Mathematical Statistics
Stockholm University

Non-parametric and parametric bootstrap techniques for arbitrary age-to-age development factor methods in stochastic claims reserving

Susanna Björkwall
Ola Hössjer
Esbjörn Ohlsson

Research Report 2008:2

ISSN 1650-0377

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
Research Report **2008:2**,
http://www.math.su.se/matstat

Non-parametric and parametric bootstrap techniques for arbitrary age-to-age development factor methods in stochastic claims reserving

Susanna Björkwall* Ola Hössjer† Esbjörn Ohlsson‡

January 2008

Abstract

In the literature, one of the the main objects of stochastic claims reserving is to find models underlying the chain-ladder method in order to analyze the variability of the outstanding claims, either analytically or by bootstrapping. In bootstrapping these models are used to find a full predictive distribution of the claims reserve, even though there is a long tradition of actuaries calculating the reserve estimate according to more complex algorithms than the chain-ladder, without explicit reference to an underlying model. In this paper we investigate existing bootstrap techniques and suggest two alternative bootstrap procedures, one non-parametric and one parametric, by which the predictive distribution of the claims reserve can be found for any age-to-age development factor method, using some rather mild model assumptions. For illustration, the procedures are applied to four different development triangles.

Keywords

Bootstrap, Chain-ladder, Development factor method, Development triangle, Dynamic financial analysis, Stochastic claims reserving.

*Mathematical Statistics, Stockholm University. E-mail: sbj@student.su.se
†Mathematical Statistics, Stockholm University.
‡Länsförsäkringar Alliance, Stockholm.

1 Introduction

The provision for outstanding claims – henceforth the claims reserve – is a major contributor to the total risk of an insurance company, especially for long-tailed lines of business. In order to estimate the risk that the provisions will not suffice to pay all claims in the end, the actuary's best estimate of the outstanding claims needs to be complemented by its predictive distribution; this is the ultimo perspective. For solvency control and risk management with Dynamic Financial Analysis we are also interested in a shorter period, say the one year risk. The reserving risk is then the risk of a negative run-off result, due to unexpectedly large claims payments, changes in inflation regime or in the discount rate in the simulated forecast year.

A well-known method for calculating the uncertainty of the claims reserve, obtained by chain-ladder, in meeting ultimate claims, or at least its mean squared error of prediction, is the one introduced by Mack (1993) and recently treated by Buchwalder *et al.* (2006) and Mack *et al.* (2006). Another popular method is bootstrapping, as introduced in this context by England & Verrall (1999) and England (2002). The latter method gives a full predictive distribution without further assumptions and can easily be used also for the purpose of finding the risk in the run-off result. Therefore, we focus on bootstrap methods here.

A standard statistical approach to claims reserving would be to first specify a model, then find an estimate of outstanding claims under the model, e.g. by maximum likelihood. Finally, the model could be used to find the precision of the estimate, possibly by bootstrapping if an analytic solution was untractable.

In practice, there is a long tradition of actuaries calculating reserve estimates without explicit reference to a model. The object of the research area called stochastic claims reserving, has mostly been to find a model and a method of giving a measure of the precision of the actuary's best estimate *post festum*, i.e. without the possibility of changing the estimate itself.

In particular the object of several papers on stochastic claims reserving has been to find a model under which the best estimate is the one given by the chain-ladder method; indeed, there has been a discussion of which method is underlying the chain-ladder, see in particular Verrall (2000), Mack & Venter (2000) and Verrall & England (2000). So even though the actuary did not use a model to pick her best estimate, these articles try to find a model that would make her work consistent with the standard approach of statistics: to specify the model before finding the estimate. In Verrall (2000) several underlying models, which produce the same reserve estimates as the chain-ladder method, are suggested, and it is also remarked on the importance of careful examination of the assumptions of the model and how the chosen model effects the outstanding claims.

In this paper we question the need to bootstrap an underlying model with claim distributions fully specified, which happens to reproduce the actuary's best estimate. Instead, we develop a bootstrap methodology for the data with as few model assumptions as possible, applicable to any age-to-age development factor method. We assume that the bootstrap procedure only depends on the mean and variance of the claims and that the chosen reserving algorithm implicitly specifies the mean structure and therefore the only additional assumption concerns the variance function. Furthermore, we discuss the non-parametric vs the parametric bootstrap and standardized vs unstandardized prediction errors. Finally, the suggested bootstrap procedures are applied to development triangles of different types.

Section 2 contains the definitions and gives an example of an age-to-age development factor method, that might be used in practise. In Section 3 the non-parametric bootstrap procedure of Pinheiro *et al.* (2003) is discussed and an alternative parametric procedure is suggested, as well as bootstrap procedures, which can be used to find the predictive distribution of any age-to-age development factor method. The double bootstrap is discussed, some details of the implementation of the bootstrap procedures are commented and finally the run-off result is defined and a sketch of a method of obtaining it's predictive distribution is provided. In Section 4 the bootstrap procedures are compared on four different development triangles.

2 A basic model

We consider data in the form of a triangle of n incremental observations $\{C_{ij}; i,j \in \nabla\}$, where ∇ denotes the upper, observational triangle $\nabla = \{i = 1,\ldots,t; j = 1,\ldots,t-i+1\}$ and C_{ij} is e.g. paid claims, the number of claims, claims incurred or some other quantity of interest of origin year i in development year j, see Table 2.1. For the time being we discuss paid claims. The actuary's goal is then to predict the sum of the delayed claim amounts in the lower, unobserved future triangle $\{C_{ij}; i,j \in \Delta\}$, where $\Delta = \{i = 2,\ldots,t; j = t-i+2,\ldots,t\}$, see Table 2.2. We write $R = \sum_\Delta C_{ij}$ for this sum, which is the outstanding claims for which the insurance company must hold a reserve.

	Development year					
Accident year	1	2	3	$\cdots$	$t-1$	t
1	C_{11}	C_{12}	C_{13}	$\cdots$	$C_{1,t-1}$	$C_{1,t}$
2	C_{21}	C_{22}	C_{23}	$\cdots$	$C_{2,t-1}$	
3	C_{31}	C_{32}	C_{33}	$\cdots$		
$\vdots$	$\vdots$	$\vdots$	$\vdots$			
$t-1$	$C_{t-1,1}$	$C_{t-1,2}$				
t	$C_{t,1}$					

Table 2.1: *The triangle ∇ of observed incremental payments.*

	Development year					
Accident year	1	2	3	$\cdots$	$t-1$	t
1						
2						$C_{2,t}$
3					$C_{3,t-1}$	$C_{3,t}$
$\vdots$					$\vdots$	$\vdots$
$t-1$			$C_{t-1,3}$	$\cdots$	$C_{t-1,t-1}$	$C_{t-1,t}$
t		$C_{t,2}$	$C_{t,3}$	$\cdots$	$C_{t,t-1}$	$C_{t,t}$

Table 2.2: *The triangle Δ of unobserved future claim costs.*

Above we have implicitly made the common assumption that claims are settled within the t observed years. In long-tailed business such as Motor TPL we often have no origin year with finalized claims; when needed, we extend the model so that the unknown claims extend beyond t in a tail of length u, i.e. over the development years $t, t+1, \ldots, t+u$, see Table 2.3. For simplicity, we use the notation Δ for the set of unobserved claims in this case, too.

In practice, the actuary has used some method to calculate an estimate of the outstanding claims R; in statistical terminology this is rather a *prediction* of R. We assume that the method gives estimates $\hat{m}_{ij}$ of the cell expectations $m_{ij} = E(C_{ij})$ for all claims in both ∇ and Δ and that these estimates are functions of our observations $\nabla C \doteq \{C_{ij}; i, j \in \nabla\}$ only. (We will use the notation ∇x to denote the ∇ collection of any variable x, and similar for Δx.) The estimate of outstanding claims is then $\hat{R} = \sum_\Delta \hat{m}_{ij}$. This is the case for age-to-age development factor methods. Note in particular that we do not assume that the reserving method is based on an explicit statistical model, in our experience this is seldom the case in practice.

Some reserving methods operate on cumulative claims $D_{ij} = \sum_{\ell=1}^{j} C_{i\ell}$ rather than incremental claims C_{ij}. Let $\mu_{ij} = E(D_{ij})$. Here is an example of an age-to-age development factor method that fits our scheme:

1. The chain-ladder method, see Taylor (2000), is used to produce development factors $\hat{f}_j$ that are estimates of $f_j = \mu_{i,j+1}/\mu_{ij}$, perhaps after excluding the oldest observations and/or sole outliers in ∇.

2. For $3 < j < t$, say, the $\hat{f}_j$'s are smoothed by some method, say exponential smoothing,

	Development year							
Accident year	1	2	3	$\cdots$	t	$t+1$	$\cdots$	$t+u$
1						$C_{1,t+1}$	$\cdots$	$C_{1,t+u}$
2					$C_{2,t}$	$C_{2,t+1}$	$\cdots$	$C_{2,t+u}$
3					$C_{3,t}$	$C_{3,t+1}$	$\cdots$	$C_{3,t+u}$
$\vdots$				$\vdots$	$\vdots$	$\vdots$		$\vdots$
$t-1$			$C_{t-1,3}$	$\cdots$	$C_{t-1,t}$	$C_{t-1,t+1}$	$\cdots$	$C_{t-1,t+u}$
t		$C_{t,2}$	$C_{t,3}$	$\cdots$	$C_{t,t}$	$C_{t,t+1}$	$\cdots$	$C_{t,t+u}$

Table 2.3: *The long tail case, with the triangle Δ of unobserved future claim costs extended with a rectangle beyond t.*

i.e. they are replaced by estimates obtained from a linear regression of $\log(\hat{f}_j - 1)$ on j. By extrapolation in the linear regression, this also yields $\hat{f}_j$ for the tail $j = t, t+1 \ldots, t+u$. The original $\hat{f}_j$'s are kept for $j \leq 3$ and the smoothed ones used for all $j > 3$.

3. Now estimates $\hat{\mu}_{ij}$ for Δ are computed as in the standard chain-ladder method.

4. Estimates of $\hat{\mu}_{ij}$ for ∇ are obtained by the process of backwards recursion described in England & Verrall (1999).

5. Finally, the obtained claim values may be discounted by some interest rate curve, or inflated by assumed claims inflation. The latter of course requires that the observations where recalculated to fixed prices in the first place.

We now have an estimator $\hat{R} = h(\nabla C)$ for some possibly quite complex function h, that might be specified only by an algorithm as in the example. Our primary object is to find the bootstrap estimate of the predictive distribution of $\hat{R}$.

3 Bootstrap methods

The basic idea of bootstrapping is to work with the *Bootstrap world* in order to make inference on the *Real world*, see Efron & Tibshirani (1993). This is done by investigating the result of B simulations in the bootstrap world and assuming that the conclusions from these are approximately valid in the real world; this is the so-called plug-in-principle, Efron & Tibshirani (1993). With the outstanding claims in consideration this means that a relation between the true outstanding claims R and its estimator $\hat{R}$ in the real world can be substituted in the bootstrap world by their bootstrap counterparts. This makes it possible to approximate the variance of the prediction error $R - \hat{R}$ as well as the predictive distribution of R.

