

Using a Bayesian Approach for Claims Reserving

by Mario V. Wüthrich

ABSTRACT

This paper applies the exponential dispersion family with its associate conjugates to the claims reserving problem. This leads to a formula for the claims reserves that is equivalent to applying credibility weights to the chain-ladder reserves and Bornhuetter-Ferguson reserves.

KEYWORDS

Claims reserves, chain-ladder method, Bornhuetter-Ferguson method, Benktander-Hovinen method, Bayesian inference, exponential dispersion model, credibility theory

1. Introduction

For pricing and tariffication of insurance contracts Bayesian ideas and techniques are well investigated and widely used in practice. For the claims reserving problem Bayesian methods are less used, although we believe that they are very useful for answering practical questions (this has already been mentioned in de Alba [6] and other sources).

In the literature, exact Bayesian models have been studied in a series of papers by Verrall [20, 21, 22], de Alba [6, 7], de Alba and Ramírez Corzo [8], Haastrup and Arjas [11], Ntzoufras and Dellaportas [18], and Scollnik [19]. Many of these results refer to explicit choices of distributions—for example, the Poisson-gamma or the (log)normal-normal cases.

The purpose of this paper is twofold:

1. It is well known in Bayesian theory that (among others) the Poisson-gamma or the normal-normal cases are specific examples of the exponential dispersion family with its associate conjugates. We show that the claims reserving problem can easily be extended to this more general family of distributions. Not surprisingly, we obtain the same results as presented in Verrall [22] and England and Verrall [10], Section 6.3, but now in our more general setup of distributions.
2. We show that for the exponential dispersion family with its associate conjugates we obtain a natural combination of two different claims reserving methods, namely the chain ladder method (see Mack [14]) and Bornhuetter and Ferguson [3] (a more detailed discussion follows below). In the special case of Poisson-gamma, this has already been discovered by England and Verrall [10], Section 6.3.

In Section 2 we define the claims reserving problem. Moreover, we introduce the exponential dispersion model with its associate conjugates and state the main results. In Section 3 we give

the conclusions comparing our Bayesian model to the classical claims reserving methods, and in Subsection 3.4 we give the link to the Bühlmann-Straub credibility model. Finally, in Section 4 we implement the theory in a practical example.

2. Exponential dispersion model with its associate conjugates applied to claims reserving

2.1. The claims reserving problem

We denote by $X_{i,j}$ incremental data. The index $i \in \{0, \dots, I\}$ denotes the accident year and $j \in \{0, \dots, J\}$ the development period ($J \leq I$). For example, $X_{i,j}$ can denote the number of claims reported in reporting period j for accident year i or it can also denote the incremental payments (i.e., claim amounts paid in development period j for accident year i). Cumulative data are denoted by

$$C_{i,j} = \sum_{k=0}^j X_{i,k}. \quad (2.1)$$

The observations up to time I are denoted by $\mathcal{D}_I = \{X_{i,j}; i + j \leq I, j \leq J\}$.

Task. Estimate $X_{i,j}$ for $i + j > I$, given the observations \mathcal{D}_I .

Terminology. We assume that $X_{i,j}$ denote incremental payments and that $C_{i,j}$ denote cumulative payments. This simplifies our language. The reader may always use a different interpretation for $X_{i,j}$.

2.2. Exponential dispersion family

In order to predict the future random variables $X_{i,j}$, $i + j > I$, one introduces stochastic models. In the present work we consider the exponential dispersion family with its associate conjugates. The exponential dispersion family is well known in generalized linear models (see for example, McCullagh and Nelder [16]), and also in its applications to the claims reserving context (see for example, England and Verrall [10] and

the references therein). On the other hand, it is also a very important family of distributions for Bayesian theory.

We formulate the exponential dispersion family with its associate conjugates directly in the framework, as we will use it for the claims reserving problem. Weights could be chosen in a more general manner; however, we choose the ones favored by Mack (see discussion in [13], Section 2).

Model Assumption 2.1. Assume we have a claims development pattern β_0, \dots, β_J with $\beta_0 > 0$, $\beta_J = 1$ and $\beta_j > \beta_{j-1}$ ($j \geq 1$). We define $\gamma_0 = \beta_0$ and $\gamma_j = \beta_j - \beta_{j-1}$ for $j \geq 1$.

(A1) Conditionally, given Θ_i , we have that $X_{i,j}$ are independent with

$$\begin{aligned} \frac{X_{i,j}}{\gamma_j \cdot \mu_0^{(i)}} &\stackrel{(d)}{\sim} dF_{i,j}^{(\Theta_i)}(x) \\ &= a \left(x, \frac{\sigma^2}{\gamma_j \cdot (\mu_0^{(i)})^2} \right) \exp \left\{ \frac{x \cdot \Theta_i - b(\Theta_i)}{\sigma^2 \cdot \gamma_j^{-1} \cdot (\mu_0^{(i)})^{-2}} \right\} d\nu(x), \end{aligned} \tag{2.2}$$

where ν is a suitable σ -finite measure on \mathbb{R} , $b \in C^2$, $\mu_0^{(i)} > 0$ and $F_{i,j}^{(\Theta_i)}$ is a probability distribution on \mathbb{R} .

