Bayesian Prediction of Outstanding Claims

Ioannis Ntzoufras and Petros Dellaportas

Department of Statistics, Athens University of Economics and Business,

76 Patission Street, 10434, Athens, Greece

e-mail:petros@aueb.gr

February 1998

Bayesian Prediction of Outstanding Claims

Ioannis Ntzoufras and Petros Dellaportas
Department of Statistics, Athens University of Economics and Business,
76 Patission Street, 10434, Athens, Greece

February 1998

Abstract

This paper deals with the prediction of the amount of outstanding claims that an insurance company will pay in the near future. We propose many competing models using Bayesian theory and Markov chain Monte Carlo methods. We use claim counts in order to expand usual log-normal model and state-space model. By this way, we incorporate information from both the outstanding claim amounts and counts data resulting to a new model formulation. We extend the commonly used models and propose a variety of modeling approaches which include log-normal, multinomial and state space Models. For each model, we provide implementation details and illustrations with real insurance data. In all the modelling approaches we use Bayesian methods and Markov chain Monte Carlo for the estimation of the posterior distributions under consideration.

Keywords: Log-Normal Model; Markov chain Monte Carlo; Missing values; Multinomial model; State Space Models;

1 Introduction

Insurance companies often do not pay the outstanding claims immediately for many reasons. Instead, the outstanding claims are being paid gradually over a period of some years. At a given time an insurance company has outstanding claims from years in the past and a major problem of interest, for example when evaluating the company, is to estimate the amount of claims that the company has to pay in the next years. Mathematically, the problem can be formulated as follows. There exist data with a structure given by Table 1. A_i, $i = 1, 2, \ldots, r$ denote the accident year and B_j, $j = 1, 2, ..., r$ denote the year that the claim was paid to the client. For example, the cell ij contains the amount Y_{ij} that the company paid with a delay of $j-1$ years for accidents happened in the year i. Moreover, the counts of claims for each cell are also given for certain insurance claims. The claim counts have the same triangular form as the claim amounts; in Table 2, n_{ij} represents claim counts that the company has paid with a delay of $j-1$ years for accidents happened in the year i and T_i denote the total number of accidents reported but not settled for each year. In some cases accidents are reported to the insurance period within a short time period, for example in car insurance. In these cases T_i estimates sufficiently enough the total number of accidents that the company will have to pay. Therefore, when we use T_i we are essentially modelling reported but not settled counts. Occurred but not yet reported counts are not considered. It is easy to assume that total number of accident is unknown and therefore also model uncertainty about claims occurred but not reported. For each year we also know the inflation factor for each cell (inf_{ij}) which is used to deinflate the claim amounts and make valid comparisons through different years.

A further comment is that claim counts may have different interpretation according to the sampling scheme. So Table 2 may represent the time that the claims were finalised (sub-payments are allowed) or the exact time that the whole corresponding claim amount was paid (no sub-payments allowed).

Modelling of outstanding claims of insurance companies has attracted attention of Tables 1 & 2
many researchers and many models and techniques have been investigated. Two models are the simplest and most broadly used. The first model is a simple log-linear Model and the second a log-normal anova type Model. These models were investigated by Taylor and Ashe (1983), Renshaw and Verrall (1994), Renshaw(1989) and Verrall (1991, 1993, 1996). Renshaw (1994) found accurate approximations of the square root prediction error for the log-normal Model. A short review of more advanced models for this problem is given by Haberman and Renshaw (1996). Recently, Verrall (1996) investigated generalized additive Models. Dynamic Models (also called state space Models) were frequently used; De Jong (1983) made a general investigation of state space Models in reserving claims problem, Verrall (1989, 1994) investigated a state space Model similar to the log-normal Models. From a Bayesian point of view Verrall (1990) produced Bayes estimates for the log-normal Model. Bayesian methods were also used by Haastrup and Arjas (1996) for estimating claim counts and amounts in individual claim data, Jewell (1989) and Alba *et al.* (1997) for estimating claim counts. Empirical Bayes estimates were also obtained by Verrall (1989, 1990). Further work in estimating claim amounts, claim counts or both

is published by Norberg (1986), Hesselager (1991), Neuhaus (1992) and Lawless (1994).

In this paper a Bayesian approach is used to investigate various models for the outstanding claims problem. The use of the Bayesian paradigm did not emanate from the need to use prior information as in Verall (1990) but, rather, from its computation flexibility that allows us to handle complex models. Markov chain Monte Carlo (MCMC) sampling strategies are used to generate samples for each posterior distribution of interest. Modelling of both claim counts and amount is not new; see Norberg (1986), Hesselager (1991), Neuhaus (1992), Haastrup and Arjas (1996) for application in individual claim data and Taylor and Ashe (1983) for aggregated data. A new innovation is the simultaneous use of data of Table 1 and Table 2 in the simple log-normal models. In our approach, we use a log normal model for the aggregated claims amounts and we add aggregated claim counts as an offset. In this way we are essentially modelling the average payment per claim finalised (or paid) in development year j of year of origin i; this quantity was noted as PPCF by Taylor and Ashe (1983). This has the advantage that increase in accidents, which is expressed via increase of claim counts in Table 2, will result to increase also in total claims in Table 1. We further model the claim counts using a multinomial distribution when T_i is known. In order to generalise the method for prediction of occurred but not yet reported claims we can assume T_i unknown and alternatively use Poisson distribution for claim counts.

