Mathematical Statistics
Stockholm University

An Overview of Stochastic Claims Reserving with an Application Using Chain Ladder, GLM and Bootstrap

Mickael Kyrlis

Examensarbete 2008:19

Postal address:
Mathematical Statistics
Dept. of Mathematics
Stockholm University
SE-106 91 Stockholm
Sweden

Internet:
http://www.math.su.se/matstat

Mathematical statistics
Stockholm University
Examensarbete 2008:19
http://www.math.su.se/matstat

An Overview of Stochastic Claims Reserving with an Application Using Chain Ladder, GLM and Bootstrap

Mickael Kyrlis*

November 2008

Abstract

Reserving is one of the most important subjects in non-life insurance mathematics. This project gives a brief overview of stochastic claims reserving, a field that has seen great development the latest decades. Attention has been given to introduce reserving in a general framework, in order to give the reader a background before describing mathematical reserving methodologies. Further on, we present a first application of Mack's chain ladder method and its assumptions. Thereafter, a modern bootstrap approach is used, with an underlying generalized linear model (GLM), using real paid claims automobile data.

Key words: *Reserving*, *Chain Ladder*, *GLM*, *Bootstrap*.

* E-mail: kyrlis@yahoo.com , Supervisor: Ola Hössjer

Acknowledgments

This work constitutes a 30 ECTS-credit Master's thesis in mathematical statistics at Stockholm University.

I am greatly indebted to my supervisor, Professor Ola Hössjer, for his outstanding collaboration, patience and guidance throughout the creation of this thesis.

I am also grateful to Anders Hedegaard Jessen from Copenhagen University, *Codan forsikring* and Niels Rietdorf from *Alpha Group insurance*, for providing me the data set and making useful comments.

Last but not least, I would like to warmly thank my family and all my friends for constantly encouraging me to complete this work.

Contents

1 Introduction

1.1 The purpose of non-life insurance

In general, insurance could be defined as a product in which the consumer (insured) is paying an amount (premium) to a company (insurer) in order to be hedged (covered) for the risk of an economic or physical loss, under certain conditions that are agreed by both parties in an insurance contract.

The concept of *risk* is the key to the existence of the insurance industry through economic history. A simple definition of risk is: "*A state of uncertainty where some of the possibilities involve a loss, catastrophe, or other undesirable outcome*" (Wikipedia). The results of these outcomes evolve economic consequences, which is exactly what insurance companies are prepared to cover for. The measurement of risks and their possible consequences to the insured and the insurer is one among many tasks an actuary is settled to fulfill.

As is well known, insurance companies are usually divided into life and non-life (general or property/casualty). Life insurance companies sell life *assurance*, annuities and pensions products while non-life insurance companies sell a wide variety of insurance contracts that are not related to life loss. Nevertheless, the majority of non-life insurance companies sell products which are related to health and sickness issues. As specified by solvency rules, the primary types of non-life insurance products are (Claims reserving manual, 1997):

- Accident & Health
- Motor Vehicle
- Aircraft
- Shipping
- Goods in Transit
- Property Damage
- General Liability
- Pecuniary Loss
- Non-proportional Treaty Reinsurance
- Proportional Treaty Reinsurance

There are different types of risks to be covered even within each category, so for economical purposes they should be subdivided into different groups of risk. Furthermore the group division has minor differences among different countries.

1.2 Loss reserving in general

An insurance company's financial condition is highly dependent on which reserving philosophy it adapts, as well as to which extent the value of the reserve is the best estimate of the business experts. The key to understand a reserving procedure in general is to understand why insurance companies have to hold reserves for the future in the first place. The basic reasons why reserves must exist are the following:

(i) Although premiums have been paid from the insured, the cover corresponding to part of these premiums that has not been settled.
(ii) There may always be claims that have been or have not been reported, even though the accidents for these claims have occurred in the past.
(iii) Various expenses such as taxes have not yet been paid.

One of actuary's responsibilities (apart from product pricing and business statistics) is to create best estimates of the reserves needed in order to fulfill some requirements such as solvency rules, company profitability, business administration (accounting to shareholders and to Inland Revenue), budgeting, claim management and control, tax rules, ratemaking, reinsurance, etc.

1.2.1 Types of Reserves

As maybe expected reserves are not only of one kind. We have to specify the categories of reserves before continuing. Briefly, the basic categories of reserves (Pantzopoulou, 2003) in a non-life insurance company are:

I. *Reserves with respect to unexpired or unearned exposure:*
- Unearned premium reserve
- Deferred acquisition costs
- Additional unexpired risk reserve

II. *Contingent reserves:*
- Catastrophe reserves
- Claims equalization reserves

III. *Reserves with respect to earned exposure:*
- Notified outstanding claims
- Incurred but not reported claims (IBNR)
- Incurred but not enough reported (IBNER)

Note that the last two types fall to the same category, i.e. "incurred but not settled" (IBNS).

1.2.2 Reserving methodology in general

In a more general sense, reserving methodology should follow the following steps, regardless of the purpose of reserving (Claims Reserving Manual, 1997):

1. Construct a model of the claim process, setting out the assumptions made.
2. Fit the model, using past observations.
3. Test the fit of the model and the assumptions, rejecting or adjusting it.
4. Use the model to make predictions about future statistics of interest.
5. Apply professional judgment and experience to choose values.

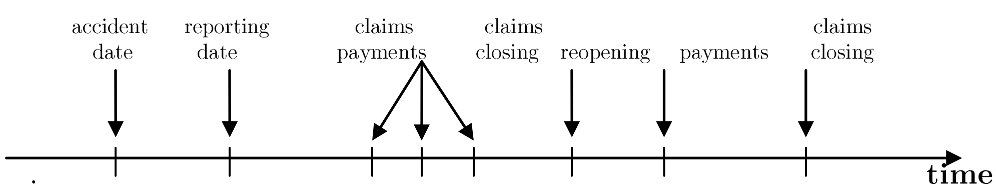
Unfortunately, in the real world the scientific methods used are often very different because the data given by insurance companies is very complex to analyze. In that matter, one must be very careful before proceeding to the fifth step. Furthermore, the availability of data is often limited.

1.3 Claims reserving

It is really important to understand how non-life insurance works, both financially and in a time-perspective (usually a year). An insurance company is receiving premiums from the policyholders through a predetermined period.
During this period, when a claim has occurred, the insurer is obligated either to pay for the claim, or to reserve the claim.
The usual problem for claims is that not all of them are reported in time (IBNR or IBNER). This fact causes a problem for insurance companies: If a claim is not reported, how do we know how much should be paid to the insured?
Other issues that can arise during the insurance period is the measurement of the *unknown* claims costs and how *known* claims cover the reserves.
Claims reserving is the actuarial process in which we want to estimate the claims amount of an insurance company's liabilities. The purpose of this procedure is to provide the insurance company with the right income statements for the balance sheets, as well as to estimate future income and expenses, if possible.
Before we start introducing all the mathematical tools, it is necessary to give a picture of the claims process in time.

1.3.1 Claims time-line

A typical claims reserving procedure is as follows:



Usually it takes several years to settle a claim. We have to take into consideration that various *delays* can come up. A reporting delay means for example, there is long waiting time from the *accident date* to the *reporting date.* Delays between the *reporting date* and the *payments* are also common, because of court decisions, the severity of the claim, the recovery of the process, etc.

Reopening means that a claim has developed unexpectedly, and has to be reanalyzed furthermore before the final payments are made.