(A2) The random vectors $(\Theta_i, (X_{i,0}, \dots, X_{i,J}))$ are independent and Θ_i are independent and identically distributed real-valued random variables with density

$$u_{\mu, \tau^2}(\theta) = d(\mu, \tau^2) \cdot \exp \left\{ \frac{\mu \cdot \theta - b(\theta)}{\tau^2} \right\} \tag{2.3}$$

with $\mu = 1$ and $\tau^2 > 0$. □

REMARKS

1. $\mu_0^{(i)}$ plays the role of the a priori expected total claim amount $E[C_{i,J}]$ for accident year i . γ_j denotes the proportion paid in development period j . Hence in Assumption (A1) we compare the payment $X_{i,j}$ to its expected value $\gamma_j \cdot \mu_0^{(i)}$ (see also Lemma 2.3 below).
2. Assumption (A1) means that the scaled sizes $Z_{i,j} = X_{i,j} / (\gamma_j \cdot \mu_0^{(i)})$ belong to the exponential

dispersion family with unknown parameter Θ_i (the parameters γ_j and $\mu_0^{(i)}$ are assumed to be known). Θ_i is a (latent) random variable [see Assumption (A2)] that describes the risk characteristics of accident year i .

3. For the moment we assume that $\mu_0^{(i)}$, γ_j , σ and τ are known. In practice, of course, this is not the case. We discuss the consequences of this fact below.
4. Assumption (A2) means that different accident years can be studied independently. Different accident years i are combined through the fact that the claims development pattern γ_j and the variance parameters σ^2 and τ^2 do not depend on i . Moreover, it is assumed that (before we have any observations $X_{i,j}$) a priori the accident years are similar. This is reflected by the fact that we choose $\mu \equiv 1$ (for the meaning of μ we also refer to Lemma 2.2).
5. A special case is obtained by choosing $F_{i,j}^{(\Theta_i)}$ as a Poisson distribution with parameter Θ_i and u_{1, τ^2} as a gamma distribution. This immediately gives the model studied by Verrall [21, 22]. Other examples are (see for example [5], Section 2.5) the binomial-beta case, gamma-gamma case, or normal-normal case.
6. Observe that $Z_{i,j}$ must be positive. This may cause problems in practical applications, since in general $Z_{i,j}$ may have both signs (see also de Alba and Ramírez Corzo [8]).

The following two lemmas are two key statements in Bayesian theory. We omit their proofs since they are fairly standard and can be found in Bernardo and Smith [2] or Bühlmann and Gisler [5] (Theorems 2.19–2.20), among other texts.

LEMMA 2.2 *The conditional distribution of Θ_i given the observations \mathcal{D}_I has density $u_{\mu_{\text{post}(i)}, \tau_{\text{post}(i)}^2}(\cdot)$ with*

$$\tau_{\text{post}(i)}^2 = \sigma^2 \cdot \left[\frac{\sigma^2}{\tau^2} + (\mu_0^{(i)})^2 \cdot \beta_{(I-i) \wedge J} \right]^{-1}, \tag{2.4}$$

$$\mu_{\text{post}(i)} = \frac{\tau_{\text{post}(i)}^2}{\sigma^2} \cdot \left[\frac{\sigma^2}{\tau^2} + (\mu_0^{(i)})^2 \cdot \beta_{(I-i) \wedge J} \cdot \bar{Z}_i \right], \quad (2.5)$$

$$\bar{Z}_i = \frac{C_{i,(I-i) \wedge J}}{\beta_{(I-i) \wedge J} \cdot \mu_0^{(i)}}, \quad (2.6)$$

where $(I - i) \wedge J$ denotes the minimum of $(I - i)$ and J .