Pinheiro *et al.* (2003) use the plug-in-principle to obtain the predictive distribution of R by a non-parametric bootstrap technique consistent with the statistical assumptions underlying the chain-ladder method in the literature. Our purpose is to modify it to a non-parametric bootstrap procedure which works for any age-to-age development factor method used in practise, e.g. the one described in the previous section. We also suggest a completely parametric approach consistent with, and as a complement to, the non-parametric procedure.

3.1 Bootstrapping data with a generalized linear model using standardized prediction errors

Some assumptions about the model structure of ∇C have to be imposed in order to bootstrap the data. In the literature a common choice is to use a generalized linear model, in particular an over-dispersed Poisson distribution with a logarithmic link function. A consequence of this underlying model is that the expected claims obtained by maximum likelihood estimation of the parameters in the generalized linear model equal the ones obtained by the chain-ladder method, if the column sums of the triangle are positive, see Renshaw & Verrall (1998). Thus, the expectations of the claims can be obtained either by maximum likelihood estimation or by the chain-ladder, while the variances, which are needed for the residuals, are given by the assumption of the generalized linear model. The bootstrap methods described by England & Verrall (1999), England (2002) and Pinheiro *et al.* (2003) are all based on generalized linear models.

The method discussed in Pinheiro *et al.* (2003) assumes the following log additive structure of the $n = t\,(t+1)/2$ incremental observations in ∇C

$$\begin{aligned}
&E(C_{ij}) = m_{ij} \quad \text{and} \quad \text{Var}(C_{ij}) = \phi m_{ij}^p \\
&log(m_{ij}) = \eta_{ij} \\
&\eta_{ij} = c + \alpha_i + \beta_j, \quad \alpha_1 = \beta_1 = 0
\end{aligned} \tag{3.1}$$

The fitted values $\nabla\hat{m}$ and the forecasts $\Delta\hat{m}$ are calculated by maximum quasi likelihood estimation of the $q = 2t - 1$ model parameters c, α_i and β_j, e.g. under the assumption of an over-dispersed Poisson distribution, i.e. $p = 1$, or a gamma distribution, i.e. $p = 2$. Estimators of the outstanding claims are then obtained by summing per accident year $\hat{R}_i = \sum_{j\in\Delta_i} \hat{m}_{ij}$, where Δ_i denotes the row corresponding to accident year i in $\Delta\hat{m}$. The estimator of the grand total is $\hat{R} = \sum_{\Delta} \hat{m}_{ij}$.

The residuals are needed for the resampling process and the common choice is to use the Pearson residuals

$$r_{ij}^{P} = \frac{C_{ij} - \hat{m}_{ij}}{\sqrt{\hat{m}_{ij}^{p}}}, \tag{3.2}$$

which should have approximately zero mean and constant variance. Pinheiro *et al.* (2003), as well as England & Verrall (1999) and England (2002), work under the assumption that the residuals are independent and identically distributed, an assumption that can be questioned, see e.g. Larsen (2007) and Appendix 1. Nevertheless, we shall adhere to this assumption.

There are two ways of adjusting the Pearson residuals. England & Verrall (1999) and England (2002) use a global adjusting factor

$$r_{ij}^{PA} = \sqrt{\frac{n}{n-q}}\, r_{ij}^{P}, \tag{3.3}$$

whereas Pinheiro *et al.* (2003) argue that the hat matrix standardized Pearson residuals are a better choice. They are given by

$$r_{ij}^{PA} = \frac{r_{ij}^{P}}{\sqrt{1 - h_{ij}}}, \tag{3.4}$$

where the h_{ij}:s are the diagonal elements of the n x n hat matrix H, which for generalized linear models is given by

$$H = X(X^{T}WX)^{-1}X^{T}W, \tag{3.5}$$

where X is an n x q design matrix and the generic elements $W_{ij,ij}$ of the n x n diagonal matrix W are

$$W_{ij,ij} = (V(m_{ij})(\frac{\partial \eta_{ij}}{\partial m_{ij}})^{2})^{-1} \tag{3.6}$$

and V is the variance function.

This choice of residual correction is in accordance with Davison & Hinkley (1997). The result of the comparison in Pinheiro *et al.* (2003) does not indicate a big difference to the correction in (3.3).

Note that the residuals are also used to produce the Pearson estimate of the unknown ϕ,

$$\hat{\phi} = \frac{1}{n-q} \sum_{\nabla} (r_{ij}^P)^2 = \frac{1}{n} \sum_{\nabla} (r_{ij}^{PA})^2 , \tag{3.7}$$

where the last equality is exact when (3.3) is used and an approximation for (3.4).

The next step is to get B new triangles of residuals ∇r^* by drawing samples with replacement from the collection of residuals in (3.3) or (3.4). This procedure means sampling from the empirical distribution function of the approximately independent and identically distributed residuals r.

Then B pseudo-triangles ∇C^* are generated by computing

$$C_{ij}^* = \hat{m}_{ij} + r_{ij}^* \sqrt{\hat{m}_{ij}^p} \quad \text{for} \quad i,j \in \nabla \tag{3.8}$$

and for these B pseudo-triangles the future values $\Delta\hat{m}^*$ are forecasted by the same method as above, i.e. by estimating the parameters of the generalized linear model. Estimators for the outstanding claims in the bootstrap world are then derived by $\hat{R}_i^* = \sum_{j\in\Delta_i} \hat{m}_{ij}^*$ and $\hat{R}^* = \sum_{\Delta} \hat{m}_{ij}^*$.

In order to get the random outcome of the true outstanding claims in the bootstrap world, i.e. $R_i^{**} = \sum_{j\in\Delta_i} C_{ij}^{**}$ and $R^{**} = \sum_{\Delta} C_{ij}^{**}$, the resampling is done once more from the empirical distribution function of the residuals to get B triangles of Δr^{**} and then solving

$$C_{ij}^{**} = \hat{m}_{ij} + r_{ij}^{**} \sqrt{\hat{m}_{ij}^p} \quad \text{for} \quad i,j \in \Delta \tag{3.9}$$

to get ΔC^{**}.

The final step is to calculate the B prediction errors and in Pinheiro *et al.* (2003) this is done by the following equations

$$\text{pe}_i^{**} = \frac{R_i^{**} - \hat{R}_i^*}{\sqrt{\widehat{Var}(R_i^{**})}} \qquad \text{and} \qquad \text{pe}^{**} = \frac{R^{**} - \hat{R}^*}{\sqrt{\widehat{Var}(R^{**})}} . \tag{3.10}$$

The predictive distributions of the outstanding claims R_i and R are then obtained by plotting

$$\tilde{R}_i^{**} = \hat{R}_i + \text{pe}_i^{**} \sqrt{\widehat{Var}(R_i)} \qquad \text{and} \qquad \tilde{R}^{**} = \hat{R} + \text{pe}^{**} \sqrt{\widehat{Var}(R)} \tag{3.11}$$

for each B.

We tacitly assume that the mean and variance of all bootstrapped quantities are conditional on the observed data ∇C. For instance, the variance of the bootstrapped outstanding claims are

$$Var(R_i^{**}) = \hat{\phi} \sum_{j \in \Delta_i} \hat{m}_{ij}^p \qquad \text{and} \qquad Var(R^{**}) = \hat{\phi} \sum_{\Delta} \hat{m}_{ij}^p \,, \tag{3.12}$$

since the variance of the bootstrapped residuals conditional on ∇C is $\hat{\phi}$ according to (3.3), (3.4) and (3.7). Since Pinheiro *et al.* (2003), as well as England (2002), consider ϕ as constant for the data, the estimates of (3.12) appearing in (3.10) are

$$\widehat{Var}(R_i^{**}) = \hat{\phi} \sum_{j \in \Delta_i} \hat{m}_{ij}^{*p} \qquad \text{and} \qquad \widehat{Var}(R^{**}) = \hat{\phi} \sum_{\Delta} \hat{m}_{ij}^{*p} \tag{3.13}$$

and hence computable from the bootstrap world data ∇C^*. Nevertheless, ϕ is unknown and therefore

$$\widehat{Var}(R_i^{**}) = \hat{\phi}^* \sum_{j \in \Delta_i} \hat{m}_{ij}^{*p} \qquad \text{and} \qquad \widehat{Var}(R^{**}) = \hat{\phi}^* \sum_{\Delta} \hat{m}_{ij}^{*p} \tag{3.14}$$

should rather be used, see Davison & Hinkley (1997). This is in analogy with the estimated variances of the true claims reserves

$$\widehat{Var}(R_i) = \hat{\phi} \sum_{j \in \Delta_i} \hat{m}_{ij}^p \qquad \text{and} \qquad \widehat{Var}(R) = \hat{\phi} \sum_{\Delta} \hat{m}_{ij}^p, \tag{3.15}$$

which are computable from the real data ∇C, as opposed to $Var(R_i)$ and $Var(R)$.

As a complement to the non-parametric procedure described above we suggest a parametric approach. In addition to the assumptions in (3.1) we assume a full distribution F, parametrised by the mean and variance, so that we may write $F = F(m_{ij}, \phi m_{ij}^p)$. Instead of resampling the residuals, we draw C_{ij}^* from $F(\hat{m}_{ij}, \hat{\phi}\hat{m}_{ij}^p)$ for all $i, j \in \nabla$ and thereby we directly get the pseudo-triangles ∇C^*. The bootstrap estimates $\hat{R}_i^* = \sum_{j \in \Delta_i} \hat{m}_{ij}^*$ and $\hat{R}^* = \sum_{\Delta} \hat{m}_{ij}^*$ are then calculated for each simulation by estimating the parameters of the generalized linear model. In order to get $R_i^{**} = \sum_{j \in \Delta_i} C_{ij}^{**}$ and $R^{**} = \sum_{\Delta} C_{ij}^{**}$ we sample once again from $F(\hat{m}_{ij}, \hat{\phi}\hat{m}_{ij}^p)$ to get C_{ij}^{**} for all $i, j \in \Delta$. Finally, the B observations of (3.10) and (3.13) are inserted into (3.11) to yield the sought predictive distribution.

These methods of bootstrapping for claims reserve uncertainty are described in Figure 1 and are referred to as the non-parametric and the parametric standardized predictive bootstrap.

England & Verrall (1999) and England (2002) use other bootstrap approaches, which are described in Appendix 2. In England (2002) the bootstrap counterparts of the outstanding claims in the real world are obtained by another simulation conditional on the one in Substage 2.1 in Figure 1. In this way the *process error* $R - E(R)$ is bootstrapped differently from Substage 2.2, while Substage 2.1 bootstraps the *estimation error* $\hat{R} - E(R)$. Thus, B triangles $\Delta\hat{m}^\dagger$ are obtained by sampling a random observation $\hat{m}_{ij}^\dagger$ from a distribution with mean $\hat{m}_{ij}^*$ and variance $\phi\,\hat{m}_{ij}^*$ for all $i, j \in \Delta$. The predictive distribution of the outstanding claims R in real world is then obtained by plotting the B values of $\tilde{R}^\dagger = \sum_\Delta m_{ij}^\dagger$. England (2002) suggests using e.g. an over-dispersed Poisson distribution, a negative binomial or a Gamma distribution as the process distribution.

