Claims Reserving When There Are Negative Values in the Development Triangle

Enrique de Alba

ITAM, Mexico
and
University of Waterloo, Canada

37 th. Actuarial Research Conference
The University of Waterloo,

August 7-10, 2002

1. INTRODUCTION

The many uncertainties involved in the payment of losses makes the estimation of the required reserves more difficult. Yet, some of the existing methods, such as the popular chain-ladder, are simple to apply. However, it has become evident that there is a need for better ways not only to estimate the reserves, but also to obtain some measures of their variability. This has led to the development of stochastic reserving models, Taylor (2000), Kass et. al. (2001), England and Verrall (2002), de Alba (2002).

The chain-ladder is used as a benchmark in several of the references mentioned above, due to its generalized use and ease of application. This facilitates comparison between methods. However, in this paper our aim in not to develop Bayesian methods that provide results close to those of the chain-ladder method. Rather, we aim at using 'objective' Bayesian methods to model both claim intensity and severity using some common assumptions and to use the resulting predictive distributions to estimate loss reserves, allowing for negative values.

In this paper we present an application of Bayesian forecasting methods to the estimation of reserves for outstanding claims. We assume that the time (number of periods) it takes for the claims to be completely paid is fixed and known, that payments are made annually and that the development of partial payments follows a stable pay-off pattern. This is in agreement with many existing models for claims reserving in non-life (general) insurance that assume, explicitly or implicitly, that the proportion of claim payments, payable in the *j-th* development period, is the same for all periods of origin, de Alba (2002). The results are applicable to any frequency of claim payments (years, quarters, etc.) and length of pay-off' period. We present a Bayesian approach to forecasting total aggregate claims (number or amount) given data on some development years for several occurrence years. Essentially the data would correspond to a typical run-off triangle used in loss reserving. We use the term claims reserving in its most general sense. In particular we are concerned with the situation when there are negative values in the development triangle of the incremental claim amounts.

We use standard notation, so that Z_{it} = incremental number (or amount) of events (claims) in the *t-th* development year corresponding to year of origin (or accident year) *i*. Thus $\{ Z_{it}; i = 1,...,k,\ t = 1,...,s \}$ where s = maximum number of years (sub periods) it takes to completely pay out the total number (or amount) of claims corresponding to a given exposure year. In this paper we **do not** assume $Z_{it} > 0$ for all $i = 1,\ldots,k$ and $t = 1,\ldots,s$. Most claims reserving methods usually assume that $s=k$ and that we know the values Z_{it} for $i + t \leq k + 1$. The known values are presented in the form of a run-off triangle, Table 1.

Negative incremental values can arise due to timing of reinsurance or salvage recoveries, or premiums being included as negative loss amounts. It could be argued that the problem is more with the data than with the methods. The data should be adjusted before applying these methods to satisfy regulatory requirements. In this respect de Alba and Bonilla

(2002) provide a list of potential adjustments. Although the estimation procedures can be applied both to incurred (paid losses and aggregate case estimates combined) or paid claims, it is probably better to use the latter, since negative values are less likely to appear. That is because case estimates are set individually and often tend to be conservative, resulting in over-estimation in the aggregate. This leads to negative incremental amounts in the later stages of development. Typically these negative values will be the result of salvage recoveries, payments from third parties, total or partial cancellation of outstanding claims, due to initial over-estimation of the loss or to possible favorable jury decision in favor of the insurer, rejection by the insurer, or plain errors.

We extend previous results using a full Bayesian model. In fact, two different models are presented: one to forecast the number of outstanding claims and one for total aggregate claims. The latter is extended from de Alba (2002) to consider negative incremental values. The model presented here allow the actuary to provide point estimates and measures of dispersion, as well as the complete distribution for the reserves.

The paper is structured as follows. Section 2 gives a brief description of previous results relevant to our approach. Section 3 introduces some Bayesian concepts and their applications in actuarial science. Section 4 describes a Bayesian model for claim amounts in the presence of negative values. Some examples are given in Section 6. All types of model are presented only in discrete time.

2. BACKGROUND

For a comprehensive, although not exhaustive, review of existing stochastic methods that can handle the existence of negative incremental values see England and Verrall (2002). Although they provide some Bayesian results, most of the methods presented there approach the problem from the point of view of frequentist or classical statistics and in the framework of generalized linear models (GLM). They provide predictions and prediction errors for the different methods discussed and show how the predictive distributions may be obtained by bootstrapping and Monte Carlo methods. From the classical viewpoint they mainly consider three models, an (over-dispersed) Poisson, a negative binomial and a Normal approximation to the latter. They also mention the standard log-Normal model which was introduced by Kremer (1982) and analyzed in detail in Verrall (1991b). They provide a Bayesian formulation for the Bornhuetter-Ferguson Technique.

