Stochastic Claims Reserving in General Insurance: Models and Methodoledgies

Jinlong HUANG and Xianyi WU*
Department of Statistics and Actuarial Science
East China Normal University

June 10, 2012

Abstract

In this article we review the main approaches extensively investigated for claims reserving in general insurance. We first introduce the models underlying the most-known Chain Ladder method and Bornhuetter-Ferguson method. Then we discuss their Bayesian versions, Generalized limear models for claims reserving and the bootstrap approaches to evaluating the variabilty of predicted/estimated reserves are reviewed also. In addition, we conclude the paper by introducing the multivariate version for claims reserving methods.

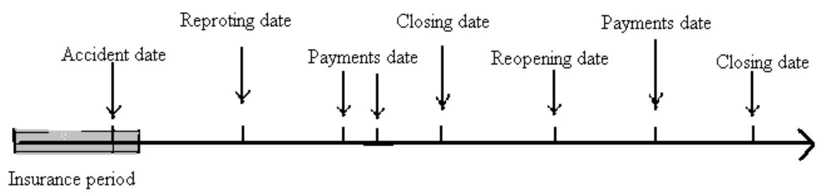
1 Introduction

1.1 General insurance and claims reserving

Insurance industries can be put in two categories: life insurance and non-life insurance. There are rather different characteristics between Non-life and life insurance products, such as terms of contracts, type of claims, risk drivers, etc., and in many countries there is a strict legal separation between these two insurance products such that any company dealing with one type of insurance products is not allowed to operate that other type of insurance products. While the term non-life insurance is known in continental Europe, it is known as General Insurance in UK and Property and Casualty Insurance in USA. A life company develops and sells such insurance products motor/car insurance (e.g., motor third party liability and motor hull), property insurance (e.g., property against fire, flooding, business interruption, etc.), liability insurance (e.g., director and officers (D&O) liability insurance) accident

*Supported by Shanghai Philosiphy and Social Science foundation under grant No. 2010BJB004.

Figure 1: Claim process.



insurance (personal and collective accident, including compulsory accident insurance and workers' compensation), health insurance, marine insurance, etc. The history of a typical non-life insurance claim can be illustrated in the following Figure 1, in which an accident occurred at a time between the insurance period and then reported some time later at the reporting date with the claim being settled at a further later time point indicated by closing date. For certain cases, due to certain reasons, e.g., new evidence is discovered for a liability insurance such that the account of the closed claim has to be reopen to handle newly raised claims.

We are here concentrated in the claims in terms of money but set aside the details in insurance operation and accounting, e.g. inflation effect that has been observed in the literature. At the evaluation date, usually the end of an accounting year, the insurer need to make down the figure of the money to be kept to cover the future payments for claims. This total amount of money is referred to as claims reserve (or loss reserve). As shown in the above figure, we will distinct incurred but not reported and reported but not settled (RBNS) claims; the meanings of which are clear from the figure. And the future payments include two part: one for the IBNR claims and one for RBNS claims.

Classically, a majority of actuaries make their loss reserving based on the data sets that are known as development triangles or aggregate data. The data sets are mostly in the form

$$\mathcal{D} = \{X_{ij}, i = 0, 1, \ldots, I, j = 0, 1, \ldots (I - i) \wedge J\}, \tag{1.1}$$

where i indicates accident years, j the development years, X_{ij} the claims amount caused by the accidents occurred in accident year i and is generally referred to as development year/period, and I the number of years in which partial or the entire claims history are observed and J is the maximum development years. Note that the calendar evaluation year is I and $I \geq J$ is generally required. The unknown X_{ij} for $j \in (I - i, J]$ is referred to as outstanding loss liabilities for $i > I - J$.

For all $i \in \{0, \cdots, I\}$, $j \in \{0, \cdots, J\}$, let $C_{i,j}$ denote the cumulative payments for accident year i of the first $j+1$ development years, i.e., $C_{ij} = \sum_{k=0}^{j} X_{ik}$.Therefore the outstanding loss liability for accident year i is

$$R_i = C_{i,J} - C_{i,I-i} = \sum_{k=I-i+1}^{J} X_{i,k}. \tag{1.2}$$

Classically, the data set $\mathcal{D}$ is taken as a set of deterministic numbers and the objective is to predict $\sum_{i=0}^{I} R_i$ or equivalently $\sum_{i=0}^{I} C_{iJ}$. In the later of last century, researchers started to model X_{ij} by random variables and hence changed the objective to estimating the mean of $\sum_{i=0}^{I} C_{iJ}$. Actually, we can generally consider the objective of claims reserving as estimating the distribution or certain functionals of the random claims $\sum_{i=0}^{I} C_{iJ}$. based on the observed data $\mathcal{D}$.

Moreover, denote the set of observations at time I for the first jth development years by

$$\mathcal{D}_j = \{X_{i,k}; i + k \leq I, 0 \leq k \leq j\}, \quad j = 0, 1, \cdots, J \tag{1.3}$$

then $\mathcal{D}_I = \mathcal{D}_J$ is all the available observations at time I. Finally, we let

$$\mathcal{F}_{i,j} = \sigma(C_{i,0}, \ldots, C_{i,j}) \quad j = 0, 1, \cdots, J-1 \tag{1.4}$$

be the filtration generated by the loss claims history of accident year i and development year j.

The existing claims reserving methods for data sets of form (1.1) include the most popular chain ladder (CL) and Bornhuetter-Ferguson (BF) methods. The stochastic methods developed after the later of last century were mainly concerned with the construction of as general as possible models that justify CL and BF methods. In all these models, the following is the fundamental assumption and is respected by all stochastic models.

Model Assumption 1.1 *1. $\{C_{i,j}, j = 0, 1, \ldots, J\}$ are mutually independent over $i = 0, 1, \ldots, I$.*

2 Stochastic Models for Chain Ladder Method

2.1 Distribution-free Chain-Ladder model

The chain ladder method may be the most popular technique for reserving because of its simplicity and distribution-free assumption. The following chain ladder model assumptions are given Mack (1993) [16].

Model Assumption 2.1 *1. There exist a set of development factors $f_0, f_1, \cdots, f_{J-1} > 0$, referred to as development factors, such that for all $0 \leq i \leq I$ and all $1 \leq j \leq J$ we have*

$$E[C_{i,j}|\mathcal{F}_{i,j-1}] = E[C_{i,j}|C_{i,j-1}] = f_{j-1}C_{i,j-1}.$$

2. There exist a set of positive numbers $\sigma_0^2, \cdots, \sigma_{J-1}^2 \geq 0$*, such that for all* $0 \leq i \leq I$ *and all* $1 \leq j \leq J$ *we have*

$$Var(C_{i,j}|\mathcal{F}_{i,j-1}) = Var(C_{i,j}|C_{i,j-1}) = \sigma_{j-1}^2 C_{i,j-1}. \tag{2.1}$$

2.1.1 Feature of the model

Under this basic assumption, it can be easily seen that

$$E[C_{i,J}|\mathcal{D}_I] = C_{i,I-i} f_{I-i} \cdots f_{J-1} \text{ and } \mathrm{Var}(C_{iJ}|\mathcal{D}_I) = C_{i,I-i} \sum_{j=I-i}^{J-1} \prod_{m=I-i}^{j-1} f_m \sigma_j^2 \prod_{n=j+1}^{J-1} f_n^2, \tag{2.2}$$

where the second equality can be verified by first noting that $\mathrm{Var}(C_{iJ}|\mathcal{D}_I) = \mathrm{Var}(C_{iJ}|\mathcal{F}_{i,I-i})$ and thus

$$\begin{aligned} \mathrm{Var}(C_{iJ}|\mathcal{D}_I) &= \mathrm{E}[\mathrm{Var}(C_{iJ}|\mathcal{F}_{i,J-1})|\mathcal{F}_{i,I-i}] + \mathrm{Var}[\mathrm{E}(C_{iJ}|\mathcal{F}_{i,J-1})|\mathcal{F}_{i,I-i}] \\ &= \sigma_{J-1}^2 \mathrm{E}[C_{i,J-1}|\mathcal{F}_{i,I-i}] + f_{J-1}^2 \mathrm{Var}(C_{iJ}|\mathcal{F}_{i,J-1}) \\ &= \sigma_{J-1}^2 C_{i,I-i} \prod_{m=I-i}^{J-2} f_m + f_{J-1}^2 \mathrm{Var}(C_{i,J-1}|\mathcal{F}_{i,I-i}), \end{aligned}$$

and then recursively applying this equality.

2.1.2 Estimation

For every $j = 1, 2, \ldots, J$, denote $\mathcal{F}_{j-1} = \bigvee_{i=0}^{I} \mathcal{F}_{i,(j-1)\wedge(I-i)}$. The effective data to estimate f_{j-1} are $C_{i,j-1}, C_{i,j}, i = 0, 2, \ldots, J-j$, satisfying a linear model

$$\begin{pmatrix} C_{0,j} \\ \vdots \\ C_{J-j,j} \end{pmatrix} = \begin{pmatrix} C_{0,j-1} \\ \vdots \\ C_{J-j,j-1} \end{pmatrix} f_j + \varepsilon,$$

where $\varepsilon = \begin{pmatrix} \varepsilon_0 & \ldots & \varepsilon_{J-j} \end{pmatrix}^T$ satisfying $\mathrm{E}(\varepsilon|\mathcal{F}_{j-1}) = 0$ and $\mathrm{Var}(\varepsilon|\mathcal{F}_{j-1}) = diag(C_{0,j-1}, \cdots, C_{J-j,j-1})$. Thus the conditional BLUE (best linear unbiased estimator) of f_{j-1} given $\mathcal{F}_{j-1}$ are

$$\widehat{f}_{j-1} = \frac{\sum_{i=0}^{I-j} C_{i,j}}{\sum_{i=0}^{I-j} C_{i,j-1}}. \tag{2.3}$$

Meanwhile, σ_j^2 can be estimated by

$$\widehat{\sigma_j}^2 = \frac{1}{I-j-1} \sum_{i=0}^{I-j-1} C_{i,j} (\frac{C_{i,j+1}}{C_{i,j}} - \widehat{f}_j)^2. \tag{2.4}$$

It is also not difficult to examine that $E\left[\widehat{f}_{I-i}\cdots\widehat{f}_{J-1}|\mathcal{F}_{i,I-i}\right] = f_{I-i}\cdots f_{J-1}$. Then an unbiased estimator of $E[C_{i,J}|\mathcal{D}_I]$ is given by

$$\widehat{C}_{i,J}^{CL} = \widehat{E}[C_{i,J}|\mathcal{D}_I] = C_{i,I-i}\widehat{f}_{I-i}\cdots\widehat{f}_{J-1} \tag{2.5}$$

Here we note that the unbiasedness does not depends on the second item of Assumption 2.1.

Remark 2.1 *From formula (2.5), it is very apparent that the estimator $\widehat{C}_{i,J}^{CL}$ might not be robust with respect to the outlier of $C_{i,I-i}$. A robustification of the chain-ladder method can be found in a recent work by Verdonck, Van Wouwe and Dhaene (2009).*

2.1.3 Estimation of the MSEP

Since the claims in different accident years are independent, it holds that

$$\mathrm{Var}\Big(\sum_{i=1}^{I} C_{iJ}|\mathcal{D}_I\Big) = \sum_{i=1}^{I}\mathrm{Var}(C_{ij}|\mathcal{D}_I)$$

The mean squared error of the prediction $C_{i,J}^{CL}$ (MSEP) is given by

$$MSEP^{CL} = \mathrm{E}\left[(\widehat{C}_{i,J}^{CL} - E[C_{i,J}|\mathcal{D}_I])^2\Big|\,\mathcal{F}_{j-1}\right] = C_{i,I-i}^2\mathrm{E}\left[(\widehat{f}_{I-i}\cdots\widehat{f}_{J-1} - f_{I-i}\cdots f_{J-1})^2\Big|\,\mathcal{F}_{j-1}\right]. \tag{2.6}$$

The following discusses the estimation of $MSEP^{CL}$

Mack's method.