England & Verrall (2006) comment on the approach of including the process error by sampling from a separate distribution, by noting that the non-parametric standardized predictive bootstrap in Pinheiro *et al.* (2003) cannot give larger extremes of the process error than the most extreme residuals observed. Nevertheless, we see no reason to assume separate distributions for the process error and the estimation error. Either we believe in the chosen distribution on the whole and use a parametric predictive bootstrap or we do not and continue to use a non-parametric predictive bootstrap.

3.2 The double bootstrap

It would be preferable to use

$$\text{pe}^{**} = \frac{R^{**} - \hat{R}^*}{\sqrt{\widehat{Var}(R^{**} - \hat{R}^*)}} \tag{3.16}$$

and

$$\tilde{R}^{**} = \hat{R} + \text{pe}^{**}\sqrt{\widehat{Var}(R - \hat{R})} \tag{3.17}$$

instead of (3.10) and (3.11), in particular if the estimation error is much larger than the process error. Although this is more complicated it can be achieved by means of a double bootstrap.

Stage 1 - Real world

· Estimate the parameters in the generalized linear model in equation (3.1).
· Forecast the future expected values $\Delta\hat{m}$ and calculate the fitted values $\nabla\hat{m}$.
· Calculate the residuals for the non-parametric bootstrap and $\hat{\phi}$ for the parametric bootstrap.
· Calculate the outstanding claims $\hat{R}_i = \sum_{j\in\Delta_i}\hat{m}_{ij}$ and $\hat{R} = \sum_{\Delta}\hat{m}_{ij}$.

Stage 2 - Bootstrap world

Substage 2.1 - The estimated outstanding claims

· Resample residuals (3.3) or (3.4) with replacement and insert into (3.8) together with $\nabla\hat{m}$ or sample from $F(\hat{m}_{ij}, \hat{\phi}\hat{m}_{ij}^{p})$ for $i, j \in \nabla$ to obtain the pseudo-reality in ∇C^*.
· Estimate the parameters in the generalized linear model with the pseudo-triangles.
· Forecast the future expected values $\Delta\hat{m}^*$.
· Calculate the estimated outstanding claims $\hat{R}_i^* = \sum_{j\in\Delta_i}\hat{m}_{ij}^*$ and $\hat{R}^* = \sum_{\Delta}\hat{m}_{ij}^*$.

Substage 2.2 - The true outstanding claims

· Resample residuals (3.3) or (3.4) with replacement and insert into (3.9) together with $\Delta\hat{m}$ or sample from $F(\hat{m}_{ij}, \hat{\phi}\hat{m}_{ij}^{p})$ for $i, j \in \Delta$ to obtain the pseudo-reality in ΔC^{**}.
· Calculate the true outstanding claims $R_i^{**} = \sum_{j\in\Delta_i} C_{ij}^{**}$ and $R^{**} = \sum_{\Delta} C_{ij}^{**}$.
· Store the prediction errors pe_i^{**} and pe^{**} in (3.10).
· Return to the beginning of the bootstrap loop and repeat it B times.

Stage 3 - Analysis of the simulations

· Obtain the predictive distribution of R_i and R, the true outstanding claims in the real world, by plotting the B values in (3.11).

Figure 1: *The procedure of the non-parametric and the parametric standardized predictive bootstrap.*

However, the computational complexity of this approach is quite prohibitive because of the nested bootstrap loop and therefore the double bootstrap is not included in our numerical study.

For each of the B bootstrap replicates, we generate $\tilde{B}$ double bootstrap claims reserves R^d and estimated claims reserves $\hat{R}^d$ in analogy with R^{**} and $\hat{R}^*$ in Section 3.1, the difference being that we use ∇C^* as our data rather than ∇C. Then

$$\widehat{Var}(R - \hat{R}) = Var(R^{**} - \hat{R}^* | \nabla C) \tag{3.18}$$

and

$$\widehat{Var}(R^{**} - \hat{R}^*) = Var(R^d - \hat{R}^d | \nabla C^*), \tag{3.19}$$

where the last variance is approximated by the sample variance of all $\tilde{B}$ double bootstrap replicates.

An alternative to (3.18) and (3.19) is to use the variance of the process and the estimation errors in (5.2) in Appendix 2, i.e.

$$\widehat{Var}(R - \hat{R}) = \widehat{Var}(R) + \widehat{Var}(\hat{R}) \tag{3.20}$$

and

$$\widehat{Var}(R^{**} - \hat{R}^*) = \widehat{Var}(R^{**}) + \widehat{Var}(\hat{R}^*), \tag{3.21}$$

where the process errors are estimated by

$$\widehat{Var}(R) = \hat{\phi} \sum_{\Delta} \hat{m}_{ij}^p \tag{3.22}$$

and

$$\widehat{Var}(R^{**}) = \hat{\phi}^* \sum_{\Delta} \hat{m}_{ij}^{*p}. \tag{3.23}$$

The estimation errors are approximated by the sample variance of the corresponding bootstrap replicates

$$\widehat{Var}(\hat{R}) = Var(\hat{R}^*) \tag{3.24}$$

and

$$\widehat{Var}(\hat{R}^*) = Var(\hat{R}^d). \tag{3.25}$$

3.3 Bootstrapping data with a simple underlying model and a reserving algorithm using unstandardized prediction errors

For the purpose of obtaining the predictive distribution of the claims reserve by bootstrapping, the assumption of a generalized linear model in (3.1) is unnecessarily strong. In practise the actuary seldom assumes any model for ∇C and ΔC, but only uses a reserving algorithm in order to estimate $\nabla \hat{m}$ and $\Delta \hat{m}$. Thus, when using the plug-in-principle we just need to make an assumption of the model that generates ∇C^* and ΔC^{**} from the data ∇C, while the reserving algorithm can be used in bootstrap world too in order to estimate $\Delta \hat{m}^*$.

We follow England & Verrall (1999), England (2002) and Pinheiro (2003) and assume independent claims C_{ij} and a variance function in terms of the means, i.e.

$$E(C_{ij}) = m_{ij} \quad \text{and} \quad \mathrm{Var}(C_{ij}) = \phi\, m_{ij}^p \tag{3.26}$$

for some $p > 0$. Thus the mean and variance of C_{ij} are still related as in (3.1), but m_{ij} need no longer satisfy the log-additive conditions in (3.1). Instead the chosen reserving algorithm implicitly specifies the structure of all m_{ij} and produces estimates of $\hat{m}_{ij}$. The bootstrap procedures are then performed as in Section 3.1 with the exception that the residuals (3.3) are used rather than (3.4). The interpretation of n and q as the number of observations and model parameters is still the same. Using the pure chain-ladder method together with the backwards recursive operation described in England & Verrall (1999) implies that $q = 2t - 1$, as for the generalized linear model in (3.1), since this procedure demands the estimation of $t-1$ development factors as well as the t starting values of the backwards recursive operation. Adding exponential smoothing of the development factors, like in the example in Section 2, can indeed complicate the determination of the number of model parameters but the correction factor in (3.3) can be considered as an approximation, although the number of parameters q typically depends on the amount of smoothing.

Standardized prediction errors may still be used, since (3.10) - (3.15) continue to hold. Indeed, it is well known that for many bootstrap procedures, resampling of standardized quantities often increases accuracy compared to using unstandardized quantities, see e.g. Hall (1995).

Nevertheless, the unstandardized prediction errors

$$\mathrm{pe}_i^{**} = R_i^{**} - \hat{R}_i^* \qquad \text{and} \qquad \mathrm{pe}^{**} = R^{**} - \hat{R}^* \tag{3.27}$$

are useful, in particular for the purpose of studying the estimation and the process errors, but also since they are always defined. On the contrary, the denominators of (3.10) may sometimes be non-positive, yielding undefined or imaginary standardized prediction errors, see Section 3.5. The predictive distributions of the outstanding claims R_i and R are then obtained by plotting

$$\tilde{R}_i^{**} = \hat{R}_i + \mathrm{pe}_i^{**} \qquad \text{and} \qquad \tilde{R}^{**} = \hat{R} + \mathrm{pe}^{**} \tag{3.28}$$

for each B. These prediction errors are used in Li (2006).

The alternative bootstrap procedures discussed above are described in detail in Figure 2 and are referred to as the non-parametric and the parametric unstandardized predictive bootstrap.

3.4 Estimation of p

In the literature the most frequent choice of dispersion parameter is $p = 1$ in order to reproduce the chain-ladder estimates under the assumption of a generalized linear model, but as indicated in the method example in Section 2, a pure chain-ladder is seldomly used in practise. Thus, another approach would be to choose the p that best fits the data.

A straightforward way of obtaining a suitable value of p is to use the unstandardized residuals

$$r_{ij} = \sqrt{\frac{n}{n-q}}\,(C_{ij} - \hat{m}_{ij})\ . \tag{3.29}$$

The following relation then holds approximatively

$$E(r_{ij}^2) \approx Var(C_{ij}) = \phi\, m_{ij}^p \tag{3.30}$$

and minimizing the function

$$f(p,\phi) = \sum_{i,j} w_{ij}\,(r_{ij}^2 - \phi\,\hat{m}_{ij}^p)^2, \tag{3.31}$$

Stage 1 - Actuary's best estimate in the real world

· Choose an appropriate reserving method e.g. the one described in Section 2.
· Forecast the future expected values $\Delta\hat{m}$ and calculate the fitted values $\nabla\hat{m}$.
· Calculate the residuals for the non-parametric bootstrap and $\hat{\phi}$ for the parametric bootstrap.
· Calculate the outstanding claims $\hat{R}_i = \sum_{j\in\Delta_i} \hat{m}_{ij}$ and $\hat{R} = \sum_{\Delta} \hat{m}_{ij}$.

Stage 2 - Bootstrap world

Substage 2.1 - The estimated outstanding claims

· Resample residuals (3.3) with replacement and insert into (3.8) together with $\nabla\hat{m}$ or sample from $F(\hat{m}_{ij}, \hat{\phi}\hat{m}_{ij}^p)$ for $i, j \in \nabla$ to obtain the pseudo-reality in ∇C^*.
· Apply the chosen reserving method to ∇C^* to obtain the forecasts $\Delta\hat{m}^*$.
· Calculate the estimated outstanding claims $\hat{R}_i^* = \sum_{j\in\Delta_i} \hat{m}_{ij}^*$ and $\hat{R}^* = \sum_{\Delta} \hat{m}_{ij}^*$.

Substage 2.2 - The true outstanding claims

· Resample residuals (3.3) with replacement and insert into (3.9) together with $\Delta\hat{m}$ or sample from $F(\hat{m}_{ij}, \hat{\phi}\hat{m}_{ij}^p)$ for $i, j \in \Delta$ to obtain the pseudo-reality in ΔC^{**}.
· Calculate the true outstanding claims $R_i^{**} = \sum_{j\in\Delta_i} C_{ij}^{**}$ and $R^{**} = \sum_{\Delta} C_{ij}^{**}$.
· Store the prediction errors pe_i^{**} and pe^{**} in (3.27).
· Return to the beginning of the bootstrap loop and repeat it B times.