REMARKS

1. The conditional distribution of the risk characteristics Θ_i given the observations \mathcal{D}_I (the posterior distribution of the latent variable Θ_i) belongs to the same family of distributions as the a priori distribution of Θ_i (before we have any observations). Thus, this meets the definition of conjugate priors.
2. The a posteriori distribution of Θ_i depends only on the observations of accident year i (due to Assumption (A2)).
3. We have assumed that the scaled observations $Z_{i,j} = X_{i,j}/(\gamma_j \cdot \mu_0^{(i)})$ have (a priori) identical distributions. However, the a posteriori distributions of $Z_{i,j}$, $i + j > I$, given \mathcal{D}_I , are different, which is reflected by $\mu_{\text{post}(i)}$ and $\tau_{\text{post}(i)}^2$.
4. Lemma 2.2 allows for an explicit calculation of the a posteriori (predictive) distributions of $(X_{i,I-i+1}, \dots, X_{i,J})$, given the observations \mathcal{D}_I (which are independent for $i = 0, 1, \dots$), namely

$$P[X_{i,I-i+1} \leq x_{I-i+1}, \dots, X_{i,J} \leq x_J \mid \mathcal{D}_I] = \int \prod_{j=I-i+1}^J F_{i,j}^{(\theta)} \left(\frac{x_j}{\gamma_j \cdot \mu_i^{(0)}} \right) \cdot u_{\mu_{\text{post}(i)}, \tau_{\text{post}(i)}^2}(\theta) d\theta. \quad (2.7)$$

Henceforth, with (2.7) we can explicitly calculate the a posteriori distributions and their moments. Moreover, this allows for simulations of the random variables. The next lemma then provides a straightforward estimate for the expected total claim amounts.

LEMMA 2.3 Under the Model Assumptions 2.1 we have

$$\mu(\Theta_i) \stackrel{\text{def.}}{=} E \left[\frac{X_{i,j}}{\gamma_j \cdot \mu_0^{(i)}} \mid \Theta_i \right] = b'(\Theta_i). \quad (2.8)$$

If $\exp\{(\mu\theta - b(\theta))/\tau^2\}$ disappears on the boundary of Θ_i for all μ, τ^2 then

$$E[X_{i,j}] = \gamma_j \cdot \mu_0^{(i)} \cdot E[\mu(\Theta_i)] = \gamma_j \cdot \mu_0^{(i)}, \quad (2.9)$$

$$\begin{aligned} \widetilde{\mu}(\Theta_i) &\stackrel{\text{def.}}{=} E[\mu(\Theta_i) \mid \mathcal{D}_I] \\ &= \alpha_i^{((I-i) \wedge J)} \cdot \bar{Z}_i + (1 - \alpha_i^{((I-i) \wedge J)}) \cdot 1, \end{aligned} \quad (2.10)$$

$$\text{where } \alpha_i^{(j)} = \frac{\beta_j}{\beta_j + \kappa_i} \text{ and } \kappa_i = \frac{\sigma^2}{(\mu_0^{(i)})^2 \cdot \tau^2}. \quad (2.11)$$

REMARKS $\widetilde{\mu}(\Theta_i)$ is a Bayesian estimator (a posteriori mean of $\mu(\Theta_i)$, given the observations \mathcal{D}_I). It is a credibility-weighted average between the a priori mean $\mu = 1$ and the observations \bar{Z}_i (defined in (2.6)). The larger the individual variation σ^2 the smaller the credibility weight; the larger the collective variability τ^2 the larger the credibility weight (for a detailed discussion on the credibility coefficient κ_i we refer to Bühlmann and Gisler [5]).

LEMMA 2.4 (Bayesian estimator for claims reserves) Choose $j = I - i < J$ and $k \in \{1, \dots, J - j\}$. Then the Bayesian estimators for $E[X_{i,j+k} \mid C_{i,0}, \dots, C_{i,j}]$ and $E[C_{i,j+k} \mid C_{i,0}, \dots, C_{i,j}]$ in Model 2.1 are as follows

$$\begin{aligned} \widetilde{X}_{i,j+k} &= \hat{E}[X_{i,j+k} \mid C_{i,0}, \dots, C_{i,j}] \\ &= \gamma_{j+k} \cdot \mu_0^{(i)} \cdot \widetilde{\mu}(\Theta_i), \end{aligned} \quad (2.12)$$

$$\begin{aligned} \widetilde{C}_{i,j+k} &= \hat{E}[C_{i,j+k} \mid C_{i,0}, \dots, C_{i,j}] \\ &= C_{i,j} + (\beta_{j+k} - \beta_j) \cdot \mu_0^{(i)} \cdot \widetilde{\mu}(\Theta_i). \end{aligned} \quad (2.13)$$

REMARK The estimators $\widetilde{\mu}(\Theta_i)$, $\widetilde{X}_{i,j+k}$ and $\widetilde{C}_{i,j}$ are unbiased, \mathcal{D}_I -measurable and minimize the quadratic loss function (see Theorem 2.5 in [5]).