Mack (1993)[16] gave the following approach for estimating the parameter estimation error. Introduce for $j \in \{I-i,\cdots,J-1\}$,

$$T_j = \widehat{f}_{I-i}\cdots\widehat{f}_{j-1}(f_j - \widehat{f}_j)f_{j+1}\cdots f_{J-1}.$$

This implies that

$$(\widehat{C}_{i,j}^{CL} - E[C_{i,j}|\mathcal{D}_I])^2 = C_{i,I-i}^2\Big(\sum_{j=I-i}^{J-1} T_j^2 + 2\sum_{I-i\le j<k\le J-1} T_jT_k\Big).$$

Note that $E[T_k|\mathcal{D}_k] = 0$ and that T_j is $\mathcal{D}_k$-measurable for $j<k$. Moreover, under the assumption on the variance of C_{ij}, see equation (2.1), We see that

$$E[T_jT_k|\mathcal{D}_k] = \begin{cases} 0, & \text{if } j<k \\ \widehat{f}_{I-i}^2\cdots\widehat{f}_{j-1}^2\frac{\sigma_j^2}{S_j^{[I-j-1]}}f_{j+1}^2\cdots f_{J-1}^2, & \text{if } j=k \end{cases} \tag{2.7}$$

where $S_j^{[k]} = \sum_{i=0}^{k} C_{ij}$. Hence we further have

$$MSEP^{CL} = \mathrm{E}\left[C_{i,I-i}^2 \sum_{j=I-i}^{J-1} \widehat{f}_{I-i}^2 \cdots \widehat{f}_{j-1}^2 \frac{\sigma_j^2}{S_j^{[I-j-1]}} f_{j+1}^2 \cdots f_{J-1}^2\right] \tag{2.8}$$

Replace the unknown parameters f_j by $\widehat{f}_j$ given in (2.3) and σ_j^2 by $\widehat{\sigma_j}^2$ given in (2.4) in (2.2) and (2.8), we get the following estimator for msep$(\widehat{C_{iJ}}^{CL}|\mathcal{D}_I)$,

$$\widehat{\mathrm{msep}}_{C_{iJ}|\mathcal{D}_I}\big(\widehat{C_{iJ}}^{CL}\big) = \big(\widehat{C_{iJ}}^{CL}\big)^2 \sum_{j=I-i}^{J-1} \frac{\widehat{\sigma}_j^2}{\widehat{f}_j^2}\Big(\frac{1}{\widehat{C_{iJ}}^{CL}} + \frac{1}{S_j^{[I-j-1]}}\Big). \tag{2.9}$$

Time Series and Enhanced Time Series Model.

The Time Series Model has been investigated by several authors, such as [2] and [21]. [6] discussed different approaches for the estimation of MSEP in the Time Series framework. Moreover, [36] proposed the Enhanced Time Series Model to describe the uncertainty whether the deterministic chain ladder factors f_j are still valid for future claims development behavior.

Model Assumption 2.2 (Time Series model) *The cumulative payments $C_{i,j}$ satisfy*

$$C_{i,j} = f_{j-1}C_{i,j-1} + \sigma_{j-1}\sqrt{C_{i,j-1}}\varepsilon_{i,j} \tag{2.10}$$

where conditional on $\mathcal{D}_0$, $\varepsilon_{i,j}$ are mutually independent with $E[\varepsilon_{i,j}|\mathcal{D}_0] = 0$, $E[\varepsilon_{i,j}^2|\mathcal{D}_0] = 1$ and $P[C_{i,j} > 0|\mathcal{D}_0] = 1$.

Define the individual development factor for accident year i and development year j by

$$F_{i,j} = \frac{C_{i,j}}{C_{i,j-1}} \tag{2.11}$$

It is easy to check that under Model Assumption 2.2, equality (2.2) also holds and further we have

$$E[F_{i,j}|C_{i,j-1}] = f_{j-1} \quad \text{and} \quad Var(F_{i,j}|C_{i,j-1}) = \frac{\sigma_{j-1}^2}{C_{i,j-1}}. \tag{2.12}$$

then the BLUE for f_j, conditional on $\mathcal{D}_j$, coincides with (2.3) and therefore (2.5) is also an unbiased estimator of $E[C_{i,J}|\mathcal{D}_I]$ in the Time Series model. Therefore the estimation error are given by

$$(\widehat{C}_{i,j}^{TS} - E[C_{i,j}|\mathcal{D}_I])^2 = C_{i,I-i}^2(\widehat{f}_{I-i}\cdots\widehat{f}_{J-1} - f_{I-i}\cdots f_{J-1})^2 \tag{2.13}$$

which coincides with (2.6).

In order to determine the (conditional) estimation error, we need to determine the volatilities $\widehat{f}_j$ around its true values f_j. [6] measures these volatilities with a conditional resampling method. That is, we estimate (2.13) by

$$C_{i,I-i}^2\Big(\prod_{k=I-i}^{J-1} E\big[\widehat{f}_k^2|\mathcal{D}_k\big] - \prod_{k=I-i}^{J-1} f_k^2\Big). \tag{2.14}$$

We therefore resample the observations $\widehat{f}_{I-i}, \cdots, \widehat{f}_{J-1}$, given $\mathcal{D}_I$. This means that for the determination of an estimator for the (conditional) estimation error we have to take into account that, given $\mathcal{D}_I$, the observations for $\widehat{f}_j$ could have been different form the observed values $\widehat{f}_j$. To regard this source of uncertainty, we degenerate a set of 'new' observation by the formula

$$C_{i,j}^* = f_{j-1}C_{i,j-1} + \sigma_{j-1}\sqrt{C_{i,j-1}}\widetilde{\varepsilon}_{i,j} \tag{2.15}$$

in the upper triangle, where $\varepsilon_{i,j}$ and $\widetilde{\varepsilon}_{i,j}$ are $i.i.d$ random variables. Therefore given $\mathcal{D}_I$, $C_{i,j}^*$ and $C_{i,j}$ have the same distribution.

From (2.3) and (2.15), we get the following representation for the resampled estimates of the development factors

$$\widehat{f}_{j-1}^* = \frac{\sum_{k=0}^{I-j} C_{k,j}^*}{\sum_{k=0}^{I-j} C_{k,j-1}} = f_{j-1} + \frac{\sigma_{j-1}}{S^{j-1}}\sum_{k=0}^{I-j}\sqrt{C_{k,j-1}}\widetilde{\varepsilon}_{i,j} \tag{2.16}$$

further given $\mathcal{D}_j$, $\widehat{f}_j^*$ and $\widehat{f}_j$ have the same distribution.

From (2.16) and the fact that the observations $C_{i,j}$ and $\widetilde{\varepsilon}_{i,j}$ are unconditionally independent, we conclude that:

1) the estimators $f_0^*, \cdots, f_{J-1}^*$ are conditionally independent w.r.t $\mathcal{D}_I$.

2) $E[\widehat{f}_{j-1}^*|\mathcal{D}_I] = f_{j-1}$ and $E[(\widehat{f}_{j-1}^*)^2|\mathcal{D}_I] = f_{j-1}^2 + \frac{\sigma_{j-1}^2}{S^{j-1}}$ for $1 \le j \le J$.

Hence (2.14) is estimated by

$$C_{i,I-i}^2\Big(\prod_{k=I-i}^{J-1}(f_k^2 + \frac{\sigma_k^2}{S^k}) - \prod_{k=I-i}^{J-1} f_k^2\Big) \tag{2.17}$$

Next, replace f_j and σ_j in (**??**) and (2.17) by their estimators, we get an estimator for the conditional MSEP.

According to (2.12), the conditional variational coefficients of $F_{i,j}$ are given by

$$V_{CO}(F_{ij}|C_{i,j-1}) = \frac{\sigma_{j-1}}{f_{j-1}}C_{i,j-1}^{-\frac{1}{2}} \to 0, \quad \text{as} \quad C_{i,j-1}\to\infty$$

which means that the risk completely disappears for very large portfolios. However, this is not the case in practice for there are always some external factors that influences a portfolio and are not diversifiable. To take this risk class into account, we refer to [36].

Model Assumption 2.3 (Enhanced Time Series model) *There exist positive constants $f_{j-1}, \sigma^2_{j-1}, a^2_{j-1}$ and random variables $\varepsilon_{i,j}$ such that for all $i \in \{0, \cdots, I\}$ and $j \in \{1, \cdots, J\}$ we have*

$$C_{i,j} = f_{j-1}C_{i,j-1} + (\sigma^2_{j-1} + a^2_j f^2_{j-1} C_{i,j-1})^{\frac{1}{2}} \sqrt{C_{i,j-1}} \varepsilon_{i,j} \tag{2.18}$$

where $\varepsilon_{i,j}$ has the same assumption as in Model Assumption 2.2. ***(Enhanced CL model or Enhanced Time Series model?)***

Theorem 2.1 *Under Model Assumption 1.1 and 2.3, we have,*
1. the conditional variational coefficients of $F_{i,j}$ is bounded from below by a^2_{j-1}, i.e.

$$V_{CO}(F_{ij}|C_{i,j-1}) \geq \lim_{C_{i,j-1}\to\infty} V_{CO}(F_{ij}|C_{i,j-1}) = a^2_{j-1}.$$

2. the conditional expectation of the ultimate claim for a single accident year coincides with (2.2)
3. the conditional process variance of the ultimate claim for a single accident year i is given by

$$\begin{aligned}\mathrm{Var}(C_{iJ}|\mathcal{D}_I) &= C^2_{i,I-i}\Big[\sum_{j=I-i}^{J-1}\prod_{n=j+1}^{J-1}\Big(1+a^2_n\Big)f^2_j\Big(\frac{\sigma^2_j}{C_{i,I-i}}\prod_{m=I-i}^{j-1} f_m + a^2_j \prod_{m=I-i}^{j} f^2_m\Big)\Big] \\ &= E[C_{ij}|\mathcal{D}_I]^2\Big[\sum_{j=I-i}^{J-1}\Big(\frac{\sigma^2_j/f^2_j}{E[C_{ij}|\mathcal{D}_I]}\Big)\prod_{n=j+1}^{J-1}(1+a^2_n)\Big].\end{aligned}$$

As to the parameter estimation, [36] points out that the sequence a_j, usually can not be estimated from the data unless the portfolio is very large. Hence in general a_j can only be estimated from the whole insurance market. The author proposed an iterative estimation for f_j and σ^2_j, that is

$$\begin{aligned}\widehat{f}_j^{(0)} &= \frac{\sum_{i=0}^{I-j-1} C_{i,j+1}}{\sum_{i=0}^{I-j-1} C_{i,j}} \\ \widehat{\sigma^2_j}^{(k)} &= \frac{1}{I-j-1}\sum_{i=0}^{I-j} C_{i,j}(F_{i,j+1} - \widehat{f}_j^{(0)})^2 - \frac{\widehat{a}^2_j \widehat{f_j^{(k-1)}}^2}{I-j-1}\Big(\sum_{i=0}^{I-j} C_{i,j} - \frac{\sum_{i=0}^{I-j} C^2_{i,j}}{\sum_{i=0}^{I-j} C_{i,j}}\Big) \\ \widehat{f}_j^{(k)} &= \frac{\sum_{i=0}^{I-j-1} C_{i,j+1}/\Delta^{(k)}_{i,j}}{\sum_{k=0}^{I-j-1} C_{kj}/\Delta^{(k)}_{i,j}}, \quad where \Delta^{(k)}_{i,j} = \widehat{\sigma^2_j}^{(k)} + \widehat{a}^2_j \widehat{f_j^{(k-1)}}^2 C_{ij}.\end{aligned}$$

2.2 Bornhuetter-Ferguson related Method

The Bornhuetter-Ferguson method (referred to as BF method below) goes back to Bornhuetter and Ferguson (1972) [4]. .

Model Assumption 2.4 *1. There exist parameters* $\mu_0, \mu_1, \cdots, \mu_I > 0$ *and* $\beta_0, \beta_1, \cdots, \beta_J > 0$ *with* $\beta_J = 1$ *such that for all* $0 \le i \le I$ *and all* $0 \le j \le J$ *we have*

$$E[C_{i0}] = \mu_i \beta_0 \text{ and } E[C_{i,j}|\mathcal{F}_{i,j-1}] = C_{i,j-1} + \mu_i(\beta_j - \beta_{j-1}) \tag{2.19}$$

2. The incremental claims $C_{i,J} - C_{i,I-i}$ *was independent of* $C_{i,0}, \cdots, C_{i,I-i}$.

Under Model Assumption 2.4, it is easy to see that

$$E[C_{i,j}] = \mu_i \beta_j \tag{2.20}$$

for all $0 \le j \le J$ and

$$E[C_{i,J}|\mathcal{D}_I] = C_{i,I-i} + E[C_{iJ} - C_{i,I-i}] = C_{i,I-i} + (1 - \beta_{I-i})\mu_i \tag{2.21}$$

Then a BF estimator for $E[C_{iJ}|\mathcal{D}_I]$ is given by

$$\widehat{C}_{i,J}^{BF} = \widehat{E}[C_{i,J}|\mathcal{D}_I] = C_{i,I-i} + (1 - \widehat{\beta}_{I-i})\mu_i \tag{2.22}$$

where $\widehat{\beta}_{I-i}$ is an appropriate estimator for β_{I-i} and μ_i is a prior estimator for the expected ultimate claims $E[C_{i,J}]$.