To provide an understandable image of claims modeling, we should first make some very important assumptions. The purpose of these assumptions is to give a realistic model of observed data, so that estimates and predictions are as close to the real world as possible. We can briefly give our considerations:

- All data must be adjusted for *inflation.* It is very important to have the proper price level both on premiums and claims amounts. Since our models are developed through time, inflation rates must be given throughout the process. We omit implementing inflation in our models. The reason is that it has no highly significant effect in our estimations (Naziropoulou, 2005). Moreover, inflation *forecasts* will not be applied in this paper.
- Only a small and conservative bias will be allowed, if the methods are going to be used in pricing a portfolio.
- All policies are assumed to start at the same *accident year*, which is not true in an insurance contract process.

1.3.2 Claims data presentation

The most widely used method to describe claims data is the *run-off triangle.* We are given data over observed claims paid per year, usually counted in thousands of a given currency. Before describing data, we have to define the rows´ and columns´ time-perspective:

- *Accident year* (rows): the year the claims occurred, i.e. the year the accident(s) occurred, leading to the corresponding claims.
- *Development year* (columns)*:* the year (or number of years) until a payment from the claims is actually settled, after the event has occurred.
- *Calendar year:* the actual year payments are made (diagonals, not shown schematically).

Let $C_{i,j}$ denote the accumulated total claims amount of accident year i, $1 \le i \le n$ which are paid or incurred up to development year j , $1 \le j \le n$.

The following table is the basic representation of the accumulated claims amount data. Note that the same figure applies for *incremental* claims $Y_{ij} = C_{i,j} - C_{i,j-1}$.

Accident	Development year (j)						
year (i)	1	2	. . .	j	n-j	n-1	n
1	$C_{1,1}$	$C_{1,2}$	. . .	$C_{1,j}$	$C_{1,n-j}$	$C_{1,n-1}$	$C_{1,n}$
2	$C_{2,1}$	$C_{2,2}$	. . .	$C_{2,j}$	$C_{2,n-j}$	$C_{2,n-1}$	$\mathbf{C_{2,n}}$
. .	. .	. .	. .	. .	. .	$\mathbf{C_{..,n-1}}$	$\mathbf{C_{..,n}}$
i	$C_{i,1}$	$C_{i,}$	. . .	. .	. .	. .	. .
n-1	$C_{n-1,1}$	$C_{n-1,2}$	$\mathbf{C_{n-1,...}}$	. . .	. . .	$\mathbf{C_{n-1,n-1}}$	$\mathbf{C_{n-1,n}}$
n	$C_{n,1}$	$\mathbf{C_{n,2}}$	$\mathbf{C_{n,3}}$	. . .	. . .	$\mathbf{C_{n,n-1}}$	$\mathbf{C_{n,n}}$

Figure 1 *Run-off triangle*

All the values from the upper left corner down to the upper right and lower left the triangle are *known* to us. Our objective is to estimate the grey shaded values of the lower right triangle, in order to calculate *outstanding claims reserves*. The formed grey area represents future claim payments which will be estimated in order to calculate the claims reserves for each *calendar year.*

There exist many methods estimating the future payments and reserves. Some of these are the *Bornheutter-Ferguson* method, the *Separation method,* the *Cape Cod* method and the *Naive Loss Ratio* method. Methods from *time series* could also be applied. Another very popular tool for estimating outstanding claims reserves is the *Chain ladder* method. Finally, due to the evolution of computers the last decades, *bootstrap* models have been the most commonly used methods in practice.

1.4 Aim of this paper

The aim of this paper is to give a slight overview of stochastic claims reserving, using some of the methods that have been used in the past, and give an introduction to the latest models that are used both in theory and in practice.

The field of claims reserving in non-life insurance is quite large with numerous models and methods, and a complete overview would be beyond the scope of a thesis. This is the reason why this paper is limited to merely a few methods. A very good attempt of overviewing claims reserving is however the work performed by Pantzopoulou, 2003.

2 Theoretical framework

2.1 Mack´s Chain ladder method

2.1.1 Basic Assumptions

The chain ladder method is the most important model in non-life insurance claims reserving. The properties of this model are simple, and the estimates are derived in a very intuitive and natural way. It requires very few distributional assumptions whatsoever. Nevertheless, we have to make some rather basic assumptions that characterize this method.

Our goal is to estimate the claims amount $C_{i,j}$ for $i+j>n+1$ where the values $C_{i,j}$ are known for $i+j \le n+1$. The final *ultimate (*total estimated amount) claims amount chain ladder estimate is given by:

(2.1) $$\hat{C}_{i,n} = C_{i,n+1-i} \cdot \hat{f}_{n+1-i} \cdot \ldots \cdot \hat{f}_{n-1} \quad , \quad 2 \le i \le n$$

where

(2.2) $$\hat{f}_j = \frac{\sum_{i=1}^{n-j} C_{i,j+1}}{\sum_{i=1}^{n-j} C_{i,j}} \qquad \text{for} \quad 1 \le j \le n-1$$

are the estimated *development factors*[1].

The *outstanding claims reserve* of accident year i is:

(2.3) $$R_i = \hat{C}_{i,n} - C_{i,n+1-i}$$

As we see $\hat{f}_j$ only depends on development year *j*. Although the individual development factors of each development year differ from each other, the ultimate claims estimate will use the same $\hat{f}_j$ for *all* accident years i. Therefore, every increase from $C_{i,j}$ to $C_{i,j+1}$ will be considered stochastic having *expected* increase from $C_{i,j}$ to $C_{i,j} \cdot f_j$ conditionally on $C_{i,j}$. This fact leads us to the first basic assumption:

(AS 1) $$E(C_{i,j+1} \mid C_{i,1}, \ldots, C_{i,j}) = C_{i,j} \cdot f_j \quad , \quad 1 \le i \le n \quad , \quad 1 \le j \le n-1,$$

[1] *Also called "age to age" factors*

where $C_{i,1},\ldots,C_{i,j}$ are all previously known values. This basic assumption states that $C_{i,j+1}$ is a random variable whose conditional expectation given $C_{i,1},\ldots,C_{i,j}$ only depends on $C_{i,j}$ multiplied by f_j, which can be seen as a *proportionality factor*. Estimated age to age factors $\hat{f}_j$ are explained as sums of known claims up to $j+1$ development years, *normalized* by sums of claims of the corresponding accident years, up to j development years. The fact that the individual age to age factors differ for each development year can be explained by rewriting the first assumption by:

(AS1´) $$E(C_{i,j+1} / C_{i,j} \mid C_{i,1},\ldots,C_{i,j}) = f_j \quad , \quad 1 \le i \le n \ \ , 1 \le j \le n-1 \ .$$

This implies that the conditional expectation of $C_{i,j+1} / C_{i,j}$ given $C_{i,j}$ is equal to f_j, and the individual factors are uncorrelated for sequential development years.

(AS2) The vectors $(C_{i,1},C_{i,2},\ldots,C_{i,n})$ and $(C_{k1},C_{k2},\ldots,C_{kn})$ are independent for *different* accident years $i \neq k$.