Stage 3 - Analysis of the simulations

· Obtain the predictive distribution of R_i and R, the true outstanding claims in the real world, by plotting the B values in (3.28).

Figure 2: *The procedure of the non-parametric and the parametric unstandardized bootstrap.*

where w_{ij} is a weight for observation C_{ij}, with respect to p and ϕ yields an estimator for p. Once a reasonable value of p is chosen and the residuals for the resampling process are defined, ϕ is estimated by (3.7). The simplest choice is to use uniform wights $w_{ij} \equiv 1$ in (3.31). Another possibility is inverse variance weighting, $w_{ij} = \widehat{Var}(r_{ij}^2)^{-1}$. In order to specify these weights, further model assumptions would be needed though.

3.5 Implementation details

There are some major problems with the process of resampling the residuals for the non-parametric bootstrap procedures. Firstly, the bootstrap world is hardly a good approximation of the real world if the claims triangle is small. Furthermore, the basic assumption of identically distributed residuals is certainly violated for $p = 1$, i.e. for an over-dispersed Poisson distribution, see Appendix 1. Depending on the chosen reserving method and the value of p, the standardized residuals in (3.2) sometimes imply a limitation of the set of triangles that can be analyzed, since the residual will be undefined or imaginary whenever a fitted value in $\nabla\hat{m}$ is non-positive. Finally, using the residuals to solve equation (3.8) sometimes results in undesirable negative increments in the pseudo-triangles.

Thus, if the claims triangle ∇C is small, a parametric bootstrap procedure seems preferable. On the other hand, if we know nothing about F and have a large triangle, a non-parametric bootstrap procedure would be our first choice. Note, however, that a parametric bootstrap procedure does not solve the problem with undefined residuals since they are needed in order to estimate ϕ as well. Furthermore, a parametric bootstrap procedure should be used if negative increments in the pseudo-triangles are unacceptable and a gamma distribution should particularly be used if it is undesirable that the increments only take on the values zero and multiples of ϕ, which is the case for the over-dispersed Poisson distribution.

The choice of prediction errors causes another problem. The standardized ones in (3.10) are sensitive to pseudo-triangles where the row sums of the outstanding claims are non-positive. An ad hoc solution is simply to cut out these pseudo-triangles from the simulation process if they are rare, another solution is to use the unstandardized prediction errors in (3.27) instead. The unstandardized ones, on the other hand, result in a predictive distribution which is more skewed to the left than the distribution obtained by the standardized prediction errors, see Section 4 for more details.

Since England & Verrall (1999), England (2002) and Pinheiro *et al.* (2003) replace the maximum likelihood estimation of the parameters in (3.1) by chain-ladder when $p = 1$, the same method is adopted here for the standardized predictive distribution in Figure 1, even though the non-positive column sums of the pseudo-triangles make the estimates disagree.

In this paper, the estimated value of p in Section 3.4 is just considered as an indicator of whether $p = 1$ or $p = 2$ should be used in the non-parametric bootstrap and whether an over-dispersed Poisson distribution or a gamma distribution should be chosen in a parametric bootstrap. The distributions of the residuals corresponding to different choices of $p \in (1, 2)$ should indeed be investigated, but this is beyond the scope of this paper.

3.6 Dynamic Financial Analysis and the one year run-off result

See Kaufmann *et al.* (2001) for an introduction to Dynamic Financial Analysis. Here the movements of the claims reserve are of particular interest. The one year run-off result is the change in the reserve during the financial year and is defined as the difference between the opening reserve at the beginning of the year and the sum of payments during the year and the closing reserve of the same portfolio at the end of the year. Thus, if we at the end of year t want to make predictions of the run-off result at the end of the unobserved year $t + 1$, and if we do not add neither a new accident year nor a new development year, we have to find the predictive distribution of

$$\hat{R}^t - \left(\sum_{i=2}^{t} C_{i,t+2-i} + \hat{R}^{t+1}\right), \tag{3.32}$$

where $\hat{R}^t$ and $\hat{R}^{t+1}$ are the estimated outstanding claims at the end of year t and $t + 1$ respectively.

One method to obtain the predictive distribution of the one year run-off result is to condition on the claims triangle ∇C. $\hat{R}^t$ is then considered fixed, while the predictive distribution of the payments corresponding to the forecast year $t+1$ is obtained by B times simulating the new diagonal $\{(i,j);\, i+j=t+2\}$ by one of the bootstrap procedures discussed above. This is done by storing e.g. the unstandardized prediction errors $\mathrm{pe}_{ij}^{**} = C_{ij}^{**} - \hat{C}_{ij}^{*}$ of each increment in the new diagonal and then adding them to the corresponding estimated values $\hat{C}_{ij}$ in the real world to obtain $\tilde{C}_{ij}^{**} = \hat{C}_{ij} + \mathrm{pe}_{ij}^{**}$. In this way B pseudo-triangles, consisting of the fixed triangle ∇C known at the end of year t and a new simulated diagonal $\tilde{C}^{**}$, are generated and the outstanding claims are then recalculated by the same reserving method as before, in order to obtain B records of $\hat{R}^{t+1\,*} = \hat{R}^{t+1}(\nabla C \cup \tilde{C}^{**})$. Finally the B values of

$$\hat{R}^t - (\sum_{i=2}^{t} \tilde{C}_{i,t+2-i}^{**} + \hat{R}^{t+1\,*})|\nabla C\,, \tag{3.33}$$

are investigated in order to estimate the predictive distribution of the one year run-off result.

De Felice & Moriconi (2003) use a similar method in order to analyze $\hat{R}^{t+1}$, but in the simulation process the oldest accident year is removed, while a new accident year, corresponding to the year $t+1$, is added to the pseudo-triangle.

4 Numerical study

The purpose of the numerical study is to compare the non-parametric and the parametric bootstrap procedures under different choices of p, F and prediction errors. Since the actuary chooses an age-to-age development factor method that fits the particular development triangle under analysis, it is difficult to find one single algorithm that works for all situations. Therefore we only use the pure chain-ladder method in the comparisons, even though the bootstrap procedures allow the use of any age-to-age development factor method as well. From now on $B = 10\,000$ simulations are used for each prediction. The upper 95 percent limits are studied due to higher robustness than, e.g., the 99.5 percentile, which is perhaps the most frequent choice in practise. The coefficients of variation are also presented.

4.1 The triangle from Taylor & Ashe (1983)

4.1.1 Comparison with Pinheiro *et al.* (2003)

First, the well-known triangle from Taylor & Ashe (1983), called Data 1 in Table 4.1, is analyzed by the non-parametric standardized predictive bootstrap procedure, i.e. the bootstrap procedure described in Pinheiro *et al.* (2003). The estimated reserves and the upper 95 percent limits for $p = 1$ and $p = 2$ are presented in Table 4.2. The second accident year is left out from the tabulation of results when $p = 1$ since a negative increment in the northeast corner of a pseudo-triangle causes a situation with an imaginary prediction error for that year. The remaining accident years are not as sensitive to negative increments as this year.

The results of the standardized predictive bootstrap procedure are in accordance with Pinheiro *et al.* (2003). As we can see, for earlier accident years, the $p = 2$ percentiles are smaller than the $p = 1$ percentiles, whereas the opposite is true for later accident years. This is natural, since most of the future claims C_{ij} of later years have large m_{ij} and hence larger variance for $p = 2$ than for $p = 1$.

	1	2	3	4	5	6	7	8	9	10
1	357 848	766 940	610 542	482 940	527 326	574 398	146 342	139 950	227 229	67 948
2	352 118	884 021	933 894	1 183 289	445 745	320 996	527 804	266 172	425 046	
3	290 507	1 001 799	926 219	1 016 654	750 816	146 923	495 992	280 405		
4	310 608	1 108 250	776 189	1 562 400	272 482	352 053	206 286			
5	443 160	693 190	991 983	769 488	504 851	470 639				
6	396 132	937 085	847 498	805 037	705 960					
7	440 832	847 631	1 131 398	1 063 269						
8	359 480	1 061 648	1 443 370							
9	376 686	986 608								
10	344 014									

Table 4.1: *Data 1 from Taylor & Ashe (1983).*

4.1.2 The choice of $\hat{\phi}$ or $\hat{\phi}^*$

We continue to use the non-parametric standardized predictive bootstrap and Data 1, but we now replace (3.13) with (3.14) in Substage 2.2 in Figure 1. Thus, we do not consider ϕ as constant for the data and therefore we replace $\hat{\phi}$ by $\hat{\phi}^*$. The results are presented in Table 4.3. As we can see, the replacement hardly affects the results.

Note that since $p = 1$ occasionally yields $\hat{m}^*_{ij} < 0$ the corresponding Pearson residuals in the bootstrap world are imaginary while $\hat{\phi}^*$ is real. Since the assumption of an over-dispersed Poisson distribution for the parametric procedure occasionally yields $\hat{m}^*_{ij} = 0$, the corresponding Pearson residuals in the bootstrap world are undefined and as a result, $\hat{\phi}^*$ is undefined as well. Thus, in the sequel we use (3.13) in all simulations.

Year	Estimated reserve	95% $p = 1$	Estimated reserve	95% $p = 2$
2	94 634		93 316	222 789
3	469 511	903 221	446 504	799 700
4	709 638	1 187 641	611 145	992 585
5	984 889	1 527 903	992 023	1 497 633
6	1 419 459	2 076 496	1 453 085	2 170 480
7	2 177 641	3 034 860	2 186 161	3 284 490
8	3 920 301	5 277 768	3 665 066	5 692 764
9	4 278 972	6 139 286	4 122 398	6 975 123
10	4 625 811	9 760 307	4 516 073	9 286 282
Total	18 680 856	23 681 062	18 085 772	23 033 968

Table 4.2: *The estimated reserves and the 95 percentiles of the non-parametric standardized predictive bootstrap with (3.13) used in Substage 2.2 of Figure 1 for Data 1. Chain-ladder is used for $p = 1$ and maximum likelihood estimation for $p = 2$.*

Year	Estimated reserve	95% $p = 1$	Estimated reserve	95% $p = 2$
2	94 634		93 316	216 698
3	469 511	889 639	446 504	796 146
4	709 638	1 186 623	611 145	978 315
5	984 889	1 533 399	992 023	1 497 722
6	1 419 459	2 082 287	1 453 085	2 136 423
7	2 177 641	3 041 716	2 186 161	3 290 061
8	3 920 301	5 290 749	3 665 066	5 738 496
9	4 278 972	6 181 331	4 122 398	6 795 927
10	4 625 811	9 328 277	4 516 073	9 476 343
Total	18 680 856	23 603 123	18 085 772	23 042 954

Table 4.3: *The estimated reserves and the 95 percentiles of the non-parametric standardized predictive bootstrap when (3.13) is replaced by (3.14) in Substage 2.2 in Figure 1 for Data 1. Chain-ladder is used for $p = 1$ and maximum likelihood estimation for $p = 2$.*

4.1.3 Maximum likelihood estimation vs chain-ladder when p = 2

The next step is to replace the maximum likelihood estimator of the model parameters by the chain-ladder for the non-parametric standardized predictive bootstrap when $p = 2$. (We already use the chain-ladder when $p = 1$, cf. Section 3.5.) Consequently, the estimated reserves in Table 4.4 are the same as when $p = 1$ in Table 4.2 whereas the percentiles in Table 4.4 are consistently higher than in Table 4.2.