England and Verrall (2002) emphasize “that some of the methods presented … are better suited for modeling paid amounts or number of claims, since incurred data, which may include over-estimation of case estimates, leading to negative incremental values, may cause problems.” After describing the stochastic basis for the chain-ladder method, they indicates that “the Normal model has the advantage that it can produce estimates for a wide range of data sets, and is less affected by the presence of negatives”

Table 1

Year of origin	Development Year 1	2	......	t	...	k-1	k
1	Z_{11}	Z_{12}	...	Z_{1t}		$Z_{1,k-1}$	Z_{1k}
2	Z_{21}	Z_{22}	...	Z_{2t}		$Z_{2,k-1}$	-
3	Z_{31}	Z_{32}	...	Z_{3t}		-	
:						-	-
k-1	$Z_{k-1,1}$	$Z_{k-1,2}$				-	-
k	Z_{k1}	-					-

The stochastic version of the chain-ladder method is defined as a generalized linear model (GLM) with an over-dispersed Poisson distribution, Renshaw and Verrall (1998). In the over-dispersed Poisson model the mean and variance are not the same. In our previous notation $m_{ij} = E(Z_{ij})$, with a variance function $V(Z_{ij}) = \phi m_{ij}$ and scale parameter $\phi > 0$, combined with the log 'link' function $log(m_{ij}) = \mu + \alpha_i + \beta_j$. Over-dispersion is achieved through ϕ. This model reproduces the estimates of the classical chain-ladder method.

Estimates of the parameters, $\hat{\mu}, \hat{\alpha}_i, \hat{\beta}_j$, are obtained by using a 'quasi-likelihood' approach. Renshaw and Verrall (1998) suggest the use of Pearson residuals in the GLM when there are negative values. They point out that it "is not applicable to all sets of data, and can break down in the presence of a sufficient number of negative incremental claims." The Poisson assumption seems inadequate for continuous variables, like claims amounts.

The negative binomial model is closely related to the previous one, Verrall (2000). The distribution in the GLM is now assumed to be a negative binomial with mean $(\lambda_j - 1)W_{i,j-1}$ and variance $\phi\lambda_j(\lambda_j - 1)W_{i,j-1}$, where $W_{ij} = \sum_{k=1}^{j} Z_{ik}$ and $\{\lambda_j : j = 2,...,n\}$ are the chain-ladder development factors. As before, ϕ is an over-dispersion parameter. This method yields essentially the same estimates as the (over-dispersed) Poisson. With a sufficient number of negative incremental claims, it is possible that some of the λ's become less than one and so the variance would not exist. It is then possible and necessary to use a Normal approximation, and the chain-ladder results can still be reproduced. It is not recommended to use the Normal approximation in all situations, mainly because real claims data are skewed, even though its application is likely to be less troublesome in practice. The normal approximation assumes the distribution is normal with the same mean as before and variance $\phi_j W_{i,j-1}$. The link function remains the same in all cases. This last model is seen to be equivalent to one proposed by Mack (1993). In addition, these models have the disadvantage that they incorporate *n* new parameters (the ϕ_j) that must also be estimated, but this is the price one must pay to estimate the reserves in the presence of negative values.

3. BAYESIAN MODELS

We do not intend to give here an extensive review of Bayesian methods. Rather we will describe them very briefly and discuss their applications in actuarial science, specifically in loss reserving. Bayesian analysis of IBNR reserves has been considered before by Jewell (1989,1990), Verrall (1990) and Haastrup and Arjas (1996). For general discussion on Bayesian theory and methods see Berger (1985), Bernardo and Smith (1994) or Zellner (1971). For a discussion of Bayesian methods in actuarial science see Klugman (1992), Makov (1996, 2001), Scollnik (2001), Ntzoufras and Dellaportas (2002) and de Alba (2002). Here, we refer only to those that can be applied to situations where $X_{it} < 0$ for some $i,t = 1,\ldots,k$.

Verrall (1990) approaches the subject of predicting outstanding claims using hierarchical Bayesian linear models, considering the fact that the chain-ladder technique is based on a linear model: the two-way analysis of variance model (ANOVA). He essentially carries out a Bayesian analysis of the two-way ANOVA model to obtain Bayes and empirical Bayes estimates. The latter are given a credibility interpretation. Two alternative formulations are considered, one with no prior information and another where he uses a specific prior distribution for the parameters.

More recently, Bayesian results are provided in England and Verrall (2002), notably for the Bornhuetter-Ferguson (B-F) technique. The Bornhuetter-Ferguson technique is useful when there is instability in the proportion of ultimate claims paid in the early development years, so that the chain-ladder technique yields unsatisfactory results. The idea is to use external information to obtain an initial estimate of ultimate claims. In the traditional B-F method use is made explicitly of perfect prior (expert) knowledge of ‘row’ parameters, ultimate claims. This is then used with the development factors of the chain-ladder technique to estimate outstanding claims. This is clearly well suited for the application of Bayesian methods when knowledge is not perfect, England and Verrall (2002). It may break down in the presence of negative values, Verrall (2002). Mack (2000) provides a summary of the technique.

Ntzoufras and Dellaportas (2002) consider various competing models using Bayesian theory and Markov chain Monte Carlo methods. Claim counts are used in order to add a further hierarchical stage in the model with log-normally distributed claim amounts. In a recent paper, de Alba (2002) presents a model for aggregate claims by separating number of claims and average claims, which are also assumed log-normally distributed. In this paper we follow essentially the approach of the latter.

A standard measure of variability is prediction error, defined as the standard deviation of the distribution of possible reserves. In the Bayesian context the usual measure of variability is the standard deviation of the predictive distribution of the reserves. This is a natural way of doing analysis in the Bayesian. In this paper our aim is to obtain not only this standard deviation, but also show the complete predictive distribution.