Recall that under Model Assumption 2.1, for any j, we have

$$E[C_{i,J}] = E[E[C_{i,J}|C_{i,J-1}]] = f_{J-1}E[C_{i,J-1}] = \cdots = \prod_{k=j}^{J-1} f_k E[C_{i,j}], \tag{2.23}$$

and hence,

$$E[C_{i,j}] = \prod_{k=j}^{J-1} f_k^{-1} E[C_{i,J}]. \tag{2.24}$$

Using (2.20), we find that $\prod_{k=j}^{J-1} f_k^{-1}$ plays the role of β_j. Therefore if we set

$$\widehat{\beta}_j^{(CL)} = \prod_{k=j}^{J-1} \widehat{f}_k^{-1}, \tag{2.25}$$

it can be readily justified that

$$\widehat{C_{i,J}}^{BF} = C_{i,I-i} + (1 - \widehat{\beta}_{I-i})\mu_i \text{ and } \widehat{C_{i,J}}^{CL} = C_{i,I-i} + (1 - \widehat{\beta}_{I-i})\widehat{C_{i,J}}^{CL}. \tag{2.26}$$

This indicates that the BF estimator and the CL estimator differ only in the estimator of the ultimate claim $C_{i,J}$.

3 Bayesian Models

There are two main directions in the framework of Bayesian theory for claim reserving. One is the exact Bayesian method, such as [1], [10],[11], [33] and [34]. The other is the credibility theory, one can refer to [12], [17], [22] and [35].

3.1 Benktander-Hovinen credibility method

This is not a Bayesian method in strict sense, but it has a similar form of credibility theory (see e.g., Bühlmann and Gisler, 2005) in non-life insurance. In this subsection, it is for the time being assumed that μ_i and β_j are known so that $\widehat{C_{i,J}}^{CL} = C_{i,I-i}/\beta_{I-i}$ and $\widehat{C_{i,J}}^{BF} = C_{i,I-i} + (1-\beta_{I-i})\mu_i$. If we let

$$u_i(c) = c\widehat{C_{i,J}}^{CL} + (1-c)\widehat{C_{i,J}}^{BF}$$

for any $c \in [0,1]$, we get a mixture of the CL estimator and BF estimator. This type of estimators for the ultimate reserves is due to Benktander (1976) and Hovinen (1981). Benktander (1976) suggested to take $c = \beta_{I-i}$ such that

$$u_i(\beta_{I-i}) = \beta_{I-i}\widehat{C_{i,J}}^{CL} + (1-\beta_{I-i})\widehat{C_{i,J}}^{BF} = C_{i,I-i} + (1-\beta_{I-i})\widehat{C_{i,J}}^{BF}.$$

It corresponds to an iterated BF estimator using the BF estimate to institute for the prior μ_i in the formula for BF estimator $\widehat{C_{i,J}}^{BF}$.

If we let $\mu_i = \kappa_i \Pi_i$, then we get the Cape-Cod Model (see Bühlmann (1983)), where κ_i reflects the average loss ratio and Π_i can be interpreted as the premium received in accident year i.

Generally, we can choose c to be the minimizer of the (unconditional) MSEP for the reserve estimator $\widehat{R_i}(c)$,

$$\mathrm{msep}_{R_i}(\widehat{R_i}(c)) = E[(R_i - \widehat{R_i}(c))^2]. \tag{3.27}$$

In order to minimize (3.27), we assume that $\mu_i, i = 0, 1, \ldots, I$ in (2.19) are independent of the data $C_{i,J}$ and $E\mu_i = EC_{i,J}$. The following theorem is due to Mack (2000) [17].

Theorem 3.1 *Under Model Assumption 2.1, the optimal credibility factor c_i^*, which minimizes the (unconditional) MSEP* (3.27) *is given by*

$$c_i^* = \frac{\beta_{I-i}}{1-\beta_{I-i}} \frac{\mathrm{Cov}(C_{i,I-i}, R_i) + \beta_{I-i}(1-\beta_{I-i})\mathrm{Var}(U_i)}{\mathrm{Var}(C_{i,I-i}) + \beta_{I-i}^2 \mathrm{Var}(U_i)}$$

3.2 Exact Bayesian Models

The Bayesian method is called exact since the Bayesian estimator $E[C_{iJ}|\mathcal{D}_I]$ is optimal in the sense that it minimizes the squared loss function (MSEP) in the class $L^2_{C_{iJ}}(\mathcal{D}_I)$, in which all estimators for C_{iJ} that are square integrable functions of the observations in $\mathcal{D}_I$, that is

$$E[C_{iJ}|\mathcal{D}_I] = \underset{Y\in L^2_{C_{iJ}}(\mathcal{D}_I)}{\text{argmin}}\ E[(C_{iJ}-Y)^2|\mathcal{D}_I].$$

Model Assumption 3.1 *1. Given Θ_i, the random variables $Y_{ij}=\frac{X_{ij}}{\mu_i\gamma_j}$, $j=0,\cdots,J$ are independent for all $i\in\{0,\cdots,I\}$,where $\gamma_j\geq 0$ and $\sum_{j=0}^J\gamma_j=1$.*

2. The pairs $(\Theta_i,(X_{i0},\cdots,X_{iJ}))$ $(i=0,\cdots,I)$ are independent and $\Theta_0,\cdots,\Theta_I$ are i.i.d.

Model Assumption 3.2 *1. $Y_{ij}=\frac{X_{ij}}{\gamma_j\mu_i}$ belong to the EDF$(\sigma^2,\omega_{ij},\Theta_i)$, i.e., Y_{ij} have density*

$$f(x|\theta_{i,j}) = a\big(x,\frac{\sigma}{\omega_{ij}}\big)\exp\Big\{\frac{x\theta_{i,j}-b(\theta_{i,j})}{\sigma/\omega_{i,j}}\Big\}, \tag{3.28}$$

where $b(\cdot)$ is twice-differentiable w.r.t θ_i, σ and ω_{ij} are some positive real-valued constants.

2. Θ_i with densities (w.r.t the Lebesgue measure)

$$u_{\mu,\tau^2}(\theta) = d(\mu,\tau^2)\exp\Big\{\frac{\mu\theta-b(\theta)}{\tau^2}\Big\}$$

For $i=0,\cdots,I$, the posterior distribution of Θ_i given $Y_{i,0},\cdots,Y_{i,I-i}$ is proportional to

$$\exp\Big\{\theta\Big[\frac{1}{\tau^2}+\sum_{j=0}^{I-i}\frac{\omega_{ij}}{\sigma^2}Y_{ij}\Big]-b(\theta)\Big[\frac{1}{\tau^2}+\sum_{j=0}^{I-i}\frac{\omega_{ij}}{\sigma^2}\Big]\Big\}$$

If $\exp\,(\mu_i\theta-b(\theta))/\tau^2$ disappears on the boundary of Θ_i for all μ_i,τ, then we have

$$\widetilde{\mu(\Theta_i)} \overset{\text{def}}{=} E[\mu(\Theta_i)|Y_{i,0},\cdots,Y_{i,I-i}] = \alpha_i\overline{Y}_i+(1-\alpha_i)\mu$$

where $\mu(\Theta_i)=E[Y_{ij}|\Theta_i]$ and

$$\alpha_i=\frac{\sum_{j=0}^{I-i}\omega_{ij}}{\sum_{j=0}^{I-i}\omega_{ij}+\sigma^2/\tau^2},\quad \overline{Y}_i=\frac{\sum_{j=0}^{I-i}\omega_{ij}Y_{ij}}{\sum_{k=0}^{I-i}\omega_{ik}}$$

Estimator 3.1 *Under Model Assumption 3.1 and 3.2, the $\mathcal{D}_I$-measurable estimators with minimized conditional MSEPs for $E[X_{i,j}|\mathcal{D}_I]$ and $E[C_{iJ}|\mathcal{D}_I]$ are, respectively, given by*

$$\widehat{X_{i,j}}^{EDF} = \gamma_j\mu_i\widetilde{\mu(\Theta_i)}$$

$$\widehat{C_{iJ}}^{EDF} = C_{i,I-i}+\sum_{j=I-i+1}^{j}\widehat{X_{i,j}}^{EDF}$$

for $1\leq i\leq I$ and $j\in\{I-i+1,\cdots,I\}$.

So far we have assumed that μ_i, γ_j, σ^2 and τ^2 are known. Under these assumptions we have that the Bayesian estimators was optimal in the sense that it minimized the (conditional) MSEP. If the parameters are not known, the problem becomes substantially more difficult and in general one loses the optimal results.

If the above parameters are unknown one can follow different approaches. Either one uses 'plug-in' estimators for these parameters or one also uses a Bayesian approach.

(a) 'Plug-in' estimator. There is no canonical way how the 'plug-in' estimator for γ_j should be constructed. In practice it is estimated by

$$\widehat{\gamma_j}^{CL} = \widehat{\beta_j}^{CL} - \widehat{\beta_{j-1}}^{CL} \tag{3.29}$$

As to the estimation of μ_i, usually one takes a plan value, a budget value or the value used for the premium calculation. For known μ_i and γ_j one can give unbiased estimators for these variance parameters.

(b) Bayesian estimator The Bayesian approach would be consistent in the sense that one applies a full Bayesian approach to all unknown model parameters. However in such a full Bayesian approach there is, in general, no analytical solution to the problem and one needs to completely rely on numerical solutions such as MCMC.

3.3 Bühlmann-Straub Credibility Model

In most of the Bayesian models, the Bayesian estimator $\widetilde{\mu(\Theta_i)}$ can not be expressed in a closed analytical form for we do not know either $f(x|\theta)$ or $\mu(\theta)$, but just the following model.

Model Assumption 3.3 *Conditionally, for all $i \in \{0, \cdots, I\}$, given Θ_i, the first two moments of random variables $Y_{ij} = \frac{X_{ij}}{\mu_i \gamma_j}$, $j = 0, \cdots, J$ are given by*

$$E[Y_{ij}|\Theta_i] = \mu(\Theta_i) \quad \text{and} \quad Var(Y_{ij}|\Theta_i) = \frac{\sigma^2(\Theta_i)}{\omega_{ij}} \tag{3.30}$$

where $\mu(\Theta_i)$ satisfies $\mu_0 \stackrel{\text{def}}{=} E[\mu(\Theta_i)] = 1$.

An usual way in actuarial science is restrict the class of possible estimators to a small class, in which estimators are linear functions of the observations $\mathbf{Y}_i = (Y_{i0}, \cdots, Y_{i,I-i})$, then we have to solve the following problem

$$\widehat{\mu(\Theta_i)}^{cred} = \underset{\widetilde{\mu} \in L(\mathbf{Y}_i, 1)}{\operatorname{argmin}} E\left[(\mu(\Theta_i) - \widetilde{\mu})^2\right]$$

where $L(\mathbf{Y}_i, 1) = \left\{\widetilde{\mu} : \widetilde{\mu} = a_{i0} + \sum_{i=0}^{I} \sum_{j=0}^{I-i} a_{ij} Y_{ij}, a_{ij} \in \mathbb{R}\right\}$.

It is easy to check that under Model Assumption 3.1 and 3.3,the best linear estimator, which is individually unbiased and which has the smallest conditional variance, is given by

$$Y_i = \frac{\sum_{j=0}^{I-i} \omega_{ij} Y_{ij}}{\sum_{j=0}^{I-i} \omega_{ij}} \tag{3.31}$$

Because of the normal equation and the fact that $E[Y_{ij}] = \mu_0$, the credibility estimator must be of the form (3.32) and satisfy (3.33)

$$\widehat{\mu(\Theta_i)}^{cred} = \alpha_i Y_{i\cdot} + (1-\alpha_i)\mu_0 \tag{3.32}$$

$$Cov(\widehat{\mu(\Theta_i)}^{cred}, Y_{i\cdot}) = \alpha_i Cov(Y_i, Y_i) = Cov(\mu(\Theta_i, Y_{i\cdot})) \tag{3.33}$$

Define $\sigma^2 = E[\sigma^2(\Theta_i)]$ and $\tau^2 = Var(\mu(\Theta_i))$ and note that

$$Var[Y_i] = \frac{\sigma^2}{\sum_{j=0}^{I-i} \omega_{ij}} + \tau^2 \quad \text{and} \quad Cov(\mu(\Theta_i), Y_i) = \tau^2$$

Denote that $\omega_{i\cdot} = \sum_{j=0}^{I-i} \omega_{ij}$, then it follows

$$\alpha_i = \frac{\tau^2}{\frac{\sigma^2}{\omega_{i\cdot}+\tau^2}} = \frac{\omega_{i\cdot}}{\omega_{i\cdot} + \frac{\sigma^2}{\tau^2}} \tag{3.34}$$

Estimator 3.2 *The inhomogeneous credibility estimator of* $\mu(\Theta_i)$ *and* C_{iJ} *are, respectively, given by*

$$\begin{aligned} \widehat{\mu(\Theta_i)}^{cred} &= \alpha_i Y_i + (1-\alpha_i)\mathbf{1} \\ \widehat{C_{iJ}}^{cred} &= C_{i,I-i} + (1-\beta_{I-i})\mu_i \widehat{\mu(\Theta_i)}^{cred} \end{aligned} \tag{3.35}$$

where Y_i *and* α_i *are given by* (3.31) *and* (3.34), *respectively.*

Theorem 3.2 *Under Model Assumption 3.1 and 3.3, with* $\omega_{ij} = \mu_i^\delta \gamma_j$ *for some appropriate* $\delta \geq 0$, *the MSEP of the inhomogeneous credibility reserving estimators is given by*

$$\text{msep}_{C_{iJ}}(\widehat{C_{iJ}}^{cred}) = \mu_i^2[(1-\beta_{I-i})\frac{\sigma^2}{\mu_i^\delta} + (1-\beta_{I-i})^2\tau^2(1-\alpha_i)]$$

for $1 \leq i \leq I$.