This is an example of bootstrapping under a model that does not produce the estimator actually employed, a model which might nevertheless be quite realistic for paid claims. We will use the chain-ladder in all remaining numerical studies, since it is popular and simple.

Year	Estimated reserve	95% $p = 2$
2	94 634	236 850
3	469 511	875 382
4	709 638	1 156 050
5	984 889	1 503 685
6	1 419 459	2 141 470
7	2 177 641	3 308 805
8	3 920 301	6 199 841
9	4 278 972	7 646 140
10	4 625 811	10 698 797
Total	18 680 856	23 991 584

Table 4.4: *The estimated reserve and the 95 percentiles of the non-parametric standardized predictive bootstrap with (3.13) used in Substage* 2.2 *in Figure 1 for Data 1. Chain-ladder is used for* $p = 2$.

4.1.4 Non-parametric bootstrap vs parametric bootstrap

For the purpose of comparing the non-parametric and the parametric bootstrap procedures we continue to use the standardized predictive bootstrap with chain-ladder for Data 1. See Table 4.5 for the upper 95 percent limits and Table 4.6 for the coefficients of variation, i.e. $\sqrt{Var(\tilde{R_i^{**}})}/\hat{R}_i$ and $\sqrt{Var(\tilde{R^{**}})}/\hat{R}$. (In the tables ODP denotes the over-dispersed Poisson distribution.)

The results of the parametric bootstrap coincide well with the results of the non-parametric bootstrap except for the last accident year. It is well-known that the chain-ladder estimate of the outstanding claims for the last accident year is extremely sensitive to outliers in the south corner of the upper triangle. If C^*_{t1} happens to be small in the pseudo-triangle then the corresponding reserve $\hat{R}^*_t$ will be small compared to R^{**}_t, which affects the prediction error in (3.10). The parametric bootstrap generates more stable C^*_{t1}:s than the non-parametric bootstrap, consequently there is a discrepancy in the results of the last accident year for the non-parametric and the parametric bootstrap procedures in Tables 4.5 - 4.6. The conclusion is that the parametric bootstrap may be preferable in some cases.

Year	Estimated Reserve	Non-parametric $p=1$	Parametric ODP	Non-parametric $p=2$	Parametric Gamma
2	94 634			236 850	220 643
3	469 511	903 221	920 956	875 382	866 833
4	709 638	1 187 641	1 215 254	1 156 050	1 162 942
5	984 889	1 527 903	1 537 266	1 503 685	1 516 868
6	1 419 459	2 076 496	2 096 805	2 141 470	2 150 441
7	2 177 641	3 034 860	3 057 599	3 308 805	3 309 838
8	3 920 301	5 277 768	5 308 472	6 199 841	6 192 286
9	4 278 972	6 139 286	6 192 655	7 646 140	7 272 012
10	4 625 811	9 760 307	9 163 520	10 698 797	9 222 470
Total	18 680 856	23 681 062	23 685 724	23 991 584	24 095 302

Table 4.5: *The estimated reserve and the 95 percentiles of the non-parametric and the parametric standardized predictive bootstrap with (3.13) used in Substage* 2.2 *in Figures 1- 2 for Data 1. Chain-ladder is used in both cases.*

4.1.5 Standardized prediction errors vs unstandardized prediction errors

From now on the unstandardized predictive bootstrap procedures are used in all tables; the results for Data 1 are presented in Tables 4.7 - 4.8. As we can see, the percentiles for the unstandardized predictive bootstrap in Table 4.7 are lower than for the standardized predictive bootstrap in Table 4.5, and the same goes for the coefficients of variation. Note that there is a large discrepancy in the coefficients of variation, in Table 4.8, for the two choices of distribution for year 2. The reason for the extreme values, when $p = 1$ or an over-dispersed Poisson distribution is assumed, is discussed in Section 4.3.

In Figures 3 (c) - (d) and 4 (c) - (d) the predictive distributions of the total claims reserve are plotted when assuming $p = 1$ for the non-parametric bootstrap procedures and an over-dispersed Poisson distribution for the parametric bootstrap procedures. The predictive distribution obtained by the unstandardized bootstrap in (c) is slightly skewed to the left compared to the one obtained by the standardized bootstrap in (d), which is almost symmetric. This follows since the process component (Figures 3 (a) and 4 (a)) has smaller variability than the estimation component (Figures 3 (b) and 4 (b)), and the latter is slightly skewed to the right. This skewness is to a large extent removed for the standardized prediction errors (3.10), because of the denominator, but not for the unstandardized prediction errors (3.27). Furthermore, from Figures 3 (a) and 4 (a), it does not seem to matter whether we use a non-parametric or parametric approach for the process error, even though England & Verrall (2006) argue that the former choice cannot give larger extremes than the most extreme residual observed. The same holds for $p = 2$ or a gamma distribution (results not shown here).

4.1.6 Estimation of p

Estimation of p by minimizing the (unweighted) sum in (3.31) yields $p = 0.7280$. Thus, $p = 1$ or an over-dispersed Poisson distribution seems to be more reasonable for this development triangle.

Year	Estimated Reserve	Non-parametric $p = 1$	Parametric ODP	Non-parametric $p = 2$	Parametric Gamma
2	94 634			76	62
3	469 511	50	50	43	42
4	709 638	37	38	32	32
5	984 889	31	31	27	28
6	1 419 459	27	27	26	26
7	2 177 641	23	23	27	26
8	3 920 301	20	20	30	29
9	4 278 972	24	25	38	35
10	4 625 811	53	50	64	48
Total	18 680 856	16	16	15	16

Table 4.6: *The estimated reserve and the coefficients of variation of the simulations (in %) of the non-parametric and the parametric standardized predictive bootstrap with (3.13) used in Substage* 2.2 *in Figures 1- 2 for Data 1. Chain-ladder is used in both cases.*

Year	Estimated Reserve	Non-parametric $p=1$	Parametric ODP	Non-parametric $p=2$	Parametric Gamma
2	94 634	274 891	252 438	168 132	167 585
3	469 511	823 274	814 256	750 175	754 646
4	709 638	1 148 468	1 125 650	1 055 135	1 064 059
5	984 889	1 486 951	1 475 088	1 414 799	1 403 919
6	1 419 459	2 040 277	2 019 093	1 995 397	1 982 611
7	2 177 641	2 983 269	2 979 860	3 043 356	3 049 215
8	3 920 301	5 201 768	5 171 112	5 579 973	5 564 848
9	4 278 972	5 916 186	5 910 048	6 363 139	6 257 000
10	4 625 811	7 755 623	7 517 443	7 387 885	7 088 050
Total	18 680 856	23 264 493	23 122 056	23 109 992	23 107 180

Table 4.7: *The estimated reserve and the 95 percentiles of the non-parametric and the parametric unstandardized predictive bootstrap when chain-ladder is used for Data 1.*

Year	Estimated Reserve	Non-parametric $p=1$	Parametric ODP	Non-parametric $p=2$	Parametric Gamma
2	94 634	121	118	52	50
3	469 511	47	46	39	38
4	709 638	38	37	31	31
5	984 889	31	31	28	27
6	1 419 459	27	26	26	26
7	2 177 641	23	23	26	26
8	3 920 301	21	21	28	27
9	4 278 972	25	25	32	32
10	4 625 811	46	44	40	38
Total	18 680 856	17	16	17	16

Table 4.8: *The estimated reserve and the coefficients of variation of the simulations (in %) of the non-parametric and the parametric unstandardized predictive bootstrap when chain-ladder is used for Data 1.*

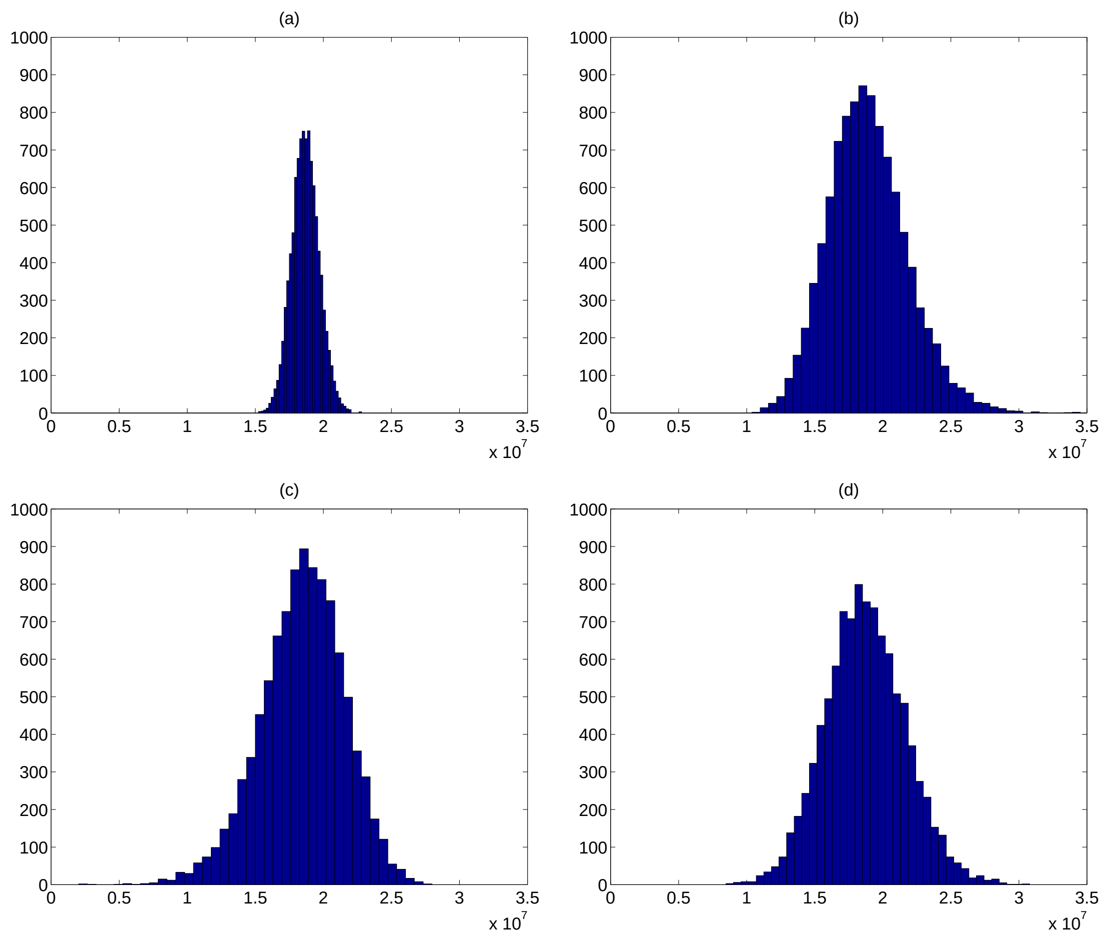

Figure 3: *Density charts of* R^{**} *(a),* $\hat{R}^{*}$ *(b) and* $\tilde{R}^{**}$ *for the unstandardized (c) and standardized (d) non-parametric predictive bootstrap procedures for Data 1 when* $p = 1$.

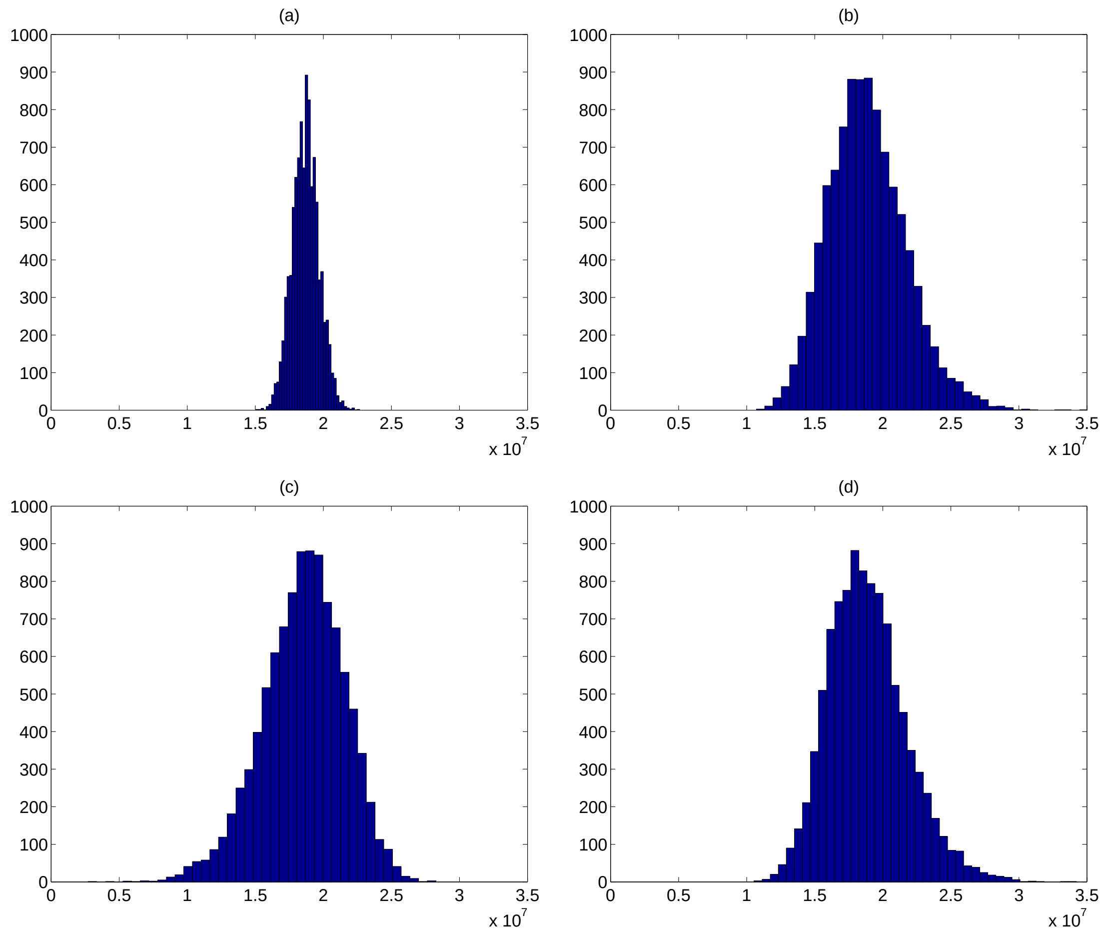

Figure 4: *Density charts of* R^{**} *(a),* $\hat{R}^{*}$ *(b) and* $\tilde{R}^{**}$ *for the unstandardized (c) and standardized (d) parametric predictive bootstrap procedures for Data 1 under the assumption of an over-dispersed Poisson distribution.*

4.2 A small triangle of claim counts

The non-parametric and the parametric unstandardized predictive bootstrap procedures are now compared on a triangle of claim counts appearing in Taylor (2000). Because of the shape of the data and in order to avoid non-positive column sums we use just the later part of the original triangle, see Table 4.9. This is reasonable since the claim counts from previous accident years are almost finalized.

	1	2	3	4	5	6	7
1989	589	210	29	17	12	4	9
1990	564	196	23	12	9	5	
1991	607	203	29	9	7		
1992	674	169	20	12			
1993	619	190	41				
1994	660	161					
1995	660						

Table 4.9: *Data 2 from Taylor (2000).*

Estimation of p yields $\hat{p} = 0.5596$, which indicates that $p = 1$ is a better choice than $p = 2$ for the non-parametric bootstrap and an over-dispersed Poisson distribution is preferable for the parametric bootstrap, as expected for claim counts. Nevertheless, the results for both choices are presented in Tables 4.10 - 4.11 and, as we can see, the results of the parametric bootstrap coincides well with the results of the non-parametric one.

The density charts of R^{**} and $\hat{R}^{*}$ are plotted in Figure 5. The variability of the estimation error is larger than the variability of the process error for Data 2 too, but the difference is not as extreme as for Data 1 in Figures 3 - 4.

4.3 A small triangle of paid claims from a short-tailed line of business

Table 4.12 shows a triangle of paid claims, provided by the Swedish insurance company *AFA Försäkring*, for the short-tailed line of business *Severance Grant*.

The results of the bootstrap procedures are presented in Tables 4.13 - 4.14. The percentiles for year 1996 are very different for the two choices of distribution. This is a consequence

Year	Estimated Reserve	Non-parametric $p=1$	Parametric ODP	Non-parametric $p=2$	Parametric Gamma
1990	8	19	18	14	14
1991	14	26	26	20	20
1992	24	40	39	34	34
1993	36	56	55	51	50
1994	65	90	89	91	90
1995	269	323	321	400	399
Total	417	500	496	555	554

Table 4.10: *The estimated reserve and the 95 percentiles of the non-parametric and the parametric unstandardized predictive bootstrap when chain-ladder is used for Data 2.*

Year	Estimated Reserve	Non-parametric $p=1$	Parametric ODP	Non-parametric $p=2$	Parametric Gamma
1990	8	74	71	43	42
1991	14	57	55	35	33
1992	24	40	39	29	28
1993	36	32	31	26	25
1994	65	23	22	25	25
1995	269	12	12	32	31
Total	417	12	12	22	21

Table 4.11: *The estimated reserve and the coefficients of variation of the simulations (in %) of the non-parametric and the parametric unstandardized predictive bootstrap when chain-ladder is used for Data 2.*

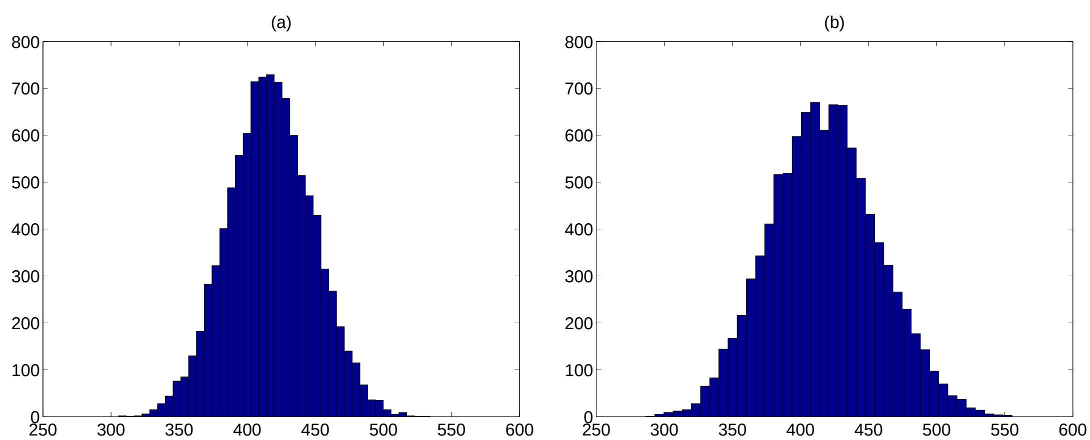

Figure 5: *Density charts of R^{**} (a) and $\hat{R}^{*}$ (b) for the unstandardized non-parametric predictive bootstrap procedures for Data 2 when $p = 1$.*

of occasional non-positive $\hat{m}^*_{ij}$ caused by the resampling process. Tables 4.15 - 4.16 show examples of pseudo-triangles when $p = 1$ for the non-parametric bootstrap procedure and an over-dispersed Poisson distribution is assumed for the parametric bootstrap procedure. By (3.27) and (3.28) these particular simulations yield $\tilde{R}^{**}_{1996} = 2\,614$ and $\tilde{R}^{**}_{1996} = 2\,876$, respectively, which is not reasonable. Thus, even though $\hat{p} = 1.1915$, a comparison of the results for $p = 1$ and $p = 2$ indicates that $p = 2$ might be a better choice for this triangle. Another alternative might be to use a truncated over-dispersed Poisson distribution to exclude zero values, but this is outside the scope of the present paper.

The density charts of R^{**} and $\hat{R}^*$ are plotted in Figure 6 and, as for previous data, the variability of the estimation error is larger than the variability of the process error.