4. A BAYESIAN MODEL FOR AGGREGATE CLAIMS

In this section we present a model for the unobserved aggregate claim amounts and hence the necessary reserves for outstanding claims. Let the random variable Z_{it} represent the value of aggregate claims in the *t-th* development year of accident year *i*, $i,t=1,...,k$. The Z_{it} are known for $i+t \leq k+1$ and we assume

$$Y_{it} = log\left(\frac{\mathbf{Z_{it}}}{\mathbf{X_{it}}} + \delta\right) = log\left(M_{it} + \delta\right), \qquad (1)$$

where we can use alternative specifications for X_{it}. The parameter δ corrects the values so as to make it possible to take logarithms. The first one of these specifications we consider is to let X_{it} be the number of closed claims in the *t-th* development year corresponding to year of origin *i*. In this case $M_{it} = Z_{it} / X_{it}$ is the corresponding average claim. This is the structure used in Taylor and Ashe (1983). The second specification we consider in this paper is to let $X_{it} = X_i$ for all $t = 1,...,k$, *i.e.* X_i is some measure of exposures in each year of origin, e.g the size of portfolio in year *i*. It is used as a standardizing measure of business volume. This is the formulation used in most of the references, e.g. Verrall (1990). A possible third specification would be to use (1) to model aggregate claim amounts without including any information on number of claims. That is, use $M_{it} = Z_{it}$ or, equivalently, let $X_{it} = 1,\ i,t = 1,...,k$. This specification is equivalent to the model of Doray (1996). Since the first formulation of the three mentioned above is more general (X_{it} depends on *t*) we shall consider it in more detail. In either case, we assume in addition that

$$Y^*_{it} = Log(M_{it} + \delta) = \mu + \alpha_i + \beta_t + \varepsilon_{ij} \qquad \varepsilon_{ij} \sim N(0,\sigma^2) \qquad (2)$$

$i=1,...,k,\quad t=1,...,k$ and $i+t \leq k+1$ so that M_{it} follows a three parameter log-normal distribution, i.e. $M_{it\,t} \sim LN(\mu + \alpha_i + \beta_t, \sigma^2, \delta)$ and

$$f(y^*_{it} \mid \mu, \alpha_i, \beta_t, \sigma^2) \propto \frac{1}{\sigma} exp[-\frac{1}{2\sigma^2}(Log(M_{it} + \delta) - \mu - \alpha_i - \beta_t)^2]\ . \qquad (3)$$

It is well known in ANOVA that certain restrictions must be imposed on the parameters in order to attain estimability. We use the alternative assumption that $\alpha_1 = \beta_1 = 0$. Also, we define $T_U = (k+1)k/2$ = number of cells with known claim information in the upper triangle; and $T_L = (k-1)k/2$ = number of cells in the lower triangle, whose claims are unknown.

It is well known that estimation in the three parameter log-normal distribution can be very unstable, Crow and Shimizu (1988). Hence we will use the 'profile' likelihood with δ replaced by its ML estimator as given in this reference (on page 123), say $\hat{\delta}$, and define $y_{it} = y^*_{it} + \hat{\delta}$. We then carry out the rest of the analysis with this value replaced in (3). Using matrix notation the model in (2) can be written as follows:

$$\underline{y} = W\underline{\theta} + \underline{\varepsilon} \qquad \underline{\varepsilon} \sim N(\underline{0}, \sigma^2 I),$$

where $\underline{y} = \{ y_{it} ; i,t = 1,...,k , i+t \le k+1\}$ is a T_U-dimension vector that contains all the observed values of Y_{it}, $\underline{\theta} = (\mu, \alpha_2,...,\alpha_k, \beta_2,...,\beta_k)'$ is the $((2k-1)\times 1)$ vector of parameters, $\underline{\varepsilon}$ is the $(T_U \times 1)$ vector of errors and W is the $(T_U \times (2k-1))$ design matrix of the model. Now

$$f(\underline{y} \mid \underline{\theta}, \sigma^2, W, \underline{x}_1, \underline{x}_2,..., \underline{x}_k, \hat{\delta}) \propto \sigma^{-T_U} exp[-\frac{1}{2\sigma^2}(\underline{y} - W\underline{\theta})'(\underline{y} - W\underline{\theta})], \qquad (4)$$

where the vectors $\underline{x}_1 = (x_{11}\ x_{12} ..\ x_{1k})'$, $\underline{x}_2 = (x_{21}\ x_{22} .. x_{2k-1})'$, …, $\underline{x}_k = (x_{k1})$ contain the known data in the triangle. We use this specification in the following sections. In the following sub-sections the reader is referred to de Alba (2002) for details, since many of the results are essentially the same. Also, in what follows we use direct Monte Carlo simulation as described in Appendix B of the aforementioned reference.