Consequence. We obtain for $I - i < J$ (see Lemma 2.3)

$$\begin{aligned}
 E[C_{i,J} | \mathcal{D}_I] &= C_{i,J-i} + \sum_{j=I-i+1}^J E[X_{i,j} | \mathcal{D}_I] \\
 &= C_{i,J-i} + \sum_{j=I-i+1}^J \gamma_j \cdot \mu_0^{(i)} \cdot E[\mu(\Theta_i) | \mathcal{D}_I] \\
 &= C_{i,J-i} + (1 - \beta_{I-i}) \cdot \mu_0^{(i)} \cdot \widetilde{\mu}(\Theta_i) \\
 &= \widetilde{C}_{i,J} \\
 &= C_{i,J-i} + (1 - \beta_{I-i}) \\
 &\quad \cdot \left[\alpha_i^{(I-i)} \cdot \frac{C_{i,J-i}}{\beta_{I-i}} + (1 - \alpha_i^{(I-i)}) \cdot \mu_0^{(i)} \right].
 \end{aligned} \tag{2.14}$$

3. Interpretation and conclusions

In the exponential dispersion family with associate conjugates (Model Assumptions 2.1) the Bayesian estimator for the expected ultimate claim $E[C_{i,J} | \mathcal{D}_I]$ at time I is given by (2.14).

Before giving an interpretation of that formula we briefly review the two (probably) most popular methods, namely the chain-ladder (CL) method (see Mack [14]) and the Bornhuetter-Ferguson (BF) method (see Bornhuetter and Ferguson [3]).

3.1. CL method

The CL method is based on the assumption that there exist development factors f_0, \dots, f_J ($f_J = 1$) such that for all $i \in \{0, \dots, I\}$ and $j \in \{1, \dots, J\}$

$$E[C_{i,j} | C_{i,0}, \dots, C_{i,j-1}] = f_{j-1} \cdot C_{i,j-1}. \tag{3.1}$$

The CL estimator of the ultimate claim $C_{i,J}$, given the observations $C_{i,0}, \dots, C_{i,j}$, is then given by ($j < J$)

$$\widehat{C}_{i,J}^{\text{CL}} = \widehat{E}[C_{i,J} | C_{i,0}, \dots, C_{i,j}] = C_{i,j} \cdot f_j \cdots f_{J-1}. \tag{3.2}$$

Define $\beta_j = \prod_{k=j}^J f_k^{-1}$. Estimate (3.2) implies

$$\widehat{C}_{i,J}^{\text{CL}} = C_{i,j} + (1 - \beta_j) \cdot \widehat{C}_{i,J}^{\text{CL}}. \tag{3.3}$$

3.2. BF method

The BF method estimates the ultimate claim by (see Mack [13])

$$\widehat{C}_{i,J}^{\text{BF}} = C_{i,j} + (1 - \beta_j) \cdot \mu_0^{(i)}, \tag{3.4}$$

where $\mu_0^{(i)}$ is an a priori estimate ignoring the data \mathcal{D}_I .

3.3. Combination of CL and BF method

We have now two extreme positions: The BF method only relies on the a priori estimate $\mu_0^{(i)}$ (ignoring the observations), whereas the CL method gives full credibility to the indication based solely on the observation $C_{i,j}$.

Benktander [1] and Hovinen [12] have made a first attempt to combine these two extreme cases. Choose $\alpha \in [0, 1]$ and define

$$\mu_0^{(i)}(\alpha) = \alpha \cdot \widehat{C}_{i,J}^{\text{CL}} + (1 - \alpha) \cdot \mu_0^{(i)}. \tag{3.5}$$

Benktander-Hovinen (BH) have chosen $\alpha = \beta_j$, which gives the BH estimate

$$\widehat{C}_{i,J}^{\text{BH}} = C_{i,j} + (1 - \beta_j) \cdot [\beta_j \cdot \widehat{C}_{i,J}^{\text{CL}} + (1 - \beta_j) \cdot \mu_0^{(i)}]. \tag{3.6}$$

Question. What is the optimal α ? Optimality is defined here as “minimizing mean square error” (see Mack [15], Section 3).

Mack [15] gives a different stochastic model (see Mack [15], (2)–(3)) under which he calculates the optimal α (see Mack [15], Theorems 2 and 3). It is of the form

$$\alpha^* = \frac{\beta_j}{\beta_j + \kappa}. \tag{3.7}$$

Henceforth, the estimator in the model considered by Mack [15] has exactly the same form as the Bayesian estimator (2.14) in our exponential dispersion model. Observe that for $I - i < J$ we have (using (2.14))

$$\begin{aligned}
 \widetilde{C}_{i,J} &= E[C_{i,J} | \mathcal{D}_I] \\
 &= C_{i,J-i} + (1 - \beta_{I-i}) \cdot [\alpha_i^{(I-i)} \cdot \widehat{C}_{i,J}^{\text{CL}} + (1 - \alpha_i^{(I-i)}) \cdot \mu_0^{(i)}].
 \end{aligned} \tag{3.8}$$

Hence we obtain in a natural way a “linear mixture” of the CL estimate and the BF estimate. It has two extreme cases:

a) Choose $\kappa_i=0$. This leads to $\alpha_i^{(I-i)}=1$ which is the CL estimate.

b) $\kappa_i = \infty$ leads to $\alpha_i^{(I-i)} = 0$ which is the BF estimate.

3.4. Linear credibility methods

Under our Model Assumptions 2.1 we can explicitly calculate the a posteriori distribution of loss for a given accident year. Moreover the a posteriori expectation of $\mu(\Theta_i)$ is linear in the observations. In general this is not the case, and one cannot explicitly calculate the a posteriori distribution. In such situations one uses a linear credibility approach, which minimizes quadratic loss functions among linear estimators.

Probably the most famous model in linear credibility theory is the Bühlmann-Straub (BS) model (see Bühlmann and Gisler [5], Chapter 4). The BS model has been used in the claims reserving context by Mack [13], Neuhaus [17] (see Section 3.4) and de Vylder [9].

In the BS model one obtains exactly the same estimate for the reserves as in our exponential dispersion model, i.e., the credibility estimator for $\mu(\Theta_i)$ in the BS model is given by (choosing an appropriate scaling, see also Mack [13])

$$\widehat{\mu(\Theta_i)}^{\text{cred}} = \alpha_i^{((I-i)\wedge J)} \cdot \bar{Z}_i + (1 - \alpha_i^{((I-i)\wedge J)}). \quad (3.9)$$

However, the credibility estimator is only the best linear approximation to $\mu(\Theta_i)$, and hence has a larger quadratic loss compared to the Bayes estimate. Moreover, it does not satisfy (2.14) and hence (3.8) (this is exactly the Bayes estimate), and it does not allow for simulation, because only the first two moments are determined by the BS model.

However, for the exponential dispersion family with associate conjugates the Bayes estimate and the credibility estimate coincide.

4. Application in practice and an example

So far we have always assumed that the following parameters are known:

1. the a priori mean $\mu_0^{(i)}$;
2. the claims development pattern β_j ;
3. the credibility coefficient κ_i , and the variance parameters σ^2 and τ^2 .

Then Lemmas 2.2 and 2.3 give the a posteriori distributions and the optimal estimators (this is a similar situation as considered in England and Verrall [10], Section 6.3). However, in practice these are often not known and need to be estimated from the data. If we replace the parameters by their estimates, then of course we lose the optimality conditions (since we have an additional error term coming from the parameter estimation). Hence we could now build a whole new theory also trying to minimize the parameter estimation error. Since this would go beyond the scope of this paper we restrict ourselves to the replacement of the parameters by appropriate estimators.

In other words this means that in practice a full analytical Bayesian formula is often not a realistic method. One way out of this dilemma is the credibility technique. Here the credibility solution is understood in replacing the unknown parameters by appropriate estimators. In Verrall [22] such estimators are called “plug-in” estimates. In a full Bayesian approach one would estimate both the exposures $\mu(\Theta_i)$ and the claims development pattern γ_j simultaneously. Such a full Bayesian approach often requires numerical methods such as the Markov Chain Monte Carlo method (see Verrall [22], de Alba [6, 7], Ntzoufras and Dellaportas [18], or Scollnik [19]).

A priori mean. As a priori mean $\mu_0^{(i)}$, one usually takes a plan value or the estimate from the premium calculation (as in the BF method). In our example below, the a priori mean is known (from budget loss ratios).

Claims development pattern. The estimation of β_j is the crucial part in which we link the different accident years. In Assumption (A2) we have assumed that the different accident years are independent, and therefore in Lemmas 2.2 and 2.3, one could not learn anything about accident year i from accident year $i' \neq i$ and vice versa.

Since we have assumed that all accident years have the same claims development pattern γ_j , we now combine the observations of the different accident years to estimate γ_j .

There is no canonical way in our model to get β_j from the data (as there is in the CL method). In practice (and in our example below) one usually uses the CL estimate for the development factors f_j : Given \mathcal{D}_I , we estimate f_j and β_j as follows (see Mack [14])

$$\hat{f}_{j-1}^{\text{CL}} = \frac{\sum_{i=0}^{I-j} C_{i,j}}{\sum_{i=0}^{I-j} C_{i,j-1}} \quad \text{and} \quad \hat{\beta}_j^{\text{CL}} = \prod_{k=j}^{J-1} 1/\hat{f}_k^{\text{CL}}. \tag{4.1}$$

It is well known that these estimates lead to an unbiased estimate in the CL model and one can estimate the mean square error of prediction for this model (see Mack [14] and Buchwalder et al. [4]). However, in our model (as in the BF model) we can neither show that $\hat{\beta}_j^{\text{CL}}$ is an appropriate estimate for the claims development pattern nor are we able to calculate the mean square error of prediction. One can easily give an estimate for the process variance (with (2.7)), but one cannot give an estimate for the estimation error since we do not even know whether β_j is estimated in an appropriate way.

Credibility coefficient. The credibility coefficient $\kappa_i = \sigma^2 \cdot (\mu_0^{(i)})^{-2} / \tau^2$ is calculated by estimating σ^2 and τ^2 ($\mu_0^{(i)}$ was already given above).

Observe that (see Theorem 2.20 in [5])

$$\text{Var} \left(\frac{X_{i,j}}{\gamma_j \cdot \mu_0^{(i)}} \middle| \Theta_i \right) = \frac{\sigma^2 \cdot b''(\Theta_i)}{(\mu_0^{(i)})^2 \cdot \gamma_j}. \tag{4.2}$$

Without loss of generality we may assume that

$$m_b \stackrel{\text{def.}}{=} E[b''(\Theta_i)] = 1. \tag{4.3}$$

Otherwise we simply multiply σ^2 and τ^2 by m_b , which in our context of an exponential dispersion family with associate conjugates leads to the same model with $b(\theta)$ replaced by $b_{(1)}(\theta) = m_b \cdot b(\theta/m_b)$. This rescaled model has then

$$\text{Var} \left(\frac{X_{i,j}}{\gamma_j \cdot \mu_0^{(i)}} \middle| \Theta_i \right) = \frac{m_b \cdot \sigma^2 \cdot b''_{(1)}(\Theta_i)}{(\mu_0^{(i)})^2 \cdot \gamma_j},$$

$$\text{with } E[b''_{(1)}(\Theta_i)] = 1, \tag{4.4}$$

$$\text{Var}(b'_{(1)}(\Theta_1)) = m_b \cdot \tau^2. \tag{4.5}$$

The credibility weights $\alpha_i^{(j)}$ do not change under this transformation since both σ^2 and τ^2 are multiplied by m_b . Hence we assume (4.3) for the rest of this work. It then follows that

$$\widehat{\sigma}^2 = \frac{1}{I} \sum_{i=0}^{I-1} \frac{1}{(I-i) \wedge J} \sum_{j=0}^{(I-i) \wedge J} (\mu_0^{(i)})^2 \cdot \gamma_j \cdot \left(\frac{X_{i,j}}{\gamma_j \cdot \mu_0^{(i)}} - \bar{Z}_i \right)^2 \tag{4.6}$$

is an unbiased estimator for σ^2 .

Define $w_i = \beta_{(I-i) \wedge J} \cdot (\mu_0^{(i)})^2$, $w_\bullet = \sum_{i=0}^{I-1} w_i$ and

$$c = \frac{I-1}{I} \left[\sum_{i=0}^{I-1} \frac{w_i}{w_\bullet} \cdot \left(1 - \frac{w_i}{w_\bullet} \right) \right]^{-1}, \tag{4.7}$$

$$\bar{Z} = \sum_{i=0}^{I-1} \frac{w_i}{w_\bullet} \cdot \bar{Z}_i, \tag{4.8}$$

$$T = \frac{I}{I-1} \sum_{i=0}^{I-1} \frac{w_i}{w_\bullet} \cdot (\bar{Z}_i - \bar{Z})^2. \tag{4.9}$$

Then

$$\widetilde{\tau}^2 = c \cdot \left(T - \frac{I \cdot \widehat{\sigma}^2}{w_\bullet} \right) \tag{4.10}$$

Table 1. Data and chain ladder parameter estimates

	Development Periods j										
	0	1	2	3	4	5	6	7	8	9	$\mu_0^{(i)}$
0	178,409	111,637	26,872	6,233	6,201	1,864	1,974	445	334	474	349,593
1	190,403	97,392	21,697	4,554	2,035	1,098	1,583	336	349		341,019
2	188,073	89,287	25,412	7,883	4,581	1,963	1,606	268			328,889
3	175,890	80,497	21,676	5,720	3,989	2,650	1,300				318,503
4	173,367	82,357	19,617	8,202	6,909	3,157					331,346
5	185,544	84,850	17,183	7,347	3,149						344,421
6	168,006	86,796	16,893	6,766							342,407
7	158,642	73,203	15,841								333,796
8	158,724	70,738									329,596
9	170,267										348,553
\hat{f}_j^{CL}	1.4925	1.0778	1.0229	1.0148	1.0070	1.0051	1.0011	1.0010	1.0014	1.0000	
$\hat{\beta}_j^{\text{CL}}$	59.0%	88.0%	94.8%	97.0%	98.4%	99.1%	99.6%	99.8%	99.9%	100.0%	

is an unbiased estimator for τ^2 (see [5] (4.26)). Since $\widetilde{\tau^2}$ could be negative, we set

$$\widehat{\tau^2} = \max\{\widetilde{\tau^2}, 0\} \quad \text{and} \quad \widehat{\kappa}_i = \widehat{\sigma^2} \cdot (\mu_0^{(i)})^{-2} / \widehat{\tau^2}. \tag{4.11}$$