Remark 3.1 *Under Model Assumption 2.3, we can also consider another type of credibility estimator, which is referred to in the literature as the homogeneous credibility. We defined the homogeneous credibility estimator of* $\mu(\Theta_i)$ *as the best estimator in the class of collectively unbiased estimators*

$$\left\{\widehat{\mu(\Theta_i)} : \widehat{\mu(\Theta_i)} = \sum_{i=0}^{I}\sum_{j=0}^{I-i} a_{ij}X_{ij}, E[\widehat{\mu(\Theta_i)}] = E[\mu(\Theta_i)] a_{ij} \in \mathbb{R}\right\}.$$

Remark 3.2 *The credibility Model can also apply to the individual development factors.*

4 Distributional Models

The Log-normal model was first considered by [13] and described in [15] and [30], section 7.3. The model considers cumulative claims and Log-normal distributions.

Model Assumption 4.1 *1. The individual development factors F_{ij} are Log-Normally distributed with deterministic parameters ξ_j and σ_j^2, that is*

$$\eta_{ij} = \log(F_{ij}) \sim N(\xi_j, \sigma_j^2) \tag{4.36}$$

for all $i \in \{0, 1, \cdots, I\}$ and $j \in \{0, 1, \cdots, J\}$, where $C_{i,-1} = 1$.

2. η_{ij} are independent for $i \in \{0, 1, \cdots, I\}$ and $j \in \{0, 1, \cdots, J\}$.

Estimator 4.1 *We estimate the parameters ξ_j and σ_j^2 as follows*

$$\widehat{\xi_j} = \frac{1}{I-j+1}\sum_{i=0}^{I-j}\log\Big(F_{ij}\Big) \quad \text{and} \quad \widehat{\sigma_j^2} = \frac{1}{I-j}\sum_{i=0}^{I-j}\Big(\log F_{ij} - \widehat{\xi_j}\Big)^2 \tag{4.37}$$

Moreover, $\widehat{\xi_j}$ and $\widehat{\sigma}_j^2$ are stochastically independent.

First we assume that the variances $\sigma_0^2, \cdots, \sigma_J^2$ are known. Define $Z_{ij} = \log C_{ij}$, hence we have

$$E[Z_{iJ}|\mathcal{D}_I] = Z_{i,I-i} + \sum_{j=I-i+1}^{J}\eta_{ij}$$

which is estimated by

$$\widehat{Z_{iJ}} = \widehat{E[Z_{iJ}|\mathcal{D}_I]} = Z_{i,I-i} + \sum_{j=I-i+1}^{J}\widehat{\xi_j}$$

Note that

$$\begin{aligned} E[C_{iJ}|\mathcal{D}_I] &= E[\exp\{Z_{iJ}\}|C_{i,I-i}] \\ &= \exp\{Z_{i,I-i}\}\prod_{j=I-i+1}^{J}E[\eta_{ij}|C_{i,I-i}] \\ &= C_{i,I-i}\exp\Big\{\sum_{j=I-i+1}^{J}\xi_j + \frac{1}{2}\sum_{j=I-i+1}^{J}\sigma_j^2\Big\} \end{aligned} \tag{4.38}$$

and

$$\begin{aligned} E[\exp\{\widehat{Z_{iJ}}\}|C_{i,I-i}] &= \exp\{Z_{i,I-i}\}E[\exp\{\sum_{j=I-i+1}^{J}\widehat{\xi_j}\}] \\ &= C_{i,I-i}\exp\Big\{\sum_{j=I-i+1}^{J}\xi_j + \frac{1}{2}\sum_{j=I-i+1}^{J}\frac{\sigma_j^2}{I-j+1}\Big\} \end{aligned} \tag{4.39}$$

thus, the next estimator is straightforward from (4.38) and (4.39).

Estimator 4.2 *Under Model Assumption 4.1 with σ_j^2 known, for $i = 1, \cdots, I$, an unbiased estimator for $E[C_{iJ}|\mathcal{D}_I]$ is given by*

$$\widehat{C_{iJ}}^{LN} = \widehat{E}[C_{iJ}|\mathcal{D}_I] = \exp\Big\{\widehat{Z_{iJ}} + \frac{1}{2}\sum_{j=I-i+1}^{J} \sigma_j^2\Big(1 - \frac{1}{I-j+1}\Big)\Big\} \tag{4.40}$$

and the conditional MSEP is given by

$$\text{msep}_{C_{iJ}|C_{i,I-i}}\Big(\widehat{C_{iJ}}^{LN}\Big) = E[C_{iJ}|C_{i,I-i}]^2\Big(\exp\{\sum_{j=I-i+1}^{J}\sigma_j^2\} + \exp\{\sum_{j=I-i+1}^{J}\frac{\sigma_j^2}{I+j-1}\} - 2\Big)$$

However, in general, the variances σ_j^2 need also be estimated from the data. We could obtain an estimator by replacing σ_j^2 by $\widehat{\sigma_j^2}$ in (4.40), but this estimator is no longer unbiased.

Estimator 4.3 *Under Model Assumption 4.1 with σ^2 unknown, the estimator for $E[C_{iJ}|\mathcal{D}_I]$ is given by*

$$\widehat{C_{iJ}}^{LN_\sigma,2} = \widehat{E}[C_{iJ}|\mathcal{D}_I] = \exp\Big\{\widehat{Z_{iJ}} + \frac{1}{2}\sum_{j=I-i+1}^{J} \widehat{\sigma_j^2}\Big(1 - \frac{1}{I-j+1}\Big)\Big\} \tag{4.41}$$

Remark 4.1 *In [10], several distribution models for incremental claims are introduced,and most of them can be regarded as a special case of generalized linear model. Hence we will illustrate them in the next section.*

5 Generalized Linear Models

The standard GLM techniques for the derivation of estimates for incremental data in a claim reserving context was first implemented by Renshaw [25] and Renshaw and Verrall [26]. A good overview on this can refer to [10] and [11].

5.1 Generalized Linear Models Framework for Claim Reserving

As the usual generalized linear model, the generalized linear model for claim reserving has three components. We illustrate them in the form of Model Assumption.

Model Assumption 5.1 *1. The increments $\{X_{i,j}, i = 0, \cdots, I, j = 0, \cdots, J\}$ of different accident years i and development year j are independent and satisfy*

$$E[X_{i,j}] = x_{i,j} \quad \text{and} \quad Var(X_{i,j}) = \frac{\phi_{i,j}}{\omega_{i,j}}V(x_{i,j}) \tag{5.42}$$

where $V(\cdot)$ is an appropriate variance function.

2. $\{x_{i,j}, i = 0, \cdots, I, j = 0, \cdots, J\}$ *can be specified by a number of unknown parameters* $\mathbf{b} = (b_1, \cdots, b_p)$ *which produce a linear predictor* $\boldsymbol{\eta} = (\eta_{i,j})_{\{i=0,\cdots,I,j=0,\cdots,J\}}$*:*

$$\eta_{i,j} = \Gamma_{i,j}\mathbf{b} \tag{5.43}$$

for appropriate (deterministic) $(1 \times p)$*-design matrices* $\Gamma_{i,j}$*.*

3. There exist a monotone link function g *such that*

$$g(x_{i,j}) = \eta_{i,j} = \Gamma_{i,j}\mathbf{b} \tag{5.44}$$

Example 5.1 *Assume that* $X_{i,j}$ *has density* (3.28) *with* $b'^{-1}(\cdot)$ *exists and we have a multiplicative structure*

$$x_{i,j} = \mu_i\gamma_j \tag{5.45}$$

with μ_i *standing for the exposure of accident year* i *and* γ_j *denotes the claim pattern over different year* j*.*

For the multiplicative structure, an straightforward choice for link function is the log function. Then we have

$$\log(x_{i,j}) = \eta_{i,j} = \log(\mu_i) + \log(\gamma_j). \tag{5.46}$$

Therefore we have $p = I + J + 1$ and

$$\mathbf{b} = (\log(\mu_1), \cdots, \log(\mu_I), \log(\gamma_0), \cdots, \log(\gamma_J))' \tag{5.47}$$

and

$$\Gamma_{i,j} = (0, \cdots, 0, e_i, 0, \cdots, 0, e_{I+j+1}, 0, \cdots, 0) \tag{5.48}$$

for $1 \leq i \leq I$ and $0 \leq j \leq J$, where the entries $e_i = 1$ and e_{I+j+1} are on the ith and $(I + j + 1)$th position respectively.

Since $X_{i,j}$ belong to Exponential Dispersion Family, we have

$$E[X_{ij}] \stackrel{\text{def}}{=} x_{ij} = b'(\theta_{ij}) \quad \text{and} \quad Var(X_{ij}) = \frac{\phi_{ij}}{\omega_{ij}}b''(\theta_{ij}) \tag{5.49}$$

then the variance function is given by

$$V(x_{ij}) = b''((b'^{-1}(x_{ij})) \tag{5.50}$$

Remark 5.1 *In [10] and [11], several incremental models are proposed, such as the Poisson model, Gaussian Model, all these models can be regarded as a special class of GLM with* $V(x_{i,j}) = x_{i,j}^p$*. For example,* $p = 0$ *we get the Gaussian Model,* $p = 1$ *we get the Poisson Model and* $p = 2$ *we get the Gamma Model.*

5.2 Parameter Estimation in the EDF

Now we use MLE on the set of observations $\mathcal{D}_I = \{X_{i,j}; i+j \leq I, 0 \leq j \leq J\}$ to estimate the unknown parameter in Example 5.1. Refer to (5.49) and (3.28), the log-likelihood function can be written as

$$
\begin{aligned}
I_{\mathcal{D}_I}(\mathbf{b}) &= \log \prod_{i+j\leq I} f(X_{ij}; \theta_{ij}, \phi_{ij}, \omega_{ij}) \\
&= \sum_{i+j\leq I} l(X_{ij}; x_{ij}, \phi_{ij}, \omega_{ij}) \\
&= \sum_{i+j\leq I} l(X_{ij}; \mu_i\gamma_j, \phi_{ij}, \omega_{ij})
\end{aligned}
$$

We maximize $I_{\mathcal{D}_I}(\mathbf{b})$ by setting the $I+J+1$ partial derivatives w.r.t the unknown parameters μ_i and γ_j equal to zero. Thus we obtain $\widehat{\mu}_i$ and $\widehat{\gamma}_j$ and hence

$$
\widehat{\mathbf{b}} = \left(\widehat{\log(\mu_1)}, \cdots, \widehat{\log(\mu_I)}, \widehat{\log(\gamma_0)}, \cdots, \widehat{\log(\gamma_J)}\right) \tag{5.51}
$$

with $\widehat{\log(\mu_i)} = \log(\widehat{\mu}_i)$ and $\widehat{\log(\gamma_j)} = \log(\widehat{\gamma}_j)$. Hence we derive the following estimator.

Estimator 5.1 *The MLE in the EDF Model 4.1 is given by*

$$
\widehat{X}_{ij}^{EDF} = \widehat{x}_{ij} = \widehat{E[X_{ij}|\mathcal{D}_I]} = \widehat{\mu}_i\widehat{\gamma}_j
$$

$$
\widehat{C}_{ij}^{EDF} = \widehat{E[C_{ij}|\mathcal{D}_I]} = C_{i,I-i} + \sum_{j=I-i+1}^{J} \widehat{X}_{ij}^{EDF}
$$

for $i+j > I$.

Remark 5.2 *Usually the calculation of MLE and MSEP need to use numerical methods, such as the Fisher's scoring method or the Newton-Raphson algorithm.*

6 Bootstrap Methods for claim reserving

The general idea behind bootstrap is to make a data resampling from the data themselves. In the actuarial literature, bootstrap methods appears, for example, in [11], [29], [31] and [32].