In the next section we briefly describe the Bayesian theory and the MCMC methodology. Section 3 presents the models under consideration. In section 4 an illustrated example with data from a Greek insurance company is presented. Finally, a brief discussion is given in section 5.

2 Bayesian Modelling via MCMC

The approach in this paper is Bayesian, and the tool adopted to implement the Bayesian paradigm is the MCMC methodology. Details of the technicalities involved for MCMC can be found, for example, in Smith and Roberts (1993), Besag and Green (1993), Gilks *et al.* (1996). To avoid extended discussion on MCMC, the following section will focus only in the modelling aspects of the problem and implementation details (form of the full conditional densities, Metropolis Hastings sampling etc.) are given in the Appendix. Technicalities used include sampling from the usual log-concave densities met in generalised linear Models (Dellaportas and Smith, 1993), MCMC strategies for problems with constrained parameter problems (Gelfand *et al.*, 1992), ways to handle the missing values problem (Gelfand *et al.*, 1990), and finally, usual transformation of the MCMC output to the actual parameters of interest. For the latter, let us give some additional information. For every Model we deal with the missing entries in Table 1 ($Y_{ij}, \quad i+j > r+1$) which are treated as parameters. Having obtained the MCMC output samples, it is straightforward to obtain the posterior sample for the parameters of interest $Y_{i.} = \sum_{j=r+2-i}^{r} Y_{ij}$ which express the outstanding claims of year i, by simply adding the generated values of $Y_{ij}, \quad i+j > r+1$. Convergence aspects were treated carefully by firstly choosing an appropriate lag interval in the MCMC output so that autocorrelation bias is minimized, and secondly by using all convergence diagnostics criteria available in CODA software (see Best *et al.*, 1995) to ensure that convergence is achieved. For details about this important aspect of MCMC methodology, see the recent review papers by Cowles and Carlin (1996) and Brooks and Roberts (1997).

3 Modelling Approaches

In this section we propose new modeling aspects expressed in terms of 4 models of the problem of outstanding claims. First, we present a Bayesian analysis of two models used in the past, the log-normal (Model 1) and the state space Model (Model 3). Second, we enhance these models by simultaneously modelling claim amounts and counts and using the total claim counts to specify appropriate parameter constraints. These modifications result to Models 2 and 4.

3.1 Model 1: Log-Normal Model

The simplest model for the data in Table 1 is a log-normal (anova type) Model. This model was investigated by Renshaw and Verrall (1994), Renshaw (1989) and Verrall (1991, 1993, 1996). Verrall (1990) produced Bayes estimates for the parameters of this Model. The Model is given by the following formulation

$$Z_{ij} = log\frac{Y_{ij}}{inf_{ij}},$$

$$Z_{ij} \sim N(\mu_{ij}, \sigma^2), \quad \mu_{ij} = \mu + \alpha_i + \beta_j, \quad i,j = 1,\ldots,r \tag{1}$$

with $N(\mu_{ij}, \sigma^2)$ denoting the normal distribution with mean μ_{ij} and variance σ^2. Note that in Taylor and Ashe (1983) and Verrall (1991, 1993, 1996) the alternative parametrization $Z_{ij} = log(Y_{ij}) - log(inf_{ij} \times E_i)$ is used, with $Z_{ij} \sim N(\mu_{ij}, \sigma^2)$, where E_i is a measure of exposure (for example size of portfolio for year i). This reparametrazation can be easily adopted for all of the following models.

Model 1 requires appropriate constraints to achieve identifiability of the parameters, so here and in the sequel models we adopt the usual sum to zero parametrization, that is, $\sum_i \alpha_i = \sum_j \beta_j = 0$. To complete the Bayesian formulation we use the following priors:

$$\mu \sim N(0, \sigma^2_\mu), \quad \alpha_i \sim N(0, \sigma^2_{\alpha_i}), \quad \beta_j \sim N(0, \sigma^2_{\beta_j}), \; i, j = 2, ..., r, \quad \tau = \sigma^{-2} \sim G(a_\tau, b_\tau)$$

with $G(a, b)$ denoting gamma distribution with mean a/b. For the kind of problems we are interested in, 'non-informative' priors are produced by using

$$\sigma^2_\mu = 1000, \quad \sigma^2_{\alpha_i} = 100, \; i = 2, ..., r, \quad \sigma^2_{\beta_j} = 100, \; j = 2, ..., r, \quad a_\tau = b_\tau = 0.001.$$

Here by 'non-informative' priors we mean distributions that carry low information about the parameters of interest. These priors are represented by distributions with great prior variance which means that they are flat over a wide range values. Instead of these distributions we may use improper prior distributions that represent total ignorance. These distributions are constructed via the use of formal rules and also called reference priors; see Bernardo and Smith (1994) and Kass and Wasserman (1996).

The main disadvantage of the above Model is that it does not use any information from the observed counts or even from the total counts. That is, any prediction of the missing claim amounts will be based only on the observed claim amounts. Therefore, a source of information for a year (or cell) such as a sudden increase of accidents will not affect the prediction of the claim amounts.