The assumption above is necessary for the unbiasedness of $\hat{f}_j$, as well for showing that accident years do not affect the chain ladder estimate, as long as there are no large changes of known claims for each calendar year and/or greater fluctuations in inflation. We can prove that all $\hat{f}_j$ for **one time-step** are unbiased using the elementary property $E(X) = E(E(X \mid Y))$. We have:

$$E(C_{i,j+1} / C_{ij}) = E(E(C_{i,j+1} / C_{ij} \mid C_{i1},\ldots,C_{ij}))$$

$$= (E(C_{i,j+1} \mid C_{i1},\ldots,C_{ij}) / C_{ij})$$

$$= E(C_{ij} \cdot f_j / C_{ij}) = E(f_j) = f_j \quad .$$

The third assumption underlying the chain-ladder method is based on basic inference principle that we should choose the estimator with the smallest variance, among several estimators. Since accident claims are independent for different accident years, the conditional variance of $C_{i,j+1}$ given past claims will depend on some averaged *weighted* variance σ_j^2 for every single development year. Hence:

(AS3) $$Var(C_{i,j+1} \mid C_{i,1},\ldots,C_{i,j}) = C_{i,j} \cdot \sigma_j^2 \ , \ \ 1 \le i \le n \ \ , \ \ 1 \le j \le n-1 ,$$

where σ_j^2 represents the conditional variance of $C_{i,j+1}$ given units of $C_{i,j}$ .The variance of $\hat{f}_j = \sum_{i=1}^{n-j} w_i(C_{i,j+1} / C_{i,j})$ is minimal when the weights w_i are inversely proportional to $Var(C_{i,j+1} / C_{i,j} \mid C_{i,1},\ldots,C_{i,j})$

Since:

(AS3´) $$Var(C_{i,j+1} / C_{i,j} \mid C_{i,1},...,C_{i,j}) = \sigma_j^2 / C_{i,j} \quad , \quad 1 \le i \le n \quad , \quad 1 \le j \le n-1$$

and $\sum_{i=1}^{n-j} w_i = 1$ the optimal weights are defined as: $w_i = C_{i,j} / \sum_{i=1}^{n-j} C_{i,j+1}$ which are derived from the Lagrange minimization method. These weights yield the estimate (2.2) of the development factors.

2.1.2 The variability of the ultimate claims and reserves

The purpose of this method is to estimate the ultimate claims $C_{i,n}$, and thereafter to find an estimate of the final reserves R_i for all accident years. The ultimate claims estimator's expectation is, with the assumptions made, equal to $E(C_{i,n})$.

We would like to have a statistically efficient estimator; i.e. the mean square error (*mse*) of $\hat{C}_{i,n}$ must be as small as possible, always conditional on previously known values. In addition, we need the *mse* of $\hat{C}_{i,n}$ for constructing confidence intervals. The *mse* is defined as:

$$mse(\hat{C}_{i,n}) = E[(C_{i,n} - \hat{C}_{i,n})^2 \mid C_{i,1}...C_{i,n-i+1}], \quad i=2,...,n.$$

For the ultimate reserve we have:

$$mse(\hat{R}_i) = E[(\hat{R}_i - R_i)^2 \mid C_{i,1}...C_{i,n-i+1}] = E[(\hat{C}_{i,n} - C_{i,n+1-i} - C_{i,n} + C_{i,n+1-i})^2 \mid C_{i,1}...C_{i,n-i+1}] =$$

$$E[(\hat{C}_{i,n} - C_{i,n})^2 \mid C_{i,1}...C_{i,n-i+1}] = mse(\hat{C}_{i,n}) \Leftrightarrow$$

(2.4) $$s.d.(\hat{R}_i) = \sqrt{mse(\hat{R}_i)} = \sqrt{mse(\hat{C}_{i,n})} = s.d.(\hat{C}_{i,n})$$

since $\hat{C}_{i,n}$ and $\check{R}_i$ are unbiased (Mack, 1993).

A stated above, the chain ladder estimates require few distributional assumptions. This means that the standard deviation is given by:

$$(2.5) \qquad (s.d.(\hat{C}_{i,n}))^2 = C_{i,n}^{\ 2} \cdot \sum_{j=n+1-i}^{n-1} (\frac{\sigma_j^{\ 2}}{f_j^{\ 2}}) \cdot \left[\frac{1}{C_{i,j}} + \frac{1}{\sum_{k=1}^{n-j} C_{kj}} \right]$$

In order to estimate $s.e(\hat{C}_{i,n})^2 = \widehat{s.d.}(\hat{C}_{i,n})^2$ we replace in (2.5) $C_{i,j}$ by $\hat{C}_{i,j}$ when $i+j>n+1$, f_j by its the estimator (2.2) , and σ_j^2 by an estimator

$$(2.6) \qquad \hat{\sigma}_j^{\ 2} = \frac{1}{n-j-1} \sum_{i=1}^{n-j} C_{ij} \left(\frac{C_{i,j+1}}{C_{i,j}} - \hat{f}_j \right)^2, 1 \le j \le n-2$$

Note that development years in (2.6) run until $n\text{-}2$ years because we cannot estimate $\sigma_{n-1}^{\ 2}$ from the observation between years $n\text{-}1$ and n. If the claims development is not finished in $n\text{-}2$, we have to extrapolate the series of $\hat{\sigma}_k^2$ for $k=n-1$ by using:

$$\hat{\sigma}_{n-1}^2 = \min(\hat{\sigma}_{n-2}^4 / \hat{\sigma}_{n-3}^2, \min(\hat{\sigma}_{n-3}^2, \hat{\sigma}_{n-2}^2))$$

These complicated formulas are derived from the basic assumptions of the chain ladder. The proof of (2.5) is very long and difficult, and it is omitted from this paper, for details see Mack (1993).

We have now established estimates of R_i and $s.e.(\check{R}_i)$, but in order to build confidence intervals[2] we can always assume a distribution for R_i. If we have enough data (which is usually **not** the case) the interval can be given by the standard normal or lognormal distribution, due to the central limit theorem. The lognormal distribution is used because claims amount data is often not symmetrical around its mean, like in the normal case. Therefore we assume:

$$R_i \in LogN(\mu_i, \sigma_i^{\ 2})$$

with unknown mean and variance. (Note that the variance σ_i^2 here is **not** the same variance as in (2.6)).
Hence, the following relations hold:

[2] *We can build confidence intervals using other methods, i.e. using Tjebyshev's inequality.*

$$E(R_i) = \exp(\mu_i + \sigma_i^2 / 2)$$
$$Var(R_i) = \exp(2\mu_i + \sigma_i^2)(\exp(\sigma_i^2) - 1) \Leftrightarrow$$

$$\sigma_i^2 = \ln(1 + Var(R_i) / E(R_i)^2)$$
$$\mu_i = \ln(E(R_i)) - \sigma_i^2 / 2$$

Using the above relations we can first estimate the reserves $\hat{E}(R_i) = \hat{R}_i$ and the variances $\hat{V}ar(R_i)$ and then use them to calculate the estimates of mean and variance of μ_i and σ_i^2. Now the confidence intervals are:

(2.7) $$\hat{R}_i \cdot \exp(\pm\lambda \cdot \hat{\sigma}_i - \hat{\sigma}_i^2 / 2) ,$$

where λ is the percentile of the chosen confidence level $(2\Phi(\lambda)-1)$, and Φ is the cumulative distribution function of a standard normal distribution.

2.2 Generalized Linear Models & chain ladder

2.2.1 Basic model

One disadvantage of the chain-ladder estimates is that it does not provide a whole predictive distribution of the outstanding claims. That is, because of the absence of a specific distribution that could possibly capture more of the variability of the claims estimates. Moreover, the mean square error is computed completely empirically based on given data. An often used approach is to employ generalized linear models (GLM). For full analysis of GLM see McGullagh and Nelder, (1989).

