_k A_{i,k}^{(k+1)}.$$
4. The incurred losses of already large claims develop according to CLM, i.e.
$$\mathsf{E}\left[L_{i,k+1}^{(k)}\middle|\mathcal{B}_{i,k}\right] = l_k L_{i,k}^{(k)}.$$
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.
$$\mathsf{E}[X_{i,\nu,k+1}^I|\mathcal{B}_{i,k}] = x_{k+1}^I \quad \text{and} \quad \mathsf{E}[X_{i,\nu,k}^P|\mathcal{B}_{i,k}] = x_k^P, \quad \text{for } N_{i,k} < \nu \le N_{i,k+1}.$$
6. Assumptions on covariances.

- 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.
- 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 disturbed.
- Except for the additional conditioning for small claims and the different upper index for large claims the formulas are almost the same as for LSRMs.

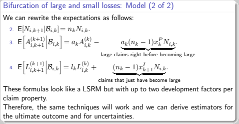
Bifurcation of large and small losses: Model (2 of 2)

We can rewrite the expectations as follows:

2. $\mathsf{E}[N_{i,k+1}|\mathcal{B}_{i,k}] = n_k N_{i,k}$.
3. $\mathsf{E}\left[A_{i,k+1}^{(k+1)}\middle|\mathcal{B}_{i,k}\right] = a_k A_{i,k}^{(k)} - \underbrace{a_k(n_k - 1)x_k^P N_{i,k}}_{\text{large claims right before becoming large}}$.
4. $\mathsf{E}\left[L_{i,k+1}^{(k+1)}\middle|\mathcal{B}_{i,k}\right] = l_k L_{i,k}^{(k)} + \underbrace{(n_k - 1)x_{k+1}^I N_{i,k}}_{\text{claims that just have become large}}$.

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. $\mathsf{E}[N_{i,k+1}|\mathcal{B}_{i,k}] = n_k N_{i,k}$.
3. $\mathsf{E}\left[A_{i,k+1}^{(k+1)}\middle|\mathcal{B}_{i,k}\right] = a_k A_{i,k}^{(k)} - \underbrace{a_k(n_k-1)x_k^P N_{i,k}}_{\text{large claims right before becoming large}}$.
4. $\mathsf{E}\left[L_{i,k+1}^{(k+1)}\middle|\mathcal{B}_{i,k}\right] = l_k L_{i,k}^{(k)} + \underbrace{(n_k-1)x_{k+1}^I N_{i,k}}_{\text{claims that just have become large}}$.

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.

2. is unchanged.

3.
$$\begin{aligned}
\mathsf{E}\left[A_{i,k+1}^{(k+1)}\middle|\mathcal{B}_{i,k}\right] &= \mathsf{E}\left[\mathsf{E}\left[A_{i,k+1}^{(k+1)}\middle|\mathcal{B}_{i,k}, A_k^{(k+1)}\right]\middle|\mathcal{B}_{i,k}\right] = a_k \mathsf{E}\left[A_k^{(k+1)}\middle|\mathcal{B}_{i,k}\right] \\
&= a_k \left(A_{i,k}^{(k)} - \mathsf{E}\left[\sum_{\nu=N_{i,k}+1}^{N_{i,k+1}} X_{i,\nu,k}^P \middle|\mathcal{B}_{i,k}\right]\right) \\
&= a_k \left(A_{i,k}^{(k)} - \mathsf{E}\left[N_{i,k+1} - N_{i,k}\middle|\mathcal{B}_{i,k}\right]\mathsf{E}\left[X_{i,N_{i,k+1},k}^P\middle|\mathcal{B}_{i,k}\right]\right) \\
&= a_k A_{i,k}^{(k)} - a_k(n_k-1)x_k^P N_{i,k}
\end{aligned}$$

4.
$$\begin{aligned}
\mathsf{E}\left[L_{i,k+1}^{(k+1)}\middle|\mathcal{B}_{i,k}\right] &= \mathsf{E}\left[L_{i,k+1}^{(k)} + \sum_{\nu=N_{i,k}+1}^{N_{i,k+1}} X_{i,\nu,k+1}^I \middle|\mathcal{B}_{i,k}\right] \\
&= l_k L_{i,k}^{(k)} + \mathsf{E}\left[N_{i,k+1} - N_{i,k}\middle|\mathcal{B}_{i,k}\right]\mathsf{E}\left[X_{i,N_{i,k+1},k}^I\middle|\mathcal{B}_{i,k}\right] \\
&= l_k L_{i,k}^{(k)} + (n_k-1)x_k^I N_{i,k}
\end{aligned}$$

Literature

[30] Ulrich Riegel.
A Bifurcation Approach for Attritional and Large Losses in Chain Ladder Calculations.
Astin Bull., 44(1):127–172, 2013.

Literature

[30] Ulrich Riegel.
A Bifurcation Approach for Attritional and Large Losses in Chain Ladder Calculations.
Astin Bull., 44(1):127–172, 2013.

Stochastic Reserving

Lecture 14

Trail exams and examples

René Dahms

ETH Zurich, Spring 2021

2 June 2021

(Last update: 26 April 2021)

2021-04-26

Stochastic Reserving

Stochastic Reserving

Lecture 14

Trail exams and examples

René Dahms

ETH Zurich, Spring 2021

2 June 2021
(Last update: 26 April 2021)

10 Examples & Trail Exam

Lecture 14: Table of contents

Preparation (only if Covid-19 allow for it)

Please bring your laptop with installed LSRMTools.

If you have problems to get the LSRMTools running ☹
be 30 minutes earlier and you will be helped. ☺ (only if Covid-19 allow for it)

Train exam

If someone of you is brave enough we can make a trail examination.