Prediction error of the expected claims development result in the chain ladder method

Michael Merz, Mario V. Wüthrich, Tübingen, Zürich

Prediction Error of the Expected Claims Development Result in the Chain Ladder Method

1 Motivation

This work is motivated by the Swiss Solvency Test [6] (SST). Non-life insurance companies have to determine various parameters and distributions in order to calculate the risk bearing capital and the target capital for the SST. One central risk driver for non-life insurance companies is the development of their claims reserves. For the solvency test one has to estimate the first two moments of the distribution of the claims development result (in the next accounting year). It is then assumed that the claims development result has a shifted lognormal distribution with parameters estimated by exactly these first two moments (see SST [6], Section 4.4.10 on p. 65).

Assuming that one has an unbiased estimator for the runoff liabilities, it is clear that the expected claims development result for the next accounting year is 0. However, since we predict future cashflows, the observations may substantially deviate from this expected value. This deviation is measured with the help of the second moment.

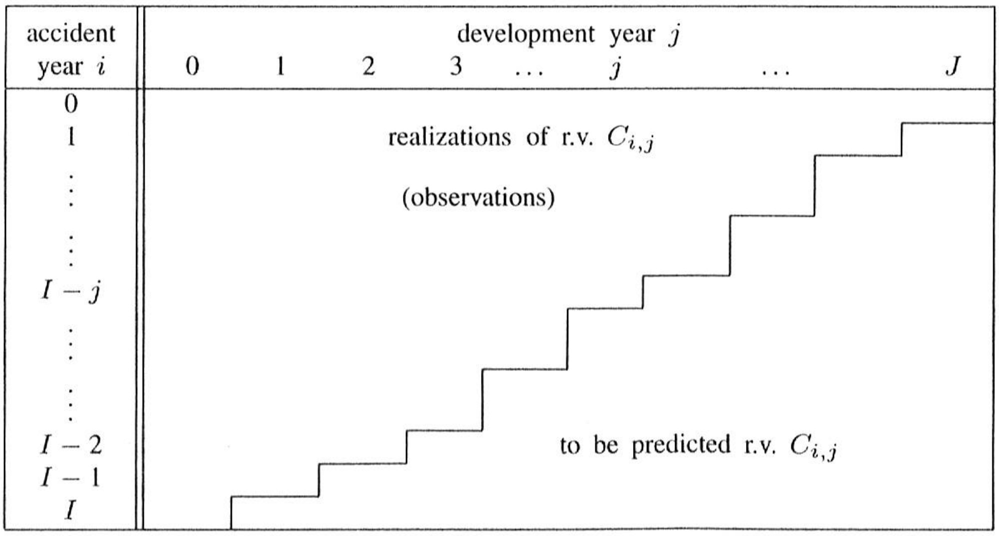
In the current version of the SST (see [6], Section 4.4.10 on p. 65) there is no underlying stochastic model defined for measuring this uncertainty. There is only an instruction that the estimated variance of the claims development result needs to have two parts, one measuring the process variance, the other one measuring the estimation error.

Our main result of the present work (Result 4.5, below) gives a mathematical approach for the estimation of the conditional mean square error of prediction (MSEP) of the expected claims development result for the next accounting year. The formula is based on Mack's classical stochastic chain ladder reserving model (Mack [3]) and its extension to the time series framework (see Buchwalder et al. [1] and Murphy [5]).

1.1 The development triangles

Assume that $C_{i,j}$ denote the cumulative payments for accident year $i \in \{0, \ldots, I\}$ until development year $j \in \{0, \ldots, J\}$. Of course, our analysis is not restricted to cumulative payments, but hold true for any meaning of $C_{i,j}$. However, to simplify our language we identify $C_{i,j}$ with cumulative payments. Moreover for simplicity, we always assume that $J = I$. Of course, all formulas similarly hold true for $I > J$ (development trapezoids).

Usually, claims development figures $C_{i,j}$ are studied in loss development triangles, which have the following structure ($I = J$):



Assume that we are at time $t = I$, i.e. we have observations

$$\mathcal{D}_I = \{C_{i,j};\ i + j \le I\}. \tag{1.1}$$

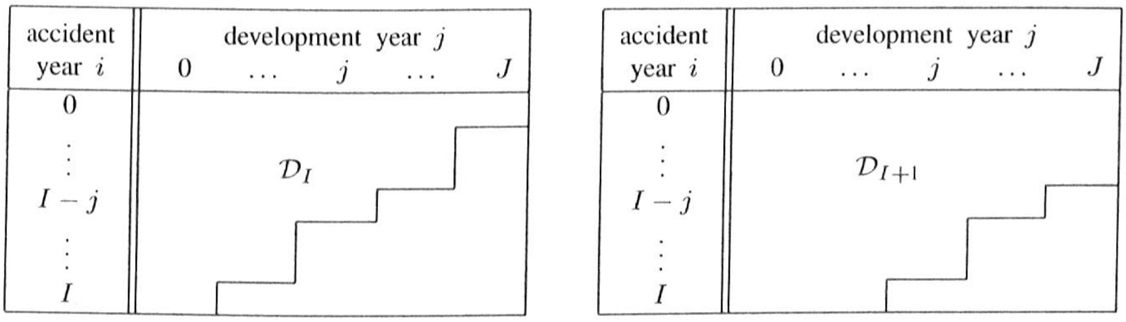
Our goal is to estimate for every accident year $i \in \{0, \ldots, I\}$

$$E\left[C_{i,J} \middle| \mathcal{D}_I\right], \tag{1.2}$$

i.e. we want to predict the (mean of the) ultimate claim $C_{i,J}$, given the information $\mathcal{D}_I$.

If we go one step ahead in time $t \mapsto t + 1$, we obtain new observations on the diagonal of our claims development rectangle. This means that at time $t + 1 = I + 1$ we have the following observations for accident years $i \le I$:

$$\mathcal{D}_{I+1} = \{C_{i,j};\ i + j \le I + 1 \text{ and } i \le I\} = \mathcal{D}_I \cup \{C_{i,I-i+1};\ i \le I\}. \tag{1.3}$$



Hence, we enlarge our σ-algebra $\sigma(\mathcal{D}_I) \to \sigma(\mathcal{D}_{I+1})$ and we look for a new estimator

$$E\left[C_{i,J} \middle| \mathcal{D}_{I+1}\right] \tag{1.4}$$

at time $I+1$. The goal of this work is to set up a stochastic model such that we are able to analyze the fluctuations of such updated predictions.

2 The chain ladder method

We study this successive prediction problem in the framework of the stochastic chain-ladder model. Mack [3] was the first one to study the distribution free stochastic model for the chain ladder reserving method. In the present work we focus on the time series version of the chain ladder model (see Buchwalder et al. [1] and Model IV in Murphy [5]). The time series version of the chain ladder model has the advantage that it defines an explicit mechanism for generating additional observations (see discussions in [1]).

