Stockholms universitet

Analysis of stochastic claims reserving uncertainty

Nicole Li

Masteruppsats i matematisk statistik
Master Thesis in Mathematical Statistics

Masteruppsats 2017:4
Försäkringsmatematik
Maj 2017

www.math.su.se

Matematisk statistik
Matematiska institutionen
Stockholms universitet
106 91 Stockholm

Matematiska institutionen

Mathematical Statistics
Stockholm University
Master Thesis **2017:4**
http://www.math.su.se

Analysis of stochastic claims reserving uncertainty

Nicole Li*

May 2017

Abstract

Traditionally, the claims reserves are settled by using deterministic methods such as chain-ladder method. However, the question of statistical quality of the reserve estimates can not be answered unless a model is found. The goal of this thesis is to analyze the uncertainty of the reserve which comes from a stochastic model called Over-dispersed Poisson model. To achieve the purpose, a data-set of one product for the past 24 years' period is used. The analysis is done by checking the underlying assumptions of the stochastic reserving model, testing different distributions and investigating the ultimate risk and the one-year reserve risk by three simulation methods. Based on the real data and the results of the back test, we can conclude that the Over-dispersed Poisson model works well. The results in all three risk analyses indicate that the Over-dispersed Poisson model works as expected when the underlying assumptions are fulfilled.

*Postal address: Mathematical Statistics, Stockholm University, SE-106 91, Sweden. E-mail: nicole.li0927@gmail.com. Supervisor: Mathias Lindholm.

Contents

1 Introduction

Insurance industries can be classified into two categories: life insurance and non-life insurance. There are different characteristics between life and non-life insurance products, such as terms of contracts, type of claims, risk drivers, etc. A typical non-life insurance claim can be illustrated in a process as in which an accident occurred at a time in the insurance period and then reported some time later at the reporting date and finally the claim being settled at a further later time point by closing date. For certain cases, due to certain reasons, for example, new evidence is discovered for a settled claim, it leads that the closed claim has to be reopened (Huang & Wu, 2012).

Claims reserve is money that is earmarked for the eventual claim payment. It is a future obligation of an insurance company. For the purpose of valuation and financial reporting, the insurers predict the ultimate payment amount for all the claims. This includes estimates of both Incurred But Not Reported (IBNR) and Reported But Not Settled (RBNS) claims. It is essential to understand the importance of setting the right claim cost since the reserve is classified as a liability in the company balance sheet and a reserve which is wrongly estimated can represent a false picture of the company's financial state to shareholder and investors. If the reserve is underestimated the insurance company will not be able to fulfil its undertakings and if it is overestimated the insurance company unnecessarily holds the excess capital instead of using it for other purposes, e.g. for investments with potentially higher return.

1.1 Problem statement and Purpose

Traditionally, the claims reserves are estimated by using deterministic methods. The chain-ladder method can be considered as the most famous deterministic method which is described as a mechanical algorithm rather than a full statistical model. The biggest advantage of this approach is its simplicity and transparency even for non-actuaries who are involved in the reserving process. However, the question of statistical quality of the reserve estimates can not be answered unless a statistical model is found underlying the algorithm.

Within the stochastic claims reserving area, many models have been discussed and applied widely in the academic literature. Each model requires different assumptions. Generalized linear model and bootstrapping of deterministic model are popular techniques in common use. The purpose of this thesis is to analyze the uncertainty of the estimated reserves which come from a stochastic model called Over-dispersed Poisson model for a Swedish insurance company. Moreover, the study aims to analyze the underlying assumptions of this stochastic model based on given data and also to investigate the ultimate risk and the one-year reserve risk where the risk estimation follows naturally from the stochastic model constructions.

1.2 Structure of the thesis

A literature review is conducted together with a presentation of the theory is given in section 2. The methodologies used in this thesis are introduced in section 3. In section 4, the empirical analysis and results are presented. Discussion and conclusion are summarized in section 5. Finally, limitations and future researches will be found in the last section.

1.3 Data material

The data used in this study is taken from a Swedish insurance company. It contains the real payments, the amount of reserves for RBNS-claims and the number of reported claims for one product for the past 24 years disposed in the traditional way according to accident year and development period. For the reason of confidentiality, the data has been distorted by multiplying all the values with the same constant. There are around 1,867,187 events in the data set excluding zero claims. In extensions to these claims we also use information from around 196,777 zero claims. Zero claims are those claims that contain neither payments nor outstanding reserves.

2 Theory

2.1 Terminology

This section introduces some basic terminologies used within insurance industry.

There are two types of outstanding claims which need to be estimated for claims reserving, RBNS (Reported But Not Settled) and IBNR (Incurred But Not Repported). RBNS is the claim that insurance has been aware of. The payment has started but the full payment has not been carried out yet. The calculation of RBNS takes consideration of the payment pattern according to the insurance company. It also requires an understanding of where claims are in the settlement process. IBNR is the event that the actual losses have been incurred, but have not officially been reported. Actuaries will estimate the potential damages to the event. Based on the actuaries' estimations the insurance company may decide to set up reserves to allocate liquid assets to those expected losses.

Ultimate loss is the total sum of insurer's payments for a fully developed loss to the insured (i.e. paid losses plus outstanding reported losses (RBNS) and incurred but not reported (IBNR) losses). In order to estimate the IBNR claims, actuaries usually start with estimating the ultimate loss using different algorithms for example Chain-Ladder or Bornhuetter-Ferguson. From the estimated ultimate loss, the IBNR can be determined.

The term "claims reserve" refers to the expected value of the future payments from RBNS and IBNR i.e. $R_{tot} = R_{RBNS} + R_{IBNR}$.

2.2 Error types

The following describes different types of the uncertainty characteristics of the outstanding claims liabilities. These uncertainties can be divided into three categories (Daykin, Pentikäinen, & Pesosnen, 1994).

Parameter error – Due to the sampling variation of the past claims data and the limited quantity of observations, the exact values of parameters are unknown. Thus, the uncertainty of parameter estimation exists.

Process error (or stochastic error) – Due to the random fluctuations of the exact amounts of the future payments even in the situation that parameters and model are correct, the process error exists.

Model error – Due to the fact that the model adopted are uncertain and only approximations to the underlying claims development mechanism.

This paper focuses on measuring the parameter error, even called estimation variance, and the process error. The model error will be considered by testing different model assumptions when applying the reserving model to the same set of claims data. The sum of the process variance and the estimation variance is called prediction variance and is a measure of the variability of the prediction. The mean squared error of the prediction (MSEP) is used for prediction variance and is defined by England and Verrall (2002) as

$$MSEP_i = E[(R_i - \hat{R}_i)^2] = Var(R_i) + Var(\hat{R}_i) \tag{1}$$

where $\hat{R}_i$ is the estimator of R_i.

In addition, $Var(R_i)$ allows for the process error and $Var(\hat{R}_i)$ allows for the parameter error.

2.3 Chain Ladder algorithm

This part is supposed to give a short description of the chain-ladder algorithm which is presented by Mack (1993a). The classical reserving method is based on the triangle of the aggregate data. Let i denote the year that accident occurs, i.e. the accident year and j denote the development year. The data is aggregate by accident year and development year. An example of run-off triangle on the aggregate data is

Known				Future payments		
$C_{1,1}$	$C_{1,2}$	$C_{1,3}$	$C_{1,4}$			
$C_{2,1}$	$C_{2,2}$	$C_{2,3}$				$C_{2,4}$
$C_{3,1}$	$C_{3,2}$				$C_{3,3}$	$C_{3,4}$
$C_{4,1}$				$C_{4,2}$	$C_{4,3}$	$C_{4,4}$

where $C_{i,j}$ denotes the total cumulative claims amount from accident year i until development year j. The claims in the north-western triangle are known values.