Assume we have n i.i.d realizations $Z_1, \cdots, Z_n$ from an an unknown distribution F and $h(F)$ is a parameter of F, such as the mean or variance. g is an unknown function of the data $Z_1, \cdots, Z_n$ which estimate $h(F)$. Our goal is to learn more about the probability distribution of $\widehat{\theta}_n \stackrel{def}{=} g(Z_1, \cdots, Z_n)$.

If we know the data generating mechanism F, we can sample new i.i.d observations from F. This would give a new value for $\widehat{\theta}_n$. Repeating this procedure several times would lead to the empirical distribution of $\widehat{\theta}_n$. However, since F is unknown, we use an estimator F_1 to reproduce observations. That is we generate new data

$$Z_1^*, \cdots, Z_n^* \quad \text{i.i.d} \overset{(d)}{\sim} F_1 \tag{6.52}$$

where F_1 is the empirical distribution of $Z_1, \cdots, Z_n$ in nonparametric case and in parametric case, $F_1 = F_{\widehat{\lambda}}$ since F is known up to a finite vector of unknown parameters λ_0 ($\widehat{\lambda}$ is an estimator of λ_0 from $Z_1, \cdots, Z_n$). The new data vector $(Z_1^*, \cdots, Z_n^*)$ is called a bootstrap sample. Then for the bootstrap sample we can calculate a new value for the estimator $\widehat{\theta}_n$,

$$\widehat{\theta}_n^* = g(Z_1^*, \cdots, Z_n^*)$$

repeating this idea several times, we get an empirical distribution $\widehat{F}_n^*$ for $\widehat{\theta}_n^*$.

In the framework of claim reserving, we are interested in the distribution of

$$h(F) = \sum_{i+j \geq I} E[X_{ij}|\mathcal{D}_I] \tag{6.53}$$

which are expected open loss liabilities/outstanding claims reserves at time I.

6.1 Log-Normal Model for Cumulative Sizes

Recall that under Model Assumption 4.1, for any $j \in \{1, \cdots, J\}$,

$$\eta_{1,j}, \cdots, \eta_{I-j,j} \overset{i.i.d}{\sim} N(\xi_j, \sigma_j^2) \tag{6.54}$$

and

$$h(F) = \sum_{i=1}^{I} C_{i,I-i}\Big(\exp\{\sum_{j=I-i+1}^{J} \xi_j + \frac{1}{2}\sum_{j=I-i+1}^{J} \sigma_j^2\} - 1\Big) \tag{6.55}$$

Since the parameters are not known they need to be estimated, and the appropriate estimators were provided by (6.1). This has led to the following estimator for $h(F)$, given the observations $\mathcal{D}_I$,

$$g(\mathcal{D}_I) = \sum_{i=1}^{I} \Big(\exp\{\sum_{j=I-i+1}^{J} \widehat{\xi}_j + \frac{1}{2}\sum_{j=I-i+1}^{J} \widehat{\sigma}_j^2(1 - \frac{1}{I-j+1})\} - 1\Big) \tag{6.56}$$

and our goal is to study the distribution of the estimator $g(\mathcal{D}_I)$.

Since we have explicit distribution assumptions to η_{ij}, we would like to apply the parametric bootstrap method. This means that we need to generate new independent observations η_{ij}^*, that is

$$\eta_{ij}^* \overset{(d)}{\sim} N(\widehat{\xi}_j, \widehat{\sigma}_j^2)$$

which leads to the bootstrap reserves

$$g^*(\mathcal{D}_I) = \sum_{i=1}^{I} \Big(\exp\{ \sum_{j=I-i+1}^{J} \widehat{\xi}_j^* + \frac{1}{2} \sum_{j=I-i+1}^{J} \widehat{\sigma_j^2}^* (1 - \frac{1}{I-j+1}) \} - 1 \Big) \tag{6.57}$$

where $\widehat{\xi_j}^*$ and $\widehat{\sigma_j^2}^*$ are estimated by with the bootstrap sample $(\eta_{i,1}^*, \cdots, \eta_{I-j,j}^*.)$ Repeating this procedure several times we obtain the empirical distribution of $g^*(\mathcal{D}_I)$, given observations $\mathcal{D}_I$.

6.2 Generalized Linear Models

In order to apply Efron's nonparametric bootstrap method to example 5.1, we once again need to find identically distributed residuals that allow for the construction of the empirical distribution F_1, see (6.52).

In the following, we assume that $\phi = \frac{\phi_{ij}}{\omega_{ij}}$ is constant and as in England and Verrall (2002, 2007), we choose the Pearson residuals given by

$$R_{ij}^{(P)}(x_{ij}) = \frac{X_{ij} - x_{ij}}{V(x_{ij})^{1/2}} \tag{6.58}$$

.

Note that the residuals have mean 0 and variance ϕ. Therefore $R_{ij}^{(P)}(x_{ij})$ is a natural object to define the bootstrap distribution. Hence, we set for $i + j \leq I$,

$$Z_{ij} = \frac{X_{ij} - \widehat{x}_{ij}}{V(\widehat{x}_{ij})^{1/2}} \tag{6.59}$$

.

These $\{Z_{ij}, i + j \leq I\}$ defines the bootstrap distribution $\widehat{F}_{\mathcal{D}_I}$. Then we resample i.i.d residuals

$$Z_{ij}^* \sim \widehat{F}_{\mathcal{D}_I}$$

and hence we define the bootstrap observations of X_{ij} by

$$X_{ij}^* = \widehat{x}_{ij} + V(\widehat{x}_{ij})^{1/2} Z_{ij}^* \tag{6.60}$$

These bootstrap observations X_{ij}^* now lead to bootstrap claims reserving triangles $\mathcal{D}_I^* = \{X_{ij}^*, i + j \leq I\}$. Using GLM methods, we calculate bootstrap estimates μ_i^*, γ_j^* and $\widehat{x}_{ij}^*$ from the bootstrap observations $\mathcal{D}_I^*$. This leads to the bootstrap claims reserves $\{X_{ij}^*, i+j > I\}$. Repeating this bootstrap sampling several times we obtain the bootstrap distribution of the claim reserves, conditioned on $\mathcal{D}_I$.

6.3 Chain Ladder Method

Under Model Assumption 2.3, the individual development factors are given by

$$F_{i,j+1} = \frac{C_{i,j+1}}{C_{ij}} = f_j + \sigma_j C_{ij}^{-1/2} \varepsilon_{i,j+1}$$

for the time being we assume that σ_j is known.

In order to apply the bootstrap method we again need to find appropriate residuals that allow for the construction of the empirical distribution F_1, from which the bootstrap observations are constructed.

Consider the following residuals for $i + j \leq I, j \geq 1$,

$$\widetilde{\varepsilon}_{ij} = \frac{F_{ij} - \widehat{f}_{j-1}}{\sigma_{j-1} C_{i,j-1}^{-1/2}} \tag{6.61}$$

where the estimator $\widehat{f}_j$ are given in (2.3). Note that $E[\widetilde{\varepsilon}_{ij}|\mathcal{B}_{j-1}] = 0$ and

$$Var(\widetilde{\varepsilon}_{ij}|\mathcal{D}_{j-1}) = 1 - \frac{C_{i,j-1}}{\sum_{i=0}^{I-j} C_{i,j-1}} < 1$$

which means that we should adjust our observable residuals $\widetilde{\varepsilon}_{ij}$ in order to obtain the correct order for the estimation error.

Define

$$Z_{ij} = \left(1 - \frac{C_{i,j-1}}{\sum_{i=0}^{I-j} C_{i,j-1}}\right)^{-1/2} \frac{F_{ij} - \widehat{f}_{j-1}}{\sigma_{j-1} C_{i,j-1}^{-1/2}} \tag{6.62}$$

where $\widehat{f}_j$ and $\widehat{\sigma}_j^2$are given in (2.3) and (1.8). These residuals $\{Z_{ij}, i + j \leq I\}$ defines the bootstrap distribution $\widehat{F}_{\mathcal{D}_I}$. Then we resample i.i.d residuals

$$Z_{ij}^* \sim \widehat{F}_{\mathcal{D}_I}$$

and hence we define the bootstrap observations $\{F_{ij}^*, i + j \leq I\}$ by

$$X_{ij}^* = \widehat{x}_{ij} + V(\widehat{x}_{ij})^{1/2} Z_{ij}^* \tag{6.63}$$

In contrast to the methods in subsections 5.2 and 5.3, the next step of reproducing bootstrap observations F_{ij}^* of F_{ij} is not straightforward.

7 Multivariate Reserving Methods

The study of Multivariate claims reserving methods are motivated by the fact that, in practice, it is quite natural to subdivide a nonlife run-off portfolio into subportfolios such that each satisfies certain homogeneity properties. They have been studied by [5], [27], [28] and [36]. Another type of multivariate claim reserving method is studied by [18] and [24], where one combines different sources of information in the same estimate.

In the following we assume that the subportfolios consist of $N \geq 1$ run-off triangles of observations of the same size. The incremental and cumulative claims of triangle $n(1 \leq n \leq N)$ for accident year i and development year j are denoted by $X_{ij}^{(n)}$ and $C_{ij}^{(n)}$.

Usually at time I, we have a total set of observations given by $\mathcal{D}_I^N = \bigcup_{n=1}^N \mathcal{D}_I^{(n)}$ where $\mathcal{D}_I^{(n)} = \{\mathcal{C}_I^{(n)} : i + j \leq I\}$ is the observations for run-off subportfolio n. And we need to predict the random variables in its complement, which is given by

$$\mathcal{D}_I^{N,c} = \{C_{ij}^{(n)} : i + j > I, 1 \leq n \leq N, i \leq I\}$$

For the derivation of the conditional MSEP of multivariate methods it is convenient to write the N subportfolios in vector form. Thus we define the N-dimensional random vectors of incremental an cumulative claims by

$$\mathbf{X}_{ij} = (X_{ij}^{(1)}, \cdots, X_{ij}^{(N)})' \qquad \text{and} \qquad \mathbf{C}_{ij} = (C_{ij}^{(1)}, \cdots, C_{ij}^{(N)})'$$

for $i \in \{0, \cdots, I\}$ and $j \in \{0, \cdots, J\}$. Moreover, we define for $k \in \{0, \cdots, J\}$,

$$\mathcal{D}_k^N = \{\mathbf{C}_{ij} : i + j \leq I, 0 \leq j \leq k\}$$

and the N-dimensional column vector consisting of 1's by $\mathbf{1} = (1, \cdots, 1)'$.

7.1 Multivariate Chain-Ladder Method

We define for $n \in (1, \cdots, N)$, $i \in (0, \cdots, I)$ and $j \in (1, \cdots, J)$ the individual development factors for accident year i and development year j by

$$F_{ij}^{(n)} = \frac{C_{ij}^{(n)}}{C_{i,j-1}^{(n)}} \qquad \text{and} \qquad \mathbf{F}_{ij} = (F_{ij}^{(1)}, \cdots, F_{ij}^{(N)})'$$

and denote by $D(\mathbf{a})$ for the $N \times N$ diagonal matrix generated by the N-dimensional vectors $\mathbf{a} = (a_1, \cdots, a_N)' \in \mathbb{R}^N$. Then we have

$$\mathbf{C}_{ij} = D(\mathbf{C}_{i,j-1})\mathbf{F}_{ij} = D(\mathbf{F}_{ij})\mathbf{C}_{i,j-1}$$

for all $j = 1, \cdots, J$ and $i = 0, \cdots, I$ The distribution-free multivariate CL model is then given by the following definition.

Model Assumption 7.1 (multivariate CL model) *1. Cumulative claims $\mathbf{C}_{ij}$ of different accident years i are independent.*

2. $(\mathbf{C}_{ij})_{j\geq 0}$ form an $N-$dimensional Markov chain. There are $N-$dimensional deterministic vectors $\mathbf{f}_j = (f_j^{(1)}, \cdots, f_j^{(N)})' > \mathbf{0}$ and symmetric positive definite $N \times N$ matrices Σ_j such that for all $0 \leq i \leq I$ and $1 \leq j \leq J$ we have

$$\begin{aligned} E[\mathbf{C}_{ij}|\mathbf{C}_{i,j-1}] &= \mathrm{D}(\mathbf{f}_{j-1})\mathbf{C}_{i,j-1} \\ \mathrm{Cov}(\mathbf{C}_{ij}, \mathbf{C}_{ij}|\mathbf{C}_{i,j-1}) &= \mathrm{D}(\mathbf{C}_{i,j-1})^{\frac{1}{2}}\Sigma_{j-1}\mathrm{D}(\mathbf{C}_{i,j-1})^{\frac{1}{2}} \end{aligned} \tag{7.64}$$

Analogical to equality (2.2), we have the following theorem.