Model Assumptions 2.1 (Chain ladder time series model)

We assume that the cumulative payments in different accident years $i \in \{0, \dots, I\}$ are independent, and that there exist constants $f_l > 0$, $\sigma_l \geq 0$ $(l = 0, \dots, J-1)$ such that for all $1 \leq j \leq J$ and $0 \leq i \leq I$ we have

$$C_{i,j} = f_{j-1} \cdot C_{i,j-1} + \sigma_{j-1} \cdot \sqrt{C_{i,j-1}} \cdot \varepsilon_{i,j}\,, \tag{2.1}$$

where $\varepsilon_{i,j}$ are independent random variables with

$$E\left[\varepsilon_{i,j}\right] = 0 \text{ and } E\left[\varepsilon_{i,j}^2\right] = 1\,. \tag{2.2}$$

□

Remarks 2.2

- Formula (2.1) defines an autoregressive structure (time series model) for the reserving problem.
- Observe that Model Assumptions 2.1 imply (see also (1.2)-(1.4))

$$E\left[C_{i,J}\middle|\mathcal{D}_I\right] = C_{i,I-i}\cdot\prod_{j=I-i}^{J-1} f_j \quad \text{and}$$
$$E\left[C_{i,J}\middle|\mathcal{D}_{I+1}\right] = C_{i,I-i+1}\cdot\prod_{j=I-i+1}^{J-1} f_j\,. \tag{2.3}$$

  This means that, as soon as the chain ladder factors f_j are known, we are able to predict the (conditionally expected) ultimate claim $C_{i,J}$ given the information $\mathcal{D}_I$ and $\mathcal{D}_{I+1}$, respectively.
- Our Model Assumptions 2.1 satisfy the model assumptions of the Mack's chain ladder model (see [3] and [1]).

As usual in the chain ladder algorithm, the age-to-age factors f_j are estimated as follows:

1. At time $t = I$ (given the information $\mathcal{D}_I$) the chain ladder factors f_j are estimated by

$$\widehat{f}_j^I = \frac{\sum_{i=0}^{I-j-1} C_{i,j+1}}{S_j^I}, \qquad \text{where} \quad S_j^I = \sum_{i=0}^{I-j-1} C_{i,j}\,. \tag{2.4}$$

2. At time $t = I+1$ (given the information $\mathcal{D}_{I+1}$) the chain ladder factors f_j are estimated by

$$\widehat{f}_j^{I+1} = \frac{\sum_{i=0}^{I-j} C_{i,j+1}}{S_j^{I+1}}, \qquad \text{where} \quad S_j^{I+1} = \sum_{i=0}^{I-j} C_{i,j}\,. \tag{2.5}$$

Mack [3] has proved that these are unbiased estimators for f_j, and moreover that $\widehat{f}_j^m$ and $\widehat{f}_l^m$ ($m = I$ or $I+1$) are uncorrelated random variables for $j \neq l$ (see Theorem 2 in [3]).

This immediately implies that, given $C_{i,I-i}$,

$$\widehat{C}_{i,j}^{I} = C_{i,I-i} \cdot \widehat{f}_{I-i}^{I} \cdot \ldots \cdot \widehat{f}_{j-2}^{I} \cdot \widehat{f}_{j-1}^{I} \tag{2.6}$$

is an unbiased estimator for $E\left[C_{i,j} \middle| \mathcal{D}_I\right]$ with $j \geq I-i$, and given $C_{i,I-i+1}$,

$$\widehat{C}_{i,j}^{I+1} = C_{i,I-i+1} \cdot \widehat{f}_{I-i+1}^{I+1} \cdot \ldots \cdot \widehat{f}_{j-2}^{I+1} \cdot \widehat{f}_{j-1}^{I+1} \tag{2.7}$$

is an unbiased estimator for $E\left[C_{i,j} \middle| \mathcal{D}_{I+1}\right]$ with $j \geq I-i+1$.

Remarks 2.3

- In the sequel we call $\widehat{f}_j^I$ and $\widehat{f}_j^{I+1}$ best estimates for the chain ladder factor f_j given the information $\mathcal{D}_I$ and $\mathcal{D}_{I+1}$, respectively.
- The realizations of the estimators $\widehat{f}_0^I, \ldots, \widehat{f}_{J-1}^I$ are known at time $t = I$, but the realizations of $\widehat{f}_0^{I+1}, \ldots, \widehat{f}_{J-1}^{I+1}$ are unknown since $C_{I,1}, \ldots, C_{I-J+1,J}$ are unknown.
- In the following we identify an empty product by 1. For example, $\widehat{C}_{i,I-i}^{I} = C_{i,I-i}$ and $\widehat{C}_{i,I-i+1}^{I+1} = C_{i,I-i+1}$, if the product (2.6) has no factors.

We have the following Lemma.

Lemma 2.4 *Under Model Assumptions 2.1 we have*

a) $C_{i,I-i+1}, \widehat{f}_{I-i+1}^{I+1}, \ldots, \widehat{f}_{J-1}^{I+1}$ *are conditionally independent w.r.t.* $\mathcal{D}_I$,

b) $$E\left[\widehat{f}_l^{I+1} \middle| \mathcal{D}_I\right] = \frac{\sum_{i=0}^{I-l-1} C_{i,l+1}}{S_l^{I+1}} + f_l \cdot \frac{C_{I-l,l}}{S_l^{I+1}} = \frac{S_l^I}{S_l^{I+1}} \cdot \widehat{f}_l^I + f_l \cdot \frac{C_{I-l,l}}{S_l^{I+1}},$$

c) $$E\left[\widehat{C}_{i,j}^{I+1} \middle| \mathcal{D}_I\right] = C_{i,I-i} \cdot f_{I-i} \cdot \prod_{l=I-i+1}^{j-1} E\left[\widehat{f}_l^{I+1} \middle| \mathcal{D}_I\right].$$