3.2 Model 2: Log-Normal & Multinomial Model

We suggest here an hierarchical model which uses both data sets in Tables 1 and 2 and has the following formulation:

$$Y_{ij} = \begin{cases} 0 & \text{when } n_{ij} = 0 \\ inf_{ij} e^{Z_{ij}} & \text{when } n_{ij} > 0 \end{cases}$$

In the case where $n_{ij} > 0$ we have

$$Z_{ij} = log \frac{Y_{ij}}{inf_{ij}},$$

$$Z_{ij} \sim N(\mu_{ij}, \sigma^2), \quad \mu_{ij} = \mu + \alpha_i + \beta_j + log(n_{ij})$$

$$\boldsymbol{n}_{i.} \sim Multinomial((\pi_1, \pi_2, \ldots, \pi_r), T_i), \quad log(\pi_j/\pi_1) = \beta_j^* \tag{2}$$

where $\boldsymbol{n}_{i.}^T = (n_{i1}, n_{i2}, \ldots, n_{ir})$ are the claim counts originated at i year and π_j is the probability for a claim count to be cleared at j year. We use corner constraints for simplification of the model (i.e. $\beta_1^* = 0$) but can be easily transformed to the equivalent model with sum to zero constraints.

The above multinomial model for the claim count is equivalent with the following log-linear model

$$n_{ij} \sim Poisson(\lambda_{ij}), \quad log(\lambda_{ij}) = \mu^* + \alpha_i^* + \beta_j^*$$

under the constraints

$$\sum_{j=1}^{r} n_{ij} = n_{i.} = T_i \;\; and \;\; \sum_{j=1}^{r} \lambda_{ij} = \lambda_{i.} = T_i$$

where μ^* and α_i^* are nuisance parameters; for more details see Agresti (1990). If we assume that T_i is unknown we allow uncertainty for occurred claims but not yet reported. In such case we can use the above log-linear model for claim counts without constraints on λ_{ij} and n_{ij}.

Formulation 2 essentially models the average payment per claim finalised/paid (PPCF=Y_{ij}/n_{ij}) rather than the aggregated claim amount per year. Moreover, it uses the additional information of total claims ($T_1, T_2, ..., T_r$ in Table 2) which may provide valuable information.

We suggest similar non-informative priors as Model 1

$$\mu \sim N(0, \sigma_\mu^2), \quad \alpha_i \sim N(0, \sigma_{\alpha_i}^2), \; i = 2, ..., r, \quad \beta_j \sim N(0, \sigma_{\beta_j}^2), \; j = 2, ..., r$$

$$\tau = \sigma^{-2} \sim G(a_\tau, b_\tau), \quad \beta_j^* \sim N(0, \sigma_{\beta_j^*}^2), \; j = 2, ..., r$$

The same values for σ_μ^2, $\sigma_{\alpha_i}^2$, $\sigma_{\beta_j}^2$ as in Model 1 can be used. For the additional parameters β_j^* we suggest $\sigma_{\beta_j^*}^2 = 100$, for $j = 2, ..., r$.

3.3 Model 3: State Space Modelling of Claim Amounts.

In this kind of problems a commonly used model is the state space (or dynamic linear) Model where the parameters depend on each other in a recursive way. A general description of MCMC in dynamic Models is given by Gamerman (1996). Carter & Kohn (1994) describe how to use Gibbs sampler for general state space Models and Carlin (1992) applies Gibbs sampler for state space Models for actuarial time series. For a short review of Bayesian methods and short description of applying MCMC in the problem of outstanding claims see Makov, Smith & Liu (1996).

The state space Model can be written as

$$Z_{ij} = log\frac{Y_{ij}}{inf_{ij}},$$
$$Z_{ij} \sim N(\mu_{ij}, \sigma^2), \qquad \mu_{ij} = \mu + \alpha_i + \beta_{ij} \tag{3}$$

with corner constraints (rather than the sum to zero constraints used in previous models)

$$\alpha_1 = \beta_{i1} = 0, \quad i = 1, 2, \ldots, r.$$

For each parameter we have the recursive equations:

$$\beta_{ij} = \beta_{i-1,j} + v_i, \quad v_i \sim N(0, \sigma_\nu^2), \quad i = 2, ..., r$$
$$\alpha_i = \alpha_{i-1} + h_i, \quad h_i \sim N(0, \sigma_h^2), \quad i = 2, ..., r.$$

In state space model we only need to define the prior distribution for the first state space parameter. The priors distributions for all other parameters are imposed 'automatically' from the state space conditions; for more details see Carlin *et al.* (1992), Carlin (1992) and Gamerman (1996).

Therefore, we take as prior for $\beta_{1j} \sim N(0, \sigma^2_{\beta_{1j}})$ and $\mu \sim N(0, \sigma^2_\mu)$ with $\sigma^2_{\beta_{1j}} = 100$ and $\sigma^2_\mu = 1000$. The corner constraints on α_i imply a mass prior probability on $\alpha_1 = 0$. The prior for the precision σ^{-2} is a gamma as in model 1. We additionally use non-informative Gamma priors for the parameters $\sigma_\nu^{-2} \sim G(a_\nu, b_\nu)$ and $\sigma_h^{-2} \sim G(a_h, b_h)$ with proposed values $a_\nu = b_\nu = a_h = b_h = 10^{-10}$. Note that $\sigma_\nu^2 = 0$ means that $\beta_{ij} = \beta_{kj}$, for every $i, j, k = 1, ..., r$ and $i \neq k$ (i.e. model 1).