Theorem 7.1 *Under Model Assumption 7.1, given $\mathcal{D}_j^N$, the conditional expectation and process variance for $\mathbf{C}_{iJ}$ are, respectively, given by*

$$\begin{aligned} E[\mathbf{C}_{ij}|\mathcal{D}_I^N] &= \prod_{l=I-i}^{j-1} D(\mathbf{f}_l)\mathbf{C}_{i,I-i} \\ \mathbf{1}'Var(\mathbf{C}_{iJ}|\mathcal{D}_I^N)\mathbf{1} &= \mathbf{1}'\Big(\sum_{j=I-i}^{J-1}\prod_{k=j+1}^{J-1} D(\mathbf{f}_k)\Sigma_{ij}^C \prod_{k=j+1}^{J-1} D(\mathbf{f}_k)\Big)\mathbf{1} \end{aligned} \tag{7.65}$$

where $\Sigma_{ij}^C = E[\mathrm{D}(\mathbf{C}_{ij})^{\frac{1}{2}}\Sigma_j\mathrm{D}(\mathbf{C}_{ij})^{\frac{1}{2}}|\mathbf{C}_{i,I-i}]$.

Under Model Assumption 7.1, conditional on $\mathcal{D}_j^N$, [23] and [28] propose the BLUE for $\mathbf{f}_j$,

$$\widehat{\mathbf{f}}_j = \Big(\sum_{i=0}^{I-j-1}\mathrm{D}(\mathbf{C}_{ij})^{\frac{1}{2}}\Sigma_j^{-1}\mathrm{D}(\mathbf{C}_{ij})^{\frac{1}{2}}\Big)^{-1}\sum_{i=0}^{I-j-1}\mathrm{D}(\mathbf{C}_{ij})^{\frac{1}{2}}\Sigma_j^{-1}\mathrm{D}(\mathbf{C}_{ij})^{\frac{1}{2}}\mathbf{F}_{i,j+1} \tag{7.66}$$

Then an unbiased estimator for $E[\mathbf{C}_{iJ}|\mathcal{D}_I^N]$ is given by

$$\widehat{\mathbf{C}_{ij}}^{CL} = (\widehat{\mathbf{C}_{ij}^{(1)}}^{CL}, \cdots, \widehat{\mathbf{C}_{ij}^{(N)}}^{CL})' = \widehat{E}[\mathbf{C}_{ij}|\mathcal{D}_I^N] = \prod_{l=I-i}^{j-1}\mathrm{D}(\widehat{\mathbf{f}}_l)\mathbf{C}_{i,I-i} \tag{7.67}$$

Using theorem 7.1 and (7.66) we have, the conditional estimation error of accident year $i \geq 0$ is given by

$$\begin{aligned} \mathbf{1}'\big(\widehat{\mathbf{C}}_{ij}^{CL} &- E[\mathbf{C}_{ij}|\mathcal{D}_I^N]\big)\big(\widehat{\mathbf{C}}_{ij}^{CL} - E[\mathbf{C}_{ij}|\mathcal{D}_I^N]\big)'\mathbf{1} \\ &= \mathbf{1}'(\prod_{j=I-i}^{J-1}\mathrm{D}(\widehat{\mathbf{f}}_j) - \prod_{j=I-i}^{J-1}\mathrm{D}(\mathbf{f}_j)\mathbf{C}_{i,I-i}\mathbf{C}'_{i,I-i}(\prod_{j=I-i}^{J-1}\mathrm{D}(\widehat{\mathbf{f}}_j) - \prod_{j=I-i}^{J-1}D(\widehat{\mathbf{f}}_j))\mathbf{1} \\ &= \mathbf{1}'\mathrm{D}(\mathbf{C}_{i,I-i})(\widehat{g}_{i|J} - g_{i|J})(\widehat{g}_{i|J} - g_{i|J})'\mathrm{D}(\mathbf{C}_{i,I-i})\mathbf{1} \end{aligned} \tag{7.68}$$

where $\widehat{g}_{i|j} = D(\widehat{\mathbf{f}}_{I-i})\cdots D(\widehat{\mathbf{f}}_{j-1})\mathbf{1}$ and $g_{i|j} = D(\mathbf{f}_j)\cdots D(\mathbf{f}_{j-1})\mathbf{1}$.

In order to determine the conditional estimation error, analogous to the univariate case, we introduce stronger model assumptions.

Model Assumption 7.2 (multivariate Time Series Model) *1. Cumulative claims C_{ij} of different accident years i are independent.*

2. There exist N-dimensional positive constants $\mathbf{f}_j = (f_j^{(1)}, \cdots, f_j^{(N)})'$ and $\boldsymbol{\sigma}_j = (\sigma_j^{(1)}, \cdots, \sigma_j^{(N)})'$ and N-dimensional random variables $\boldsymbol{\varepsilon}_{i,j+1} = (\varepsilon_{i,j+1}^{(1)}, \cdots, \varepsilon_{i,j+1}^{(N)})$ such that for $0 \le i \le I$ and $0 \le j \le J-1$, we have

$$\mathbf{C}_{i,j+1} = \mathrm{D}(\mathbf{f}_j)\mathbf{C}_{i,j} + \mathrm{D}(\mathbf{C}_{ij})^{1/2}\mathrm{D}(\boldsymbol{\varepsilon}_{i,j+1})\boldsymbol{\sigma}_j$$

3. The random variables $\boldsymbol{\varepsilon}_{i,j+1}$ are independent with $E[\boldsymbol{\varepsilon}_{i,j+1}] = \mathbf{0}$ and positive definite

$$Cov(\boldsymbol{\varepsilon}_{i,j+1}, \boldsymbol{\varepsilon}_{i,j+1}) = E[\boldsymbol{\varepsilon}_{i,j+1}\boldsymbol{\varepsilon}_{i,j+1}^{'}] = \begin{pmatrix} 1 & \rho_j^{(1,2)} & \cdots & \rho_j^{(1,N)} \\ \rho_j^{(2,1)} & 1 & \cdots & \rho_j^{(2,N)} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_j^{(N,1)} & \rho_j^{(N,2)} & \cdots & 1 \end{pmatrix}$$

where $\rho_j^{(n,m)} \in (-1, 1)$ for $n \ne m$.

We now describe the conditional resampling approach in the multivariate setup, that is, we conditionally resample $\widehat{\mathbf{f}}_{I-i}, \cdots, \widehat{\mathbf{f}}_{J-1}$ given the triangle $\mathcal{D}_I^N$. Hence we generate 'new' observations $\widetilde{\mathbf{C}}_{i,j+1}$ for $i \in \{0, \cdots, I\}$ and $j \in \{0, \cdots, J-1\}$ using the approach

$$\widetilde{\mathbf{C}}_{i,j+1} = \mathrm{D}(\mathbf{f}_j)\mathbf{C}_{i,j} + \mathrm{D}(\mathbf{C}_{ij})^{1/2}\mathrm{D}(\widetilde{\varepsilon}_{i,j+1})\boldsymbol{\sigma}_j \tag{7.69}$$

where $\widetilde{\varepsilon}_{i,j+1}$,$\boldsymbol{\varepsilon}_{i,j+1}$ are i.i.d copies, given $\mathcal{D}_0^N$. This means that $\mathbf{C}_{i,j}$ acts as a deterministic volume measure and we resample successively the next observation $\widetilde{\mathbf{C}}_{i,j+1} \overset{(d)}{=} \mathbf{C}i, j+1$,given $\mathcal{D}_0^N$.

In the spirit of conditional resampling approach this leads to the following resampled estimates of the multivariate development factors

$$\begin{aligned} \widehat{\mathbf{f}}_j^{(*)} &= \Big(\sum_{i=0}^{I-j-1} \mathrm{D}(\mathbf{C}_{ij})^{1/2}\Sigma_j^{-1}\mathrm{D}(\mathbf{C}_{ij})^{1/2} \Big)^{-1} \sum_{i=0}^{I-j-1} \mathrm{D}(\mathbf{C}_{ij})^{1/2}\Sigma_j^{-1}\mathrm{D}(\mathbf{C}_{ij})^{1/2}\mathrm{D}(\mathbf{C}_{ij})^{-1}\widetilde{\mathbf{C}}_{i,j+1} \\ &= \mathbf{f}_j + \sum_{i=0}^{I-j-1} \mathrm{D}(\mathbf{C}_{ij})^{1/2}\Sigma_j^{-1}\mathrm{D}(\mathbf{C}_{ij})^{1/2}\Big)^{-1} \sum_{i=0}^{I-j-1} \mathrm{D}(\mathbf{C}_{ij})^{1/2}\Sigma_j^{-1}\mathrm{D}(\widetilde{\varepsilon}_{i,j+1})\boldsymbol{\sigma}_j. \end{aligned} \tag{7.70}$$

As in the univariate time series model, we denote the conditional probability measure of the resampled estimators by $P^{(*)}_{\mathcal{D}_I^N}$. This way we obtain the following lemma.

Lemma 7.1 *Under the resampling assumption, we have*
1. the estimators $\widehat{\mathbf{f}}_0^{(*)}, \cdots, \widehat{\mathbf{f}}_{J-1}^{(*)}$ are independent and unbiased for $\mathbf{f}_j$ under the probability measure $P^{(*)}_{\mathcal{D}_I^N}$.
2. $E^{(*)}_{\mathcal{D}_I^N}[\widehat{f}_j^{(n)}\widehat{f}_j^{(m)}] = f_j^{(n)}f_j^{(m)} + W_j(n,m)$, where $W_j(n,m)$ is the entry (n,m) of the $N \times N$ matrix defined by $W_j = \Big(\sum_{i=0}^{I-j-1} \mathrm{D}(\mathbf{C}_{ij})^{1/2}\Sigma_j^{-1}\mathrm{D}(\mathbf{C}_{ij})^{1/2} \Big)^{-1}$

Using the above lemma, we obtain the following estimator for (7.68),

$$\mathbf{1}'\mathrm{D}(\mathbf{C}_{i,I-i})E^*_{\mathcal{D}_I^N}\big[(\widehat{g}_{i|J}-g_{i|J})(\widehat{g}_{i|J}-g_{i|J})'\big]\mathrm{D}(\mathbf{C}_{i,I-i})\mathbf{1}=\mathbf{1}'\mathrm{D}(\mathbf{C}_{i,I-i})\big(\Delta_{i,J}^{(n,m)}\big)_{1\le n,m\le N}\mathrm{D}(\mathbf{C}_{i,I-i})'\mathbf{1} \tag{7.71}$$

where the (n,m) entry of $\big(\Delta_{i,J}^{(n,m)}\big)_{1\le n,m\le N}$ is given by

$$\Delta_{i,J}^{(n,m)}=\prod_{j=I-i}^{J-1}\big(f_j^{(n)}f_j^{(m)}+W_j(n,m)\big)-\prod_{j=I-i}^{J-1}f_j^{(n)}f_j^{(m)} \tag{7.72}$$

7.2 Multivariate Additive Loss Reserving Method

The multivariate ALR model was proposed by [14] and [28]. The multivariate ALR method is based on incremental claims, and hence is more in the spirit of the (univariate) GLM models. In this subsection we closely follow [19].