Proof of Lemma 2.4. a) Given $\mathcal{D}_I$, $C_{i,I-i+1} = f_{I-i} \cdot C_{i,I-i} + \sigma_{I-i} \cdot \sqrt{C_{i,I-i}} \cdot \varepsilon_{i,I-i+1}$ is a function (random variable) in $\varepsilon_{i,I-i+1}$ and for $l = I-i+1, \ldots, J-1$

$$\begin{aligned}\widehat{f}_l^{I+1} &= \frac{\sum_{i=0}^{I-l} C_{i,l+1}}{S_l^{I+1}} = \frac{\sum_{i=0}^{I-l-1} C_{i,l+1}}{S_l^{I+1}} + \frac{C_{I-l,l+1}}{S_l^{I+1}} \\ &= \frac{S_l^I}{S_l^{I+1}} \cdot \widehat{f}_l^I + \frac{f_l \cdot C_{I-l,l} + \sigma_l \cdot \sqrt{C_{I-l,l}} \cdot \varepsilon_{I-l,l+1}}{S_l^{I+1}}\end{aligned} \tag{2.8}$$

is a function (random variable) in $\varepsilon_{I-l,l+1}$. By our model assumptions the random variables $\varepsilon_{i,I-i+1}, \varepsilon_{i-1,I-i+2}, \ldots, \varepsilon_{I-J+1,J}$ are independent, hence claim a) is proved.
b) From Model Assumptions 2.1 and (2.8) we have that

$$E\left[\widehat{f}_l^{I+1}\middle|\mathcal{D}_I\right] = \frac{\sum_{i=0}^{I-l-1} C_{i,l+1}}{S_l^{I+1}} + f_l \cdot \frac{C_{I-l,l}}{S_l^{I+1}} . \tag{2.9}$$

c) Follows from a) and our Model Assumptions 2.1.
This completes the proof of Lemma 2.4. □

3 Claims development result

The outstanding claims liabilities at time $t = I$ for accident year $i \in \{1, \ldots, I\}$ are given by

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

these are the outstanding claims payments at time I. Analogously, at time $t+1 = I+1$ the outstanding liabilities for accident year i are given by

$$R_i^{I+1} = C_{i,J} - C_{i,I-i+1} . \tag{3.2}$$

Using the chain ladder framework (Model Assumptions 2.1), given $C_{i,I-i}$,

$$\widehat{R}_i^{\mathcal{D}_I} = \widehat{C}_{i,J}^I - C_{i,I-i} \qquad (1 \le i \le I), \tag{3.3}$$

is an unbiased estimator for $E\left[R_i^I\middle|\mathcal{D}_I\right]$ and, given $C_{i,I-i+1}$,

$$\widehat{R}_i^{\mathcal{D}_{I+1}} = \widehat{C}_{i,J}^{I+1} - C_{i,I-i+1} \qquad (1 \le i \le I), \tag{3.4}$$

is an unbiased estimator for $E\left[R_i^{I+1}\middle|\mathcal{D}_{I+1}\right]$.

Definition 3.1 (Claims development result)

The claims development result for accident year $i \in \{1, \ldots, I\}$ in accounting year $(I, I+1]$ is given by

$$\mathrm{CDR}_i(I+1) = E\left[R_i^I\middle|\mathcal{D}_I\right] - \left(X_{i,I-i+1} + E\left[R_i^{I+1}\middle|\mathcal{D}_{I+1}\right]\right) , \tag{3.5}$$

with incremental payments given by $X_{i,I-i+1} = C_{i,I-i+1} - C_{i,I-i}$. □

Corollary 3.2 (Best estimator) *$CDR_i(I+1)$ is a $\mathcal{D}_{I+1}$-measurable random variable with conditional mean 0 given $\mathcal{D}_I$. Moreover,*

$$CDR_i(I+1) = E\left[C_{i,J} \middle| \mathcal{D}_I\right] - E\left[C_{i,J} \middle| \mathcal{D}_{I+1}\right]$$
$$= \left(E\left[C_{i,I-i+1} \middle| C_{i,I-i}\right] - C_{i,I-i+1}\right) \cdot \prod_{j=I-i+1}^{J-1} f_j \,. \tag{3.6}$$

Proof. This is the martingale property of successive predictions. □

Since the claims development factors f_j are unknown, they are replaced by their estimators at time I and $I+1$, respectively. The observed claims development result at time $I+1$ for accident year i (the estimator of the claims development result at time $I+1$ viewed from time I) is given by ($1 \le i \le I$)

$$\widehat{\mathrm{CDR}}_i(I+1) = \widehat{R}_i^{\mathcal{D}_I} - \left(X_{i,I-i+1} + \widehat{R}_i^{\mathcal{D}_{I+1}}\right) = \widehat{C}_{i,J}^{I} - \widehat{C}_{i,J}^{I+1} \,. \tag{3.7}$$

Lemma 3.3 *Under Model Assumptions 2.1 the expected claims development result in the chain ladder method for unknown claims development factors f_j is given by*

$$E\left[\widehat{CDR}_i(I+1) \middle| \mathcal{D}_I\right] \tag{3.8}$$
$$= C_{i,I-i} \cdot \left(\prod_{j=I-i}^{J-1} \widehat{f}_j^I - f_{I-i} \cdot \prod_{j=I-i+1}^{J-1} \left(\frac{S_j^I}{S_j^{I+1}} \cdot \widehat{f}_j^I + f_j \cdot \frac{C_{I-j,j}}{S_j^{I+1}} \right) \right) . \tag{3.9}$$

Proof. The proof easily follows from the definition of the chain ladder estimator and Lemma 2.4. □

Remarks 3.4

Lemma 3.3 gives the expected claims development result (viewed from time $t = I$) if we estimate the chain ladder factors f_j by their best estimators $\widehat{f}_j^I$ and $\widehat{f}_j^{I+1}$ at time I and at time $I+1$, respectively.
Observe that on the right-hand side of (3.8) we have the unknown parameters f_j. If these are replaced by their best estimators $\widehat{f}_j^I$ at time I, the right-hand side of (3.8) becomes 0. In this sense we have best estimates for the reserves, but pay attention to the fact that using estimates for true parameters leads to a bias, which says that the estimated claims reserves are not martingales. This is in contrast to the best estimate reserves for known chain ladder factors f_j (see

Corollary 3.2). Observe, using $S_j^{I+1} = S_j^I + C_{I-j,j}$, that (3.8) can be rewritten as follows

$$E\left[\widehat{\mathrm{CDR}}_i(I+1)\middle|\mathcal{D}_I\right] = \widehat{C}_{i,J}^I \cdot \left(1 - \frac{f_{I-i}}{\widehat{f}_{I-i}^I} \cdot \prod_{j=I-i+1}^{J-1} \left(1 + \left(f_j - \widehat{f}_j^I\right) \cdot \frac{C_{I-j,j}}{\widehat{f}_j^I \cdot S_j^{I+1}}\right)\right). \tag{3.10}$$

From this we see that (3.10) is 0 either when $f_j = \widehat{f}_j^I$ or if we average over $\widehat{f}_j^I$, since $\widehat{f}_j^I$ are unbiased, uncorrelated estimates for f_j.