The idea of the chain-ladder algorithm is that all accident years have the similar behaviours and the cumulative claims amounts are approximately the claims amounts at the last development year multiplying a factor i.e

$$C_{i,j+1} = f_j C_{i,j} \tag{2}$$

The factor f_j is called chain-ladder (CL) factors, age-to-age factors or link ratios and is calculated as

$$\hat{f}_j = \frac{\sum_{i=1}^{I-j-1} C_{i,j+1}}{\sum_{i=1}^{I-j-1} C_{i,j}} \tag{3}$$

With CL factors we predict the ultimate claim $C_{i,J}$ for $i + J > I$ at time $t = I$ by

$$\hat{C}_{i,J} = C_{i,I-i} \prod_{j=I-i}^{J-1} \hat{f}_j \tag{4}$$

Thus, the reserve at time $t = I$ for accident year $i > I - J$ is

$$\hat{R}_i = \hat{C}_{i,J} - C_{i,I-i} \tag{5}$$

By aggregating over all accident years the outstanding loss liability of the past exposure claims is

$$\hat{R} = \sum_{i=I-J+1}^{J} \hat{R}_i \tag{6}$$

2.4 Negative incremental claims

Sometimes there exists negative values in the development triangle of incremental claim amounts. Typically they are consequences of the already reported claims which are cancelled due to the initial overestimation of the loss or the insurance company receives payments from third parties (Gould, 2008). Some methods of claims reserving may produce reserve estimates even there exists negative values. However, for some methods, the results would be unreliable when the negative increment values involves.

There has been suggested different solutions on how to handle the negative incremental claims. One method involves adding one positive constant on all increment claims. The constant needs to be subtracted after the analysis is completed. This method provides suitable results as long as there are not too many negative claims (Gould, 2008). However, this procedure makes the variability of the result depend on the constant added earlier. Another method to adjust the negative payments is to cancel matching positive amounts if there exists. If no matching positive amounts exist then the amount is subtracted from previous payments.

2.5 Claims development result

Wüthrich, Merz & Lysenko (2007) and Wüthrich (2015) give a definition of the claims development result.

$$CDR_{i,t+1} = R_i^t - (X_i^{t+1} + R_i^{t+1}) = \hat{C}_{i,J}^t - \hat{C}_{i,J}^{t+1}. \tag{7}$$

where R_i^t is the best estimate reserves at time $t \geq I$ for accident year $i > t-J$. $\hat{C}_{j,J}^t$ is the ultimate claim of accident year i at time $t \geq I$. $X_{i,t-i+1}$ is the payment between t and $t+1$.

If the claims development result is positive we have a gain because we have overestimated the outstanding loss liabilities at time t, otherwise we have a loss.

The uncertainty in this position measured by MSEP with given information

D:

$$MSEP_{CDR_{i,t+1}|D} = E[(CDR_{i,t+1} - \widehat{CDR}_{i,t+1})^2|D] \tag{8}$$

3 Methodology

3.1 Stochastic models associated with the CL algorithm

The most weakness of the chain-ladder method is that it is a deterministic algorithm hence the uncertainty of the reserve estimate can not be determined. To amend this shortcoming, the stochastic models associated with the chain-ladder algorithm have been developed. The variation of the estimate from the models is possible to be analyzed. In this section, the stochastic models underlying the chain-ladder algorithm will be presented as well as the underlying assumptions for each model. In this section we also aim to show that these stochastic models indeed provide the same estimates as the chain-ladder method.

3.1.1 Mack's distribution-free chain-ladder model

Thomas Mack (1993a) proposed a distribution free stochastic model which is probably the most common stochastic reserving method. The model produces equivalent results to the chain-ladder algorithm. Recall, $C_{i,j}$ represents the total claims amount. The basic assumptions underlying the stochastic Mack method are the same as the ones for the deterministic chain-ladder.

Model Assumptions:
CL1:
There exists development factors $f_1, ..., f_{I-1} > 0$ so that

$$E[C_{i,j+1} \mid C_{i,1}, ..., C_{i,j}] = f_j C_{i,j} \tag{9}$$

where I is the total development year and the CL factor $E[\hat{f}_j] = f_j$ is calculated by formula (3) above. The CL factor $\hat{f}_j$ is the same for all accident years within development year j.

CL2:

$$\{C_{i,1},...,C_{i,I}\},\{C_{k,1},...,C_{k,I}\},\ i\neq k,\ \text{are independent.} \tag{10}$$

CL3:

There exists $\alpha_j > 0$ so that

$$Var(C_{i,j+1}|C_{i,1},...,C_{i,j}) = C_{i,j}\alpha_j^2 \tag{11}$$

where

$$\hat{\alpha}_j^2 = \frac{1}{I-j-1}\sum_{i}^{I-j} C_{i,j}\left(\frac{C_{i,j+1}}{C_{i,j}} - \hat{f}_j\right)^2 \quad 1 \leq j \leq I-2 \tag{12}$$

is an unbiased estimator of $\alpha_j^2, 1 \leq j \leq I-2$. For estimating α_{I-1}^2 Mack suggests

$$\hat{\alpha}_{I-1}^2 = \min(\hat{\alpha}_{I-2}^4/\hat{\alpha}_{I-3}^2, \min(\hat{\alpha}_{I-3}^2, \hat{\alpha}_{I-2}^2)). \tag{13}$$

As with equation (5) the reserve is given by

$$R_i = C_{i,I} - C_{i,I+1-i}, \tag{14}$$

Further, Mack (1993a) defines the mean squared error of the estimator $\hat{C}_{i,I}$ of $C_{i,I}$ by $mse(\hat{C}_{i,I})$, note that this differs from MSEP from equation (1), which is defined to be

$$\begin{aligned} mse(\hat{R}_i) = mse(\hat{C}_{i,I}) &= E[(\hat{C_{i,I}} - C_{i,I})^2|D] \\ &= Var(C_{i,I}|D) + (E[C_{i,I}|D) - \hat{C}_{i,I})^2 \end{aligned} \tag{15}$$

where $D = \{C_{ik}|i+k \leq I+1\}$, i.e the set of data observed so far. The first term $Var(C_{i,I}|D)$ represents the process risk and the second term $(E[C_{i,I}|D) - \hat{C}_{i,I})^2$ represents the parameter risk.

Formulas used to check the chain-ladder Assumptions

Spearman's rank correlation coefficient

Let r_{ik}, $1 \leq i \leq I-k$, denote the rank of $C_{i,k+1}/C_{ik}$ obtained in this way, $1 \leq r_{ik} \leq I-k$. Then doing the same with the preceding development factors $C_{ik}/C_{i,k-1}$, $1 \leq i \leq I-k$, leaving out $C_{I-+1-k,k}/C_{I+1-k,k-1}$ for which the subsequent development factor has not yet been observed. Let s_{ik}, $1 \leq i \leq I-k$, be the ranks obtained in this way $1 \leq r_{ik} \leq I-k$. Then the Spearman's rank correlation coefficient T_k is defined as

$$T_k = 1 - 6 \sum_{i=1}^{I-k} (r_{ik} - s_{ik})^2 / ((I-k)^3 - I + k) \tag{16}$$

and the $(I-k-1)$-weighted average of T_ks is defined as

$$\begin{aligned} T &= \sum_{k=2}^{I-2} (I-k-1) T_k / \sum_{k=2}^{I-2} (I-k-1) \\ &= \sum_{k=2}^{I-2} \frac{I-k-1}{(I-2)(I-3)/2} T_k \end{aligned} \tag{17}$$

3.1.2 Poisson and over-dispersed Poisson model

Renshaw & Verrall (1998) noticed the link between the chain-ladder algorithm and Poisson distribution. They were the first to implement the model using standard methodology for a full statistical review. The model is in form of a generalized linear model with a log-link function. Let $y_{i,j}$ denote the incremental claims, remember that the total claims amount $C_{i,j} = \sum_{k=1}^{j} y_{i,j}$. An example run-off triangle of incremental claims:

$y_{1,1}$ $y_{1,2}$ $y_{1,3}$ $y_{1,4}$

$y_{2,1}$ $y_{2,2}$ $y_{2,3}$

$y_{3,1}$ $y_{3,2}$

$y_{4,1}$

Model Assumptions:
The Over-dispersed Poisson model assumes that the incremental claims amounts $y_{i,j}$ are independent Over-dispersed Poisson distributed with mean m_{ij} and variance ϕm_{ij}. The scale parameter ϕ is unknown and estimated from the data. The mean can be written with a log-link function.

$$\log m_{ij} = \mu + \alpha_i + \beta_j \tag{18}$$

where α_1=β_1=0, $i = 1, ..., I$ connects to the accident years and $j = 1, ..., J$ connects to the development years.