The GLM theory allows for a random variable $Y_{i,j}$'s stochastic component to have parameters from a larger family of distributions, instead of assuming the normal distribution as in the classical linear model. Thus, the density function of $Y_{i,j}$ can be expressed in the form of an exponential dispersion family as:

$$f_{Y_{ij}}(y_{ij};\theta_{ij},\varphi) = \exp\left\{\frac{(y_{ij} \cdot \theta_{ij} - b(\theta_{ij}))}{a(\varphi)} - c(y_{ij},\varphi)\right\}$$

where $\theta_{i,j}$ is the canonical parameter, φ is the dispersion parameter and $b(\vartheta_{ij})$, $a(\varphi)$, $c(y_{ij}, \varphi)$ are functions of θ_{ij} , φ and $y_{i,j}$. We also define a link function g that relates the expected value of $Y_{i,j}$ to a linear combination of the explanatory variables used in the regression models. The link function is defined by requiring that

$\eta_{ij}=g(m_{ij})$ where $m_{ij}=E(Y_{ij})$.

Note that the GLM assumption requires that the incremental claim amounts are independent. We can now assume some distributions of the variable Y_{ij}. The most commonly used distributions for claims amounts are overdispersed Poisson, gamma or lognormal. By assigning a distribution to the claims data we can test whether a model fits data to a satisfying level. When we assume a logarithmic link function $g(m_{ij})=\log(m_{ij})$ and a Poisson or gamma distribution for the incremental claims Y_{ij} we have the following structure:

$$E(Y_{i,j}) = m_{ij} \ , \ Var(Y_{i,j}) = \varphi \cdot m_{ij}^{p}$$

and

$$\log(m_{ij}) = \eta_{ij}$$
$$\eta_{ij} = c + \alpha_i + \beta_j$$
$$\alpha_1 = \beta_1 = 0$$

where $p=1$ for Poisson and $p=2$ for gamma. This model gives a total of ***2n-1*** regression variables **c,** $\boldsymbol{\alpha_2,\dots,\alpha_n}$, $\boldsymbol{\beta_2,\dots,\beta_n}$. Note that the last equality is set to zero to avoid overparametrization.

2.2.2 ML-equations and parameter estimation

To get the maximum-likelihood estimations of the regression parameters c, a_i and β_j we first define the log-likelihood function of the joint density of $Y = \{Y_{i,j}; 2 \le i + j \le n + 1\}$ as:

$$\ell(\{y_{ij}\};\{\theta_{ij}\};\varphi) = \log\left[\prod_{i,j} f_\theta(y_{ij};\theta_{ij},\varphi)\right] = \frac{1}{a(\varphi)}\sum_{i,j}(y_{ij}\theta_{ij} - b(\theta_{ij})) + \sum_{i,j} c(y_{ij},\varphi) \quad (2.8)$$

where the parameters θ_{ij} are functions of $\eta_{i,j}$ depending on the chosen probability model. The dispersion parameter $a(\varphi)$, does not influence the maximum likelihood values. Using the relations in section 2.2.1 with $g(m_{ij}) = \eta_{ij}$, $m_{ij} = b'(\theta_{ij})$ and taking the first derivative with respect to $\boldsymbol{\beta_j}$, we have:

$$\frac{\partial \ell}{\partial \beta_j} = \sum_{i=1}^{n+1-j} \frac{\partial \ell}{\partial \theta_{ij}} \frac{\partial \theta_{ij}}{\partial \beta_j} = \sum_{i=1}^{n+1-j} (y_i - b'(\theta_{ij}))\frac{\partial \theta_{ij}}{\partial \beta_j} = \sum_{i=1}^{n+1-j} (y_{ij} - b'(\theta_{ij}))\frac{\partial \theta_{ij}}{\partial m_{ij}} \frac{\partial m_{ij}}{\partial \eta_{ij}} \frac{\partial \eta_{ij}}{\partial \beta_j}$$

Then, we know that $m_{ij} = b'(\theta_{ij}) \Leftrightarrow \partial m_{ij} / \partial \theta_{ij} = b''(\theta_{ij})$, and

$$v(m_{ij}) = b''(\theta_{ij}) \Leftrightarrow \partial\theta_{ij} / \partial m_{ij} = 1 / v(m_{ij})$$

where $v(m_{ij}) = m_{ij}^{p}$ is the variance function.

Furthermore we have:

$$\partial m_{ij} / \partial\eta_{ij} = 1 / [g'(m_{ij})]$$
$$\partial\eta_{ij} / \partial\beta_{j} = 1$$

And finally, using all the above relations we get the ML-equations by:

$$\sum_{i} \frac{y_{ij} - m_{ij}}{v(m_{ij})g'(m_{ij})} = 0 \qquad j = 2,...,n$$

Analogously an additional set of n ML-equations is obtained by differentiating with respect to $c,\ a_2,..,a_n$.

2.3 Bootstrap methodology in stochastic claims reserving

A natural way of quantifying uncertainty when having too little data is using the *bootstrap.* A simple definition of bootstrap is: "a computer-intensive, general purpose approach to statistical inference, falling within a broader class of resampling methods" (Wikipedia). Furthermore as we shall see that the lognormal confidence intervals in (2.7) are of very high uncertainty, since they assume a lognormal distribution of the reserves.

The idea of bootstrap is to create "false" data to make inference and compare with the actual real data (Efron & Tibshirani, 1995). In stochastic claims reserving we can in this way compare the true outstanding claims reserves and the reserves estimator, and most importantly approximate the variance of the prediction error, and the predictive distribution of the reserves (Björkwall, et al., 2008). Some of the first papers using bootstrap in claims reserving are by England & Verall, (1999) and by Pinheiro,(2003)

2.3.1 Non-parametric bootstrap with GLM

In order to begin resampling the data, we have to use the GLM model introduced in section 2.2.1.

As described in (Björkwall et al, 2008), the proper way to resample the triangles is to resample the *residuals* with replacement, and not to pick and replace from the original data. The reason is that the original data is ordered in accident and development years, while the residuals are standardized. The most commonly used residual is the (unscaled) Pearson residual defined (using the same notation as in section 2.2.1,):

(2.9) $$r_{ij}^{p} = \frac{Y_{ij} - \hat{m}_{ij}}{\sqrt{\hat{m}_{ij}{}^{p}}} \quad ,2 \le i + j \le n + 1$$

where $\hat{m}_{ij} = \hat{c} + \hat{\alpha}_i + \hat{\beta}_j$ is the maximum likelihood estimate of m_{ij} .