4 Conditional prediction error in the expected claims development result

In this section we want to study the volatility of the expected claims development result (3.8) viewed from time $t = I$. If we know the true chain ladder factors f_j, this is simply the process variance (since future cashflows are random variables). If the chain ladder factors f_j are unknown, they are estimated by their best estimates $\widehat{f}_j^I$ and $\widehat{f}_j^{I+1}$ at time I and at time $I+1$, respectively. This implies that in addition to the process variance, we also have an estimation error. We study the estimation error of the expected claims development result viewed from time $t = I$.

4.1 Single accident years

We first derive estimators for the uncertainties in the claims reserves and the claims development result for a single accident year i. When aggregating over accident years the calculations and estimators become slightly more complicated (see Section 4.2, below).

Definition 4.1 (Conditional mean square error of prediction)

The conditional mean square error of prediction at time $t = I$ for the expected claims development result of a single accident year $i \in \{1, \ldots, I\}$ is defined as

$$\mathrm{MSEP}\left(E\left[\widehat{\mathrm{CDR}}_i(I+1)\middle|\mathcal{D}_I\right]\right) = E\left[\left(\mathrm{CDR}_i(I+1) - E\left[\widehat{\mathrm{CDR}}_i(I+1)\middle|\mathcal{D}_I\right]\right)^2\middle|\mathcal{D}_I\right]. \tag{4.1}$$

□

Remark. In the present work we study (as a first step) the MSEP of the expected claims development result (3.8). In a second step, one could also study the MSEP of the claims development result (3.7). Here we omit these calculations, the formulas for the claims development results (3.7) get slightly more complicated since one has a covariance term.
As usual, the MSEP is separated into two parts: a) (conditional) process variance and b) (conditional) estimation error. Using Corollary 3.2 we obtain:

$$\mathrm{MSEP}\left(E\left[\widehat{\mathrm{CDR}}_i(I+1)\middle|\mathcal{D}_I\right]\right) = \underbrace{\mathrm{Var}\left(\mathrm{CDR}_i(I+1)\middle|\mathcal{D}_I\right)}_{\text{process variance}} + \underbrace{\left(E\left[\widehat{\mathrm{CDR}}_i(I+1)\middle|\mathcal{D}_I\right]\right)^2}_{\text{estimation error}}. \qquad (4.2)$$

The process variance describes the pure random part of our claims reserving problem (since we deal with stochastic processes). The estimation error determines the uncertainty which comes from the estimation of the "true" chain ladder factors f_j.

4.1.1 (Conditional) Process variance

We start with the study of the first term on the right-hand side of (4.2), the process variance.

Lemma 4.2 (Conditional process variance for a single accident year)
Under Model Assumptions 2.1, the conditional process variance for the claims development result of accident year $i \in \{1, \ldots, I\}$ in accounting year $(I, I+1]$, given the observations $\mathcal{D}_I$, is given by

$$Var\left(CDR_i(I+1) \mid \mathcal{D}_I\right) = E\left[C_{i,J}\middle|\mathcal{D}_I\right]^2 \cdot \frac{\sigma_{I-i}^2/f_{I-i}^2}{C_{i,I-i}}. \qquad (4.3)$$

Remarks 4.3

- An estimator for the conditional process variance is obtained by replacing the right-hand side of (4.3) by the best estimator at time I

$$\widehat{\mathrm{Var}}\left(\mathrm{CDR}_i(I+1) \mid \mathcal{D}_I\right) = \left(\widehat{C}_{i,J}^I\right)^2 \cdot \frac{\left(\widehat{\sigma}_{I-i}^I\right)^2 / \left(\widehat{f}_{I-i}^I\right)^2}{C_{i,I-i}}, \qquad (4.4)$$

with

$$\left(\widehat{\sigma}_{j-1}^{I}\right)^2 = \frac{1}{I-j} \cdot \sum_{i=0}^{I-j} C_{i,j-1} \cdot \left(\frac{C_{i,j}}{C_{i,j-1}} - \widehat{f}_{j-1}^{I}\right)^2. \tag{4.5}$$

- Observe that the expression in (4.3) is the conditional process variance for one single development/accounting year. Aggregating these expressions over single accident years in an appropriate way leads to the well-known formula for the process variance of the chain ladder reserves (see proof of Theorem 3 in [3]).
- Since claims in different accident years are independent, we can easily aggregate the conditional process variance over different accident years to obtain the conditional process variance of the claims development result for the whole runoff portfolio,

$$\mathrm{Var}\left(\sum_{i=1}^{I} \mathrm{CDR}_i(I+1) \,\middle|\, \mathcal{D}_I\right) = \sum_{i=1}^{I} \mathrm{Var}\left(\mathrm{CDR}_i(I+1) \mid \mathcal{D}_I\right). \tag{4.6}$$

Proof of Lemma 4.2. Using Model Assumptions 2.1 and Corollary 3.2 we obtain

$$\begin{aligned} \mathrm{Var}\left(\mathrm{CDR}_i(I+1)\middle|\, \mathcal{D}_I\right) &= \mathrm{Var}\left(E\left[C_{i,J}\middle|\, \mathcal{D}_I\right] - E\left[C_{i,J}\middle|\, \mathcal{D}_{I+1}\right]\middle|\, \mathcal{D}_I\right) \\ &= \mathrm{Var}\left(E\left[C_{i,J}\middle|\, \mathcal{D}_{I+1}\right]\middle|\, \mathcal{D}_I\right) \\ &= \prod_{j=I-i+1}^{J-1} f_j^2 \cdot \mathrm{Var}\left(C_{i,I-i+1}\middle|\, \mathcal{D}_I\right) \\ &= C_{i,I-i} \cdot \prod_{j=I-i+1}^{J-1} f_j^2 \cdot \sigma_{I-i}^2 . \end{aligned} \tag{4.7}$$

This completes the proof of Lemma 4.2. □

4.1.2 Estimation error in the expected claims development result

Now, we treat the second term on the right-hand side of (4.2), i.e. we want to determine the volatility of the parameter estimators in the expected claims development result for a conditional approach (Approach 3 in Buchwalder et al. [1]). Approach 3 in [1] gives an answer to the question how one should estimate the (conditional) volatility of $\widehat{f}^2_{I-i} \cdots \widehat{f}^2_{J-1}$ around $f^2_{I-i} \cdots f^2_{J-1}$, namely by studing the volatility of the following terms

$$E\left[\widehat{f}^2_{I-i}\Big|\mathcal{B}_{I-i}\right] \cdots E\left[\widehat{f}^2_{J-1}\Big|\mathcal{B}_{J-1}\right], \tag{4.8}$$

where $\mathcal{B}_j = \{C_{i,k} \in \mathcal{D}_I;\ k \leq j\}$. Notice that the $\widehat{f}_j$'s are (unconditionally) not independent (see Mack et al. [4]), but (4.8) leads to a product structure. For other versions to measure the estimation error we refer to Buchwalder et al. [1] and the discussion papers by Mack et al. [4], Gisler [2] and Venter [8].
Notice that from Lemma 3.3 we have that

$$\begin{aligned} &E\left[\widehat{\mathrm{CDR}}_i(I+1)\Big|\mathcal{D}_I\right] \\ &= C_{i,I-i} \cdot \left(\prod_{j=I-i}^{J-1} \widehat{f}^I_j - f_{I-i} \cdot \prod_{j=I-i+1}^{J-1} \left(\frac{S^I_j}{S^{I+1}_j} \cdot \widehat{f}^I_j + f_j \cdot \frac{C_{I-j,j}}{S^{I+1}_j}\right)\right). \end{aligned} \tag{4.9}$$