Further, the model also assumes that the sum of the incremental claims in column j is non-negative. Note that some negative incremental claims are allowed. Mathematically,

$$\sum_{i=1}^{n-j+1} y_{ij} \geq 0 \quad \text{for all } j \tag{19}$$

The estimates of the parameters in the model given in (18) and (19) can be obtained by MLE (Wüthrich, 2015). The estimates of the aggregate payments of claims in each accident year i can be obtained from the sums:

$$\hat{C}_{in} = C_{i,n-i+1} + \sum_{j=n-i+2}^{n} e^{\hat{\mu}+\hat{\alpha_i}+\hat{\beta_j}} (i = 1, 2, ..., n) \tag{20}$$

where $\hat{\mu}$, $\hat{\alpha_i}$ and $\hat{\beta_j}$ are the maximum likelihood estimates of the parameters.

Further, we prove that the generalized linear model with Poisson distribution gives exactly the same results as the chain-ladder algorithm under assumptions.

Let $p_{(i),j}$ denote the probability that a claim with accident year i is reported in development year j. Then

$$p_{(i),j} = \frac{p_j}{\sum_{k=1}^{n-i+1} p_k} \tag{21}$$

where p_j is the probability that a claim is reported in development year j and $\sum_{k=1}^{n} p_k = 1$. Note that the right side in equation (21) is independent of i. This last assumption implies that we assume that all claims are reported by the end of development year n.

Using the multinomial distribution, the conditional likelihood L_c is given by

$$L_c = \prod_{i=1}^{n} \left(\frac{C_{i,n-i+1}!}{\prod_{j=1}^{n-i+1} y_{ij}!} \prod_{j=1}^{n-i+1} p_{(i)j}^{y_{ij}}\right) \tag{22}$$

and consider the estimate for the accident year which has been reported up to development year index j, the estimate of ultimate cumulative claims for accident year $n-j+1$ is

$$\hat{C}_{n-j+1,n} = \frac{C_{n-j+1,j}}{1-\sum_{k=j+1}^{n}\hat{p}_k}. \tag{23}$$

This can be compared with the chain-ladder estimates:

$$\hat{C}_{n-j+1,n} = C_{n-j+1,j}\hat{\lambda}_{j+1}\hat{\lambda}_{j+2}...\hat{\lambda}_n \tag{24}$$

where

$$\hat{\lambda}_j = \frac{\sum_{i=1}^{n-j+1} C_{ij}}{\sum_{i=1}^{n-j+1} C_{i,j-1}}. \tag{25}$$

Rosenberg (1990) derived a technique for obtaining the estimates $\hat{p}_j$ which is recursive.

$$\begin{aligned}\hat{p}_j &= \frac{y_{1j}+y_{2j}+...+y_{n-j+1,j}}{C_{1n}+\frac{C_{2,n-1}}{1-\hat{p}_n}+...+\frac{C_{n-j+1}}{1-\hat{p}_{j+1}-...-\hat{p}_n}} \\ &= \frac{y_{1j}+y_{2j}+...+y_{n-j+1.j}}{C_{1n}+\hat{C}_{2,n}+...+\hat{C}_{n-j+1,n}}\end{aligned} \tag{26}$$

Note that $\hat{p}_n = \frac{y_{in}}{C_{1n}}$.

From equations (23) and (24) it can be seen that

$$\hat{\lambda}_{j+1}\hat{\lambda}_{j+2}...\hat{\lambda}_n = \frac{1}{1-\hat{p}_{j+1}-\hat{p}_{j+2}-...-\hat{p}_n} \tag{27}$$

and

$$\hat{\lambda}_j\hat{\lambda}_{j+1}...\hat{\lambda}_n = \frac{1}{1-\hat{p}_j-\hat{p}_{j+1}-...-\hat{p}_n} \tag{28}$$

Thus

$$\hat{\lambda}_j\hat{\lambda}_{j+1}...\hat{\lambda}_n = \frac{1}{\frac{1}{\hat{\lambda}_{j+1}\hat{\lambda}_{j+2}...\hat{\lambda}_n}-\hat{p}_j} \tag{29}$$

and

$$\hat{\lambda}_j = \frac{1}{1-\hat{p}_j\hat{\lambda}_{j+1}\hat{\lambda}_{j+2}...\hat{\lambda}_n} \tag{30}$$

For $j = n$ and since

$$\begin{aligned}
\hat{\lambda}_n &= \frac{1}{1-\hat{p}_n} \qquad \text{from (30)} \\
&= \frac{1}{1-\frac{y_{1n}}{C_{1n}}} \qquad \text{from (26)} \\
&= \frac{C_{in}}{C_{1n}-y_{1n}} \\
&= \frac{C_{1n}}{C_{1,n-1}}
\end{aligned} \tag{31}$$

We rewrite the equation (26) as

$$\hat{p}_j = \frac{y_{1j} + y_{2j} + ... + y_{n-j+1,j}}{C_{1n} + C_{2,n-1}\hat{\lambda}_n + ... + C_{n-j+1,j}\hat{\lambda}_{j+1}...\hat{\lambda}_n} \tag{32}$$

Substituting into equation (30)

$$\hat{\lambda}_j = \frac{1}{1 - \frac{y_{1j}+y_{2j}+...+y_{n-j+1,j}}{C_{1n}+C_{2,n-1}\hat{\lambda}_n+...+C_{n-j+1,j}\hat{\lambda}_{j+1}...\hat{\lambda}_n}\hat{\lambda}_{j+1}\hat{\lambda}_{j+2}...\hat{\lambda}_n} \tag{33}$$

and

$$C_{in} + C_{2,n-1}\hat{\lambda}_n + ... + C_{n-j+1,j}\hat{\lambda}_{j+1}...\hat{\lambda}_n = \hat{\lambda}_{j+1}...\hat{\lambda}_n \sum_{i=1}^{n-j+1} C_{ij}. \tag{34}$$

Hence, $\hat{\lambda}_j$ can be rewritten as

$$\begin{aligned}
\hat{\lambda}_j &= \frac{\sum_{i=1}^{n-j+1} C_{ij}}{\sum_{i=1}^{n-j+1} C_{ij} - (y_{1j} + y_{2j} + ... + y_{n-j+1,j})} \\
&= \frac{\sum_{i=1}^{n-j+1} C_{ij}}{\sum_{i=1}^{n-j+1} C_{i,j-1}}
\end{aligned} \tag{35}$$

which is the chain-ladder estimate.

England & Verrall (2002) give a formula for MSEP which is

$$\begin{aligned}
MSEP = &\sum_{i,j\in\Delta} \phi\hat{m}_{ij} + \sum_{i,j\in\Delta} \hat{m}_{ij}^2 Var(\hat{\eta}_{ij}) \\
&+ 2\sum_{i_1j_1\in\Delta, i_2j_2\in\Delta} \hat{m}_{i_1j_1}\hat{m}_{i_2j_2} Cov(\hat{\eta}_{i_1j_1}, \hat{\eta}_{i_2j_2}).
\end{aligned} \tag{36}$$

where ϕ is the dispersion parameter and $\hat{m}_{ij} = \exp(\hat{\eta}_{ij})$. Note that ϕm_{ij} corresponds to the plug-in estimation of the theoretical process variance according to the Over-dispersed Poisson model. The rest of expression gives the estimation error.