The residual is chosen not to be corrected for degrees of freedom as there is no significant difference using such residuals, when the number of accident years is fairly large (Pinheiro, 2003). Note that the residuals are assumed to be approximately independent and identically distributed. We therefore estimate the dispersion parameter, which is needed to evaluate the variance, as:

(2.10) $$\hat{\varphi} = \frac{1}{(n(n+1)/2) - (2n-1)} \sum (r_{ij}^{p})^{2}$$

where *(n(n+1)/2)* is number of observations and *(2n-1)* is the number of parameters. The bootstrap procedure then starts when we draw B samples $r_{ij}^{p^*}$ from the residuals in (2.9) with replacement, building B new residual triangles. Then, we calculate B bootstrap triangles of claims amounts using:

(2.11) $$Y_{ij}^{*} = \hat{m}_{ij} + r_{ij}^{*}\sqrt{\hat{m}_{ij}^{p}} \quad , 2 \le i + j \le n + 1$$

We then develop the resampled upper triangle using the chain ladder method, getting resampled forecasts of outstanding claims. This yields upper triangle outstanding claims $\hat{R}_i^*, i = 2,\ldots,n$ for each accident year i and $\hat{R}^*$ for the total outstanding claim.
We then resample future values $Y_{i,j}{}^{*}$ by resampling from (2.9) and repeating the bootstrap procedure for the lower triangle. In this way we get the resampled true outstanding claims in the bootstrap procedure and *unstandardized* prediction errors.

$$pe_i^{*} = R_i^{*} - \hat{R}_i^{*}$$

for each accident year i and similarly R^{*} for the resampled total outstanding claim and prediction errors:

$$pe^{*} = R^{*} - \hat{R}^{*}$$

The reason we use unstandardized prediction errors is that that they are defined in all cases. It is shown (Bjorkwall, et al.,2008) that standardized prediction errors often give negative values, giving imaginary numbers.

Finally, the predictive distribution of the outstanding claims is obtained from the B resampled $\tilde{R}_i^* = \hat{R}_i + pe_i^*$ for accident year i and similarly the B resampled $\tilde{R}^* = \hat{R} + pe^*$ for the total claim.

3 Data analysis

Our dataset consists of 55 observations of third party liability (TPL) automobile data of *paid* claims amounts.

3.1 Chain Ladder results

We can begin our data analysis by first looking at the data itself:

Accident year	Development year 1	2	3	4	5	6	7	8	9	10
1	451,288	339,519	333,371	144,988	93,243	45,511	25,217	20,406	31,482	1,729
2	448,627	512,882	168,467	130,674	56,044	33,397	56,071	26,522	14,346	
3	693,574	497,737	202,272	120,753	125,046	37,154	27,608	17,864		
4	652,043	546,406	244,474	200,896	106,802	106,753	63,688			
5	566,082	503,970	217,838	145,181	165,519	91,313				
6	606,606	562,543	227,374	153,551	132,743					
7	536,976	472,525	154,205	150,564						
8	554,833	590,880	300,964							
9	537,238	701,111								
10	684,944									

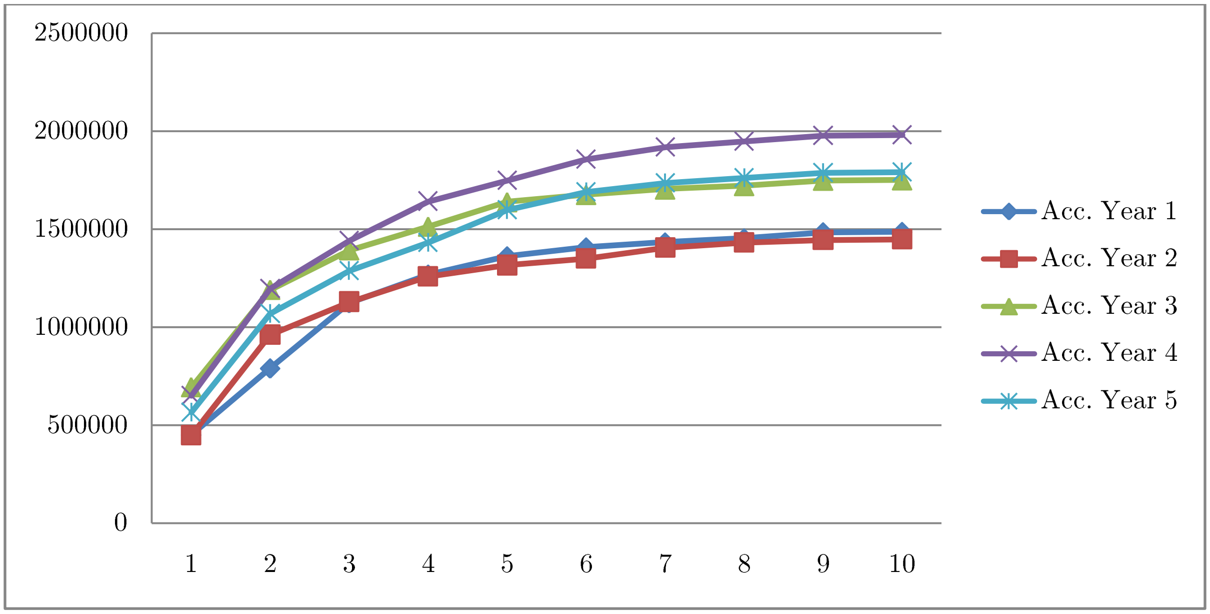
Table 3.1, *Incremental claims amounts*

At a first glance, the increments in table 3.1 seem to have reasonable size and are not negative. Nevertheless, we can be critical about some possible outliers (marked in the triangle). In accident year 1, development year 10, the amount is 1729 which is a very small increment that will surely influence our estimations. Some data points in development year 2 are also considerable, i.e. we see that the amounts are larger than in accident year 1, which is usually not the case in this kind of data. If we add each development year progressively we get the cumulated data:

Acc. year	Development year 1	2	3	4	5	6	7	8	9	10
1	451,288	790,807	1,124,178	1,269,166	1,362,409	1,407,920	1,433,137	1,453,543	1,485,025	1,486,754
2	448,627	961,509	1,129,976	1,260,650	1,316,694	1,350,091	1,406,162	1,432,684	1,447,030	
3	693,574	1,191,311	1,393,583	1,514,336	1,639,382	1,676,536	1,704,144	1,722,008		
4	652,043	1,198,449	1,442,923	1,643,819	1,750,621	1,857,374	1,921,062			
5	566,082	1,070,052	1,287,890	1,433,071	1,598,590	1,689,903				
6	606,606	1,169,149	1,396,523	1,550,074	1,682,817					
7	536,976	1,009,501	1,163,706	1,314,270						
8	554,833	1,145,713	1,446,677							
9	537,238	1,238,349								
10	684,944									

Table 3.2: *Claims amounts* $C_{i,j}$ *cumulated*

To get an even clearer picture of how the amounts are increasing we look at the following diagram, only for accident years 1-5 (this is because the lines intersect, and we cannot observe clearly):



All accident years are following a trend, (looks like exponential trend). From this trend and from the chain ladder formulas, we first calculate the *age to age factors* of **each** data point as $\hat{f}_{ij} = \dfrac{C_{i,j+1}}{C_{i,j}}$, and then we use formula (2.2) which is needed to progress the triangle later on, getting the following table:

Accident	**Development year**									
year	1	2	3	4	5	6	7	8	9	10
1	1.75233	1.42156	1.12897	1.07347	1.0334	1.01791	1.01424	1.02166	1.00116	-
2	2.14323	1.17521	1.11564	1.04446	1.02536	1.04153	1.01886	1.01001		
3	1.71764	1.16979	1.08665	1.08257	1.02266	1.01647	1.01048			
4	1.83799	1.20399	1.13923	1.06497	1.06098	1.03429				
5	1.89028	1.20358	1.11273	1.1155	1.05712					
6	1.92736	1.19448	1.10995	1.08564						
7	1.87997	1.15275	1.12938							
8	2.06497	1.26269								
9	2.30503									
10	-									
Chain Ladder	1.93666	1.2166	1.11709	1.07835	1.04097	1.02743	1.01426	1.01588	1.00116	

Table 3.3 Development factors of each data point & Chain ladder development factors

In table 3.3 we see that there are three factors exceeding 2, giving a somewhat unexpectedly large increase. This can be considered unusual but it is absolutely rational: Most of the development factors in the beginning of claims development (years 1 to 3) are much higher than 1,and then drop to values close to 1, which means that most of the expenses of unsettled claims are reported or known to the insurer the first years. This does not have to be a rule of thumb, but it can be one of many reasons.