For the calculation of the conditional estimation error, we now need to determine the volatility of $\widehat{f}^I_j$, since at that stage, the estimators $\widehat{f}^I_j$ of the chain ladder factors f_j are random variables. To address this source of uncertainty we proceed as usual in statistics: Given $\mathcal{D}_I$, we generate a set of "new" observations (resampling of the next step in the time series, given $C_{i,j-1}$):

$$Z_{i,j} = f_{j-1} \cdot C_{i,j-1} + \sigma_{j-1} \cdot \sqrt{C_{i,j-1}} \cdot \widetilde{\varepsilon}_{i,j}, \tag{4.10}$$

where $\varepsilon_{i,j}$, $\widetilde{\varepsilon}_{i,j}$ are independent and identically distributed. Observe that, given $C_{i,j-1}$ $Z_{i,j} \stackrel{(d)}{=} C_{i,j}$.
This means that we generate new observations $Z_{i,j}$ on the set $\mathcal{D}_I$. These new observations (on the conditional structure) lead to a set of "new" realisations for the estimated claims development factors (since we do not want to overload the notation, we do not introduce a new notation for these resampled development factors $\widehat{f}^I_j$)

$$\begin{aligned} \widehat{f}^I_j &= \frac{\sum_{i=0}^{I-j-1} Z_{i,j+1}}{\sum_{i=0}^{I-j-1} C_{i,j}} \\ &= f_j + \frac{\sigma_j}{S^I_j} \cdot \sum_{i=0}^{I-j-1} \sqrt{C_{i,j}} \cdot \widetilde{\varepsilon}_{i,j+1} \qquad (0 \leq j \leq J-1). \end{aligned} \tag{4.11}$$

Our goal is to study the volatility of these new observations $\widehat{f}_j$ given by (4.11), conditioned on $\mathcal{D}_I$. Unlike the observations $\{C_{i,j};\ i+j \le I\}$ the observations $\{Z_{i,j};\ i+j \le I\}$ and also the new realisations $\widehat{f}_0^I, \ldots, \widehat{f}_{J-1}^I$ are random variables given the upper triangle $\mathcal{D}_I$. Furthermore, the observations $C_{i,j}$ and the random variables $\widetilde{\varepsilon}_{i,j}$ are unconditionally independent. This is exactly Approach 3 in [1] as described above, which leads to a multiplicative structure for the derivation of an estimate for the conditional parameter error.
From (4.11) we see that the new realisations satisfy the following (averaging over possible outcomes):

1) the estimators $\widehat{f}_0^I, \ldots, \widehat{f}_{J-1}^I$ are conditionally independent w.r.t. $\mathcal{D}_I$,

2) $E\left[\widehat{f}_{j-1}^I \middle| \mathcal{D}_I\right] = f_{j-1}$ for $1 \le j \le J$ and

3) $\operatorname{Var}\left(\widehat{f}_{j-1}^I \middle| \mathcal{D}_I\right) = \frac{\sigma_{j-1}^2}{S_{j-1}^I}$ for $1 \le j \le J$.

Henceforth with 1)–3), we estimate the (conditional) volatility of the second term in (4.2) by (see also (4.9))

$$
\begin{aligned}
&C_{i,I-i}^2 \cdot E\left[\left(\prod_{j=I-i}^{J-1} \widehat{f}_j^I - f_{I-i} \cdot \prod_{j=I-i+1}^{J-1}\left(\frac{S_j^I}{S_j^{I+1}} \cdot \widehat{f}_j^I + f_j \cdot \frac{C_{I-j,j}}{S_j^{I+1}}\right)\right)^2 \middle| \mathcal{D}_I\right] \\
&= C_{i,I-i}^2 \cdot \left(\prod_{j=I-i}^{J-1} E\left[\left(\widehat{f}_j^I\right)^2 \middle| \mathcal{D}_I\right]\right. \\
&\qquad + f_{I-i}^2 \cdot \prod_{j=I-i+1}^{J-1} E\left[\left(\frac{S_j^I}{S_j^{I+1}} \cdot \widehat{f}_j^I + f_j \cdot \frac{C_{I-j,j}}{S_j^{I+1}}\right)^2 \middle| \mathcal{D}_I\right] \\
&\quad \left. - 2 \cdot E\left[\widehat{f}_{I-i}^I \cdot f_{I-i} \cdot \prod_{j=I-i+1}^{J-1} \widehat{f}_j^I \cdot \left(\frac{S_j^I}{S_j^{I+1}} \cdot \widehat{f}_j^I + f_j \cdot \frac{C_{I-j,j}}{S_j^{I+1}}\right) \middle| \mathcal{D}_I\right]\right).
\end{aligned}
\tag{4.12}
$$

In other words we average $\left(E\left[\widehat{\mathrm{CDR}}_i(I+1) \middle| \mathcal{D}_I\right]\right)^2$, given the information $\mathcal{D}_I$, to obtain an estimate for the (conditional) estimation error (i.e. the second term in (4.2)). Setting

$$
\alpha_j = \frac{S_j^I}{S_j^{I+1}} \in [0, 1], \tag{4.13}
$$

and using

$$f_j \cdot \frac{C_{I-j,j}}{S_j^{I+1}} = \frac{S_j^{I+1} \cdot f_j - S_j^I \cdot f_j}{S_j^{I+1}}, \tag{4.14}$$

(4.12) multiplied by $C_{i,I-i}^{-2}$ becomes

$$\prod_{j=I-i}^{J-1} \left[\text{Var}\left(\widehat{f}_j^I \middle| \mathcal{D}_I \right) + f_j^2 \right] + f_{I-i}^2 \cdot \prod_{j=I-i+1}^{J-1} \left[\alpha_j^2 \cdot \text{Var}\left(\widehat{f}_j^I \middle| \mathcal{D}_I \right) + f_j^2 \right]$$
$$- 2 \cdot f_{I-i}^2 \cdot \prod_{j=I-i+1}^{J-1} \left[\alpha_j \cdot \text{Var}\left(\widehat{f}_j^I \middle| \mathcal{D}_I \right) + f_j^2 \right]. \tag{4.15}$$