As in Model 1 we note that this model does not use any information from claim counts (Table 2).

3.4 Model 4: State Space Modelling of Average Claim Amount per Accident.

Here we generalise the state space Model (3) by incorporating information from data in Table 2. We suggest the Model

$$Y_{ij} = \begin{cases} 0 & \text{when } n_{ij} = 0 \\ inf_{ij} e^{Z_{ij}} & \text{when } n_{ij} > 0 \end{cases}$$

In the case where $n_{ij} > 0$ we have

$$Z_{ij} = log \frac{Y_{ij}}{inf_{ij}},$$

$$Z_{ij} \sim N(\mu_{ij}, \sigma^2), \quad \mu_{ij} = \mu + \alpha_i + \beta_{ij} + log(n_{ij})$$

$$\boldsymbol{n}_{i.} \sim Multinomial((\pi_1, \pi_2, \ldots, \pi_r), T_i), \quad log(\pi_j/\pi_1) = \beta_j^*, \quad \beta_1 = 0 \tag{4}$$

and the recursive associations

$$\beta_{ij} = \beta_{i-1,j} + v_i, \quad v_i \sim N(0, \sigma_\nu^2), \quad i = 2, ..., r$$

$$\alpha_i = \alpha_{i-1} + h_i, \quad h_i \sim N(0, \sigma_h^2), \quad i = 2, ..., r.$$

The priors can be defined similarly as in Model 2 and 3. We can also use state-space formulation for the claim counts but this will unecessary complicate the model and will not affect very much the results since we are interested for estimation of the claim amounts. Moreover, we may have convergence difficulties and non-identifiable parameters.

4 Illustrated Example

The following data are from a major Greek Insurance company. Tables 3-5 give the claim amounts, the claim counts, the total counts and the inflation factor of each year. The analysis of the data above was initiated by deinflating the data in Table 3 using the inflation factors in Table 5. Therefore, the resulting predictive amounts presented in this section should be multiplied by the corresponding inflation factor to represent amounts for a specific year (for example multiply by 257/100=2.57 to get the inflated amount for year 1996).

Tables 3-5

Tables 6-8

Posterior summaries of Models 1-4 are given in Tables 7 and 8. Note the striking difference of our proposed models 2 and 4 when compared with the existing approaches 1 and 3 for outstanding claim amounts for 1991 and 1992. This deviation is easily explainable if we examine carefully the data in Table 4. The remaining outstanding claims for 1991 are only 132 and account for the 1.05% of the total claim counts (12,601). However, this percentage is comparably much smaller than the corresponding outstanding claim counts of 1989 and 1990 which was 3.03% and 4.48% respectively. This decrease is being taken into account by our models and the produced estimates for 1991 are appropriately adjusted.

Table 9 gives the posterior summaries for variance components for all models. The state space Model for claim amounts (Model 3) resulted in very similar posterior densities with the simple log-normal Model (Model 1) because σ_ν^2 was estimated to be very small. Table 9

5 Discussion

In this paper we analysed the well known problem of outstanding claims of insurance companies. All models where investigated using Bayesian theory and MCMC methodology.

The models fitted can be divided in two categories. The first category models the claim amounts and does not incorporate any information from the claim counts data. The second category contains models that use the average claim amount per accident and therefore information from the claim counts; this is the approach we recommend, and we believe that is the most sensible and robust for predictions in case of policy changes or abnormal fluctuations.

For the data we analysed, we noticed that the state space Model for claim amounts (Model 3) did not differ very much from the simple log-normal Model. This is due to the small posterior values of σ_ν^2 (mean=0.0379, median=0.0002) and may imply that no interaction is needed when modelling the total claim amounts. On the other hand, when we incorporated the claim counts we resulted in posterior density of σ_ν^2 (mean=0.1249, median=0.1197) which gives evidence for a non-zero interaction factor. The posterior expected value of σ^2 in Model 4 was very low comparatively to the other models.

The above models can be generalised by adding other factors in the log-normal model. For example we may assume that the variance of the log-normal model should depend on the claim counts of the corresponding cell. This may not extremely improve the model since we are taking logarithms of the claim amount which is already a way for adjustment of the variance of the error term. These more complicated models are beyond the aim of this paper which illustrates the way claim counts can be incorporated in simple log-normal models using Bayesian methodology and Markov chain Monte Carlo techniques.

Finally, we would like to mention that the Bayesian paradigm used in this paper did not utilize the advantage of using informative prior densities. By illustrating our results with non-informative priors, we only provide a yardstick for comparison with other approaches. However, any prior knowledges can be incorporated in our models using usual quantification arguments.

6 Acknowledgments

We would like to thank Prof. Nickos Frangos for introducing us to the problem and providing us the data and the referees for their useful comments on the paper.

Address for Correspondence

Dr. Petros Dellaportas, Department of Statistics, Athens University of Economics and Business, 76 Patission Street, 10434, Athens, Greece.