The following table shows known (cumulated) and predicted values (2.1) of future accumulated claims using the same age to age factors for each accident year as explained in chapter 2. Then, the corresponding variability between accident years (which is needed to obtain the standard errors between development years) was calculated using formula 2.6, and extrapolating for the last two development years:

Accident	**Development year**									
year	1	2	3	4	5	6	7	8	9	10
1	451288	790807	1124178	1269166	1362409	1407920	1433137	1453543	1485025	1,486,754
2	448627	961509	1129976	1260650	1316694	1350091	1406162	1432684	1447030	1,448,715
3	693574	1191311	1393583	1514336	1639382	1676536	1704144	1722008	1749278	1,751,387
4	652043	1198449	1442923	1643819	1750621	1857374	1921062	1948970	1979834	1,981,700
5	566082	1070052	1287890	1433071	1598590	1689903	1736459	1761686	1789584	1,791,061
6	606606	1169149	1396523	1550074	1682817	1749973	1798184	1824307	1853197	1,856,619
7	536976	1009501	1163706	1314270	1416478	1473005	1513586	1535574	1559892	1,563,619
8	554833	1145713	1446677	1616673	1742398	1811932	1861850	1888898	1918811	1,922,669
9	537238	1238349	1514508	1692475	1824095	1896889	1949147	1977464	2008779	2,002,268
10	684944	1333266	1630593	1822200	1963909	2042282	2098546	2129033	2162749	2,144,804
Variance										
(2.6)	17680	5622	347	643	413	171	18	49	18	

Table 3.4 Cumulative claims, predicted cumulative claims $C_{i,j}$ and variances from (2,6)

Now we can calculate the *claims ultimate* and the *outstanding claims reserves* by (2.3), and using formula (2.5), we calculate the standard error of the estimations. We can also calculate an estimate of the coefficient of variation for each accident year:

i	$\hat{C}_{i,10}$	$\check{R}_{i,}$	s.e ($\check{R}_i$)	s.e ($\check{R}_i$)/$\check{R}_i$
1	1,486,754	-	-	-
2	1,448,715	1,685	8,790	521.7%
3	1,751,387	29,379	19,305	65.7%
4	1,981,700	60,638	22,835	37.7%
5	1,791,061	101,158	31,188	30.8%
6	1,856,619	173,802	47,011	27.0%
7	1,563,619	249,349	56,684	22.7%
8	1,922,669	475,992	71,230	15.0%
9	2,002,268	763,919	146,344	19.2%
10	2,144,804	1,459,860	252,247	17.3%
Total	17,949,593	3,315,779	354,818	10.7%

Table 3.5 Chain ladder estimates

Most of the values in the last column are between 10-70 %, which is a rather pleasant fact, and shows that the variations of the reserves are not too large. We see that the standard error in accident year 2 is very large (nearly 5 times larger) compared to the reserve estimate. This result comes along with the fact discussed in table 3.1. The claims amount in accident year 1, development year 10 was very small, and affected the variability of the estimate. Therefore this estimate *cannot* be trusted in any way due to its uncertainty.

Now we can proceed building confidence intervals as discussed in chapter 2. The lognormal approach gives the following confidence intervals, using a 95% confidence level:

i	$C_{i,10}$	Ri	lower limit	upper limit
1	1,486,754	-	-	-
2	1,448,715	1,685	9	11,402
3	1,751,387	29,379	7,588	79,444
4	1,981,700	60,638	27,795	115,859
5	1,791,061	101,158	53,552	174,496
6	1,856,619	173,802	99,663	282,431
7	1,563,619	249,349	156,600	377,519
8	1,922,669	475,992	351,653	630,183
9	2,002,268	763,919	517,155	1,088,483
10	2,144,804	1,459,860	1,027,834	2,013,368
	Total	3,315,779	2,674,758	4,063,889

Table 3.6 Reserve estimates using lognormal confidence intervals

We clearly see that the intervals are of high uncertainty, especially in the first accident years. Even the intervals through accident years 3-10 are very large (almost double and half amounts). This is due to the high confidence level, but mostly the large intervals are a result of the lack of data (only 55 observations) in this case. Note also that the lower limit for accident year 2 is equal to 9 which is extremely low.

The conclusion is that we cannot trust these confidence intervals for making business decisions. If we lower the confidence level to 80% we get much narrower intervals, but then arises the usual problem of using such high risk level for a business line. This is not appropriate and we have to use another safer approach for the variability of the chain-ladder estimates, as well for the confidence intervals of the reserves.

3.2 Bootstrap and GLM results

We execute the bootstrap process with **B=10000** and having *p=1* for overdispersed Poisson and *p=2* for gamma. Reserves were estimated using GLM (or just taking the usual chain ladder estimates in the overdispersed Poisson case). The upper limits are picked from the bootstrap density of the predictive distributions of outstanding claims, using 95% and 99.5% percentiles. We also compute quantile-quantile plots for the fitted residuals in both cases. Note that the overall percentile is **not** the sum of the accident year percentiles of the reserves.

Furthermore we calculated in table 3.9 the coefficient of variation (CV) of the reserves estimates, i.e.: $CV(\tilde{R}_i^*) = \sqrt{Var(\tilde{R}_i^*)}/E(\tilde{R}_i^*)$

Acc. Year *i*	Resevre estimate	95% percentile	99,5% percentile
1	0	0	0
2	1,685	9,520	16,556
3	29,379	60,462	89,025
4	60,638	105,103	134,303
5	101,158	156,979	194,423
6	173,802	244,812	290,676
7	249,349	334,844	392,811
8	475,992	600,056	682,064
9	763,919	932,079	1,039,896
10	1,459,860	1,754,934	1,916,761
Total	**3,315,779**	**3,798,645**	**4,022,859**

Table 3.7: *Reserves and bootstrap percentiles estimates when p=1*

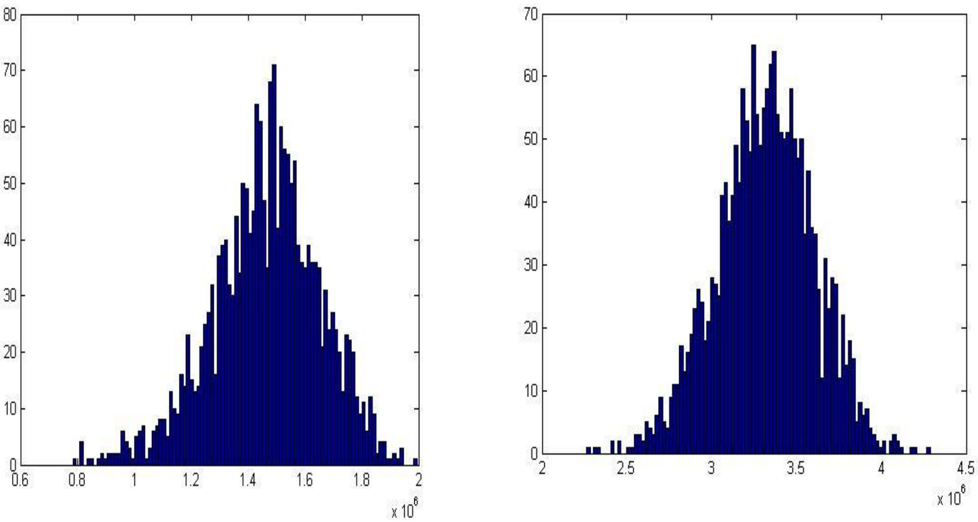


Figure 2*: Densities of* $\tilde{R}^*_{10}$ *(left) and* $\tilde{R}^*$ *(right) using non-parametric bootstrap samples with unstandardized prediction errors and p=1.*