4.1 LOSS RESERVING USING CLAIMS PER PERIOD

We want to estimate (or obtain the distribution of) aggregate claims for accident year *i* given information on at least one year that has fully developed and perhaps on *m* previous completely known accident years. Let $Z^*_{it} = \sum_{j=1}^{t} Z_{ij}$ for $1 \le t \le k$. Hence, in the run-off triangle setup, we are really interested in estimating Z^*_{ik} $i=2,...,k$, given Z^*_{1k}, X_{it} and Z_{it}, $i=1,...,k$ $t=1,...,k$, with $i+t \le k+1$. Conditioning on X_{it} is particularly important since later we combine the results obtained here with the marginal posterior distribution of the X_{it}, $i,t = 1,…,k$, to estimate outstanding aggregate claims. Now let $R_i = Z^*_{ik} - Z^*_{ia_i}$, for $i=2,...,k$ with $a_i=k-i+1$, so that $Z^*_{ia_i}$ is the accumulation of Z_{it} up to the latest development period and R_i = the total of the aggregate claims process for the development years for which it is unknown, both corresponding to business year *i*.

Hence, using (4) and the same assumptions about the distribution of the number of claims and notation as in Section 4 in de Alba (2002), as well as from independence of the number of claims and the average claim per cell, the joint *pdf* is

$$f(\underline{y},\underline{x}_1,...,\underline{x}_k \mid \underline{\theta},n_1,n_2,...,n_k,\underline{p},\sigma,W,\hat{\delta}) = f(\underline{y} \mid \underline{\theta},\sigma,W,\hat{\delta}) \times \prod_{i=1}^{k} f_{k-i+1}(\underline{x}_i \mid n_i,\underline{p}) \quad ,$$

where $f_{k-i+1}(\underline{x}_i \mid n_i,\underline{p})$ is where denotes a *(k-i+1)*-dimensional multinomial *pf* . We assume the parameters are independent a-priori and specify non-informative priors, de Alba (2002). The joint posterior distribution is then seen to be

$$f(\underline{\theta},\sigma,n_2,...,n_k,\underline{p} \middle| D,\hat{\delta}) \propto \sigma^{-(T_U+1)} exp[-\frac{1}{2\sigma^2}(\underline{y}-W\underline{\theta})'(\underline{y}-W\underline{\theta})] \times \frac{n_1!}{\prod_{t=1}^{k} x_{1t}!} \prod_{t=1}^{k} p_t^{x_{1t}} \times$$

$$\frac{(n_2-1)!}{x_{nk}!(x_2^*-1)!}(1-p_{k-1}^*)^{x_{2k}} p_{k-1}^{*x_2^*} \frac{x_2^*!}{\prod_{t=1}^{k-1} x_{2t}!} \prod_{t=1}^{k-1} \left(\frac{p_t}{p_{k-1}^*} \right)^{x_{2t}} \times$$

$$\frac{(n_3-1)!}{(n_3-x_3^*)!(x_3^*-1)!}(1-p_{k-2}^*)^{n_3-x_3^*} p_{k-2}^{*x_3^*} \frac{x_3^*!}{\prod_{t=1}^{k-2} x_{3t}!} \prod_{t=1}^{k-2} \left(\frac{p_t}{p_{k-2}^*} \right)^{x_{3t}} \times$$

$$\vdots$$

$$\frac{(n_k-1)!}{(n_k-x_{k1})!(x_k^*-1)!}(1-p_1^*)^{n_k-x_{k1}} p_1^{*x_k} \frac{x_k^*!}{x_{k1}!} \prod_{t=1}^{1} \left(\frac{p_1}{p_1^*} \right)^{x_{k1}} . \qquad (5)$$

where D represents all the known information included in the posterior distribution, i.e. $D = \{ x_{11},..x_{1k},x_1^*,x_2^*,...,x_k^*,n_1,W,\underline{y,\hat{\delta}} \}$. We can rewrite (5) as

$$f(\underline{\theta},\sigma,n_2,...,n_k,\underline{p} \middle| D) \propto f(\underline{\theta},\sigma \middle| D) \times \prod_{i=2}^{k} f(n_i \middle| \underline{p},D) \times f(\underline{p} \middle| D),$$

where

$$f(\underline{\theta},\sigma \middle| D) = f(\underline{\theta} \middle| \sigma,D) \times f(\sigma \middle| D).$$

Since $\underline{\theta} = (\mu,\alpha_2,...,\alpha_k,\beta_2,...,\beta_k)'$, and from the first factor in (5), we can write

$$f(\underline{\theta} \middle| \sigma,D) \propto \sigma^{-(2k-1)} \exp[-\frac{1}{2\sigma^2}(\underline{\theta}-\underline{\hat{\theta}})'W'W(\underline{\theta}-\underline{\hat{\theta}})]$$

and

$$f(\sigma \mid D) \propto \sigma^{-(T_U-2k+2)} \exp[-\frac{1}{2\sigma^2}(\underline{y}-W\underline{\hat{\theta}})'(\underline{y}-W\underline{\hat{\theta}})],$$

with $\underline{\hat{\theta}} = (W'W)^{-1}W'\underline{y}$. This is the 'square-root inverted-Gamma' distribution, Bernardo and Smith (1994, page 119). Furthermore, recalling a_i=k-i+1,

$$f(n_i|\underline{p},D) = \frac{(n_i-1)!}{(x_i^*-1)!(n_i-x_i^*)!}(p_{a_i}^*)^{x_i^*}(1-p_{a_i}^*)^{n_i-x_i^*},$$