Appendix

A Model 2

Conditional posterior distributions for Model 2 are given in detail as a guidance for application of MCMC methods for any model based approach. The procedure is similar for other models. Iterative samples from these conditional densities provide, after some burn-in period and by using an appropriate sample lag (see Table 9), the required posterior samples.

The full posterior of model 2 is given by

$$f(\mu, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\sigma^{-2}}, \beta^*, \boldsymbol{N_{i+j>r+1}}, \boldsymbol{Z_{i+j>r+1}} | z_{i+j\leq r+1}, \boldsymbol{N}_{i+j\leq r+1}) \propto$$

$$f(\boldsymbol{z}|\mu, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{N}) f(\boldsymbol{N}|\boldsymbol{\beta^*}) f(\mu) f(\boldsymbol{\alpha}) f(\boldsymbol{\beta}) f(\boldsymbol{\sigma^{-2}}) f(\boldsymbol{\beta^*}) \tag{5}$$

where $\boldsymbol{N_{i+j>r+1}}, \boldsymbol{Z_{i+j>r+1}}$ are the missing claim counts and inflation adjusted logarithms of amounts, respectively; $\boldsymbol{N_{i+j\leq r+1}}, \boldsymbol{Z_{i+j\leq r+1}}$ are the data for claim counts and inflation adjusted log amounts respectively; $\boldsymbol{z}, \boldsymbol{N}$ are the matrices containing both data and missing values; $\boldsymbol{\beta}^T = (\beta_2, \ldots, \beta_r)$, $\boldsymbol{\alpha}^T = (\alpha_2, \ldots, \alpha_r)$ and $\boldsymbol{\beta}^{*T} = (\beta_2^*, \ldots, \beta_r^*)$. Note that due to sum to zero constraints the parameters α_1 and β_1 are substituted from $-\alpha_2 \ldots - \alpha_r$ and $-\beta_2 \ldots - \beta_r$ and therefore the posterior distribution of these parameters are simply functions of the sampled values for $\alpha_2, \ldots, \alpha_r$ and $\beta_2, \ldots, \beta_r$. In computer programs is convenient to calculate also constrained parameters α_1 and β_1 and therefore each time we sample on of the other parameters these two values must be updated.

Note that

$$f(\boldsymbol{z}|\mu, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{N}) = f(\boldsymbol{z}_{n_{ij}\neq 0}|\mu, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{N}_{n_{ij}\neq 0}) f(\boldsymbol{z}_{n_{ij}=0}|\mu, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{N}_{n_{ij}=0})$$

where $\boldsymbol{z}_{n_{ij}\neq 0}$ is the vector of claim amounts that correspond to non-zero claim counts and $\boldsymbol{z}_{n_{ij}=0}$ is the vector that correspond to zero claim counts.

$$f(\boldsymbol{z}|\mu, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{N}) = \begin{cases} f(\boldsymbol{z}_{n_{ij}\neq 0}|\mu, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{N}_{n_{ij}\neq 0}) & \text{when } y_{ij} = 0 \text{ for all } n_{ij} = 0 \\ 0 & otherwise \end{cases}$$

$$f(\boldsymbol{z}_{n_{ij}\neq 0}|\mu, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{N}_{n_{ij}\neq 0}) = (2\pi\sigma^2)^{-d/2} exp\left(\sum_{i=1}^{r}\sum_{j=1}^{r} I(n_{ij}\neq 0)\{Z_{ij} - \mu - \alpha_i - \beta_j - log(n_{ij})\}^2\right) \tag{6}$$

where $I(A)$ is an indicator variable taking value one if A condition is satisfied and zero otherwise.

$$f(\boldsymbol{N}|\boldsymbol{\beta}^*) = \frac{\prod_{i=1}^{r} (T_i!)}{\prod_{i=1}^{r} \prod_{j=1}^{r} (n_{ij}!)} \prod_{i=1}^{r} \prod_{j=1}^{r} (\pi_j^{n_{ij}})$$

and can be written as

$$f(\boldsymbol{N}|\boldsymbol{\beta}^*) = exp\left(\sum_{i=1}^{r} log(T_i!) - \sum_{i=1}^{r} \sum_{j=1}^{r} log(n_{ij}!) + \sum_{j=2}^{r} n_{.j}\beta_j^* - n_{..}log(\sum_{k=1}^{r} e^{\beta_k^*}) \right)$$

where $n_{..} = \sum_{i=1}^{r} \sum_{j=1}^{r} n_{ij} = \sum_{i=1}^{r} T_i$ and $n_{.j} = \sum_{i=1}^{r} n_{ij}$.

Usually all data claim counts are large and therefore the corresponding missing counts are positive. In this case the full conditional posterior distributions are simplified. The full conditional distribution of this usual case is presented below.

1. The posterior full conditional distribution for μ is given by

$$f(\mu|.) = N\left(\frac{Z_{..} - ln_{..}}{r^2 + \sigma^2/\sigma_\mu^2}, \frac{\sigma^2}{r^2 + \sigma^2/\sigma_\mu^2} \right) \tag{7}$$

where $ln_{..} = \sum_{i=1}^{r} \sum_{j=1}^{r} log(n_{ij})$ and $Z_{..} = \sum_{i=1}^{r} \sum_{j=1}^{r} Z_{ij}$.