REMARK One may view as a major deficiency of the present model that we lose the optimalities when replacing the unknown parameters by their estimates. On the other hand the following formula shows that it can be very useful in practice: define $\alpha_{i,*}^{(j)} = \widehat{\beta}_j^{\text{CL}} / (\widehat{\beta}_j^{\text{CL}} + \widehat{\kappa}_i)$. For $I - i < J$, equation (2.14) leads to the following estimate of the ultimate claim payments:

$$\begin{aligned} \widetilde{C}_{i,J}^* &= C_{i,J-i} + (1 - \widehat{\beta}_{I-i}^{\text{CL}}) \cdot \left(\alpha_{i,*}^{(I-i)} \cdot \frac{C_{i,J-i}}{\widehat{\beta}_{I-i}^{\text{CL}}} + (1 - \alpha_{i,*}^{(I-i)}) \cdot \mu_0^{(i)} \right) \\ &= \alpha_{i,*}^{(I-i)} \cdot \widehat{C}_{i,J}^{\text{CL}} + (1 - \alpha_{i,*}^{(I-i)}) \cdot \widehat{C}_{i,J}^{\text{BF}}. \end{aligned} \tag{4.12}$$

In the last step above we have assumed that both f_j and β_j are estimated by (4.1) (this is the usual choice done in practice for the CL and the BF methods).

REMARKS

- Our estimate $\widetilde{C}_{i,J}^*$ for the ultimate claim payments is a credibility weighted average of the CL estimate and the BF estimate. The credibility weight is determined by the development pattern β_j , the a priori estimate $\mu_0^{(i)}$ and the

variances σ^2 and τ^2 of the processes. Since it is increasing in β_j we give higher credibility to the CL estimate for older accident years.

- Combining the CL estimate and the BF estimate is a very old problem in claims reserving. In some insurance companies there are rules of thumb for when to choose which estimate (see also Mack [15]). Equation (4.12) gives a natural way to combine the CL and the BF estimates.

4.1. Example

The observed incremental payments $X_{i,j}$, $i, j \in \{0, \dots, 9\}$, are given in Table 1.

This is a rather homogeneous data set, with fast development. After two years, almost 90% of the total claim amount is paid. On the other hand it also looks long-tailed since we still observe some payments after seven years (in the present work we do not bother about choosing tail factors for the CL method).

Using our parameter estimates from above we obtain the following estimates for σ^2 and τ^2

$$\widehat{\sigma^2} = (10,119)^2 \quad \text{and} \quad \widehat{\tau^2} = (60)^2.$$

The credibility coefficients, the credibility weights and the estimates for the ultimate claim payments are now determined with the help of (4.12). We obtain the results given in Table 2.

Table 2. Resulting reserves

i	$\hat{\beta}_{1-i}^{CL}$	$\hat{\kappa}_i$	$\alpha_{i,*}^{(1-i)}$	Estimated Ultimate Claim Payments			Estimated Claim Reserves		
				$\widehat{C}_{i,J}^{CL}$	$\widehat{C}_{i,J}^{BF}$	$\widetilde{C}_{i,J}^*$	CL	BF	$[\widetilde{C}_{i,J}^* - C_{i,J-i}]$
0	100.0%	23.3%	81.1%	334,444	334,444	334,444	0	0	0
1	99.9%	24.5%	80.3%	319,900	319,929	319,905	454	484	460
2	99.8%	26.3%	79.1%	319,860	319,882	319,865	788	810	792
3	99.6%	28.1%	78.0%	292,758	292,849	292,778	1,036	1,127	1,056
4	99.1%	26.0%	79.3%	296,167	296,471	296,230	2,559	2,863	2,622
5	98.4%	24.0%	80.4%	302,767	303,413	302,894	4,695	5,341	4,821
6	97.0%	24.3%	80.0%	287,044	288,700	287,376	8,584	10,239	8,915
7	94.8%	25.6%	78.8%	261,161	264,909	261,957	13,475	17,223	14,271
8	88.0%	26.2%	77.0%	260,759	269,021	262,656	31,297	39,559	33,194
9	59.0%	23.5%	71.5%	288,791	313,319	295,772	118,524	143,052	125,504
Total							181,412	220,697	191,637

Conclusions of the example.