i = $1,2,\ldots,k$, and from Appendix A in de Alba (2002)

$$f(\underline{p}\,|\,D) = \frac{n_1!}{\prod_{t=1}^{k} x_{1t}!}\prod_{t=1}^{k} p_t^{x_{1t}} \times \frac{x_2^*!}{\prod_{t=1}^{k-1} x_{2t}!}\prod_{t=1}^{k-1}\left(\frac{p_t}{p_{k-1}^*}\right)^{x_{2t}} \times$$

$$\ldots\times\frac{x_3^*!}{\prod_{t=1}^{k-2} x_{3t}!}\prod_{t=1}^{k-2}\left(\frac{p_t}{p_{k-2}^*}\right)^{x_{3t}} \times\ldots\times\frac{x_k^*!}{x_{k1}!}\prod_{t=1}^{1}\left(\frac{p_1}{p_1^*}\right)^{x_{k1}}.$$

To compute the reserves for the outstanding aggregate claims we need to estimate the lower portion of the triangle. We do this by obtaining the mean and variance of the predictive distribution. Hence for each cell we have:

$$E(Z_{it}\,|\,D) = E(M_{it}x_{it}\,|\,D) \;\; = E(M_{it}\,|\,D)E(x_{it}\,|\,D),$$

because of the independence of $\underline{M}$ and $(\underline{x}_1,\ldots,\underline{x}_k)$, where $\underline{M} = \{M_{it};\; i,t = 1,\ldots,k\;, i+t \le k+1\}$. Then the Bayes estimate of outstanding claims for year of business i is $\sum_{t>k-i+1} E(Z_{it}\,|\,D)$. The Bayes 'estimator' of the variance (the predictive variance) for that same year is too cumbersome to derive. Hence we use direct simulation from the posterior distributions to generate a set of N randomly generated values for the number of claims in each cell of the (unobserved) lower right triangle $x_{it}^{(j)}$, $i = 2,\ldots,k$, $t > k-i+1$, for j=$1,\ldots,N$. Then, also for $j = 1,\ldots,N$, we first generate random values of the $y_{it}^{(j)}$ and from them, for the average payment $M_{it}^{(j)} = exp\{y_{it}^{(j)}\} - \hat{\delta}$ for each one of those same cells, and finally for the corresponding pending aggregate loss payment $Z_{it}^{(j)} = x_{it}^{(j)} * M_{it}^{(j)}$, $i = 2,\ldots,k$ and $t > k-i+1$. These values include both parameter variability and process variability. Thus we can compute a random value of the total required reserves $R^{(j)} = \sum_{i,t} Z_{it}^{(j)}$. The mean and variance can be computed as

$$\sigma_R^2 = \frac{1}{N}\sum_{j=1}^{N}\frac{(R^{(j)}-\overline{R})^2}{N} \quad and \quad \overline{R} = \frac{1}{N}\sum_{j=1}^{N} R^{(j)} \quad .$$

The standard deviation σ_R thus obtained is an 'estimate' for the prediction error of the number of claims to be paid. The simulation process has the added advantage that it is not necessary to obtain explicitly the covariances that may exist between parameters. They are dealt with implicitly.

4.2 LOSS RESERVING USING A MEASURE OF EXPOSURE PER ACCIDENT YEAR

In this sub-section we present the second specification for equation (1). Now in it let $X_{it} = X_i =$ some exposure factor or the (known) exposures in each year of origin, Verrall (1990), Renshaw and Verrall (1994), England and Verrall (1999). Since X_i, $i = 1,\ldots,k$, are no longer random variables we only need to model M_{it} as before. The expressions for $Var(Z_{it} \mid D)$ and $Cov(Z_{is}, Z_{it} \mid D)$ will simplify somewhat, but they will still be cumbersome to compute. Hence we obtain the predictive distribution by simulation but we only need to generate the samples for the $M_{it}^{(j)} = exp\{ y_{it}^{(j)} \} - \hat{\delta}$ and the required reserves per cell will now be $Z_{it}^{(j)} = X_i * M_{it}^{(j)}$, all $i = 2,..,k$, $t > k - i + 1$. Using this procedure we can obtain the predictive distribution of the reserves that will be comparable to those given in the references, but with the advantage of having the complete predictive distribution. Notice that since the global model is arrived at using all the information on $\underline{p}$ and this vector does not appear when using $X_{it} = X_i$ = exposures, then there is no need to model X_{it} when it is defined this way. Further simplification can be attained if there is no information on claims per cell or exposure. In that case we would have $X_{it} = 1$ for all *i* and *t*.

5. APPLICATION

In this section we present two sets of data that conatain negative values, Table 2 and Table 3. These have been extensively used to estimate the reserves by different methods. Table 2 comes from Mack (1994) and has only one negative value. Table 3 is taken from Verrall (1991b) and includes three negative values. Tables 4 and 5 present the results of applying different methods to these sets. In Table 4 we compare the results of applying the chain-ladder method, the over-dispersed Poisson and our Bayesian simulation. As expected the required reserves are exactly the same for the first two. We also include the Bootstrap estimator of the standard deviation as given in England and Verrall (2002). This last reference also provides the reserves estimated with the B-F method, which are 50,002; lower than the others. All the standard devistions are well the same range, with the Bayesian one slightly larger. The more striking difference is in the reserves. The

Bayesian result is much higher. It is near the one that one would obtain by applying a straightforward lognormal model with some adjustments for the negative value, Mack (1994). Clearly, the chain-ladder does not seem to be affected by the negative value. Figure 1 shows the predictive distribution of the reserves for Year of Origin 2 (Row 2) and for the total. They are clearly skewed, specially the first one. Analysis for the other Rows show similar results, but they are not included for the sake of brevity. This may be one cause for this difference.