Model Assumption 7.3 (multivariate ALR time series model) *1. Incremental claims* $\mathbf{X}_{ij}$ *of different accident years* i *are independent.*

2. There exist $(N\times N)$*-dimensional deterministic positive definite matrices* $V_0,\cdots,V_I$ *and* N*-dimensional constants* $(j=1,\cdots,J)$

$$\mathbf{m}_j=(m_j^{(1)},\cdots,m_j^{(N)})' \quad \text{and} \quad \boldsymbol{\sigma}_{j-1}=(\sigma_{j-1}^{(1)},\cdots,\sigma_{j-1}^{(N)})'$$

with $\sigma_{j-1}^{(N)}>0$ *for all* $n=1,\cdots,N$ *as well as* $N-$*dimensional random variables* $\boldsymbol{\varepsilon}_{ij}$ *such that*

$$\mathbf{X}_{ij}=V_i\mathbf{m}_j+V_i^{1/2}D(\boldsymbol{\varepsilon}_{ij})\boldsymbol{\sigma}_{j-1}.$$

3. The random variables $\boldsymbol{\varepsilon}_{i,j+1}$ *are with the same assumption as in Model Assumption 7.2.*

Theorem 7.2 *Under Model Assumption 7.3, we have for all* $i\in\{1,\cdots,I\}$,

$$\begin{aligned} E[\mathbf{C}_{iJ}|\mathcal{D}_I^N] &= E[\mathbf{C}_{iJ}|\mathbf{C}_{i,I-i}]=\mathbf{C}_{i,I-i}+V_i\sum_{j=I-i+1}^{J}\mathbf{m}_j \\ \mathbf{1}'Var(\mathbf{C}_{iJ}|\mathcal{D}_I^N)\mathbf{1} &= \mathbf{1}'V_i^{1/2}\sum_{j=I-i+1}^{J}\Sigma_{j-1}V_i^{1/2}\mathbf{1} \end{aligned} \tag{7.73}$$

In most practical applications we have to estimate the ratios $\mathbf{m}_j$ from the data in the upper triangle. [28] proposed the following unbiased estimator.

$$\widehat{\mathbf{m}}_j=\Big(\sum_{i=0}^{I-j}V_i^{\frac{1}{2}}\Sigma_{j-1}^{-1}V_i^{\frac{1}{2}}\Big)^{-1}\sum_{i=0}^{I-j}(V_i^{\frac{1}{2}}\Sigma_{j-1}^{-1}V_i^{\frac{1}{2}})V_i\mathbf{X}_{ij} \tag{7.74}$$

Then the multivariate ALR estimator for $E[\mathbf{C}_{ij}|\mathcal{D}_I^N]$ is, for $i+j>I$, given by

$$\widehat{\mathbf{C}_{ij}}^{AD} = (\widehat{C_{ij}^{(1)}}^{AD}, \cdots, \widehat{C_{ij}^{(N)}}^{AD})' = \widehat{E}[\mathbf{C}_{iJ}|\mathcal{D}_I^N] = \mathbf{C}_{i,I-i} + V_i \sum_{j=I-i+1}^{J} \widehat{\mathbf{m}}_j \tag{7.75}$$

Remark 7.1 *In* (7.73)-(7.74)*, we assume* Σ_{j-1}^{-1} *are known, otherwise* $\mathbf{m_j}$ *and* Σ_{j-1}^{-1} *should be estimated iteratively,*

$$\widehat{\mathbf{m}}_j^{(k)} = \Big(\sum_{i=0}^{I-j} V_i^{\frac{1}{2}}\big(\widehat{\Sigma}_{j-1}^{(k)}\big)^{-1} V_i^{\frac{1}{2}}\Big)^{-1} \sum_{i=0}^{I-j} (V_i^{\frac{1}{2}}\big(\widehat{\Sigma}_{j-1}^{(k)}\big)^{-1} V_i^{\frac{1}{2}}) V_i \mathbf{X}_{ij}$$

$$\widehat{\Sigma}_{j-1}^{(k)} = \frac{1}{I-j} \sum_{i=0}^{I-j} V_i^{-1/2}\Big(\mathbf{X}_{ij} - V_i \widehat{\mathbf{m}}_j^{(k-1)}\Big)\Big(\mathbf{X}_{ij} - V_i \widehat{\mathbf{m}}_j^{(k-1)}\Big)' V_i^{-1/2}$$

Lemma 7.2 *The estimator for the conditional MSEP of the ultimate claim for single accident year* i *is given by*

$$\begin{aligned}\widehat{\mathrm{msep}}_{\sum_n C_{ij}^{(n)}|\mathcal{D}_I^N}\Big(\sum_{n=1}^{N} \widehat{C_{ij}^{(n)}}^{AD}\Big) &= \mathbf{1}' V_i^{1/2} \sum_{j=I-i+1}^{J} \widehat{\Sigma}_{j-1} V_i^{1/2} \mathbf{1} \\ &\quad + \mathbf{1}' V_i \sum_{j=I-i+1}^{J} \Big(\sum_{l=0}^{I-j} V_l^{1/2} \widehat{\Sigma}_{j-1} V_l^{1/2}\Big)^{-1} V_i \mathbf{1}\end{aligned}$$

and the estimator for the conditional MSEP of the ultimate claim for aggregated accident year is given by

$$\begin{aligned}\widehat{\mathrm{msep}}_{\sum_{i,n} C_{ij}^{(n)}|\mathcal{D}_I^N}\Big(\sum_{i=1}^{I}\sum_{n=1}^{N} \widehat{C_{ij}^{(n)}}^{AD}\Big) &= \sum_{i=1}^{I} \widehat{\mathrm{msep}}_{\sum_n C_{ij}^{(n)}|\mathcal{D}_I^N}\Big(\sum_{n=1}^{N} \widehat{C_{ij}^{(n)}}^{AD}\Big) \\ &\quad + 2 \sum_{1\le i<k\le I} \mathbf{1}' V_i \sum_{j=I-i+1}^{J} \Big(\sum_{l=0}^{I-j} V_l^{1/2} \widehat{\Sigma}_{j-1} V_l^{1/2}\Big)^{-1} V_k \mathbf{1}\end{aligned}$$

Remark 7.2 *In [20], Merz and Wüthrich combine Model 7.2 and 7.3 into one model. The consideration of such a combination is motivated by the fact that, in general, not all subportfolios satisfy the same homogeneity assumptions and/or sometimes we have a prior information for some selected portfolios.*

7.3 Munich Chain-Ladder

In practice, one often has the situation that different sources of information are available to predict the ultimate claim. For example, one has cumulative payments data and claims-incurred data to estimate the ultimate claim. Usually the CL method is applied to both data independently.

In the following we denote by superscript Pa for the payment data and superscript In for the claims-incurred data. Moreover, we define $\widetilde{\mathcal{D}}_j = \mathcal{D}_j^{Pa} \bigcup \mathcal{D}_j^{In}$ for $j = 0, 1, \cdots, J$.

If we assume both the data satisfy Model Assumption 2.1, we can independently predict the ultimate claim of accident year i by

$$\widehat{C_{iJ}^{Pa}}^{CL} = C_{i,I-i}^{Pa} \prod_{j=I-i}^{J-1} \frac{\sum_{i=0}^{I-j-1} C_{i,j+1}^{Pa}}{\sum_{i=0}^{I-j-1} C_{i,j}^{Pa}} \quad \text{and} \quad \widehat{C_{iJ}^{In}}^{CL} = C_{i,I-i}^{In} \prod_{j=I-i}^{J-1} \frac{\sum_{i=0}^{I-j-1} C_{i,j+1}^{In}}{\sum_{i=0}^{I-j-1} C_{i,j}^{In}}$$

Since we predict the same random variable twice, namely $C_{iJ} = C_{iJ}^{Pa} = C_{iJ}^{In}$, we expect the two estimators are close to each other. The crucial idea in the MCL model proposed by [24] is to combine the information coming from cumulative payments and claims incurred data. This is done using the paid/incurred rations. That is, we consider

$$Q_{ij} = \frac{C_{ij}^{Pa}}{C_{ij}^{In}}$$

Model Assumption 7.4 (MCL Model) *1. Both the cumulative payments C_{ij}^{Pa} and the Claim incurred C_{ij}^{In} satisfy Model assumption 2.1 with parameters $(f_j^{Pa}, \sigma_j^{Pa})$ and $(f_j^{In}, \sigma_j^{In})$, respectively. Further we assume C_{ij}^{Pa} and C_{ij}^{In} are independent of different accident years.*

2. There are constants λ^{Pa}, λ^{In} such that for all $0 \le i \le I$ and $0 \le j \le J-1$ we have

$$E\Big[\frac{C_{i,j+1}^{Pa}}{C_{ij}^{Pa}}|\widetilde{\mathcal{D}}_j\Big] = f_j^{Pa} + \lambda^{Pa} Var\Big(\frac{C_{i,j+1}^{Pa}}{C_{ij}^{Pa}}|\widetilde{\mathcal{D}}_j^{Pa}\Big) \frac{Q_{ij}^{-1} - E[Q_{ij}^{-1}|\mathcal{D}_j^{Pa}]}{Var(Q_{ij}^{-1}|\mathcal{D}_j^{Pa})} \tag{7.76}$$

$$E\Big[\frac{C_{i,j+1}^{In}}{C_{ij}^{In}}|\widetilde{\mathcal{D}}_j\Big] = f_j^{In} + \lambda^{In} Var\Big(\frac{C_{i,j+1}^{In}}{C_{ij}^{In}}|\widetilde{\mathcal{D}}_j^{In}\Big) \frac{Q_{ij}^{-1} - E[Q_{ij}^{-1}|\mathcal{D}_j^{In}]}{Var(Q_{ij}^{-1}|\mathcal{D}_j^{In})} \tag{7.77}$$

In this subsection we consider the approach proposed by [18]. That is, we have to solve the optimization problems

$$\widehat{Z_{ij}^{Pa}} = \underset{Z \in L(C_{ij}^{In},1)}{\text{argmin}} \; E[(E[C_{ij}^{Pa}|\widetilde{\mathcal{D}}_{j-1}] - Z)^2|\mathcal{D}_{j-1}^{Pa}]$$

and

$$\widehat{Z_{ij}^{In}} = \underset{Z \in L(C_{ij}^{Pa},1)}{\text{argmin}} \; E[(E[C_{ij}^{In}|\widetilde{\mathcal{D}}_{j-1}] - Z)^2|\mathcal{D}_{j-1}^{In}]$$

where $L(C_{ij}, 1) = \{a_{i,0} + a_{i,1}C_{i,j-1} : a_{i,0}, a_{i,1} \in \mathbb{R}\}$. The estimators $\widehat{Z_{ij}^{Pa}}$ are called best affine-linear one-step estimators of $E[C_{ij}^{Pa}|\widetilde{\mathcal{B}}_j]$ given $\mathcal{D}_j^{Pa}$ and $\mathcal{D}_j^{In}$. It can be shown that the estimators $\widehat{Z_{ij}^{Pa}}$ exist and are unique almost surely. Further, $\widehat{Z_{ij}^{Pa}}$ satisfies the (conditional) normal equations

$$E[(E[C_{ij}^{Pa}|\widetilde{\mathcal{D}}_{j-1}] - \widehat{Z_{ij}^{Pa}})1|\mathcal{D}_{j-1}^{Pa}] = 1 \quad a.s. \tag{7.78}$$

$$E[(E[C_{ij}^{Pa}|\widetilde{\mathcal{D}}_{j-1}] - \widehat{Z_{ij}^{Pa}})C_{i,j-1}^{In}|\mathcal{D}_{j-1}^{Pa}] = 1 \quad a.s. \tag{7.79}$$

The same conclusions apply analogously to $\widehat{Z_{ij}^{In}}$. Based on (7.78) and (7.79), we obtain the following representation for $\widehat{Z_{ij}^{Pa}}$ and $\widehat{Z_{ij}^{In}}$.

Theorem 7.3 *Under Model Assumption 7.4, the best affine-linear one-step estimators* $\widehat{Z_{ij}^{Pa}}$ *and* $\widehat{Z_{ij}^{In}}$ *given* $\mathcal{D}_j^{Pa}$ *and* $\mathcal{D}_j^{In}$*, are given by*

$$\widehat{Z_{ij}^{Pa}} = f_{j-1}^{Pa}C_{i,j-1}^{Pa} + \lambda^{Pa}Var\Big(C_{ij}^{Pa}|\mathcal{D}_{j-1}^{Pa}\Big)^{\frac{1}{2}}\frac{C_{i,j-1}^{In} - E[C_{i,j-1}^{In}|\mathcal{D}_{j-1}^{Pa}]}{Var(C_{i,j-1}^{In}|\mathcal{D}_{j-1}^{Pa})^{\frac{1}{2}}}$$

and

$$\widehat{Z_{ij}^{In}} = f_{j-1}^{In}C_{i,j-1}^{In} + \lambda^{In}Var\Big(C_{ij}^{In}|\mathcal{D}_{j-1}^{In}\Big)^{\frac{1}{2}}\frac{C_{i,j-1}^{Pa} - E[C_{i,j-1}^{Pa}|\mathcal{D}_{j-1}^{In}]}{Var(C_{i,j-1}^{Pa}|\mathcal{D}_{j-1}^{In})^{\frac{1}{2}}}$$

where $\lambda^{Pa} = \mathrm{Cor}(C_{ij}^{Pa}, C_{i,j-1}^{In}|\mathcal{D}_{j-1}^{Pa})$ *and* $\lambda^{In} = \mathrm{Cor}(C_{ij}^{In}, C_{i,j-1}^{Pa}|\mathcal{B}_{j-1}^{In})$.

Obviously, the estimator $\widehat{Z_{ij}^{Pa}}$ and $\widehat{Z_{ij}^{In}}$ are unbiased for $E[C_{ij}^{Pa}|\widetilde{\mathcal{D}}_{j-1}]$ and $E[C_{ij}^{Pa}|\widetilde{\mathcal{D}}_{j-1}]$, respectively.

In order to perform the MCL method we need to estimate/predict the two correlation coefficients λ^{Pa} and λ^{In} as well as $E[Q_{i,j-1}|\mathcal{D}_{j-1}^{In}]$, $E[Q_{i,j-1}^{-1}|\mathcal{D}_{j-1}^{Pa}]$, $Var[Q_{i,j-1}|\mathcal{B}_{j-1}^{In}]$ and $Var[Q_{i,j-1}^{-1}|\mathcal{D}_{j-1}^{Pa}]$.