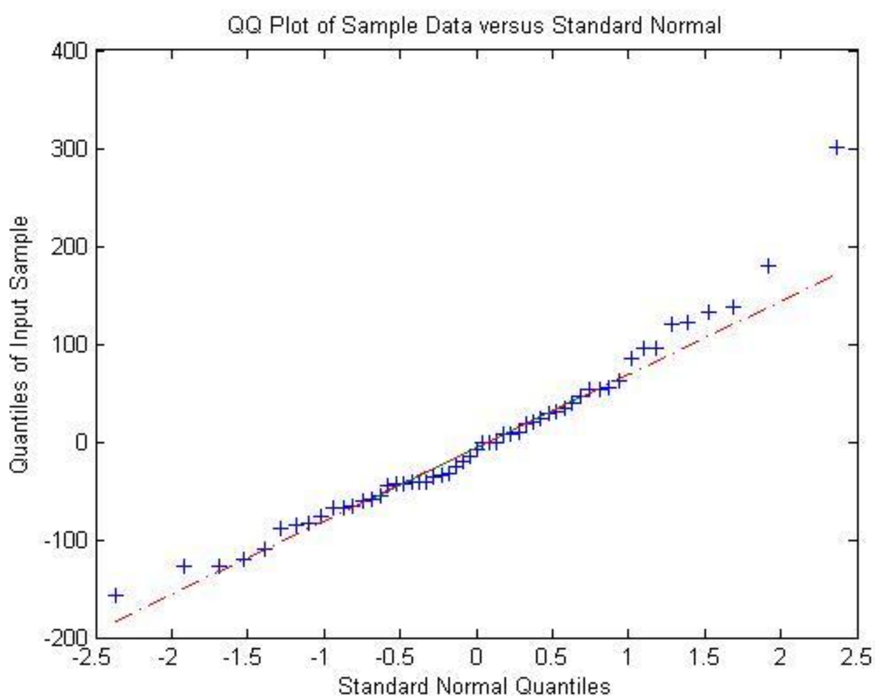


Figure 3: *QQ-plot of standardized residuals vs. quantiles of standard normal distribution when p=1.*

Acc. Year i	Resevre estimate	95% percentile	99,5% percentile
1	0	0	0
2	1,571	2,460	2,959
3	24,722	37,201	43,889
4	67,535	92,210	105,522
5	113,886	151,246	171,270
6	170,718	222,775	250,853
7	240,019	311,892	353,585
8	486,253	632,025	703,942
9	761,674	1,002,114	1,121,986
10	1,492,799	2,110,779	2,411,014
Total	**3,359,178**	**4,089,321**	**4,436,611**

Table 3.8*: Reserves estimates and bootstrap percentiles estimates when p=2*

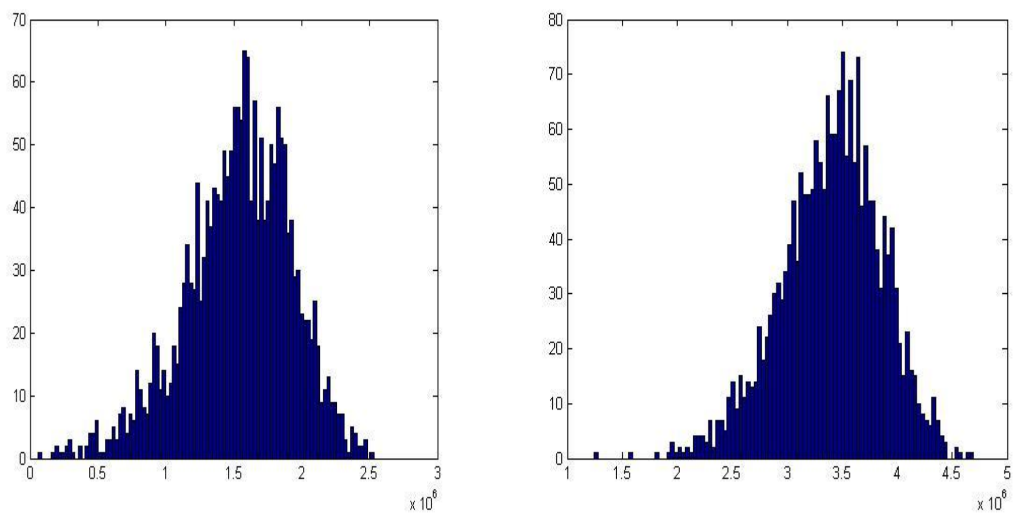


Figure 4: *Densities of $\tilde{R}^*_{10}$ (left) and $\tilde{R}^*$ (right) using non-parametric bootstrap samples with unstandardized prediction errors and p=2.*

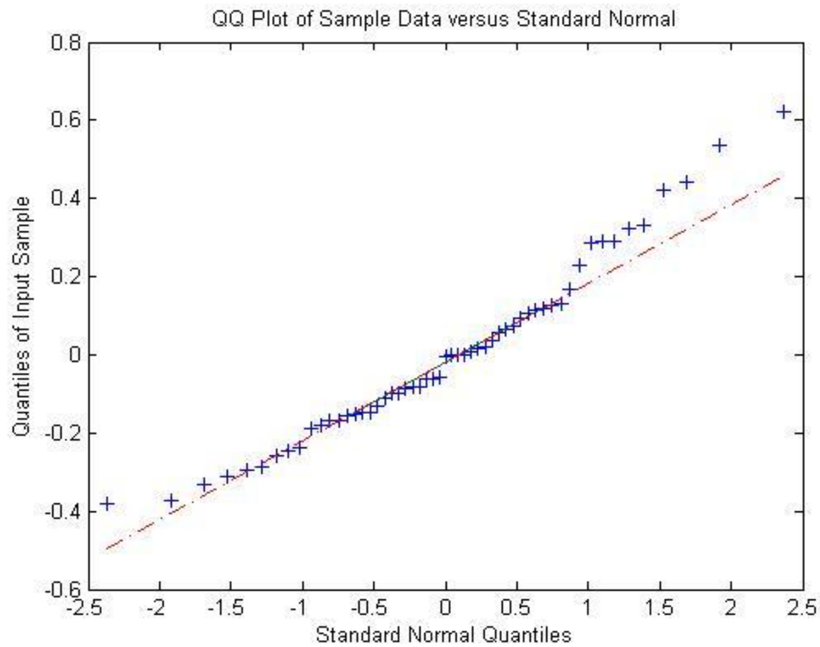


Figure 5: *QQ-plot of simulated standardized residuals vs. quantiles of standard normal distribution when p=2.*

i	*p=1*	*p=2*
2	317.6%	37.2%
3	64.8%	29.7%
4	43.6%	22.9%
5	32.2%	20.3%
6	24.5%	19.0%
7	20.0%	19.4%
8	15.6%	20.0%
9	13.4%	20.9%
10	13.1%	26.8%
Total	8.5%	13.4%

Table 3.9 *CV in % for bootstrap reserve estimates*

3.2.1 Comments on results

When we compare the estimated reserves in tables 3.7 and 3.8 for Poisson and gamma respectively, we see that the values are quite close, in particular the total reserve which is the most important number. For some accident years the reserves are lower when $p=2$, and in other accident years the opposite is true. Therefore there is no clear conclusion in this comparison. Analogous results can we see in Björkwall et al. (2008).