Table 5 provides the results of applying the the same three methods as above, but now we include also those of the straightforward three parameter log-normal model without using the 'variance correction' when estimating the claim amounts per cell. This is the column labeled log-Normal. We also include the result of ignoring the cells with the negative values, i.e. treating them as missing values. In this case there no big differences, the largest being the latter. It is interesting to note that some of the methods yield negative reserves for some of the accident years, but the total is positive. Figure 2 shows the predicitve distribution for accident year 2 (top panel), accident year 6 (middle panel) and for the total (bottom panel). They are all much less skewed than in the previous example. This may expplain why the differences between the results of the different methods are smaller.

The Bayesian method presented here constitutes an appealing alternative to claims reserving methods in the presence of negative values in incremental claims for some cells of the development triangle. Further analysis is needed to clarify some of the differences, which may be warranted when the data is very skewed, as in the first example. On the other hand, this method will not break down even in the presence of a considerable number of negative values. Another point for research will be to avoid the use of the profile likelihood and carry out a fully Bayesian analysis of the problem.

Table 2

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	5012	3257	2638	898	1734	2642	1828	599	54	172
2	106	4179	1111	5270	3116	1817	-103	673	535	
3	3410	5582	4881	2268	2594	3479	649	603		
4	5655	5900	4211	5500	2159	2658	984			
5	1092	8473	6271	6333	3786	225				
6	1513	4932	5257	1233	2917					
7	557	3463	6926	1368						
8	1351	5596	6165							
9	3133	2262								
10	2063									

Source: Mack (1994)

Table 3

	1	2	3	4	5	6	7	8	9	10	11	12
1	290,089	266,666	314,364	468,721	264,735	269,916	125,922	540,684	120,757	58,963	50,837	151,645
2	401,574	648,101	673,897	656,985	458,421	373,010	31,541	279,066	98,551	177,200	-422,178	
3	251,430	373,741	1,827,086	-429,298	801,041	746,157	109,788	212,418	101,225	-3,883		
4	48,924	213,108	644,118	248,680	1,202,333	311,357	1,067,149	697,658	650,711			
5	62,782	278,404	880,618	611,843	243,380	335,226	205,508	164,632				
6	10,684	109,837	189,684	581,492	69,177	323,129	207,976					
7	271,613	290,244	587,769	660,187	681,626	413,425						
8	151,219	183,554	485,830	431,524	427,587							
9	97,658	141,952	369,009	450,971								
10	51,843	119,089	530,706									
11	145,703	421,333										
12	21,019											

Source: Verrall (1991b)

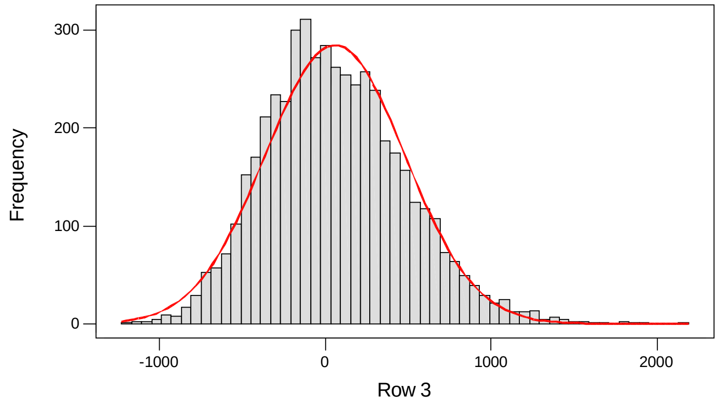
Table 4

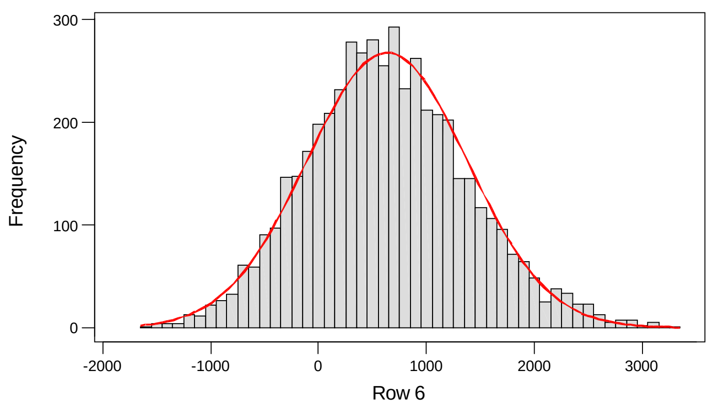
Comparison of Results. Mack Data						
	Chain-L	Over-dispersed Poisson			Bayesian	
Row	Reserves	Reserves	Std. Dev.	Bootstrap	Reserves	Std. Dev.
2	154	154	556	695	189	759
3	617	617	1,120	1,343	1,515	1,548
4	1,636	1,636	1,775	1,992	3,551	2,318
5	2,747	2,747	2,231	2,377	3,577	2,396
6	3,649	3,649	2,440	2,563	4,283	2,833
7	5,435	5,435	3,124	3,093	5,065	3,252
8	10,907	10,907	5,032	5,135	12,654	6,204
9	10,650	10,650	6,075	6,018	14,262	7,790
10	16,339	16,339	12,987	13,644	24,364	15,458
TOTAL	52,135	52,135	18,193	19,267	69,459	21,788