This last expression is equal to

$$\prod_{j=I-i}^{J-1} f_j^2 \cdot \left\{ \prod_{j=I-i}^{J-1} \left[\frac{\sigma_j^2/f_j^2}{S_j^I} + 1 \right] + \prod_{j=I-i+1}^{J-1} \left[\alpha_j^2 \cdot \frac{\sigma_j^2/f_j^2}{S_j^I} + 1 \right] \right.$$
$$\left. - 2 \cdot \prod_{j=I-i+1}^{J-1} \left[\alpha_j \cdot \frac{\sigma_j^2/f_j^2}{S_j^I} + 1 \right] \right\}. \tag{4.16}$$

(4.16) gives the estimator for the estimation error which is in the spirit of Buchwalder et al. [1]. This expression is simplified by the following linear approximation: For x_j small we have

$$\prod_j (x_j + 1) \approx 1 + \sum_j x_j. \tag{4.17}$$

Hence the expression in (4.16) is approximated by

$$\prod_{j=I-i}^{J-1} f_j^2 \cdot \left\{ \frac{\sigma_{I-i}^2/f_{I-i}^2}{S_{I-i}^I} + \sum_{j=I-i+1}^{J-1} (1-\alpha_j)^2 \cdot \frac{\sigma_j^2/f_j^2}{S_j^I} \right\}$$
$$= \prod_{j=I-i}^{J-1} f_j^2 \cdot \left\{ \frac{\sigma_{I-i}^2/f_{I-i}^2}{S_{I-i}^I} + \sum_{j=I-i+1}^{J-1} \left(\frac{C_{I-j,j}}{S_j^{I+1}} \right)^2 \cdot \frac{\sigma_j^2/f_j^2}{S_j^I} \right\}. \tag{4.18}$$

This leads to the following estimator for the conditional estimation error:

Result 4.4 (Conditional estimation error estimator for a single accident year) *We have the following estimator for the conditional estimation error of the expected claims development result for a single accident year $i \in \{1, \ldots, I\}$ in accounting year $(I, I+1]$*

$$\widehat{E}\left[\left(E\left[\widehat{CDR}_i(I+1) \mid \mathcal{D}_I\right]\right)^2 \mid \mathcal{D}_I\right] = \left(\widehat{C}_{i,J}^I\right)^2 \cdot \widehat{\Delta}_{i,J}^I \tag{4.19}$$

with

$$\widehat{\Delta}_{i,J}^I = \frac{\left(\widehat{\sigma}_{I-i}^I\right)^2 / \left(\widehat{f}_{I-i}^I\right)^2}{S_{I-i}^I} + \sum_{j=I-i+1}^{J-1} \left(\frac{C_{I-j,j}}{S_j^{I+1}}\right)^2 \cdot \frac{\left(\widehat{\sigma}_j^I\right)^2 / \left(\widehat{f}_j^I\right)^2}{S_j^I}. \tag{4.20}$$

$\widehat{f}_j^I$ and $\left(\widehat{\sigma}_j^I\right)^2$ were defined in (2.4) and (4.5), respectively.

4.2 *Aggregated accident years*

In this section we derive estimators for the uncertainties in the claims reserves estimators and the claims development result for aggregated accident years. Consider $i \neq k$. Using the independence of cumulative payments in different accident years and Corollary 3.2 we obtain

$$\begin{aligned} &\mathrm{MSEP}\left(E\left[\widehat{\mathrm{CDR}}_i(I+1) + \widehat{\mathrm{CDR}}_k(I+1)\middle|\mathcal{D}_I\right]\right) \\ &\quad = \mathrm{Var}\left(\mathrm{CDR}_i(I+1)\middle|\mathcal{D}_I\right) + \mathrm{Var}\left(\mathrm{CDR}_k(I+1)\middle|\mathcal{D}_I\right) \\ &\qquad + \left(E\left[\widehat{\mathrm{CDR}}_i(I+1)\middle|\mathcal{D}_I\right] + E\left[\widehat{\mathrm{CDR}}_k(I+1)\middle|\mathcal{D}_I\right]\right)^2. \end{aligned} \tag{4.21}$$

This means that we can easily calculate the conditional process variance for aggregated accident years. The calculation of the conditional estimation error is more sophisticated since we use the same observations for both accident years to estimate the ultimate claims $C_{i,J}$ and $C_{k,J}$. For the last term in the equality above we obtain

$$\left(E\left[\widehat{\mathrm{CDR}}_i(I+1)\middle|\mathcal{D}_I\right] + E\left[\widehat{\mathrm{CDR}}_k(I+1)\middle|\mathcal{D}_I\right]\right)^2 \tag{4.22}$$

$$\begin{aligned} &= \left(E\left[\widehat{\mathrm{CDR}}_i(I+1)\middle|\mathcal{D}_I\right]\right)^2 + \left(E\left[\widehat{\mathrm{CDR}}_k(I+1)\middle|\mathcal{D}_I\right]\right)^2 \\ &\quad + 2 \cdot E\left[\widehat{\mathrm{CDR}}_i(I+1)\middle|\mathcal{D}_I\right] \cdot E\left[\widehat{\mathrm{CDR}}_k(I+1)\middle|\mathcal{D}_I\right]. \end{aligned} \tag{4.23}$$

Henceforth, in addition to the terms obtained for the conditional estimation error for a single accident year, we obtain a covariance term between different accident years, which is reflected by the cross-product in the equality above. We have

$$
\begin{aligned}
& E\left[\widehat{\mathrm{CDR}}_i(I+1)\middle|\mathcal{D}_I\right] \cdot E\left[\widehat{\mathrm{CDR}}_k(I+1)\middle|\mathcal{D}_I\right] \\
&= C_{i,I-i} \cdot \left(\prod_{j=I-i}^{J-1} \widehat{f}_j^I - f_{I-i} \cdot \prod_{j=I-i+1}^{J-1} \left(\frac{S_j^I}{S_j^{I+1}} \cdot \widehat{f}_j^I + f_j \cdot \frac{C_{I-j,j}}{S_j^{I+1}}\right)\right) \\
&\quad \cdot C_{k,I-k} \cdot \left(\prod_{j=I-k}^{J-1} \widehat{f}_j^I - f_{I-k} \cdot \prod_{j=I-k+1}^{J-1} \left(\frac{S_j^I}{S_j^{I+1}} \cdot \widehat{f}_j^I + f_j \cdot \frac{C_{I-j,j}}{S_j^{I+1}}\right)\right).
\end{aligned}
\tag{4.24}
$$