Regarding the percentiles we can clearly see that when $p=2$, in the first accident years, the amounts are closer to the reserve estimates compared to the $p=1$ case. The percentiles for accident years 2-4 do not exceed the reserves as much in table 3.7 as in 3.8. At the last 3 accident years and in the total reserves we see that when $p=2$ the percentiles are higher. This can be explained by the different form of the variance functions $v(m)$ when p differs. When $p=2$, larger variability is allowed for cells with large m_{ij} . Such cells are more frequently represented in the lower triangle for later accident years, which leads to higher quantiles.

The claim amount increment $Y_{1,10}$ is relatively very low compared to the rest of the data which is surely a reason for the results of the first accident years.

The density charts in figures 2 and 4 show that the distribution of the bootstrapped reserves estimates are slightly skewed, having somewhat heavy left tails. The basic reason for the skewness is that the prediction errors are not standardized, see Björkwall et al. (2008).

When we plot the residual distributions in figures 3 and 5 against a standard normal distribution we see that we don't get a perfectly linear relationship, and many values are outlined.

The coefficient of variation in table 3.9 shows a rather expected result: in accident year 2 , where we have only one claim amount in the last development year, CV is nearly 10 times larger when $p=1$, compared to $p=2$. This happens due to, as explain above, when cells have low $m_{i,j}$, the variance function for p=1 is comparatively larger than for $p=2$. In this sense we can consider that perhaps the model with $p=2$ *fits* the given data better than in the $p=1$ case.

4 Conclusions & discussion

The aim of this work is to overview the reserving concept in non-life insurance and to compare some of the models already used by many, with the latest research articles in the field.

The main problem with reserving techniques *in practice* is that they do not follow underlying models, but actuaries use internal methods of reserving. Both Mack's chain ladder method and the bootstrap method require few model assumptions so that the results can mimic reality as well as possible. Furthermore, Mack's lognormal assumption for the confidence intervals is very strong, and its consequences could be analyzed further, both theoretically and in practice.

When we use non-parametric bootstrap we define unstandardized prediction errors, and as stated in Björkwall et al. (2008), resampling from unstandardized quantities is often not as accurate as standardized quantities, despite that standardized quantities can produce imaginary prediction errors, due to negative denominators. A suggested research in Björkwall et al. (2008) is to use a "double bootstrap" for real *and* simulated data sets.

Regarding the residuals a future study could be to investigate their distribution for different choices of p in the GLM model. The QQ-plots presented is just an indication of which distribution the residuals follow, and as said above the standard normal case could be a good choice.

Furthermore, as well known, the numerical results and the different methods are always dependent on what kind of data we have. As stated previously, some data points need to be investigated in more detail, before making strong assumptions and suggesting various models.

Finally, our data set consisted of only 55 observations, which is relatively small. Insurance groups often divide data into quarter year or even in months, which can probably give even better results.

5 References

Björkwall S., Hössjer O. and Ohlsson E. (2008): *Non-parametric and parametric bootstrap techniques for arbitrary age-to-age development factor methods in stochastic claims reserving*, Research report, Mathematical Statistics, Stockholm University.

Claims Reserving Manual, vol. 2. Institute of Actuaries, London, link: http://www.actuaries.org.uk/general_insurance/documents/crm

Dahl P., (2003): *Introduction to Reserving*,

Efron, B., Tibshirani, R.J., (1993): *An Introduction to the Bootstrap. Chapman and Hall, London.*

England P., Verrall R., (1999): *Analytic and bootstrap estimates in claims reserving*, Insurance: Mathematics and economics 25, 281-293.

Mack T., (1993): *Measuring the variability of chain ladder reserve estimates*, Claims Reserving manual vol.2 Institute of Actuaries, London.

Mack, T., (1993): *Distribution free calculation of the standard error of chain ladder reserve estimates*. ASTIN Bulletin 23 (2), 213–225.

Mack, T., (1994): *Which stochastic model is underlying the chain ladder model?* Insurance: Mathematics and Economics 15, 133–138.

McCullagh, P., Nelder, J., (1989*):* *Generalized Linear Models*, 2nd ed. Chapman and Hall, London.

Naziropoulou A. (2005): *Simulering av inflations-och återförsäkringsinverkan på avsättning för oreglerade skador med hänsyn till Solvens II*. Examensarbete 2005:7. Master thesis, Mathematical Statistics, Stockholm University.

Ohlsson E. , Johansson B.,(2008): *Non-life Insurance Pricing with Generalized Linear Models*, Lecture Notes, Mathematical statistics, Stockholm University.

Pantzopoulou E., (2003): *The concept of Reserving and Reserving methodologies in General Insurance*, Master thesis, Department of Actuarial Science and Statistics, City University, London.

Pinheiro P.,Andrade S. and Centeno ,M.D.L., (2003): *Bootstrap Methodology in Claims Reserving*, The journal of risk and insurance, vol. 70, No.4 , 701-714.

Rietdorf N., (2008): *Chain Ladder Reservering*, unpublished working paper.

Wikipedia, link: www.wikepedia.org

Wiser R.F, (2005): *Loss Reserving*, Casualty Actuarial Society (CAS).

6 Appendix

(a)

Standardized Pearson residuals when p=1

-34.116	-157.755	300.727	38.700	-14.925	-41.267	-67.602	-1.027	54.289	0
-20.613	120.826	-58.113	8.518	-127.441	-87.945	95.945	45.893	-54.998	
179.534	-36.119	-66.770	-85.393	28.575	-109.325	-83.224	-40.794		
24.118	-60.222	-40.742	62.966	-65.623	122.268	54.759			
-7.793	-43.414	-45.097	-31.745	138.494	96.189				
17.784	9.643	-42.782	-24.765	29.947					
53.257	7.035	-120.732	34.515						
-75.515	20.789	85.529							
-127.791	132.041								
0									

Standardized Pearson residuals when p=2

-0.085	-0.260	0.621	0.072	-0.063	-0.148	-0.372	-0.099	0.332	0
0.001	0.231	-0.098	0.064	-0.380	-0.312	0.538	0.289	-0.332	
0.442	0.114	0.009	-0.084	0.289	-0.286	-0.294	-0.190		
-0.061	-0.153	-0.155	0.056	-0.237	0.421	0.128			
-0.112	-0.149	-0.180	-0.169	0.287	0.324				
0.019	0.017	-0.083	-0.058	0.105					
0.093	0.035	-0.247	0.119						
-0.130	-0.003	0.133							
-0.168	0.168								
0									

(b)

GLM parameters for *p=1, 2*

Parameter	p=1	p=2
c	13.005	13.036
α2	-0.026	-0.096
α3	0.164	-0.026
α4	0.287	0.341
α5	0.186	0.256
α6	0.222	0.188
α7	0.050	-0.004
α8	0.257	0.256
α9	0.298	0.269
α10	0.366	0.328
β2	0.065	0.073
β3	-0.803	-0.802
β4	-1.222	-1.221
β5	-1.513	-1.527
β6	-2.086	-2.150
β7	-2.447	-2.436
β8	-3.074	-3.008
β9	-2.953	-2.965
β10	-5.550	-5.580