- We see that the estimated $\hat{\kappa}_i = (\widehat{\sigma}^2 / (\mu_0^{(i)})^2) \cdot (\widehat{\tau}^2)^{-1}$ is around 25% in our example. This means that a claims development factor $\hat{\beta}_j^{CL}$ of 25% gives already a credibility weight of 50% to the observation. Observe also that the credibility weight is always smaller than 1.
- The a priori estimate $\mu_0^{(i)}$ is rather conservative since the BF estimate is always larger than the CL estimate. Of course, this can have various reasons, which are not further discussed here. Our estimate $\widetilde{C}_{i,J}^*$ then lies between the CL and the BF estimates. Since the credibility weights $\alpha_{i,*}^{(1-i)}$ are larger than 50%, our estimate is closer to the CL estimate.

References

[1] Benktander, G., “An Approach to Credibility in Calculating IBNR for Casualty Excess Reinsurance,” *Actuarial Review*, April 1976, p. 7.

[2] Bernardo, J. M., and A. F. M. Smith, *Bayesian Theory*, New York: Wiley, 1994.

[3] Bornhuetter, R. L., and R. E. Ferguson, “The Actuary and IBNR,” *Proceedings of the Casualty Actuarial Society* 59, 1972, pp. 181–195.

[4] Buchwalder, M., H. Bühlmann, M. Merz, and M. V. Wüthrich, “The Mean Square Error of Prediction in the Chain Ladder Reserving Method (Mack and Murphy revisited),” *ASTIN Bulletin* 36, 2006, pp. 521–571.

[5] Bühlmann, H., and A. Gisler, *A Course in Credibility Theory and Its Applications*, New York: Springer, 2005.

[6] de Alba, E., “Bayesian Estimation of Outstanding Claim Reserves,” *North American Actuarial Journal* 6:4, 2002, pp. 1–20.

[7] de Alba, E., “Claims Reserving When There Are Negative Values in the Runoff Triangle: Bayesian Analysis Using the Three-Parameter Log-Normal Distribution,” *North American Actuarial Journal* 10:3, 2006, pp. 45–59.

[8] de Alba, E., and M. A. Ramírez Corzo, “Bayesian Claims Reserving When There Are Negative Values in the Runoff Triangle,” *2006.1 Proceedings of Instituto Tecnológico Autónomo de México*, Actuarial Research Clearing House, 2006, http://www.soa.org/news-and-publications/files/pdf/arch06v40n1_VII.pdf.

[9] De Vylder, F., “Estimation of IBNR Claims by Credibility Theory,” *Insurance: Mathematics and Economics* 1, 1982, pp. 35–40.

[10] England, P. D., and R. J. Verrall, “Stochastic Claims Reserving in General Insurance,” *British Actuarial Journal* 8, 2002, pp. 443–518.

[11] Haastруп, S., and E. Arjas, “Claims Reserving in Continuous Time: A Nonparametric Bayesian Approach,” *ASTIN Bulletin* 26, 1996, pp. 139–164.

[12] Hovinen, E., “Additive and Continuous IBNR,” *ASTIN Colloquium*, Loen, Norway, 1981.

[13] Mack, T., “Improved Estimation of IBNR Claims by Credibility,” *Insurance: Mathematics and Economics* 9, 1990, pp. 51–57.

[14] Mack, T., “Distribution-Free Calculation of the Standard Error of Chain Ladder Reserve Estimates,” *ASTIN Bulletin* 23, 1993, pp. 213–225.

[15] Mack, T., “Credible Claims Reserves: The Benktander Method,” *ASTIN Bulletin* 30, 2000, pp. 333–347.

- [16] McCullagh, P., and J. A. Nelder, *Generalized Linear Models* (2nd edition), London: Chapman and Hall, 1989.
- [17] Neuhaus, W., “Another Pragmatic Loss Reserving Method or Bornhuetter/Ferguson Revisited,” *Scandinavian Actuarial Journal*, 1992, pp. 151–162.
- [18] Ntzoufras, I., and P. Dellaportas, “Bayesian Modelling of Outstanding Liabilities Incorporating Claim Count Uncertainty,” *North American Actuarial Journal* 6:1, 2002, pp. 113–128.
- [19] Scollnik, D. P. M., “Implementation of Four Models for Outstanding Liabilities in WinBUGS: A Discussion of a Paper by Ntzoufras and Dellaportas,” *North American Actuarial Journal* 6:1, 2002, pp. 113–128.
- [20] Verrall, R. J., “Bayesian and Empirical Bayes Estimation for the Chain Ladder Model,” *ASTIN Bulletin* 20, 1990, pp. 217–238.
- [21] Verrall, R. J., “An Investigation into Stochastic Claims Reserving Models and Chain-Ladder Technique,” *Insurance: Mathematics and Economics* 26, 2000, pp. 91–99.
- [22] Verrall, R. J., “A Bayesian Generalized Linear Model for the Bornhuetter-Ferguson Method of Claims Reserving,” *North American Actuarial Journal* 8:3, 2004, pp. 67–89.