3.2 Predictive distribution of claims reserve

In this part we aim to present simulation method for finding the whole predictive distribution of the reserve. There are two main methods which are currently used for simulation, bootstrapping and Bayesian methods using Markov Chain Monte Carlo techniques. In this paper only the bootstrapping method will be adopted.

Bootstrapping is a simple simulation-based approach for obtaining the information from a sample of data without any knowledge about the underlying data. By randomly drawing, with replacement, from observed data, a new data set which is consistent with the underlying distribution will be created. Bootstrapping is useful only when the underlying model is correctly fitted to the data. Bootstrapping is applied to the data which are required to be independent and identically distributed. In this thesis, we combine bootstrapping method and claims payments' distribution since the distribution is a requirement in our researched model.

3.3 Best Estimate

Best estimate is a term used in insurance companies. When we sum over all the expected payments from RBNS and IBNR claims we get an estimate of the corresponding reserves, i.e. R_{RBNS} and R_{IBNR}. The total reserve R_{tot} is the sum of R_{RBNS} and R_{IBNR}.

3.4 Ultimate risk

The ultimate risk presents the risk of the ultimate loss, in other words, the uncertainty of the sum of the paid loss, RBNS loss and IBNR loss. With the simulation method, we obtain a matrix of claims losses. The true total ultimate reserve R can be determined from the matrix. Further we estimate the claims reserve $\hat{R}$ using the stochastic model. The ultimate risk is

$$U_k = \hat{R}_k - R_k \tag{37}$$

where U_k is the outcome in simulation number k, $k = 1, ..., B$. B is the total number of simulations.

3.5 One-year-reserve risk

Quantifying uncertainty with ranges of possible values and associated probabilities (i.e., with probability distributions) helps everyone understand the risks involved. Wüthrich, Merz & Lysenko (2007) and Wüthrich (2015) publish papers about the one-year reserve risk. They perform an analytic formula for the MSEP of the claims development result, see formula (8). Here we present the work which is done by Ohlsson and Lauzeningks (2009) with a general simulation approach. This approach will be adopted in this analysis.

Let X^{t+1} denote the ultimo cost which consists all payments made during time t and $t+1$ for all accident years, R^t denote the opening reserve at the time t for all accident years and R^{t+1} is the closing reserve estimate at time $t+1$ for all accident years. Then the technical run-off result is given by

$$T = R^t - X^{t+1} - R^{t+1} \tag{38}$$

This is same as the CDR definition given in equation (7).

In the Solvency II directive the solvency capital requirement (SCR) for the standard formula is based on the metric set as 99,5% Value-at-Risk over a one-year horizon.

$$p = Pr[L(T) \leq VaR] = 99.5\% \tag{39}$$

This means that the SCR of the reserve is the 99,5% quantile of the loss distribution L of T, which is defined as $L(T) = T$. We use simulation method to obtain the probability distribution by the following three steps.

Step 1: The opening reserve

We start by estimating R^t, here the opening reserve is not considered stochastic.

Step 2: Generating the new year
The next step is to generate the claims paid in the next year for each accident year i.e. we need to generate a new diagonal in the development triangle from a payment distribution with estimated parameters. By take the sum along the new diagonal we get the ultimo cost denoted X^{t+1}.

Step 3: The closing reserve
Estimate the reserve R^{t+1} with reserving method at time point $t+1$.

3.6 Simulation

This part describes approaches for simulations related to risk analysis. In order to test the assumption of payment distribution, we compare Poisson distribution with negative binomial distribution. Note that we simulate payments from different distributions but always fit an Over-dispersed Poisson model.

3.6.1 Method 1

Method 1 is related to the ultimate risk. We need to generate incremental claims payments. We seek a two dimension matrix, the dimensions correspond to the number of the accident years observed. Thus, our matrix is 10 $\times$ 10. We use the parameters in table 1 as the start value in the generalized linear model and simulate a payment matrix from Poisson distribution and negative binomial distribution respectively. The matrix is regarded as observed, thus the observed total ultimate "true" reserve R_k in simulation number k is the sum of the elements in the southeast triangle in the matrix. Further, we pretend that only northwest triangle is known and predict the southeast triangle with the Over-dispersed Poisson model. By sum up all payments in the new predicted southwest triangle, we obtain the estimated total claims reserve $\hat{R}_k$ and the ultimate risk U_k in equation (37) for simulation number k. Repeat the process B times.

3.6.2 Method 2

Method 2 is adopted in order to obtain the one-year reserve risk described in section 3.5.

Step 1
Here we still seek a 10×10 matrix of incremental payments. We still use the parameters in table 1 to simulate a matrix from Poisson distribution and negative binomial distribution respectively. By using the northwest triangle from the matrix we calculate the total claims reserve which is the sum of the reserves for accident years up to development year 10. This is $\hat{R}^0$ which is considered non-stochastic.

Step 2
Pretend that only northwest triangle is known and fit an Over-dispersed Poisson model. Simulate a new diagonal of payments from different distributions with the new estimated parameters. Let this correspond to simulation number i and denote the sum of payments with X_i^1.

Step 3
Given the northwest triangle and simulated payments for next year, predict the claims reserves for all accident year up to development year 10 with the Over-dispersed model. Denote the sum of the all reserves with R_i^1 corresponding to simulation number i.

Step 4
Calculate $T_i = R^0 - (X_i^1 + R_i^1)$.

Repeat step 2 to 4 B times and obtain a distribution for T by using the simulated $T_1, ..., T_B$. Moreover, from the simulated matrix in step 1, a true value of

$$T_{true} = R_{true}^0 - (X_{true}^1 + R_{true}^1) \tag{40}$$

can be determined.

3.6.3 Method 3

In the third method, we still need to estimate R^0, X^1 and R^1. To be different from method 2, the R^0 is not fixed.

Step 1
Simulate a 10×10 matrix by using the parameters in table 1 from Poisson distribution and negative binomial distribution respectively.

Step 2
Using the northwest triangle in the matrix to predict claims reserves for all accident years up to development year 10. Let this correspond to simulation number i and denote the sum of the claims reserves with R_i^0.

Step 3
From the simulated matrix in step 1 calculate the total payment for the next year and denote it with X_i^1.

Step 4
Given the northwest triangle and the payments for next year predict the claims reserves for all accident years up to development year 10. Denote the sum of all claims reserves with R_i^1.

Step 5
Calculate $T_i = R_i^0 - (X_i^1 + R_i^1)$

Repeat step 1 to 5 B times and obtain a distribution for T by using simulated $T_1, ..., T_B$. In this simulation method, we do not use a stochastic claims reserving method, but we have simulated data and see how the expected prediction of reserves behaves.

4 Analysis & Findings

The data set used in the analysis is a set of claims data for one product from a Swedish insurance company. The insurance claims are organized by accident years and development years. It contains the amount of payment for each reported claim, the amount of reserves for RBNS-claims and the number of reported claims for last 10 years. When fitting the stochastic chain-ladder models, only data containing the amount of payments will be used.

Checking the chain-ladder assumptions

We begin with checking the chain-ladder assumptions against the data following the processes presented by Mack (1993b). Here we use the last 10 years' not full-developed data. As has been pointed out before, the three basic implicit chain-ladder assumptions are:

(1) $E[C_{i,k+1} \mid C_{i1}, ..., C_{ik}] = C_{ik} f_k$
(2) Independence of accident years
(3) $Var(C_{i,k+1} \mid C_{i1}, ..., C_{ik}) = C_{ik}\alpha_k^2$

First, we look at the assumption (1) and (3) for a fixed k and for $i = 1, ...I$. There, $C_{ik}, 1 \leq i \leq I$, are considered as given non-random values and the equation of assumption (1) can be interpreted as an ordinary regression model of the type

$$y_i = a + bx_i + \epsilon_i, \tag{41}$$

where a and b are the regression coefficients and ϵ_i is the error term with $E[\epsilon_i] = 0$. In our special case, we have a=0, b = f_k i.e. development factor, $k = 1, ..., 8$. Further, we have observations of the independent variable $y_i = C_{i,k+1}$ at the points $x_i = C_{ik}$ for $i = 1, ..., I - k$. Therefore, we can estimate the regression coefficient $b = f_k$ by the usual least squares method.