2. **[a]** For $i = 2, ..., r$ we have

$$f(\alpha_i|.) = N\left(\frac{Z_{i.} - Z_{1.} - (ln_{i.} - ln_{1.}) - r \sum_{k\neq 1,i} \alpha_k}{2r + \sigma^2/\sigma_{\alpha_i}^2}, \frac{\sigma^2}{2r + \sigma^2/\sigma_{\alpha_i}^2} \right) \tag{8}$$

where $ln_{i.} = \sum_{j=1}^{r} log(n_{ij})$ and $Z_{i.} = \sum_{j=1}^{r} Z_{ij}$.

[b] $\alpha_1 = -\sum_{i=2}^{r} \alpha_i$, due to sum to zero constraints.

3. **[a]** For $j = 2, ..., r$ we have

$$f(\beta_j|.) = N\left(\frac{Z_{.j} - Z_{.1} - (ln_{.j} - ln_{.1}) - r \sum_{k\neq 1,j} \beta_k}{2r + \sigma^2/\sigma_{\beta_j}^2}, \frac{\sigma^2}{2r + \sigma^2/\sigma_{\beta_j}^2} \right) \tag{9}$$

where $ln_{.j} = \sum_{i=1}^{r} log(n_{ij})$ and $Z_{.j} = \sum_{i=1}^{r} Z_{ij}$

[b] $\beta_1 = -\sum_{j=2}^{r} \beta_j$, due to sum to zero constraints.

4. The full posterior conditional distribution of the $\tau = \sigma^{-2}$ is given by

$$f(\tau|.) = Gamma(a_\tau + r^2/2, b_\tau + SS/2), \quad SS = \sum_{i=1}^{r} \sum_{j=1}^{r} (Z_{ij} - \mu_{ij})^2, \tag{10}$$

with $\mu_{ij} = \mu + \alpha_i + \beta_j + log(n_{ij})$.

5. The missing amounts Z_{ij}, are given by

$$f(Z_{ij}|.) = N(\mu_{ij}, \sigma^2), \quad i = 2, 3, ..., r, \quad j = r - i + 2, ..., r, \tag{11}$$

with μ_{ij} as above.

6. **[a]** For $\beta_j^*, \quad j = 2, ..., r$ we have that

$$f(\beta_j^*|.) \propto exp\left(\beta_j^* n_{.j} - n_{..} log(\sum_{k=1}^{r} e^{\beta_k^*}) - 0.5\beta_j^{*2}/\sigma^2_{\beta_j^*}\right) \tag{12}$$

where $n_{.j} = \sum\limits_{i=1}^{r} n_{ij}$.

[b] Set

$$\beta_1^* = 0. \tag{13}$$

Note that in order to receive a sample from (12) we either use Metropolis-Hastings algorithm or Gilks and Wild (1992) algorithm for log concave distributions; see Dellaportas and Smith (1993). Both methods provide approximately identical results.

7. The posterior of the missing counts n_{ij} for $j = r - i + 2, ..., r - 1$ $i = 3, ...r$ is rather complicated since

$$f(n_{ij}|.) \propto f(\boldsymbol{z}|\mu, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{N}) f(\boldsymbol{N}|\boldsymbol{\beta}^*).$$

The constraint $T_i = \sum\limits_{j=1}^{r} n_{ij}$ reduces the above posterior to

$$f(n_{ij}|.) \propto f(Z_{ij}|\mu, \alpha_i, \beta_j, n_{ij}) f(Z_{ir}|\mu, \alpha_i, \beta_r, n_{ir}) f(n_{ij}|\beta_j^*) f(n_{ir}|\beta_r^*)$$

where n_{ir} is calculated as $n_{ir} = T_i - \sum\limits_{j=1}^{r-1} n_{ij}$.

In more detail the above posterior is given by

$$f(n_{ij}|.) \propto \frac{\pi_j^{n_{ij}}}{n_{ij}!} \frac{\pi_r^{\Omega_{ij} - n_{ij}}}{(\Omega_{ij} - n_{ij})!} \exp\left(g_{ij}(n_{ij}) + g_{ir}(\Omega_{ij} - n_{ij})\right)$$

where $\Omega_{ij} = T_i - \sum\limits_{k \neq j, r} n_{ik}$ and $g_{ij}(n_{ij}) = -0.5\sigma^{-2}(Z_{ij} - \mu - \alpha_i - \beta_j - log(n_{ij}))^2$.

We sample the missing counts using the following metropolis step. Propose missing n'_{ij} and $n'_{ir} = \Omega_{ij} - n'_{ij}$, with $i = 3, \ldots, r$ and $j = r - i + 2, \ldots, r$ from

$$Binomial(p_j, \Omega_{ij}),$$

where $p_j = \frac{\pi_j}{\pi_j + \pi_r} = (1 + exp(\beta^*_r - \beta^*_j))^{-1}$.

Accept the proposed move with probability

$$a = min\{1, exp(g_{ij}(n'_{ij}) + g_{ir}(\Omega_{ij} - n'_{ij}) - g_{ij}(n_{ij}) - g_{ir}(\Omega_{ij} - n_{ij}))\}. \qquad (14)$$

References

[1] Agresti A. (1990). Categorical Data Analysis. *John Wiley and Sons*, USA.