Assume $i < k$, hence $I - i > I - k$. If we resample the observed chain ladder factors $\widehat{f}_j^I$ analogously to the construction in Section 4.1.2, we obtain resampled values which are conditionally independent, given $\mathcal{D}_I$. Hence, for these resampled values $\widehat{f}_j$, given $\mathcal{D}_I$, we have

$$
\begin{aligned}
& E\Bigg[\left(\prod_{j=I-i}^{J-1} \widehat{f}_j^I - f_{I-i} \cdot \prod_{j=I-i+1}^{J-1} \left(\frac{S_j^I}{S_j^{I+1}} \cdot \widehat{f}_j^I + f_j \cdot \frac{C_{I-j,j}}{S_j^{I+1}}\right)\right) \\
&\quad \cdot \left(\prod_{j=I-k}^{J-1} \widehat{f}_j^I - f_{I-k} \cdot \prod_{j=I-k+1}^{J-1} \left(\frac{S_j^I}{S_j^{I+1}} \cdot \widehat{f}_j^I + f_j \cdot \frac{C_{I-j,j}}{S_j^{I+1}}\right)\right) \Bigg| \mathcal{D}_I\Bigg] \\
&= \Bigg(\prod_{j=I-i}^{J-1} E\left[\left(\widehat{f}_j^I\right)^2 \middle| \mathcal{D}_I\right] \\
&\quad + f_{I-i}^2 \cdot \prod_{j=I-i+1}^{J-1} E\left[\left(\frac{S_j^I}{S_j^{I+1}} \cdot \widehat{f}_j^I + f_j \cdot \frac{C_{I-j,j}}{S_j^{I+1}}\right)^2 \middle| \mathcal{D}_I\right] \\
&\quad - 2 \cdot f_{I-i}^2 \cdot \prod_{j=I-i+1}^{J-1} E\left[\widehat{f}_j^I \cdot \left(\frac{S_j^I}{S_j^{I+1}} \cdot \widehat{f}_j^I + f_j \cdot \frac{C_{I-j,j}}{S_j^{I+1}}\right) \middle| \mathcal{D}_I\right]\Bigg) \cdot \prod_{j=I-k}^{I-i-1} f_j .
\end{aligned}
\tag{4.25}
$$

Observe that the first two terms on the right-hand side of (4.25) are exactly the same as the first two terms on the right-hand side of (4.12). The expression in

(4.25) is linearly approximated (as in Section 4.1.2) by

$$\prod_{j=I-k}^{I-i-1} f_j \cdot \prod_{j=I-i}^{J-1} f_j^2 \cdot \left\{ \frac{\sigma_{I-i}^2 / f_{I-i}^2}{S_{I-i}^I} + \sum_{j=I-i+1}^{J-1} \left(\frac{C_{I-j,j}}{S_j^{I+1}} \right)^2 \cdot \frac{\sigma_j^2 / f_j^2}{S_j^I} \right\}. \quad (4.26)$$

This leads to the following estimator for the covariance term in the conditional estimation error. For $i < k$ we set

$$\widehat{E}\left[E\left[\widehat{\mathrm{CDR}}_i(I+1) \middle| \mathcal{D}_I \right] \cdot E\left[\widehat{\mathrm{CDR}}_k(I+1) \middle| \mathcal{D}_I \right] \middle| \mathcal{D}_I \right] = \widehat{C}_{i,J}^I \cdot \widehat{C}_{k,J}^I \cdot \widehat{\Delta}_{i,J}^I, \quad (4.27)$$

where $\widehat{\Delta}_{i,J}^I$ was defined in (4.20). This leads to the following estimator for the conditional mean square error of prediction for aggregated accident years:

Result 4.5 (Conditional MSEP of the expected claims development result) *We estimate the conditional MSEP for the expected claims development result in accounting year* $(I, I+1]$ *by*

$$\widehat{MSEP}\left(E\left[\sum_{i=1}^{I} \widehat{CDR}_i(I+1) \middle| \mathcal{D}_I \right] \right) = \sum_{i=1}^{I} \widehat{MSEP}\left(E\left[\widehat{CDR}_i(I+1) \middle| \mathcal{D}_I \right] \right) + 2 \cdot \sum_{i<k} \widehat{C}_{i,J}^I \cdot \widehat{C}_{k,J}^I \cdot \widehat{\Delta}_{i,J}^I, \quad (4.28)$$

with conditional MSEP of the expected claims development result of a single accident year $i \in \{1, \ldots, I\}$ *in accounting year* $(I, I+1]$ *given by*

$$\widehat{MSEP}\left(E\left[\widehat{CDR}_i(I+1) \middle| \mathcal{D}_I \right] \right) = \widehat{Var}\left(CDR_i(I+1) \mid \mathcal{D}_I \right) + \left(\widehat{C}_{i,J}^I \right)^2 \cdot \widehat{\Delta}_{i,J}^I, \quad (4.29)$$

see also (4.4) *and* (4.19)-(4.20).

5 Example

For the example we use the Taylor-Ashe [7] data, which was also used by Verrall [9], [10] and Mack [3] (cf. Table 1 in [3]).

The estimators $\widehat{f}_j^I$ for the chain ladder factors f_j show (cf. Table 2) that, on the one hand, a large amount is paid within the first three development years, but,

AY	development period j									
i	0	1	2	3	4	5	6	7	8	9
0	357848	1124788	1735330	2218270	2745596	3319994	3466336	3606286	3833515	3901463
1	352118	1236139	2170033	3353322	3799067	4120063	4647867	4914039	5339085	
2	290507	1292306	2218525	3235179	3985995	4132918	4628910	4909315		
3	310608	1418858	2195047	3757447	4029929	4381982	4588268			
4	443160	1136350	2128333	2897821	3402672	3873311				
5	396132	1333217	2180715	2985752	3691712					
6	440832	1288463	2419861	3483130						
7	359480	1421128	2864498							
8	376686	1363294								
9	344014									

Table 1: Run-off-triangle (cumulative payments)

	0	1	2	3	4	5	6	7	8
$\widehat{f}_j^I$	3.49061	1.74733	1.45741	1.17385	1.10382	1.08627	1.05387	1.07656	1.01772
$(\widehat{\sigma}_j^I)^2$	160280.33	37736.86	41965.21	15182.90	13731.32	8185.77	446.62	1147.37	446.62

Table 2: Estimators $\widehat{f}_j^I$ and $(\widehat{\sigma}_j^I)^2$ for the parameters f_j and σ_j^2, respectively.

on the other hand, the data is also longtailed, since we still observe substantial payments in late development periods. For the estimation of $(\sigma_8^I)^2$ we use the formula given in [3], before Theorem 3.

Hence, using the estimates $\widehat{f}_j^I$ and $(\widehat{\sigma}_j^I)^2$ we find estimators $\sum_{i=1}^I \widehat{R}_i^{\mathcal{D}_I}$ for the aggregated claims liabilities $\sum_{i=1}^I R_i^I$ at time $t = I$ and the corresponding conditional errors for the expected development result in accounting year $(I, I+1]$ (see Table 3).

In Table 3 the process standard deviation, $\sqrt{\text{estimation error}}$ and standard error of prediction for the expected development result in accounting year $(I, I+1]$ and the aggregated ultimate loss over all accident years, respectively, are given. The coefficient of variation Vco is always measured relative to the estimated outstanding liabilities (reserves). Moreover, the estimates for the prediction errors for the aggregated ultimate loss are provided (they are taken from [1]).

	$\sum_i \widehat{R}_i^{\mathcal{D}_I}$	process std.dev.	Vco	$\sqrt{\text{estim. error}}$	Vco	$\sqrt{\text{MSEP}}$	Vco
Dev. Result	18'680'856	1'335'912	7.15%	1'064'436	5.70%	1'708'123	9.14%
Ultimate	18'680'856	1'878'292	10.05%	1'569'349	8.40%	2'447'618	13.10%

Table 3: Process standard deviation, $\sqrt{\text{estimation error}}$ and standard error of prediction for the expected development result and the aggregated ultimate loss.

We observe that the coefficient of variation of the estimated expected claims development result $E\left[\widehat{\text{CDR}}_i(I+1)\middle|\mathcal{D}_I\right]$ within the time interval $(I, I+1]$ is 9.14%. This means that the uncertainty of the claims development result relative to the total reserves is about 9% (in the chain ladder method). The total uncertainty of the claims reserves (claims development until ultimate) is about 13%. Hence we see that the first claims development period has about the same uncertainty as the sum of all claims development periods after time $I+1$ (note that $\sqrt{2}\cdot 9\% \approx 13\%$).
Moreover, we obtain a split of the uncertainty into process variance and estimation error. This is one of the crucial decompositions in the new solvency guidelines, i.e. one needs to quantify which part of the uncertainty comes from the randomness of our stochastic processes and which part comes from estimation errors. Estimation errors can often be understood as an answer to the question "how good can an actuary predict the true parameters in his model" (if we believe in a certain model). For our example the coefficient of the estimation error (in this chain ladder model) is about 5.7%. If we compare this value to the default values for parameter risk given in the SST ([6], Section 8.4.6), we see that our numerical value is within a reasonable range.

5.1 Conclusion and Outlook

We have studied the uncertainties of the expected claims development result in the chain ladder model. For this expected claims development result we have found two estimators for the natural decomposition of the prediction error into (conditional) process error and (conditional) estimation error.

For future research our developments raise two interesting questions:

1) We have derived estimators for expected claims development results (3.8). In a next step, we would like to obtain similar results for the claims development results (3.7).

2) Observe that the coefficient of variation of the estimation error goes to zero the more observations we have (see (4.20)). But we all know that in practice this is not the case, since we try to predict future cashflows with the help of past information. Since things may always change in the future (e.g. change in jurisdiction), we would like to have a model, where the coefficient of variation of the parameter error is bounded below by some positive value.

Bibliography

[1] Buchwalder, M., Bühlmann H., Merz, M., Wüthrich, M.V. (2006). The mean square error of prediction in the chain ladder reserving method (Mack and Murphy revisited). Astin Bulletin 36, 2, 521–542.

[2] Gisler, A. (2006). The estimation error in the chain-ladder reserving method: a Bayesian approach. Astin Bulletin 36, 2, 554–565.

[3] Mack, T. (1993). Distribution-free calculation of the standard error of chain ladder reserve estimates. Astin Bulletin 23, 2, 213–225.

[4] Mack, T., Quarg, G., Braun, C. (2006). The mean square error of prediction in the chain ladder reserving method - a comment. Astin Bulletin 36, 2, 543–552.

[5] Murphy, D.M. (1994). Unbiased loss development factors. Proc. CAS Vol. LXXXI, 154–222.

[6] Swiss Solvency Test (2006). Technisches Dokument zum Swiss Solvency Test, Bundesamt für Privatversicherungen, Version 02. Oktober 2006. Available under: `http://www.bpv.admin.ch/themen/00506/00552/index.html?lang=de`

[7] Taylor, G.C., Ashe, F.R. (1983). Second moments of estimates of outstanding claims. J. Econometrics 23, 37–61.

[8] Venter, G. (2006). Discussion of mean square error of prediction in the chain ladder reserving method. Astin Bulletin 36, 2, 566–571.

[9] Verrall, R.J. (1990). Bayes and empirical Bayes estimation for the chain ladder model. Astin Bulletin 20, 217–243.

[10] Verrall, R.J. (1991). On the estimation of reserves from loglinear models. Insurance: Math. Econom. 10, 75–80.

Michael Merz
University Tübingen
Faculty of Economics
D-72074 Tübingen
michael.merz@uni-tuebingen.de

Mario V. Wüthrich
ETH Zürich
Department of Mathematics
CH-8092 Zürich
wueth@math.ethz.ch

Abstract

Using the chain ladder method we estimate the total ultimate claim amounts at time I and time $I+1$ (successive best estimate predictions for the ultimate claims when updating the information from time I to time $I+1$). The claims development result is then defined to be the difference of these two best estimators. We analyze the volatility of this updating procedure for both the process variance and the estimation error in the parameters of the chain ladder model.

Zusammenfassung

Unter Verwendung der Chain Ladder-Methode werden die Schadenaufwände zum Zeitpunkt I und zum Zeitpunkt $I+1$ geschätzt (sukzessive Best Estimate-Schätzungen für die Schadenaufwände unter Verwendung der Information zum Zeitpunkt I resp. $I+1$). Das Abwicklungsergebnis ist definiert als die Differenz dieser beiden Schätzungen. Wir untersuchen die Volatilität dieser Aktualisierungsprozedur, sowohl für die Prozessvarianz als auch für den Schätzfehler mit Hilfe der Parameter des Chain Ladder-Modells.

Résumé

Nous estimons le montant total des sinistres (ultimates) au temps I et au temps $I+1$ en utilisant la méthode «chain-ladder» (estimations de type «best estimate» successives des sinistres en actualisant l'information du temps I au temps $I+1$). Le résultat du développement des sinistres est défini par la différence de ces deux estimations de type «best-estimate». Nous analysons la volatilité de ce procédé d'évaluation pour la variance du processus ainsi que pour l'erreur d'estimation des paramètres du modèle «chain ladder».