We calculate the development factors $f_k = f_{k1}$ according to formula (3) together with the alternative development factors f_{k0} given by

$$f_{k0} = \frac{\sum_{i=1}^{I-k} C_{ik}C_{i,k+1}}{\sum_{i=1}^{I-k} C_{ik}^2} \tag{42}$$

and f_{k2} according to

$$f_{k2} = \frac{1}{I-k}\sum_{i=1}^{I-k}\frac{C_{i,k+1}}{C_{ik}} \tag{43}$$

The check of the linearity and the variance assumption is usually done by carefully inspecting plots of the data and the residuals, see figures 7 to 14:

1. $C_{i,k+1}$ against C_{ik}.
In order to see if there exists an approximately linear relationship around a straight line through the origin with slope f_k. (Assumption 1)

2. Weighted residuals $e_{ik} = \frac{C_{i,k+1}-C_{ik}\hat{f}_k}{\sqrt{C_{ik}}}$ against C_{ik}.
In order to see if the residuals do not show any specific trend but appear purely random. (Assumption 3)

Mack (1993a) recommends also to compare all three residual plots if the usual chain-ladder factors $f_k = f_{k1}$ need to be replaced by alternative factors:

Plot 0: $C_{i,k+1} - C_{ik}f_{k0}$ against C_{ik},
Plot 1: $(C_{i,k+1} - C_{ik}f_{k1})/\sqrt{C_{ik}}$ against C_{ik},
Plot 2: $(C_{i,k+1} - C_{ik}f_{k2})/C_{ik}$ against C_{ik},

Since the regression plots in figures 7 to 14 clearly show that the regression line with slope f_k captures the direction of the data points very well. The weighted residuals are quite random in the residual plots, we decide to keep the usual development factors $f_k = f_{k1}$ and do not show figures of Plot 0 and Plot 2 discussed above. To sum it up, we conclude that the first and the third assumptions of chain-ladder method are fulfilled.

Next, we can carry through the tests for accident year influences and for correlations between subsequent development factors which is associated with assumption (2). We start with the table of development factor which is calculated by $f_{ik} = \frac{C_{i,k+1}}{C_{i,k}}$, see table 6 in appendix. We perform r_ik and s_ik which denote the rank of $\frac{C_{i,k+1}}{C_{ik}}$ and $\frac{C_{ik}}{C_{i,k-1}}$ respectively. Furthermore, we obtain the rank factors r_{ik} and s_{ik} according to the size of the development factors, then leave out the last element and rank the column again. Results are summarized in table 7 and 8. By formula (16) and (17), the Spearman's rank correlation coefficient T_k and the $(I-k-1)$-weighted average of T_ks can be calculated, see also table 9. The results of the weighted average T is 0.012. The 50% confidence limits for T are ± 0.127. Thus, T is within its 50%-interval and the hypothesis of having uncorrelated development factors is not rejected.

To test for accident year effects first recall the development factors f_{ik}, see table 6. We have to subdivide each column F_k into the subset SF_k of 'smaller' factors below the median of F_k and into the subset LF_k of 'larger' factors above the median. This can be done with the help of the rank columns r_{ik} established above. Replacing a small rank with S, a large rank with L and a mean rank with $*$, see table 10. We count L_j and S_j, which present the number of "L"s and "S"s respectively for each diagonal. With notations $Z_j = min(L_j, S_j)$, $n = L_j + S_j$ and $m = (n-1)/2$ we obtain $E[Z_j] = \frac{n}{2} - \binom{n-1}{m}\frac{n}{2^n}$ and $Var(Z_j) = \frac{n(n-1)}{4} - \binom{n-1}{m}\frac{n(n-1)}{2^n} + E[Z_j] - (E[Z_j]^2)$, see table 11. The test statistic $Z = \sum Z_j = 8$ is not outside its 95%-range (9.25 - 2 $\times$ 1.615, 9.25 + 2 $\times$ 1.615) = (6.02, 12.48). Hence, the null-hypothesis of not having significant accident year influences is not rejected.

We have checked the chain-ladder assumptions against the data and it clearly confirms that the chain-ladder algorithm is appropriate for our data. Hence, it also indicates that stochastic models associated with the chain-ladder algorithm can be adopted in this analysis.

Model fitting

Since the three assumptions of chain-ladder technique are fulfilled for the data, the next step is to fit the Over-dispersed Poisson model on the same data-set as above. As has been mentioned in Methodology section, the Over-dispersed Poisson model underlies the chain-ladder technique, which means that the estimates of reserve come from an Over-dispersed Poisson model and Mack's distribution free model should be identical if the underlying assumptions of Over-dispersed model are fulfilled. When we fit the Over-dispersed

Poisson model on the data in this part, we assume that all underlying assumptions are fulfilled.

The maximum likelihood estimates of parameters are summarized in table 1. Parameters αs and βs relate to accident period and development period respectively. Dispersion parameter is taken to be 145340.

	Estimate	Std. Err	P-value
μ	19.42708	0.02238	<2e-16
α_2	0.18181	0.03023	6.67e-7
α_3	0.27660	0.02961	3.70e-11
α_4	0.35827	0.02911	1.82e-14
α_5	0.44514	0.02860	<2e-16
α_6	0.56361	0.02797	<2e-16
α_7	0.69214	0.02735	<2e-16
α_8	0.74864	0.02710	<2e-16
α_9	0.80569	0.02686	<2e-16
α_{10}	0.88105	0.02685	<2e-16
β_2	-2.73912	0.02447	<2e-16
β_3	-6.91683	0.20821	<2e-16
β_4	-8.60811	0.53242	<2e-16
β_5	-12.1989	3.57232	1.60e-3
β_6	-12.0342	3.70744	2.53e-3
β_7	-26.3276	3274.99	9.94e-1
β_8	-26.2750	3783.62	9.95e-1
β_9	-26.2116	4635.78	9.96e-1
β_{10}	-26.1176	6559.39	9.97e-1

Table 1: Results of parameter estimations

Recall the generalized linear model in equation (18), the future incremental payments can be obtained. The estimate of the total reserve is identical as Mack's model. We plot also the residuals $e_{ij} = C_{ij} - \hat{C}_{ij}$ and the weighted residuals $e_{ij} = \frac{C_{ij} - \hat{C}_{ij}}{\sqrt{C_{ij}}}$ of the northwest triangle in figures 1 and 2. X-axis corresponds the observations in the triangle. Residuals in figure 1 are quite random and tend to be zero when index is larger then 40. Weighted residuals are within a certain range and the most are near zero.

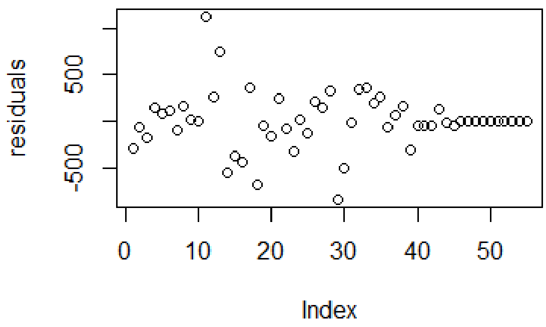

Figure 1: Residuals of incremental payment in northwest triangle

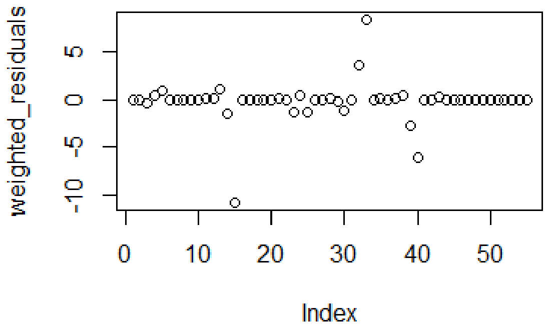

Figure 2: Weighted residuals of incremental payment in northwest triangle

Back test

To validate the model further we look at the performance of the Over-dispersed Poisson model when using prior years and then compare the payment outcomes with the true outcomes. We decide to take the first 13 years data with full development. To begin with, we fit the model one more time on the first 13 years data-set (the northwest triangle) and get new parameter estimations. Moreover, we predict the reserves for each accident year. The results are presented in table 2 and the reserves before accident year 7 are zero. MSEP is calculated by formula (1). The total reserve from the Over-dispersed Poisson model is overestimated by 6% and the standard error is 1.54e6.

Accident year i	Reserve (True outcomes)	Reserve (OPD)	MSEP
7	438	141	8.8e4
8	6000	647	2.9e7
9	1192	2856	2.8e6
10	6035	7288	1.6e6
11	57298	29386	7.8e8
12	36050	66385	9.2e8
13	12111000	12844425	5.4e11
Total	12218013	12951095	

Table 2: Comparison of predicted reserves to true outcomes.

Based on the real data, we can conclude that the Over-dispersed Poisson model works well.

Risk

We compare the ultimate risk and the one-year reserve risk by simulation methods described in section 3.6. The number of simulations is 15000 in all analysis. To analyze the assumption of Poisson distribution for the incremental payments, we assume also that the incremental payments are negative binomial distributed. Note that we simulate from Poisson and negative binomial distributions respectively, but always fit an Over-dispersed Poisson model. When we simulate from a Poisson distribution, all model assumptions are fulfilled. For simulations from the negative binomial distribution, which is also known as Poisson-Gamma mixture, we can view the negative binomial as a Poisson(λ) distribution, where λ is itself a random variable, distributed as a gamma distribution with shape and scale parameters. Thus, we first simulate from a Gamma distribution with the shape parameter equal to the squared expected value divided by the variance and the scale parameter equal to the variance divided by the expected value. Then we simulate from a Poisson distribution with the parameter λ.

By looking at the ultimate risk, we find that these two distributions give almost identical expected reserves. The mean value for Poisson distribution is closer to zero, which indicates that the Over-dispersed Poisson model behaves better when all assumptions are fulfilled. However, the standard deviation is lower for the negative binomial distribution compared to the Poisson distribution, see figure 3. It is interesting to compare ultimate risk from the simulation method with England and Verrall's formula (36). Note that the prediction error is a square root of MSEP. From table 3, it can be clearly seen that the ultimate risk from simulation method 1 is much smaller than the prediction error from the formula no matter which distribution is adopted. However, the percent value of prediction error itself is not extremely large to the total estimated reserve.

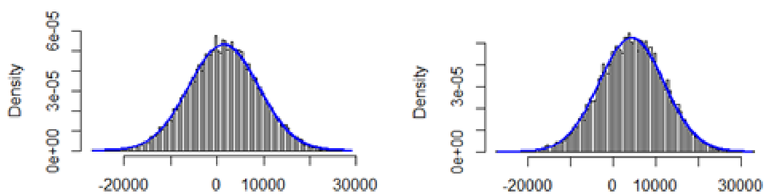

Figure 3: Distribution of ultimate risk, $\hat{R}-R$, simulated from Over-dispersed Poisson distribution (left) and negative binomial distribution (right).

<table>
<tr><td rowspan="2">England & Verrall's formula</td><td colspan="2">Reserve</td><td colspan="2">Prediction error</td><td colspan="2">Prediction error %</td></tr>
<tr><td colspan="2">44298604</td><td colspan="2">2837938</td><td colspan="2">6%</td></tr>
<tr><td rowspan="3">Ultimate risk (simulation metod 1)</td><td colspan="3">Poisson</td><td colspan="3">Negative-binomial</td></tr>
<tr><td>Reserve</td><td>Mean</td><td>Sd</td><td>Reserve</td><td>Mean</td><td>Sd</td></tr>
<tr><td>44295743</td><td>377</td><td>7589</td><td>44295745</td><td>4414</td><td>7521</td></tr>
</table>

Table 3: Results for MSEP and ultimate risk

Moreover, we obtain the one-year reserve risk according to formula (38). We follow the process that has been described in simulation method 2. The distributions of one-year reserve risk are presented in figure 4, results are summarized in table 4 and the empirical cumulative distribution function is in figure 5. The values of true outcome in table 4 are determined by formula (40). Poisson distribution has lower mean value, standard deviation and Value at Risk. Difference between the mean value and the true outcome is quite large for negative binomial distribution. Since the mean values from both distributions are non positive, we have reserve losses. In summary, the Over-dispersed Poisson model still behaves better when all assumptions are fulfilled.

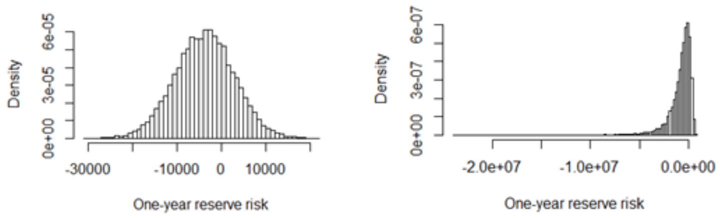

Figure 4: Distribution of one-year reserve risk simulated from Poisson distribution (left) and negative binomial distribution (right). Simulation method 2

	Poisson distribution			Negative-Binomial distribution		
True outcome	-1748			-1046069		
R^0	44298688			44298593		
	Mean	Sd	99.5%	Mean	Sd	99.5%
$E[R^0 - (\hat{X}^1 + \hat{R}^1)]$	-3641	6679	13280	--685920	968956	513944

Table 4: Results for the one-year reserve risk, simulation method 2

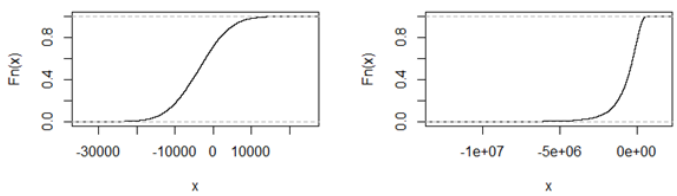

Figure 5: Empirical cumulative distribution function of one-year reserve risk simulated from Poisson distribution (left) and negative binomial distribution (right)

Finally, we adopt simulation method 3 to obtain the risk uncertainty. Remember, here the reserve R^0 is not considered to be fixed. The reserving model still works better when all underlying assumptions are fulfilled, see figure 6 and table 5.

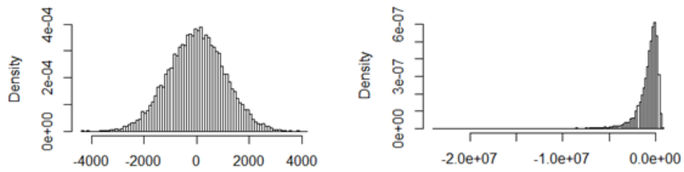

Figure 6: Distribution of one-year reserve risk, Poisson distribution (left) and negative binomial distribution (right). Simulation method 3

	Poisson distribution			Negative-Binomial distribution		
Expected total reserve	44298598			44298635		
	Mean	Sd	99.5%	Mean	Sd	99.5%
$E[\hat{R}^0 - (\hat{X}^1 + \hat{R}^1)]$	-4	1074	2763	-848977	1193026	492169

Table 5: Results for the one-year reserve risk, simulation method 3

5 Discussion and Conclusion

Our motivation to write this paper is to analyze the uncertainty of reserve estimate which comes from a stochastic reserving model called Over-dispersed Poisson model. In the Theory part we have presented that the reserve contains paid losses, outstanding reported losses and incurred but not reported losses. Daykin, Pentikäinen & Pesosnen (1994) states different types of uncertainty. They are parameter error, process error and model error. We analyzed these three uncertainty types.

Renshaw & Verrall (1998) implemented the model in form of a generalized linear model with a log-link function and noticed the link between the chain-ladder algorithm and the Poisson distribution, which is called Over-dispersed Poisson model. We first have checked the chain-ladder assumptions on our real data since the Over-dispersed Poisson model is associated with the chain-ladder algorithm. We have also showed that Over-dispersed Poisson model gives an identical reserve estimate as the chain-ladder method. With the help of maximum likelihood functions, we have determined the parameter estimations and fitted the Over-dispersed Poisson model on real data. Residuals are randomly placed and the weighted residuals are within a certain interval. By back test on the first 13 years data with full development of true payments, it has clearly showed that the Over-dispersed model overestimates reserves 6% , see table 2. Based on the real data, we conclude that the Over-dispersed model works well.

In order to analyze the uncertainty of reserve estimates we have adopted three simulation methods. The results from simulations with Poisson distribution assumption are compared to negative binomial distribution assumption. These two distributions provide almost identical reserves, which is reasonable since the negative binomial distribution is over-dispersed relative to the Poisson. However, they provide large differences in the risk analysis. It can be partly explained by the fact that there are more parameters in negative binomial distribution, which leads to larger prediction uncertainties. Partly, connecting to the Over-dispersed Poisson model, one of the basic model assumptions is broken when a negative binomial distribution is adopted. Hence, the Over-dispersed Poisson model does not behave well.

Give a look at the results of all three risk analyses, the Poisson distribution has a mean value near to zero in the ultimate risk calculation. However, it is more unstable since the standard deviation is larger. We also have compared the prediction error which is a square root of MSEP from England & Verrall's formula. The simulation method offers a smaller value than the prediction error from the formula. However, the percent value of the prediction error itself is not extremely large compared to the total estimated reserve. For the one-year reserve risk, results of simulations from Poisson distribution indicate lower values in mean value, standard deviation and Value at Risk. Difference between simulation method 2 and 3 is that method 2 aims to examine how good the stochastic model is for the one-year outcome reserve risk. However, a stochastic model is not a indispensable condition in simulation method 3, but we have simulated data and see how the expected prediction of reserves behaves.

In summary, we have showed that the reserve risks behave better when the underlying assumptions of Over-dispersed Poisson model are fulfilled. In the other word, the Over-dispersed model works as expected given that the underlying assumptions are fulfilled. The results can become poor if the underlying assumptions are broken.

6 Limitations and future researches

There are several ways to extend and improve our analysis. First, the data in our study is limited to payments due to the property of the insurance product. One would have to include data of reserves in further researches. One could investigate how different stresses on the estimated parameters affect the reserve if the data of reserves is involved.

Secondly, Renshaw & Verrall (1998) assume the incremental payments are Poisson distributed in the model. Beside the Poisson distribution, we have also assumed and analyzed that the incremental payments are negative binomial distributed. Kremer (1982) used a log-normal distribution, which is a most common distribution used for claims payments. Other distributions could be analyzed in the further analysis.

7 References

References

[1] Dahl, P. (2003) Introduction to Reserving. *Stockholms University*

[2] Daykin, C. D., Pentikäinen, T. & Pesosnen, M. (1994) Practical Risk Theory for Actuaries. *London: Chapman & Hall.*

[3] England, P. D. & Verrall, R. J. (2002) Stochastic claims reserving in general insurance. *Institute of Actuaries and Faculty of Actuaries*

[4] Gould, I. L. (2008) Stochastic chain-ladder models in nonlife insurance. *The University of Bergen*

[5] Huang, J. & Wu, X. (2012) Stochastiv Claims Reserving in General Insurance: Models and Methodologies. *Department of Statistics and Actuarial Science, East China Normal University*

[6] Kremer, E. (1982) IBNR-claims and the two-way model of ANOVA. *Scand. Act. J* 1, 47-55.

[7] Mack, T. (1993a) Distribution-free calculation of the standard error of chain ladder reserve estimates. *ASTIN BULLETIN,* 23:2.

[8] Mack, T. (1993b) Measuring the Variability of Chain Ladder Reserve Estimates. *The Casualty Actuarial Society*

[9] Niman, P. (2007) Stochastic claims reserving of a Swedish life insurance portfolio. *The Swedish Actuarial Society.*

[10] Ohlsson, E. & Lauzeningks, J. (2009). The one-year non-life insurance risk. *Insurance: Mathematics and Economics,* 45(2):203 – 208.

[11] Pinheiro, P. J. R. & Andrade e Silva, J. M. & Centeno, M. L. (2000). Bootstrap methodology in claim reserving. *Centre for Applied Maths, Forecasting & Economic Decision, ISEG Lisboa*

[12] Renshaw, A. E. & Verrall, R.J. (1998) A stochastic model underlying the chain-ladder technique. *B.A.J.4,IV,* 903-923.

[13] Rosenberg, P. S. (1990) A simple correlation of AIDS surveillance data for reporting delays. *J. AIDS,* 3, 49-54.

[14] Wüthrich, M. V. (2015). Non-Life Insurance: Mathematics & Statistics. *Department of Mathematics, ETH*

[15] Wüthrich, M. V., Merz, M. & Lysenko, N. (2007). Uncertainty in the claims development result in the chain ladder method. *ETH, Zürich*

8 Appendix

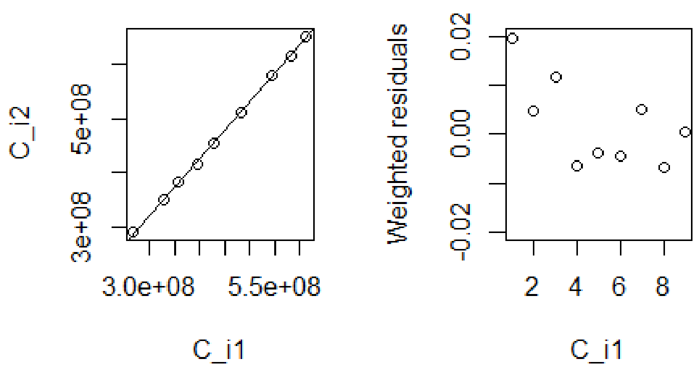

Figure 7: Regression and residuals C_{i2} against C_{i1} (left) and weighted residuals (right).

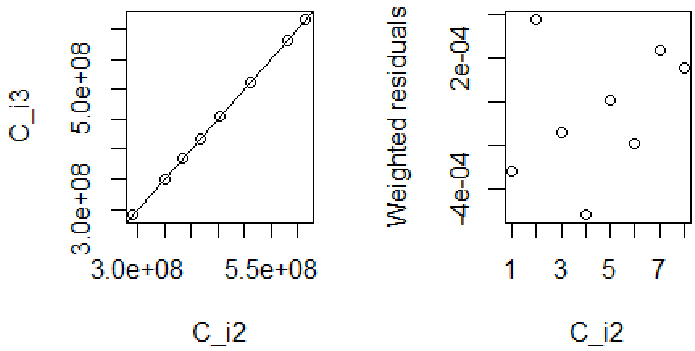

Figure 8: Regression and residuals C_{i3} against C_{i2} (left) and weighted residuals (right).

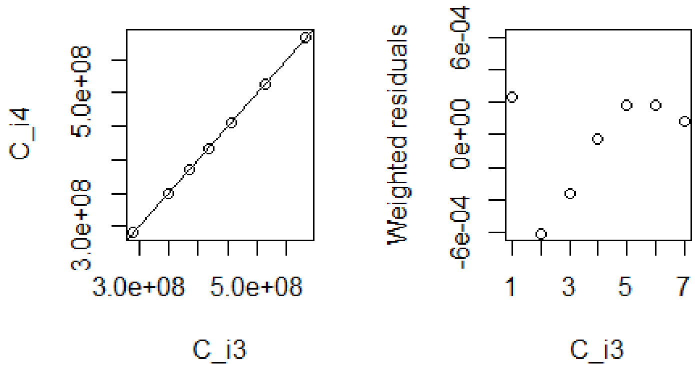

Figure 9: Regression and residuals C_{i4} against C_{i3} (left) and weighted residuals (right).

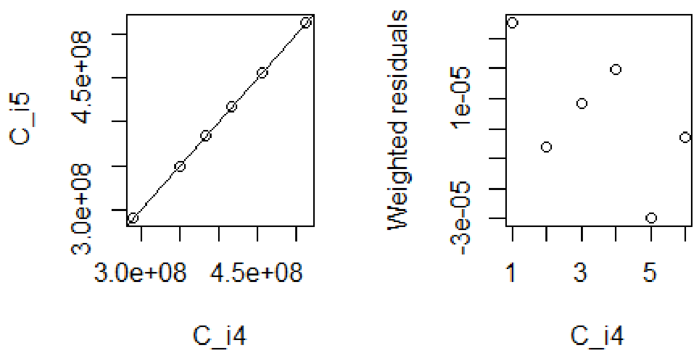

Figure 10: Regression and residuals C_{i5} against C_{i4} (left) and weighted residuals (right).

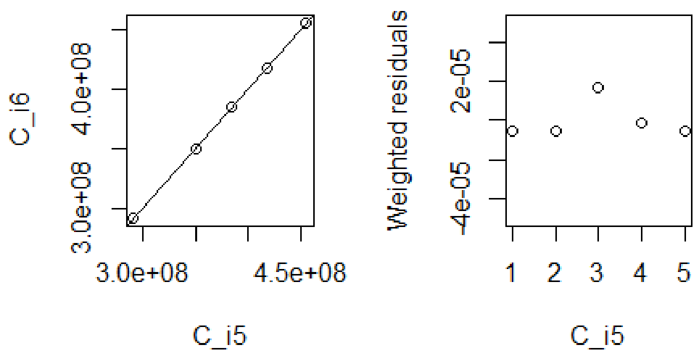

Figure 11: Regression and residuals C_{i6} against C_{i5} (left) and weighted residuals (right).

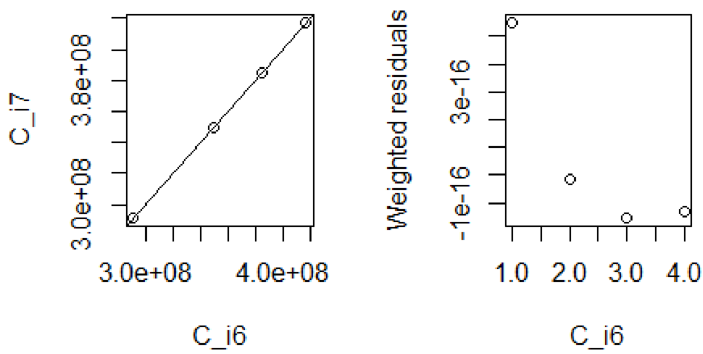

Figure 12: Regression and residuals C_{i7} against C_{i6} (left) and weighted residuals (right).

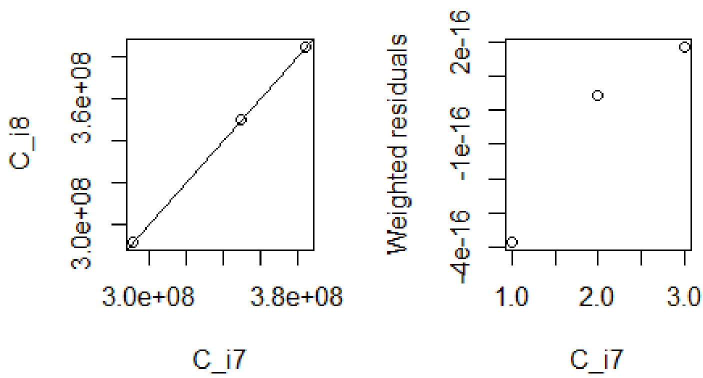

Figure 13: Regression and residuals C_{i8} against C_{i7} (left) and weighted residuals (right).

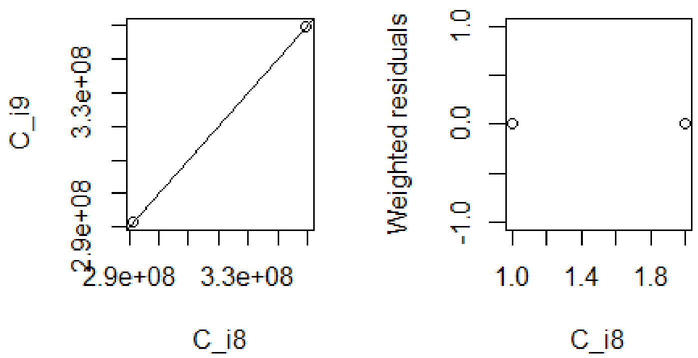

Figure 14: Regression and residuals C_{i9} against C_{i8} (left) and weighted residuals (right).

	F_1	F_2	F_3	F_4	F_5	F_6	F_7	F_8	F_9
i=1	1.0834173	1.0006354	1.0004246	1.0000380	1.0000000	1	1	1	1
i=2	1.0684349	1.0013337	0.9995802	0.9999966	1.0000000	1	1	1	
i=3	1.0754817	1.0008155	0.9998331	1.0000115	1.0000224	1	1		
i=4	1.0570732	1.0004382	1.0001660	1.0000223	1.0000046	1			
i=5	1.0598077	1.0009638	1.0003792	0.9999732	1.0000001				
i=6	1.0592190	1.0007658	1.0003801	1.0000000					
i=7	1.0688169	1.0011918	1.0002811						
i=8	1.0569352	1.0011137							
i=9	1.0641125								

Table 6: Development factors f_{ik}.

	r_{i1}	r_{i2}	r_{i3}	r_{i4}	r_{i5}	r_{i6}	r_{i7}	r_{i8}
i=1	9	2	7	6	1	1	1	1
i=2	6	8	1	2	1	1	1	1
i=3	8	4	2	4	5	1	1	
i=4	2	1	3	5	4	1		
i=5	4	5	5	1	3			
i=6	3	3	6	3				
i=7	7	7	4					
i=8	1	6						
i=9	5							

Table 7: Results of rank factor r_{ik}.

	s_{i2}	s_{i3}	s_{i4}	s_{i5}	s_{i6}	s_{i7}	s_{i8}
i=1	8	2	6	5	1	1	1
i=2	5	7	1	2	1	1	1
i=3	7	4	2	3	4	1	
i=4	2	1	3	4	3		
i=5	4	5	4	1			
i=6	3	3	5				
i=7	6	6					
i=8	1						
i=9							

Table 8: Results of rank factor s_{ik}.

k	2	3	4	5	6	7	8
T_k	0.024	-0.464	0.371	-0.250	-0.300	1	1
I-k-1	7	6	5	4	3	2	1

Table 9: Spearman's rank coefficients T_k.

	j=1	j=2	j=3	j=4	j=5	j=6	j=7	j=8	j=9
j=1	L	S	L	L	S	*	*	*	*
j=1	L	L	S	S	S	*	*	*	
j=1	L	S	S	L	L	*	*		
j=1	S	S	S	L	L	*			
j=1	S	L	L	S	*				
j=1	S	S	L	S					
j=1	L	L	*						
j=1	S	L							
j=1	*								

Table 10: Results of large rank, small rank and mean rank

j	S_j	L_j	Z_j	n	m	$E[Z_j]$	$Var(Z_j)$
2	1	1	1	2	0	0.5	0.25
3	0	3	0	3	1	0.75	0.1875
4	3	1	1	4	1	1.25	0.4375
5	5	0	0	5	2	1.5625	0.3711
6	3	2	2	5	2	1.5625	0.3711
7	1	4	1	5	2	1.5625	0.3711
8	2	3	2	5	2	1.5625	0.3711
9	1	1	1	2	0	0.5	0.25
Total			8			9.25	2.6094

Table 11: Results of expected value and variance of Z_j