[2] de Alba E., Moreno T.M. and Juares M. (1997). Bayesian Estimation of IBNR Reserves *Technical Report*, I.T.A.M., Mexico.

[3] Bernardo J.M. and Smith A.F.M. (1994). Bayesian Theory. *John Willey and Sons*, UK.

[4] Besag J. and Green P.J. (1993). Spatial Statistics and Bayesian Computation. *Journal of the Royal Statistical Society*, Series B, **55**, 25–37.

[5] Best N., Cowles M.K. and Vines K. (1995). CODA: Convergence Diagnostics and Output Analysis Software for Gibbs Sampling Output. *MRC Biostatistics Unit*, Institute of Public Health, Cambridge, UK.

[6] Brooks S.P. and Roberts G.O. (1997). Assessing Convergence of Markov Chain Monte Carlo. *Technical Report*, University of Bristol, UK.

[7] Carlin B.P. (1992). State Space Modelling of Non-standard Actuarial Time Series. *Insurance: Mathematics and Economics*, **11**, 209–222.

[8] Carlin B.P. Polson N.G. and Stoffer D.S.(1992). State Space Modelling of Non-standard Actuarial Time Series. *Journal of the American Statistical Association*, **87**, 493–500.

[9] Carter C.K. and Kohn R.(1994). On Gibbs Sampler for State Space Models. *Biometrika*, **81**, 541–553.

[10] Cowles M.K. and Carlin B.P. (1996). Markov Chain Monte Carlo Diagnostics: A Comparative Review. *Journal of the American Statistical Association*, **91**, 883-904.

[11] De Jong P. and Zehnwirth B. (1983). Claims Reserving, State Space Models and Kalman Filter. *The Journal of the Institute of Actuaries*, **110**, 157–181.

[12] Dellaportas P. and Smith A.F.M. (1993). Bayesian Inference for Generalized Linear and Proportional Hazards Models via Gibbs Sampling. *Applied Statistics*, **42**, 3, 443–359.

[13] Gamerman D. (1996). Markov Chain Monte Carlo for Dynamic Generalized Linear Models. *Technical Report*, Instututo de Mathematica, Universidade Federal do Rio de Janeiro, Brazil.

[14] Gelfand A.E., Hills S.E., Racine-Poon A. and Smith A.F.M. (1990). Illustration of Bayesian Inference in Normal Data Models using Gibbs Sampling. *Journal of the American Statistical Association*, **85**, 972–985.

[15] Gelfand A.E., Smith A.F.M. and Lee T-M. (1992). Bayesian Analysis of Constrained Parameters and Truncated Data Problems Using Gibbs Sampling. *Journal of the American Statistical Association*, **87**, 418, 523–532.

[16] Gilks W.R., Richardson S. and Spiegelhalter D.J (1996). Markov Chain Monte Carlo in Practice. *Chapman and Hall*, UK.

[17] Gilks W.R. and Wild P. (1992). Adaptive Rejection Sampling for Gibbs Sampling. *Applied Statistics* **41**, 337–348.

[18] Haastrup S. and Arjas E. (1996) Claims Reserving in Continuous time: An Non-parametric Bayesian Approach. *Astin Bulletin*, **26**, 139–164.

[19] Haberman S. and Renshaw A.E. (1996). Generalized Linear Models and Actuarial Science. *The Statistician*, **45**, 407–436.

[20] Hesselager O. (1991) Prediction of Outstanding Claims: An Hierarchical Credibility Approach. *Scandinavian Actuarial Journal*, **1991**, 25–47.

[21] Jewell W.S.(1989) Predicting IBNYR Events and Delays. *Astin Bulletin*, **19**, 25–55.

[22] Kass R.E. and Wasserman L. (1996). The Selection of Prior Distributions by Formal Rules. *Journal of American Statistical Association*, **91**, 1343–1370.

[23] Lawless J.F. (1994). Adjustments for Reporting Delays and the Prediction of Occurred but not Reported Events. *Canadian Journal of Statistics*, **22**, 15-31.

[24] Makov U.E., Smith A.F.M. and Liu Y.H. (1996). Bayesian Methods in Actuarial Science. *The Statistician*, **45**, 503–515.

[25] Neuhaus W. (1992). IBNR Models with Random Delay Distributions. *Scandinavian Actuarial Journal*, **1992**, 97–107.

[26] Norberg R. (1986). A Contribution to Modelling of IBNR Claims. *Scandinavian Actuarial Journal*, **1986**, 155–203.

[27] Renshaw A.E. (1989). Chain Ladder and Interactive Modelling (Claims Reserving and GLIM). *The Journal of the Institute of Actuaries*, **116**, 559–585.

[28] Renshaw A. (1994). On the Second Moment Properties and the Implementation of Certain GLIM Based Stochastic Claims Reserving Models. *Actuarial Research Paper 65, Department of Actuarial and Statistics, City University*, London, UK.

[29] Renshaw A. and Verrall R. (1994). A Stochastic Model Underlying the Chain-Ladder Technique. *Proceedings XXV ASTIN Colloquium*, Cannes, France.

[30] Smith A.F.M. and Roberts G.O (1993). Bayesian Computation via the Gibbs Sampler and Related Markov Chain Monte Carlo Methods, (with discussion). *Journal of the Royal Statistical Society*, Series B, **55**, 3–24.