For the derivation of reasonable estimates we assume that $E[Q_{i,j-1}^{-1}|\mathcal{D}_{j-1}^{Pa}]$ and $E[Q_{i,j-1}|\mathcal{D}_{j-1}^{In}]$ as well as $Var[Q_{i,j-1}^{-1}|\mathcal{D}_{j-1}^{Pa}]$ and $Var[Q_{i,j-1}|\mathcal{D}_{j-1}^{In}]$ are constants depending only on $j = 1, \cdots, J$. We set

$$\widehat{q}_j = \frac{1}{\sum_{i=0}^{I-j} C_{ij}^{In}} \sum_{i=0}^{I-j} C_{ij}^{In} Q_{ij} = \frac{\sum_{i=0}^{I-j} C_{ij}^{Pa}}{\sum_{i=0}^{I-j} C_{ij}^{In}}$$

$$\widehat{q_j^{-1}} = \frac{1}{\sum_{i=0}^{I-j} C_{ij}^{Pa}} \sum_{i=0}^{I-j} C_{ij}^{Pa} Q_{ij} = \frac{\sum_{i=0}^{I-j} C_{ij}^{In}}{\sum_{i=0}^{I-j} C_{ij}^{Pa}} = \widehat{q}_j^{\,-1}$$

$$\widehat{Var}(Q_{ij}|\mathcal{D}_j^{In}) = \Big(\sum_{i=0}^{I-j} C_{ij}^{In} - \frac{\sum_{i=0}^{I-j}(C_{ij}^{In})^2}{\sum_{i=0}^{I-j} C_{ij}^{In}}\Big)^{-1} \sum_{i=0}^{I-j} C_{ij}^{In}(Q_{ij} - \widehat{q}_j)^2$$

and

$$\widehat{Var}(Q_{ij}^{-1}|\mathcal{D}_j^{Pa}) = \Big(\sum_{i=0}^{I-j} C_{ij}^{Pa} - \frac{\sum_{i=0}^{I-j}(C_{ij}^{Pa})^2}{\sum_{i=0}^{I-j} C_{ij}^{Pa}}\Big)^{-1} \sum_{i=0}^{I-j} C_{ij}^{Pa}(Q_{ij}^{-1} - \widehat{q_j^{-1}})^2$$

for all $0 \le i \le I$ and $0 \le j \le J-1$.

Lemma 7.3 *Under the assumption that* $E[Q_{ij}^{-1}|\mathcal{B}_j^{Pa}]$, $E[Q_{ij}|\mathcal{D}_j^{In}]$, $\mathrm{Cov}[Q_{ij}^{-1}|\mathcal{D}_j^{Pa}]$ *and* $\mathrm{Cov}[Q_{ij}|\mathcal{D}_j^{In}]$ *are constants depending only on* j *for* $0 \le j \le J$*, we have*

1. given $\mathcal{D}_j^{In}$, $\widehat{q_j}$ is an unbiased estimator for $E[Q_{ij}|\mathcal{D}_j^{In}]$ and $\widehat{Var}(Q_{ij}|\mathcal{D}_j^{In})$ is an unbiased estimator for $Var(Q_{ij}|\mathcal{D}_j^{In})$, and

2. given $\mathcal{D}_j^{Pa}$, $\widehat{q_j^{-1}}$ is an unbiased estimator for $E[Q_{ij}^{-1}|\mathcal{B}_j^{Pa}]$ and $\widehat{Var}(Q_{ij}^{-1}|\mathcal{D}_j^{Pa})$ is an unbiased estimator for $Var(Q_{ij}^{-1}|\mathcal{D}_j^{Pa})$.

There remains to estimate the $\widetilde{\mathcal{D}}_{j-1}-$measurable correlation coefficients λ^{Pa} and λ^{In}. They are estimated by

$$\widehat{\lambda^{Pa}} = \frac{\sum_{1\le i+j\le I} \widetilde{Q}_{i,j-1}^{-1}\widetilde{F}_{ij}^{Pa}}{\sum_{1\le i+j\le I}(\widetilde{Q}_{i,j-1}^{-1})^2} \quad \text{and} \quad \widehat{\lambda^{In}} = \frac{\sum_{1\le i+j\le I} \widetilde{Q}_{i,j-1}\widetilde{F}_{ij}^{In}}{\sum_{1\le i+j\le I}(\widetilde{Q}_{i,j-1})^2}$$

respectively with

$$\widetilde{Q}_{i,j-1} = \frac{Q_{i,j-1} - \widehat{q_{j-1}}}{\widehat{Var}(Q_{i,j-1}|\mathcal{D}_{j-1}^{In})^{1/2}} \quad \text{and} \quad \widetilde{Q}_{i,j-1}^{-1} = \frac{Q_{i,j-1}^{-1} - \widehat{q_{j-1}^{-1}}}{\widehat{Var}(Q_{i,j-1}^{-1}|\mathcal{D}_{j-1}^{In})^{1/2}}$$

$$\widetilde{F}_{i,j-1}^{Pa} = \frac{C_{i,j-1}^{Pa} - \widehat{f_{j-1}^{Pa}}C_{i,j-1}^{Pa}}{\widehat{\sigma_{j-1}^{Pa}}(C_{i,j-1}^{Pa})^{1/2}} \quad \text{and} \quad \widetilde{F}_{i,j-1}^{In} = \frac{C_{i,j-1}^{In} - \widehat{f_{j-1}^{In}}C_{i,j-1}^{In}}{\widehat{\sigma_{j-1}^{In}}(C_{i,j-1}^{In})^{1/2}}$$

Hence the predictors of the MCL method are given by

Estimator 7.1 *The MCL estimators are given iteratively by*

$$\widehat{E}[C_{ij}^{Pa}|\widetilde{\mathcal{D}}_I] = \widehat{E}[C_{i,j-1}^{Pa}|\widetilde{\mathcal{D}}_I]\Big(\widehat{f_{j-1}^{Pa}} + \widehat{\lambda^{Pa}}\frac{\widehat{\sigma_{j-1}^{Pa}}}{\widehat{E}[C_{i,j-1}^{Pa}|\widetilde{\mathcal{D}}_I]^{1/2}}\widetilde{Q}_{i,j-1}^{-1}\Big)$$

$$\widehat{E}[C_{ij}^{In}|\widetilde{\mathcal{D}}_I] = \widehat{E}[C_{i,j-1}^{In}|\widetilde{\mathcal{D}}_I]\Big(\widehat{f_{j-1}^{In}} + \widehat{\lambda^{In}}\frac{\widehat{\sigma_{j-1}^{In}}}{\widehat{E}[C_{i,j-1}^{In}|\widetilde{\mathcal{D}}_I]^{1/2}}\widetilde{Q}_{i,j-1}^{-1}\Big)$$

for $i+j > I$, where we set $\widehat{E}[C_{i,I-i}^{Pa}|\widetilde{\mathcal{D}}_I] = C_{i,I-i}^{Pa}$ and $\widehat{E}[C_{i,I-i}^{In}|\widetilde{\mathcal{D}}_I] = C_{i,I-i}^{In}$.

References

[1] Alba de, E. (2002). *Bayesian estimation of outstanding claim reserves.* North American Actuarial J. 6/4, 1-20.

[2] Barnett, G. and Zehnwirth, B. (2000). *Best estimates for Reserves.* Proc. CAS, Vol.LXXXVII, 245-321.

[3] Benktander, G. (1976). An approach to credibility in calculating IBNR for casualty excess reinsurance. The Actuarial Review, April 1976, page 7.

[4] Bornhuetter, R. L., and Ferguson, R. E. (1972). *The actuary and IBNR.* Proc. CAS 59, 181-195.

[5] Brehm, P. J. (2002). *Correlation and the aggregation of unpaid loss distributions.* CAS Forum (Fall), 1-24.

[6] 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.

[7] Bühlmann, H. and Gisler, A. (2005), *A course in credibility theory and its applications.* Springer, Berlin.

[8] England, P. D. and Verrall, R, J. (1999). *Analytical and bootstrap estimates of prediction error in claim reserving.* Insurance: Math. Econom. 25/3, 281-293.

[9] England, P. D. and Verrall, R, J. (2001). *A flexible framework for stochastic claim reserving.* Proc. CAS, Vol. LXXXVIII, 1-38.

[10] England, P. D. and Verrall, R, J. (2002). *Stochastic claim reserving in general insurance.* British Actuarial J. 8/3, 443-518.

[11] England, P. D. and Verrall, R, J. (2007). *Predictive distributions of outstanding liabilities in general insurance.* Ann. Actuarial Science 1/2, 221-270.

[12] Gisler, A. and Wüthrich, M. V. (2007). *Credibility for the chain ladder reserving method.* Conference paper, 37th Astin Colloquium 2007, Orlando, USA.

[13] Hertig, J. (1985). *A statistical approach to the IBNR-reserves in marine insurance.* ASTIN Bull. 15/2, 171-183.

[14] Hess, K. T., Schmidt, K.D. and Zocher, M. (2006). *Multivariate loss prediction in the multivariate additive model.* Insurance: Math. Econom. 39/2,185-191.

[15] Jong De, P. (2006). *Forecasting runoff triangles.* North American Actuarial J. 10/2,28-38.

[16] Mack, T. (1993). *Distribution-free calculation of the standard error of chain-ladder reserve estimates.* ASTIN Bull. 23, 213-225.

[17] Mack, T. (2000). *Credible claim reserves: the Benktander method.* Astin Bulletin 30 (2), 333-347.

[18] Merz, M. and Wüthrich, M. V. (2006). *A credibility approach to the Munich chain-ladder method.* Blätter DGVFM XXXVII, 619-628.

[19] Merz, M. and Wüthrich, M. V. (2009a). *Prediction error of the multivariate addtive loss reserving method for dependent lines of business.* Variance vol 3(1), 131-151.

[20] Merz, M. and Wüthrich, M. V. (2009b). *Combining chain ladder and additive loss reserving method for dependent lines of business.* Variance vol 3(2),270-291.

[21] Murphy, D. M. (1994). *Unbiased loss development factors.* Proc. CAS, vol LXXXI. 154-222.

[22] Neuhaus, W. (1992). *Another pragmatic loss reserving method or Bornhuetter-Ferguson revisited.* Scand. Actuarial J. 2, 151-162.

[23] Pröhl, C. and Schmidt, K. D. (2005). *Multivariate chain-ladder.* Dresdner Schriften zur Versicherungs mathematik 3/2005.

[24] Quary, G. and Mack, T. (2004). *Munich chain ladder.* Blätter DGVFM XXVI,597-630.

[25] Renshaw, A. E. (1995). *Claim reserving by joint modelling.* Astin Colloquium 1995, Leuven, Belgium.

[26] Renshaw, A. E. and errall, R, J.(1998). *A stochastic model underlying the chain ladder technique.* British Actuarial J. 4/4, 903-923.

[27] Schmidt, K. D. (2006a). *Methods and models of loss reserving based on run-off triangles: a unifying survey.* CAS Forum (Fall), 269-317.

[28] Schmidt, K. D. (2006b). *Optimal and additive loss reserving for dependent lines of business.* CAS Forum (Fall), 319-351.

[29] Taylor, G. C. (1987). *Regression model in claims analysis I: theory.* Proc. CAS, Vol. LXXIV, 354-383.

[30] Taylor, G. (2000). *Loss reserving: an actuarial perspective.* Kluwer Academic Publishers, Boston.

[31] Taylor G. and McGuire, G. (2005). *Synchronous bootstrapping of seemingly unrelated regressions.* Conference papaer, 36th Astin Colloquium 2005, Zürich, Switzerland.

[32] Taylor G. and McGuire, G. (2007). *A synchronous bootstrap to account for dependencies between lines of business in the estimation of loss reserve prediction error.* North American Actuarial J. 8/1, 37-44.

[33] Verrall, R, J. (1990). *Bayes and empirical bayes estimation for the chain ladder method.* Asin Bulletin 20/2, 217-243.

[34] Verrall, R, J. (2004). *A bayesian generalized linear model for the Bornhuetter-Ferguson method of claim reserving.* North American Actuarial J. 8/3, 67-89.

[35] Vylder De, F. (1982). *Estimation of claim IBNR claims by credibility theory.* Insurance: Math. Econom. 1/1, 35-40.

[36] Wüthrich, M.V. (2008). Prediction error in the chain ladder method. *Insurance: Mathematics and Economics*, **42**, 378-388.

[37] Wüthrich M.V. and Merz, M. (2008). *Stochastic claims reserving methods in insurance.* John Wiley & Sons.

[38] Verdonck, T., Van Wouwe, M. and Dhaene, A. (2009). A robustification of the chain-ladder method. *North American Actuarial Journal*, **13** (2): 280-298.