Table 5

Comparison of Results. Verrall Data					
Row	Bayesian	Chain-ladder	Log-Normal	ODP	Missing
2	179,230	184,720	193,396	184,720	185,913
3	58,630	(21,405)	(12,061)	(21,405)	306,927
4	62,530	87,020	532,171	87,020	473,254
5	863,030	238,643	157,646	238,643	408,801
6	634,900	328,846	(92,416)	328,846	319,098
7	224,900	1,052,768	1,373,008	1,052,768	1,248,633
8	1,830,100	1,027,397	939,318	1,027,397	1,191,910
9	1,352,100	1,206,533	1,034,681	1,206,533	1,184,983
10	791,900	1,347,809	1,067,784	1,347,809	1,270,170
11	1,913,500	3,616,144	3,368,769	3,616,144	4,078,328
12	1,492,200	398,872	1,375,010	398,872	557,668
TOTAL	9,403,100	9,467,347	9,937,308	9,467,347	11,225,685

Figure 2


300
200
100
0
Frequency
-1000
0
1000
2000
Row 3


300
200
100
0
Frequency
-2000
-1000
0
1000
2000
3000
Row 6

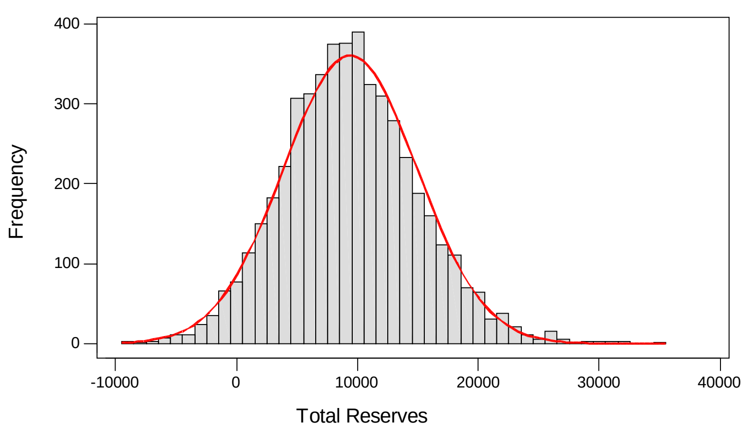
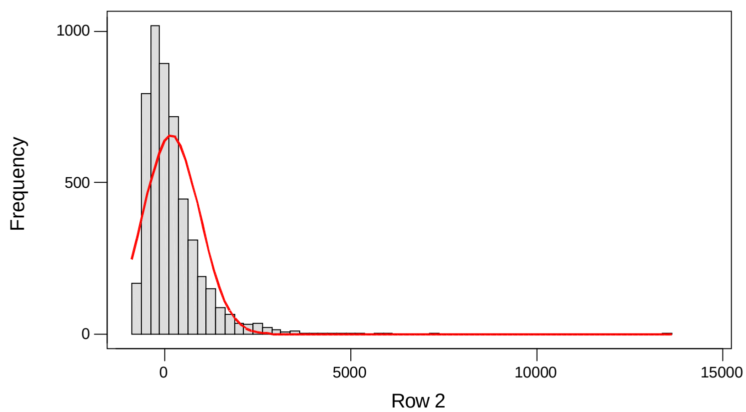
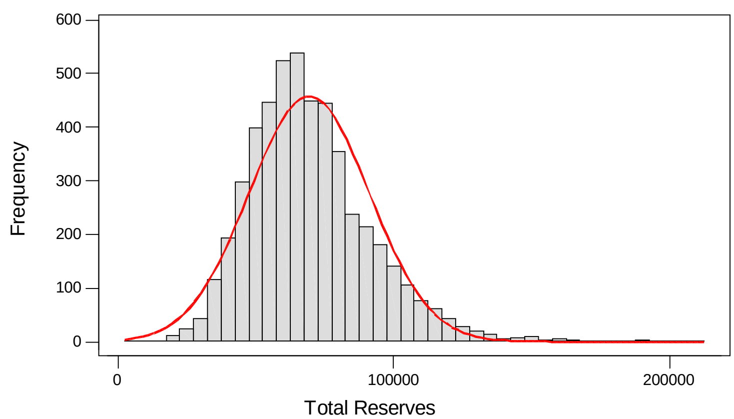

400
300
200
100
0
Frequency
-10000
0
10000
20000
30000
40000
Total Reserves

Figure 1


1000
500
0
Frequency
0
5000
10000
15000
Row 2


600
500
400
300
200
100
0
Frequency
0
100000
200000
Total Reserves

REFERENCES

(1) de Alba, E. (2002), Bayesian Estimation of Outstanding Claims Reserves, *North American Actuarial Journal* 6, forthcoming.

(2) de Alba, E. and R. Bonilla (2002), "Un Modelo Para el Tratamiento de Valores Negativos en el Triángulo de Desarrollo Utilizado en la Estimación de Reservas para SONR", *Transactions of the 27th International Congress of Actuaries*, Cancún, México.

(3) de Alba, E. and Mendoza, M. (1996), Discrete Bayesian Models for Forecasting with Stable Seasonal Patterns, *Advances in Econometrics*, Vol. 11, Part B., 267-281.

(4) Berger, J.O. (1985), *Statistical Decision Theory and Bayesian Analysis*, 2nd. Ed., Springer-Verlag, New York.

(5) Bernardo, J.M. and A.F.M. Smith (1994), *Bayesian Theory*, John Wiley & Sons, New York.

(6) Brown, R. L. (1993), *Introduction to Ratemaking and Loss Reserving for Property and Casualty Insurance*, ACTEX Publications.

(7) Crow, E.L. and Shimizu, K. (1988), *Lognormal Distributions. Theory and Applications*, Marcel Dekker, New York.

(8) Doray, L.G. (1996), UMVUE of the IBNR reserve in a lognormal linear regression model, *Insurance: Mathematics and Economics 18*, 43-57, Elsevier Science B.V.

(9) England, P. and R. Verrall (1999), Analytic and bootstrap estimates of prediction errors in claims reserving, *Insurance: Mathematics and Economics*, Vol. 25, 281-293.

(10) England, P. and R. Verrall (2002), Stochastic Claims Reserving in General Insurance, Institute of Actuaries and Faculty of Actuaries, 1-76.

(11) Haastrup, S. and E. Arjas (1996), Claims Reserving in Continuous Time; A Nonparametric Bayesian Approach, *ASTIN Bulletin* 26(2), 139-164.

(12) Jewell, W.S. (1989), Predicting IBNYR Events and Delays. I Continuous Time, *ASTIN Bulletin* 19(1), 25-56.

(13) Jewell, W.S. (1990), Predicting IBNYR Events and Delays. II DiscreteTime, *ASTIN Bulletin* 20(1), 93-111.

(14) Kass, R., Goovaerts, M., Dhaene, J. and Denuit, M. (2001), Modern Actuarial Risk Theory, Kluwer Academic Publishers.

(15) Klugman, S.A. (1992) *Bayesian Statistics in Actuarial Science*, Kluwer: Boston.

(16) Kremer, (1982), IBNR claims and the two-way model of ANOVA, *Scandinavian Actuarial Journal*, 47-55.

(17) Mack, T. (1993), Distribution-Free calculation of the standard error of chain ladder reserve structure, *ASTIN Bulletin* 23, 213-225.

(18) Mack, T. (1994), Which Stochastic Model is Underlying the Chain Ladder Method?, *Insurance: Mathematics and Economics*, Vol. 15, 133-138.

(19) Makov, U.E., A.F.M. Smith & Y.-H. Liu (1996), "Bayesian Methods in Actuarial Science", *The Statistician*, Vol. 45,4, pp. 503-515.

(20) Makov, U.E. (2001), "Principal applications of Bayesian Methods in Actuarial Science: A Perspective", *North American Actuarial Journal* 5(4), 53-73.

(21) Ntzoufras, I. and Dellaportas, P. (2002), Bayesian Modelling of Outstanding Liabilities Incorporating Claim Count Uncertainty, *North American Actuarial Journal* 6(1), 113-136.

(22) Renshaw, A.E. and R. Verrall (1994), A stochastic model underlying the chain-ladder technique, *Proceedings XXV ASTIN Colloquium*, Cannes.

(23) Renshaw, A.E. and R. Verrall (1998), A stochastic model underlying the chain-ladder technique, *British Acturial Journal* 4,(IV) 905-923.

(24) Scollnik, D.P.M. (2001), Actuarial Modeling With MCMC and BUGS, *North American Actuarial Journal* 5(2), 96-124.

(25) Taylor, G.C. (1986), *Claim Reserving in Non-Life Insurance*, Elsevier Science Publishers, New York.

(26) Taylor, G.C. (2000), *Claim Reserving. An Actuarial Perspective*, Elsevier Science Publishers, New York.

(27) Verrall, R. (1990), Bayes and Empirical Bayes Estimation for the Chain Ladder Model, *ASTIN Bulletin* 20(2), 217-243.

(28) Verrall, R. (1991a), On the estimation of reserves from loglinear models, *Insurance: Mathematics and Economics*, Vol. 10, 75-80.

(29) Verrall, R. (1991b), Negative Incremental Claims: Chain ladder and their Models, *Journal of the Institute of Actuaries* 120(I), 171-183.

(30) Verrall, R. (2000), An investigation into stochastic claims reserving models and the chain-ladder technique, *Insurance: Mathematics and Economics*, Vol. 26, 91-99.

(31) Zellner, A. (1971), An Introduction to Bayesian Inference in Econometrics, Wiley, New York.