4.4 A large triangle of paid claims from a long-tailed line of business

Finally the two bootstrap procedures are applied to a development triangle for Motor TPL, a typically long-tailed line of business, where there are still unreported claims. Due to an

	1	2	3	4	5	6	7
1995	48 545	56 786	32 659	12 973	4 005	1 696	490
1996	58 294	79 824	38 287	15 957	4 617	1 427	
1997	73 859	73 237	35 281	13 960	3 854		
1998	65 707	67 632	32 832	12 158			
1999	92 901	80 931	36 508				
2000	66 834	47 630					
2001	45 838						

Table 4.12: *Data 3 provided by the Swedish insurance company AFA Försäkring.*

Year	Estimated Reserve	Non-parametric $p=1$	Parametric ODP	Non-parametric $p=2$	Parametric Gamma
1996	621	2 369	2 124	873	862
1997	2 408	5 377	5 382	3 128	3 116
1998	6 317	10 763	10 823	8 027	7 960
1999	25 536	34 668	34 673	32 242	32 163
2000	46 196	59 249	58 820	58 910	58 395
2001	82 821	107 213	105 455	110 188	108 440
Total	163 898	195 586	195 097	195 876	193 573

Table 4.13: *The estimated reserve and the 95 percentiles of the non-parametric and the parametric unstandardized predictive bootstrap when chain-ladder is used for Data 3.*

Year	Estimated Reserve	Non-parametric $p = 1$	Parametric ODP	Non-parametric $p = 2$	Parametric Gamma
1996	621	173	169	26	25
1997	2 408	77	74	19	18
1998	6 317	44	42	17	16
1999	25 536	22	22	17	16
2000	46 196	17	17	17	17
2001	82 821	17	17	21	20
Total	163 898	12	12	12	12

Table 4.14: *The estimated reserve and the coefficients of variation of the simulations (in %) of the non-parametric and the parametric unstandardized predictive bootstrap when chain-ladder is used for Data 3.*

					757
				2 007	-746
			604	3 300	326
		16 480	6 116	2 838	1 108
	30 487	11 674	2 527	1 924	-948
47 952	16 537	14 875	2 315	640	-368

					-1 236
				1 949	-1 367
			2 653	1 561	-1 095
		17 954	3 478	2 046	-1 435
	26 073	11 364	2 201	1 295	-908
40 247	18 232	7 947	1 539	905	-635

Table 4.15: *An example of pseudo-triangles when $p = 1$; the left triangle is ΔC^{**} and the right triangle is $\Delta\hat{m}^*$.*

					2 255
				1 503	2 255
			3 758	3 006	0
		18 788	2 255	3 006	2 255
	30 061	12 024	3 006	752	752
42 085	27 055	9 770	2 255	2 255	752

					0
				1 712	0
			4 919	1 716	0
		16 892	5 385	1 878	0
	28 929	12 041	3 838	1 339	0
55 292	28 841	12 004	3 827	1 335	0

Table 4.16: *An example of pseudo-triangles when an over-dispersed Poisson distribution is assumed; the left triangle is ΔC^{**} and the right triangle is $\Delta\hat{m}^*$.*

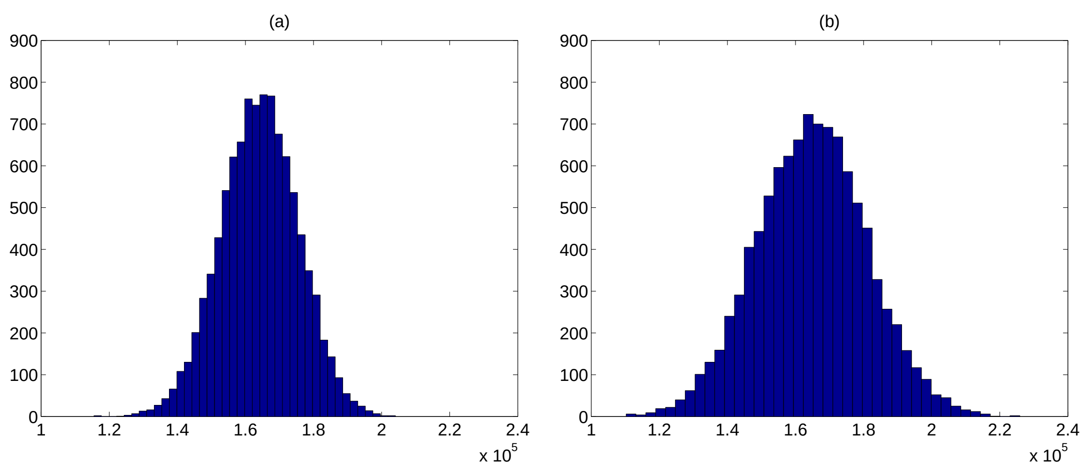

Figure 6: *Density charts of R^{**} (a) and $\hat{R}^*$ (b) for the unstandardized non-parametric predictive bootstrap procedures for Data 3 when $p = 1$.*

outlier in the oldest accident year (1987) we exclude this year from the original triangle in Naziropoulou (2005), see Table 4.17 for Data 4.

Estimation of p yields $\hat{p} = 0.7773$ and the results of the bootstrap procedures are presented in Tables 4.18 - 4.19. The conclusions are the same as in the earlier examples. The density charts of R^{**} and $\hat{R}^{*}$ are plotted in Figure 7 and for Data 4 the variability of the estimation error is again larger than the variability of the process error.

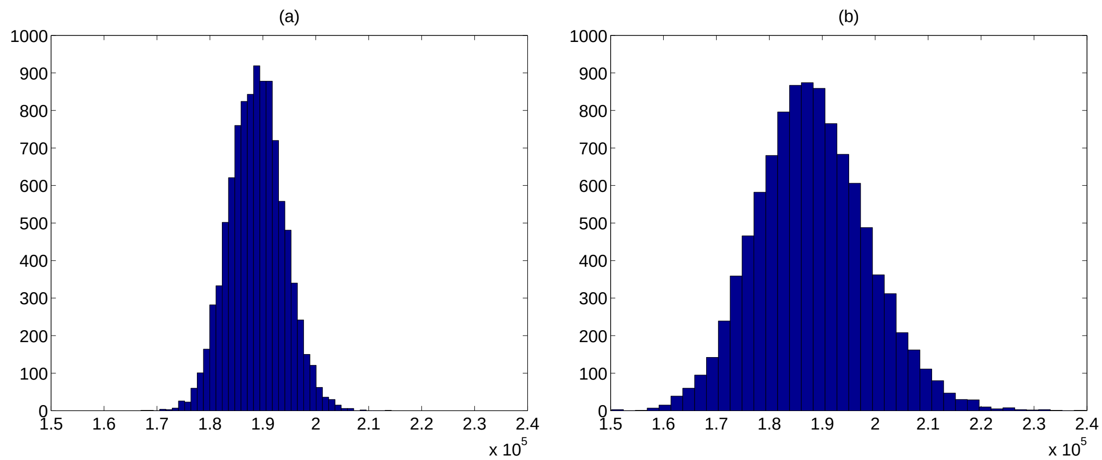

Figure 7: *Density charts of R^{**} (a) and $\hat{R}^{*}$ (b) for the unstandardized non-parametric predictive bootstrap procedures for Data 4 when $p = 1$.*

5 Conclusions

So far most papers concerning bootstrapping for claims reserve uncertainty focus on obtaining the predictive distribution for the chain-ladder method by assuming underlying models, which reproduce the chain-ladder estimates. However, the assumption of an underlying model is generally not made in practise for the purpose of estimating the claims reserve, since the actuary rather uses somewhat complex reserving algorithms, without reference to statistical models. In this paper we suggest using either a non-parametric or a parametric bootstrap methodology with as few model assumptions as possible in order to make the bootstrap pro-

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1988	13 286	9 064	2 260	1 271	1 295	1 007	1 484	150	1 110	598	780	1 262	1 470	350	881	496	170
1989	12 428	9 740	2 387	1 751	1 261	902	1 054	1 086	1 378	1 983	634	1 129	1 346	700	844	1 142	
1990	13 292	8 996	2 615	1 493	1 462	834	1 102	734	1 297	1 160	781	2 021	997	416	417		
1991	13 174	9 023	2 476	1 586	1 361	1 056	758	955	972	1 468	1 029	2 483	599	996			
1992	12 300	8 562	2 444	1 282	1 444	637	1 474	1 368	944	1 328	1 013	1 250	1 009				
1993	12 710	7 747	2 242	2 164	1 478	1 263	1 069	2 160	962	3 870	803	475					
1994	11 935	8 340	2 814	1 870	1 464	1 107	1 221	1 214	1 617	1 310	1 591						
1995	11 959	9 377	2 804	2 488	1 746	1 466	3 168	1 832	1 763	2 051							
1996	11 518	8 953	3 269	1 865	1 522	1 753	1 770	1 717	2 084								
1997	11 621	8 233	3 705	2 091	2 080	1 697	1 800	2 418									
1998	12 416	8 518	2 670	1 951	1 861	1 365	1 874										
1999	12 957	8 917	3 172	2 550	2 141	2 116											
2000	12 964	10 432	3 060	2 382	1 606												
2001	14 959	12 404	4 017	2 663													
2002	16 890	11 899	3 633														
2003	17 167	11 629															
2004	17 658																

Table 4.17: *Data 4 from Naziropoulou (2005).*

Year	Estimated Reserve	Non-parametric $p = 1$	Parametric ODP	Non-parametric $p = 2$	Parametric Gamma
1989	184	551	608	311	314
1990	1 000	1 785	1 810	1 528	1 538
1991	1 765	2 783	2 773	2 523	2 530
1992	2 250	3 401	3 386	3 084	3 080
1993	3 586	5 010	5 002	4 798	4 819
1994	4 947	6 611	6 563	6 576	6 580
1995	6 811	8 805	8 761	9 014	8 952
1996	8 245	10 607	10 523	10 902	10 886
1997	9 865	12 444	12 460	13 060	12 988
1998	10 797	13 455	13 493	14 131	14 245
1999	13 529	16 623	16 531	17 759	17 764
2000	14 933	18 240	18 179	19 716	19 661
2001	19 798	23 719	23 700	26 008	26 280
2002	22 920	27 141	27 176	30 525	30 771
2003	26 757	31 598	31 539	36 447	36 359
2004	40 854	48 032	48 283	61 070	60 315
Total	188 242	207 770	207 461	218 375	217 784

Table 4.18: *The estimated reserve and the 95 percentiles of the non-parametric and the parametric unstandardized predictive bootstrap when chain-ladder is used for Data 4.*

Year	Estimated Reserve	Non-parametric $p = 1$	Parametric ODP	Non-parametric $p = 2$	Parametric Gamma
1989	184	131	129	48	47
1990	1 000	49	49	35	34
1991	1 765	36	36	28	28
1992	2 250	31	31	24	24
1993	3 586	25	24	22	22
1994	4 947	21	21	22	21
1995	6 811	18	18	21	20
1996	8 245	16	16	20	20
1997	9 865	15	15	20	20
1998	10 797	14	15	20	20
1999	13 529	13	13	20	20
2000	14 933	13	13	21	20
2001	19 798	12	12	21	21
2002	22 920	11	11	22	22
2003	26 757	11	11	24	24
2004	40 854	11	11	34	33
Total	188 242	6	6	11	10

Table 4.19: *The estimated reserve and the coefficients of variation of the simulations (in %) of the non-parametric and the parametric unstandardized predictive bootstrap when chain-ladder is used for Data 4.*

cedures more consistent with the actuary's way of working. It is assumed that the bootstrap procedures only depend on the mean and variance of the claims, while the actuary's choice of reserving algorithm implicitly specifies the mean structure. Consequently, the suggested bootstrap procedures can be used to obtain the predictive distribution of any age-to-age development factor method. The non-parametric and the parametric bootstrap procedures are compared to techniques described in Pinheiro *et al.* (2003), as well as in England (2002), and finally they are applied to four development triangles of different types.

We have seen that the results of the parametric standardized predictive bootstrap are consistent with the results of its non-parametric counterpart in Pinheiro *et al.* (2003). Furthermore, the unstandardized predictive bootstrap procedures have revealed that the variability of the estimation error, when chain-ladder is used, is larger than the variability of the process error for all four investigated development triangles and for the two largest of them the difference is considerable. Finally, our simulation results are almost the same for the non-parametric and the parametric approach.

Since resampling of standardized quantities often increases accuracy compared to using unstandardized quantities, the standardized predictive bootstrap is in theory preferable to the unstandardized one. We have seen that the standardized case yields higher estimated risk, seemingly due to the fact that it makes the distribution more symmetric than the unstandardized case, where the predictive distribution is skewed to the left. A disadvantage of the standardized predictive bootstrap is that the denominators of (3.10) may sometimes be non-positive, yielding undefined or imaginary prediction errors. In principle, this could be corrected by the double bootstrap, which provides a better estimation of the variance since it includes the estimation error as well as the process error. Therefore, it would be interesting, in a future paper, to analyze the behaviour of the double bootstrap method both for simulated and real data sets.

Finally, a somewhat surprising result of the numerical studies is that the estimation error is consistently larger than the process error. This could be the case for further study.

References

Buchwalder, M., Bühlmann, H., Merz, M. & Wüthrich, M. (2006): *The Mean Square Error of Prediction in the Chain Ladder Reserving Method (Mack and Murphy revisited).* ASTIN Bulletin, 36(2), 521-542.

Davison, A. C. & Hinkley, D. V. (1997): *Bootstrap Methods and their Applications.* Cambridge University Press.

De Felice, M. & Moriconi, F. (2003): *Risk Based Capital in P&C Loss Reserving or Stressing the Triangle.* Working paper. Available online at: http://www.math.ethz.ch/finance/PC_DFM.pdf

Efron, B. & Tibshirani, R.J. (1993): *An Introduction to the Bootstrap.* Chapman & Hall, N. Y.

England, P. (2002): *Addendum to "Analytic and Bootstrap Estimates of Prediction Error in Claims Reserving".* Insurance: Mathematics and Economics, 31, 461-466.

England, P. & Verrall, R. (1999): *Analytic and Bootstrap Estimates of Prediction Errors in Claims Reserving.* Insurance: Mathematics and Economics, 25, 281-293.

England, P. & Verrall, R. (2006): *Predictive Distributions of Outstanding Liabilities in General Insurance.* Annals of Actuarial Science, 1, II, 221-270.

Hall, P. (1995): *The Bootstrap and Edgeworth Expansion.* Springer.

Kaufmann, R., Gadmer, A. & Klett, R. (2001): *Introduction to Dynamic Financial Analysis.* ASTIN Bulletin, 31(1), 213-249.

Larsen, C. R. (2007): *An Individual Claims Reserving Model.* ASTIN Bulletin, 37(1), 113-132.

Li, J. (2006): *Comparison of Stochastic Reserving Methods.* Australian Actuarial Journal, volume 12, issue 4, 489-569.

Mack, T. (1993): *Distribution Free Calculation of the Standard Error of Chain Ladder Reserve Estimates.* ASTIN Bulletin, 23(2), 213-225.

Mack, T., Quarg, G. & Braun, C. (2006): *The Mean Square Error of Prediction in the Chain Ladder Reserving Method - a Comment.* ASTIN Bulletin, 36(2), 543-552.

Mack, T. & Venter, G. (2000): *A Comparison of Stochastic Models that Reproduce Chain Ladder Reserve Estimates.* Insurance: Mathematics and Economics, 26, 101-107.

McCullagh, P. & Nelder, J.A. (1989): *Generalized Linear Models.* London: Chapman and Hall.

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. Available online at: http://www.math.su.se/matstat/reports/serieb/2005/rep7/report.pdf (In Swedish.)

Pinheiro, P.J.R., Andrade e Silva, J.M. & Centeno, M.d.L (2003): *Bootstrap Methodology in Claim Reserving.* The Journal of Risk and Insurance, 4, 701-714.

Renshaw, A.E. & Verrall, R.J. (1998): *A Stochastic Model Underlying the Chain-Ladder Technique.* British Actuarial Journal, 4, IV, 903-923.

Taylor, G. (2000): *Loss Reserving - An Actuarial Perspective.* Boston: Kluwer Academic Press.

Taylor, G. & Ashe, F.R. (1983): *Second Moments of Estimates of Outstanding Claims.* Journal of Econometrics, 23, 37-61.

Verrall, R. (2000): *An Investigation into Stochastic Claims Reserving Models and the Chain-ladder Technique.* Insurance: Mathematics and Economics, 26, 91-99.

Verrall, R. & England, P. (2000): *Comments on: "A Comparison of Stochastic Models that reproduce Chain Ladder Reserve Estimates", by Mack and Venter.* Insurance: Mathematics and Economics, 26, 109-111.

Appendix 1

The basic assumption of the resampling process of the non-parametric bootstrap is independent and identically distributed residuals. We will now motivate that the model in (3.1) gives approximately identically distributed residuals r_{ij} for the majority of residuals (3.2) or (3.3) in the upper triangle (not close to any of the corners) when $p = 2$ (gamma distribution), but not for $p = 1$ (over-dispersed Poisson distribution). By large triangles we mean that $t \to \infty$ and hence also $n \to \infty$. For each *fixed* ij, $\hat{m}_{ij}$ is a consistent estimate of m_{ij} as n grows, and $q/n \to 0$. Hence, for large n, the residuals can be written as

$$r_{ij} = \frac{C_{ij} - m_{ij}}{\sqrt{m_{ij}^p}}.$$

Since the moment generating function of a $\Gamma(\alpha, \beta)$ distribution is $M(t) = (1 - \beta\, t)^{-\alpha}$ and $p = 2$ is equivalent to $C_{ij} \in \Gamma(\frac{1}{\phi}, \phi\, m_{ij})$, the residuals r_{ij} are identically distributed according to

$$M_{r_{ij}}(t) = e^{-t}\, M_{C_{ij}}(\frac{t}{m_{ij}}) = e^{-t}(1 - \phi\, t)^{-\frac{1}{\phi}}.$$

The moment generating function of a $Po(\lambda)$ distribution is $M(t) = e^{\lambda(e^t - 1)}$, but since $p = 1$ implies an over-dispersed Poisson distribution we need a help variable X_{ij} in order to find the distribution of the residuals. The underlying model is fulfilled if $C_{ij} = \phi\, X_{ij}$, $X_{ij} \in Po(\frac{m_{ij}}{\phi})$ and the residuals are distributed according to

$$M_{r_{ij}}(t) = e^{-t\sqrt{m_{ij}}}\, M_{C_{ij}}\,(\frac{t}{\sqrt{m_{ij}}}) = e^{-t\sqrt{m_{ij}}}\, M_{X_{ij}}\,(\frac{\phi\, t}{\sqrt{m_{ij}}}) = e^{-t\sqrt{m_{ij}}}\, e^{\frac{m_{ij}}{\phi}(e^{\frac{\phi\, t}{\sqrt{m_{ij}}}} - 1)}.$$

The distributions of the residuals r_{ij} depend on m_{ij} and consequently the residuals cannot be identically distributed.

Appendix 2

In order to find the variability of the claims reserve obtained by the chain-ladder method England & Verrall (1999) assume the model structure in (3.1) and on the basis of the standard error of prediction of a single future value C_{ij} in ΔC, i.e.

$$SEP(C_{ij}) = \sqrt{\widehat{E}(C_{ij} - \hat{C}_{ij})^2} \cong \sqrt{\widehat{Var}(C_{ij}) + \widehat{Var}(\hat{C}_{ij})}, \tag{5.1}$$

an expression for the standard error of prediction of the total claims reserve is derived as

$$\begin{aligned} SEP(R) &= \sqrt{\widehat{Var}(R - \hat{R})} \approx \sqrt{\widehat{Var}(R) + \widehat{Var}(\hat{R})} \qquad (5.2) \\ &= \sqrt{\widehat{Var}(\sum_{\Delta} C_{ij}) + \widehat{Var}(\sum_{\Delta} \hat{C}_{ij})} \\ &\approx \sqrt{\sum_{\Delta} \hat{\phi}\, \hat{m}_{ij}^{p} + \sum_{\Delta} \hat{m}_{ij}\, \widehat{Var}(\hat{\eta}_{ij}) + 2 \sum_{\Delta,\, i_1 j_1 \neq i_2 j_2} \hat{m}_{i_1 j_1}\, \hat{m}_{i_2 j_2}\, \widehat{Cov}(\hat{\eta}_{i_1 j_1}\, \hat{\eta}_{i_2 j_2})}\,, \end{aligned}$$

where $\hat{\eta}_{ij}$ is the estimate of η_{ij} appearing in (3.1). The first term provides for the variance of the process error and can easily be estimated analytically, while the two last terms, providing for the variance of the estimation error, can be obtained by bootstrapping. When $p = 1$, England & Verrall (1999) replace equation (5.2) by the bootstrap standard error of prediction

$$SEP_{bs}(R) = \sqrt{\hat{\phi}\, \hat{R} + (SE_{bs}(\hat{R}^*))^2}, \tag{5.3}$$

where $SE_{bs}(\hat{R}^*)$ is the standard error of the B simulated values of $\hat{R}^*$ obtained by the non-parametric standardized bootstrap procedure in Substage 2.1 in Figure 1. However, England & Verrall (1999) substitute the maximum likelihood estimates of the model parameters in Figure 1 by the chain-ladder method.

In order to obtain a complete predictive distribution England (2002) extended the method in England & Verrall (1999) by replacing the analytic calculation of the process error by another simulation conditional on the bootstrap simulation. The process error is included to the B triangles $\Delta\hat{m}^*$ by sampling a random observation from a process distribution with mean $\hat{m}_{ij}^*$ and variance $\phi\, \hat{m}_{ij}^*$ to obtain the future claims $\Delta m^{\dagger}$. The predictive distribution of the outstanding claims is then obtained by plotting the B values of $\tilde{R}^{\dagger} = \sum_{\Delta} m_{ij}^{\dagger}$ and finally the standard deviation of the simulations gives the standard error of prediction of the outstanding claims.

England (2002) presents no justification of this procedure, but sampling from over-dispersed Poisson distributions with mean $\hat{m}^*_{ij}$ and variance $\phi\,\hat{m}^*_{ij}$ will indeed provide us with a predictive distribution of R consistent with (5.3). Since

$$E(R^\dagger|\,\Delta\hat{m}^*) = \sum_\Delta E(m^\dagger_{ij}|\,\Delta\hat{m}^*) = \sum_\Delta \hat{m}^*_{ij} = \hat{R}^*$$

and

$$\mathrm{Var}(R^\dagger|\,\Delta\hat{m}^*) = \sum_\Delta Var(m^\dagger_{ij}|\,\Delta\hat{m}^*) = \sum_\Delta \hat{\phi}\,\hat{m}^*_{ij} = \hat{\phi}\,\hat{R}^*$$

the variance of the simulated predictive distribution is

$$\begin{aligned}\mathrm{Var}(R^\dagger) &= E[\mathrm{Var}(R^\dagger|\,\Delta\hat{m}^*)] + \mathrm{Var}[E(R^\dagger|\,\Delta\hat{m}^*)] \\ &= E(\hat{\phi}\,\hat{R}^*) + \mathrm{Var}(\hat{R}^*) = \hat{\phi}\,E(\hat{R}^*) + \mathrm{Var}(\hat{R}^*) \approx \hat{\phi}\,\hat{R} + \mathrm{Var}(\hat{R}^*),\end{aligned}$$

where, in the last step, we used $E(\hat{R}^*) \approx \hat{R}$ and (3.12).