[31] Taylor G.C. and Ashe F.R. (1983). Second Moments of estimates of Outstanding Claims. *Journal of Econometrics*, **23**, 37–61.

[32] Verrall R. (1989). State Space Representation of the Chain Ladder Linear Model. *The Journal of the Institute of Actuaries*, **116**, 589–610.

[33] Verrall R. (1990). Bayes and Empirical Bayes Estimation for the Chain Ladder Model. *Astin Bulletin*, **20**, 217–243.

[34] Verrall R. (1991). Chain Ladder and Maximum Likelihood. *The Journal of the Institute of Actuaries*, **118** , 489–499.

[35] Verrall R. (1993). Negative Incremental Claims: Chain Ladder and Linear Models. *The Journal of the Institute of Actuaries*, **120**, 171–183.

[36] Verrall R. (1994). A Method for Modelling Varying-off Evolutions in Claims Reserving. *Astin Bulletin*, **24**, 325–332.

[37] Verrall R. (1996). Claims Reserving and Generalized Additive Models. *Insurance: Mathematics and Economics*, **19** , 31–43.

List of Tables

		B				
		B_1	B_2	$\cdots$	B_{r-1}	B_r
	A_1	Y_{11}	Y_{12}	$\cdots$	$Y_{1,r-1}$	Y_{1r}
	A_2	Y_{21}	Y_{22}	$\cdots$	$Y_{2,r-1}$	
A	$\vdots$	$\vdots$	$\vdots$			
	A_{r-1}	$Y_{r-1,1}$	$Y_{r-1,2}$			
	A_r	Y_{r1}				

Table 1:

		B					
		B_1	B_2	$\cdots$	B_{r-1}	B_r	
	A_1	n_{11}	n_{12}	$\cdots$	$n_{1,r-1}$	n_{1r}	T_1
	A_2	n_{21}	n_{22}	$\cdots$	$n_{2,r-1}$		T_2
A	$\vdots$	$\vdots$	$\vdots$				$\vdots$
	A_{r-1}	$n_{r-1,1}$	$n_{r-1,2}$				T_{r-1}
	A_r	n_{r1}					T_r

Table 2:

		B						
	Year	1	2	3	4	5	6	7
	1989	527003	220645	130250	84085	72182	21656	49868
	1990	715247	341364	166001	99845	108648	91958	
	1991	1166119	428365	166410	155376	191644		
A	**1992**	1686294	647331	335290	427069			
	1993	2780948	961010	444610				
	1994	3619446	1328151					
	1995	4002087						

Table 3:

	Year	B 1	2	3	4	5	6	7	Total
	1989	6622	1943	489	138	61	223	66	9542
	1990	6943	2133	632	154	162	390		10496
	1991	8610	2216	736	651	256			12601
A	**1992**	9791	3167	1570	624				15565
	1993	11722	3192	1773					17735
	1994	13684	3664						19746
	1995	13068							18600

Table 4:

Year	1989	1990	1991	1992	1993	1994	1995	1996
Inflation (%)	100.0	120.4	143.9	166.6	190.6	214.2	235.6	257.0

Table 5:

	Year											
Model	1990		1991		1992		1993		1994		1995	
1	1107	(17)	1374	(22)	1904	(69)	2505	(118)	3026	(238)	3112	(556)
2	1105	(19)	1322	(6)	1787	(48)	2400	(121)	2892	(271)	2950	(698)
3	1103	(13)	1379	(27)	1896	(68)	2533	(167)	3013	(282)	3091	(475)
4	1101	(3)	1330	(2)	1789	(8)	2433	(31)	2752	(59)	2767	(168)

Table 6:

	Year											
Model	1990		1991		1992		1993		1994		1995	
1	34	(17)	65	(22)	215	(69)	409	(118)	773	(238)	1413	(555)
2	32	(19)	13	(6)	97	(48)	304	(121)	639	(271)	1251	(698)
3	30	(13)	70	(27)	206	(68)	436	(167)	760	(282)	1393	(475)
4	28	(28)	21	(20)	99	(85)	336	(31)	498	(59)	1068	(168)

Table 7:

Model	Year 1990		1991		1992		1993		1994		1995		Total	
1	1222	(338)	679	(177)	470	(140)	299	(110)	152	(59)	88	(54)	2909	(670)
2	1085	(450)	582	(215)	375	(171)	191	(109)	66	(40)	37	(29)	2336	(806)
3	1166	(289)	677	(225)	496	(227)	310	(179)	161	(154)	85	(99)	2895	(834)
4	937	(136)	456	(70)	353	(78)	179	(43)	87	(24)	41	(20)	2052	(226)

Table 8:

Posterior	Model Parameters							
	σ^2				σ_ν^2		σ_h^2	
Value	Model 1	Model 2	Model 3	Model 4	Model 3	Model 4	Model 3	Model 4
mean	0.0893	0.1366	0.0623	0.00008	0.0379	0.1249	0.1091	0.0150
median	0.0816	0.1231	0.0603	0.00008	0.0002	0.1197	0.0777	0.0112
st.dev.	0.0409	0.0596	0.0399	0.00002	0.0777	0.0324	0.1227	0.0145

Table 9: