Kenya Journal of Sciences Series A Vol. 13 No. 1, 2008

Estimation of IBNR Claims Reserves Using Linear Models

Author:
Patrick G. O. Weke
School of Mathematics, University of Nairobi, Kenya
Email: pweke@uonbi.ac.ke

Abstract
Stochastic models for triangular data are derived and applied to claims reserving data. The standard actuarial technique, the chain ladder technique is given a sound statistical foundation and considered as a linear model. The chain ladder technique and the two-way analysis of variance are employed for purposes of estimating and predicting the IBNR claims reserves.

***Keywords**: IBNR Claims Reserves, Chain Ladder Model, Linear Model, Outstanding Claims.*

INTRODUCTION

If claims runoff triangles are to be analysed statistically, as a data analysis exercise, it is desirable to express them as linear models. If the claims are analysed using a model for each row, then it may be straightforward to write down a linear model. The use of linear models to analyse the data by row can give useful insights into the nature of the data, but it is the linear model which is close to the chain ladder technique that is of greatest interest to actuaries. This linear model, whose connection with the chain ladder technique was first identified by Kremer (1982), is described in sections 3 and 4.

The data is assumed to be lognormally distributed and is first logged before a linear model is applied. The transformation from the raw data to the logged data is, obviously, straightforward, but the reverse transformation, once the analysis has been carried out, is not simple. This is dealt with in section 5. The process is represented in the following diagram (Fig. 1):

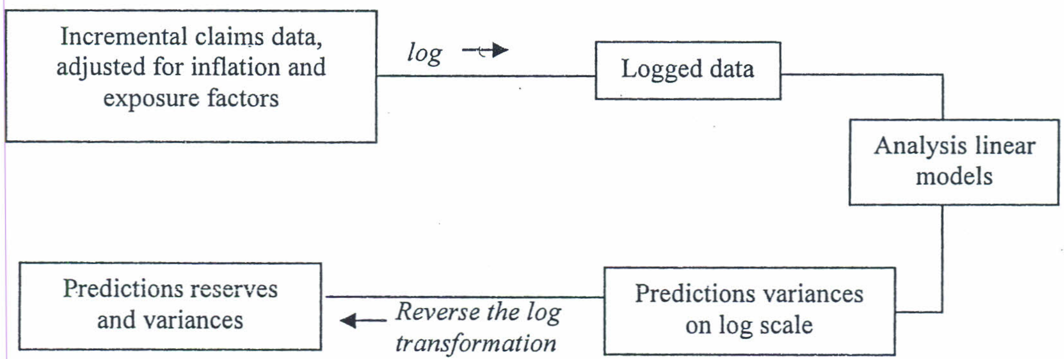

Figure 1. Procedure for analyzing claims data using loglinear models

Prediction from linear models when the data are lognormally distributed was first considered by Finney (1942). Finney considered a sample of independently, identically distributed data, and the theory was generalized to a sample of independently, but not necessarily identically distributed data by Bradu and Mundlak (1970). Subsequent papers by Renshaw (1989), Verrall (1989), and Weke (2006) have considered the properties of the estimators in more detail. The techniques outlined in this paper have been implemented in GLIM (Baker and Nelder, 1978) and the results presented.

2. LINEAR MODELS

The linear model to be considered is

$$y = X\beta + \varepsilon \tag{2.1}$$

where y is a data vector of length n, X is an nxp design matrix, β is a column vector consisting of n unknown parameters, and ε is an error vector of length n. The error vector ε is assumed to have mean zero and variance-covariance matrix Σ..

The minimum variance linear unbiased estimators of the parameters, β, are the weighted least-squares estimators, $\hat{\beta}$, where

$$\hat{\beta} = \left(X'\Sigma^{-1}X\right)^{-1} X'\Sigma^{-1}y. \tag{2.2}$$

If the errors, ε, are assumed to be jointly normally distributed, then the estimators, $\hat{\beta}$, are also the maximum likelihood estimators. Since a logarithmic transformation will be applied to the data, the reverse transformation to estimate actual claims will depend on the estimation method being used. One estimator can be obtained by simply substituting the estimators into the equations. This is used in the lemmas which show the similarity between the chain ladder technique and a certain linear model. However, these estimators, and indeed the maximum likelihood estimators, are biased, and it may be better to use unbiased estimators. If the errors are assumed to be uncorrelated with equal variance then equation (2.1) simplifies to

$$\hat{\beta} = (X'X)^{-1} X'y \tag{2.3}$$

which is a form which will also be used.

The distributional properties of the maximum likelihood estimators, $\hat{\beta}$, are well-known. Assuming that the errors are independently, identically distributed with variance σ^2,

$$\hat{\beta} \sim N\left(\beta, \sigma^2 (X'X)^{-1}\right). \tag{2.4}$$

3. THE CHAIN LADDER TECHNIQUE AS A LINEAR MODEL

Kremer (1982) showed that the chain ladder technique is very similar to a two-way analysis of variance and investigated the properties of the estimators. This section describes the connection between the actuarial chain ladder technique and the statistical analysis of variance method. Assuming a triangular data set (without loss of generality) the cumulative claims data, to which the chain ladder technique is applied, are

$$\{C_{ij} : i = 1,...,t; j = 1,...,t-i+1\} \tag{3.1}$$

The differenced data, to which the analysis of variance model is applied, are

$$\{Z_{ij} : i = 1,...,t; j = 1,...,t-i+1\} \tag{3.2}$$

where $Z_{ij} = C_{ij} - C_{i,j-1}, \qquad j \geq 2$

$Z_{i1} = C_{i1}.$

The chain ladder technique is based on the model

$$E\left[C_{ij}\right] = \lambda_j C_{i,j-1}; \qquad j = 2,...,t. \tag{3.3}$$

The parameter λ_j is estimated by $\hat{\lambda}_j$, where

$$\hat{\lambda}_j = \frac{\sum_{i=1}^{t-j+1} C_{ij}}{\sum_{i=1}^{t-j+1} C_{i,j-1}}. \tag{3.4}$$

The expected ultimate loss, $E\left[C_{ij}\right]$, is estimated by multiplying the latest loss, $C_{i,i-j+1}$, by the appropriate estimated λ – values:

$$\text{estimate of } E[C_{it}] = \left(\prod_{j=t-i+2}^{t} \hat{\lambda}_j\right) C_{i,i-j+1}. \tag{3.5}$$

The chain ladder technique produces forecasts which have a row effect and a column effect. The column effect is obviously due to the parameters $\{\lambda_j; j = 2,...,t\}$. There is also a row effect since the estimates for each row depend not only on the parameters $\{\lambda_j; j = 2,...,t\}$, but also on the row being considered. The latest cumulative claims, $C_{i,t-i+1}$, can be considered as the row effect. This leads to consideration of other models which have row and column effects, in particular the two-way analysis of variance model. The connection is first made with a multiplicative model (see Weke, 2003). This uses the non-cumulative data, Z_{ij}, and models them according to:

Kenya Journal of Sciences Series A Vol. 13 No. 1, 2008

$$E\left[Z_{ij}\right]=U_iS_j \tag{3.6}$$

where U_i is a parameter for row i, and S_j is a parameter for row j.
A multiplicative error structure is assumed and also

$$\sum_{j+1}^{t} S_j = 1. \tag{3.7}$$

In this model, S_j is the expected proportion of ultimate claims which occur in the j th development year; and U_i is the expected total ultimate claim amount for business year i (neglecting any tail factor). The estimates of U_i will be compared with the estimates of $E[C_{it}]$ in equation (3.5) and S_j and λ_j will be related to each other.

The analysis of variance estimators are based on the model (3.6) and the chain ladder technique is based on the model (3.3).

In terms of the models, ignoring for the moment the estimation of the parameters, this simply represents a reparameterisation.

Under the chain ladder model, the expected claim total for business year i is

$$\prod_{j=t-i+2}^{t} \lambda_j C_{i,t-i+1} \tag{3.8}$$

and the expected claim amount in development year $t-i+2$ is

$$\lambda_{t-i+2}C_{i,t-i+1} - C_{i,t-i+1}. \tag{3.9}$$

The equivalent quantities under the multiplicative model (3.6) are

$$U_i \tag{3.10}$$

and

$$U_iS_{t-i+2}. \tag{3.11}$$

Equating (3.8) and (3.9) with (3.10) and (3.11), respectively, gives

$$S_{t-i+2} = \frac{\lambda_{t-i+2} - 1}{\prod_{j=t-i+2}^{t} \lambda_j}.$$

The expected claim amount for development year $t-i+3$ under each model is

$$\lambda_{t-i+3}\lambda_{t-i+2}C_{i,t-i+1} - \lambda_{t-i+2}C_{i,t-i+1} \tag{3.12}$$

and $$U_i S_{t-i+3} \tag{3.13}$$

which gives

$$S_{t-i+3} = \frac{\lambda_{t-i+3} - 1}{\prod_{j=t-i+3}^{t} \lambda_j}.$$

In general, the expected proportion of ultimate claims can be written in the form

$$S_j = \frac{\lambda_j - 1}{\prod_{l=j}^{t} \lambda_l}. \tag{3.14}$$

Considering year of business t, the expected total claim amount under each model is

$$\left[\prod_{j=2}^{t} \lambda_j\right] C_{t1}$$

and $$U_t.$$

The claim amount in development year 1, C_{t1}, is modeled by $U_t S_1$, and so it can be seen that

$$S_1 = \frac{1}{\prod_{l=2}^{t} \lambda_l}. \tag{3.15}$$

To summarize, the chain ladder model (3.3) is equivalent to the multiplicative model given by equation (3.6) with the following relationships between the parameters:

$$S_1 = \left(\prod_{l=2}^{t} \lambda_l\right)^{-1}$$

$$S_j = \left(\prod_{l=j}^{t} \lambda_l\right)^{-1} (\lambda_j - 1)$$

$$U_i = E(C_{it}).$$

Equations (3.4) and (3.5) give the estimates of $\{\lambda_j : j = 2,...,t\}$ and $E(C_{it})$. Estimators of $\{S_i : i = 2,...,t\}$ and $\{U_j : j = 2,...,t\}$ can be obtained by applying a linear model to the logged incremental claims data. Taking logs of both sides of equation (3.6), and assuming that the incremental claims are positive, results into

$$E(Y_{ij}) = \mu + \alpha_i + \beta_j \qquad (3.16)$$

where $Y_{ij} = \log Z_{ij}$ denotes the cumulative claims in development year j in respect of accident year i, and the errors now have an additive structure and are assumed to have mean zero. The errors will be assumed to be identically distributed with variance σ^2, although this distributional assumption can be relaxed. Kremer (1982) defines μ as the mean of the $\log U_i$'s and $\log S_j$'s, so that the restriction

$$\sum_{i=1}^{t} \alpha_i = \sum_{j=1}^{t} \beta_j = 0$$

is imposed.

An alternative assumption is that $\alpha_1 = \beta_1 = 0$. In this case

$$\alpha_i = \log U_i - \log U_1 \qquad (3.17)$$

$$\beta_j = \log S_j - \log S_1 \qquad (3.18)$$

$$\mu = \log U_1 + \log S_1 . \qquad (3.19)$$

The latter set of assumptions is more appropriate for the more sophisticated techniques. However, prediction and estimation of the claims is unaffected by the choice of the assumptions.

The assumption that error terms, ε_{ij}, are independently, identically distributed with variance σ^2 will be used, so that the estimators are given by equation (2.3). Now equation (3.16) can be written in the form of equation (2.1). Suppose, for example, there are three years of data then

$$\begin{bmatrix} y_{11} \\ y_{12} \\ y_{21} \\ y_{13} \\ y_{22} \\ y_{31} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_2 \\ \alpha_3 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \varepsilon_{21} \\ \varepsilon_{13} \\ \varepsilon_{22} \\ \varepsilon_{31} \end{bmatrix} \tag{3.20}$$

clearly gives the form of the parameter vector and the design matrix.

The following lemma, due to Kremer (1982), gives the normal equations for the chain ladder linear model.

3.1 Lemma

For n years data, the best linear unbiased estimators of μ, α_i, β_j are the solutions of

$$\hat{\alpha}_i = \frac{1}{t-i+1}\sum_{j=1}^{t-i+1}\left(Y_{ij} - \frac{1}{t-j+1}\sum_{l=1}^{t-j+1}\left(Y_{lj} - \hat{\alpha}_l\right)\right); \quad i = 2,\dots,t \tag{3.21}$$

$$\hat{\beta}_j = \frac{1}{t-j+1}\sum_{i=1}^{t-j+1}\left(Y_{ij} - \frac{1}{t-i+1}\sum_{l=1}^{t-i+1}\left(Y_{il} - \hat{\beta}_l\right)\right); \quad j = 2,\dots,t \tag{3.22}$$

$$\hat{\mu} = \frac{1}{t(t+1)}\sum_{i=1}^{t}\sum_{j=1}^{t-i+1}\left(Y_{ij} - \hat{\alpha}_i - \hat{\beta}_j\right). \tag{3.23}$$

Proof:

The normal equations of (2.3) lead to

$$(t-i+1)\hat{\mu} + (t-i+1)\hat{\alpha}_i + \sum_{j=2}^{t-i+1}\hat{\beta}_j = \sum_{j=1}^{t-i+1} Y_{ij}; \quad i = 2,\dots,t \tag{3.24}$$

$$(t-j+1)\hat{\mu} + \sum_{i=2}^{t-j+1}\hat{\alpha}_i + (t-j+1)\hat{\beta}_j = \sum_{i=1}^{t-j+1} Y_{ij}; \quad j = 2,\dots,t \tag{3.25}$$

$$\frac{t(t+1)}{2}\hat{\mu} + \sum_{i=2}^{t}(t-i+1)\hat{\alpha}_i + \sum_{j=2}^{t}(t-j+1)\hat{\beta}_j = \sum_{i=1}^{t}\sum_{j=1}^{t-i+1} Y_{ij} \tag{3.26}$$

Noting that $\hat{\alpha}_1 = \hat{\beta}_1 = 0$, equations (3.26) and (3.23) are equivalent. Also equations (3.24) and (3.25) can be written as

$$\hat{\alpha}_i = \frac{1}{t-i+1}\sum_{j=1}^{t-i+1}\left(Y_{ij} - \hat{\beta}_j\right) - \hat{\mu} \qquad (3.27)$$

$$\hat{\beta}_i = \frac{1}{t-j+1}\sum_{i=1}^{t-j+1}\left(Y_{ij} - \hat{\beta}_j\right) - \hat{\mu} \qquad (3.28)$$

Substituting equation (3.27) into equation (3.28) and vice versa gives equations (3.21) and (3.22).

4. RELATIONSHIP BETWEEN THE ESTIMATORS OF THE LINEAR MODEL AND THE CHAIN LADDER MODEL

The previous section derived the relationship between the parameters of the multiplicative model and the chain ladder technique. The parameters are estimated in different ways according to which method is used, and this section is devoted to examining the relationships between the estimators of the parameters.

This section contains two lemmas. The first deals with the estimation of S_j and U_i - the parameters of the multiplicative model – using the chain ladder technique. The second lemma derives the estimators of S_j and U_i using the two-way analysis of variance model. The two sets of estimators are then shown to be similar. Thus, it will be shown that the chain ladder method will produce results which are similar to those produced by the analysis of variance method. The latter has been studied in great depth in statistical literature and the method has the advantage of a great deal of theoretical background. The theory of analysis of variance will be applied to insurance data, bearing in mind that the main method in use in the industry is the chain ladder method.

4.1 Lemma

If

$$S_j = \frac{\lambda_j - 1}{\prod_{l=j}^{t}\lambda_l}; \quad j = 2,\ldots,t \qquad (4.1)$$

and λ_j is estimated by $\tilde{\lambda}_j$, where

$$\tilde{\lambda}_j = \frac{\sum_{i=1}^{t-j+1} C_{ij}}{\sum_{i=1}^{t-j+1} C_{ij-1}} \qquad (4.2)$$

then the estimators of S_j, $\tilde{S}_j$, satisfy the relationship

$$\tilde{S}_j = \frac{\sum_{i=1}^{t-j+1} Y_{ij}}{\sum_{i=1}^{t-j+1} C_{i,t-i+1} \Big/ \left(1 - \sum_{l=t-i+2}^{t} \tilde{S}_l\right)}; \quad j = 2,\ldots,t. \tag{4.3}$$

Also, the estimate of U_i is $\tilde{U}_i$, where

$$\tilde{U}_i = \frac{\sum_{j=1}^{t-i+1} Z_{ij}}{\sum_{j=1}^{t-i+1} \tilde{S}_j}. \tag{4.4}$$

Proof:
Equations (4.1) and (4.2) imply that

$$\tilde{S}_j = \frac{\tilde{\lambda}_j - 1}{\prod_{l=j}^{t} \tilde{\lambda}_l} = \frac{\sum_{i=1}^{t-j+1} C_{ij} - \sum_{i=1}^{t-j+1} C_{ij-1}}{\sum_{i=1}^{t-j+1} C_{ij} \prod_{l=j}^{t} \tilde{\lambda}_l} \tag{4.5}$$

Now, it can be shown by induction that (see Kremer, 1982)

$$\sum_{i=1}^{t-j+1} C_{ij} = \sum_{i=1}^{t-j+1} C_{i;t-i+1} \Big/ \prod_{l=j+1}^{t-i+1} \tilde{\lambda}_l. \tag{4.6}$$

Substituting equation (4.6) into equation (4.5) gives

$$\tilde{S}_j = \frac{\sum_{i=1}^{t-j+1} Z_{ij}}{\left(\sum_{i=1}^{t-j+1} C_{i,t-i+1} \Big/ \prod_{l=j+1}^{t-i+1} \tilde{\lambda}_l\right) \prod_{l=j+1}^{t} \tilde{\lambda}_l} = \frac{\sum_{i=1}^{t-j+1} Z_{ij}}{\sum_{i=1}^{t-j+1} C_{i,t-i+1} \prod_{l=t-i+2}^{t} \tilde{\lambda}_l} \tag{4.7}$$

It can also be shown by induction that

$$\left[\prod_{l=k}^{t} \lambda_l\right]^{-1} = 1 - \sum_{l=k}^{t} S_l.$$

This is true for $k = 2$ by virtue of (3.15) and the relationship

$$1-\sum_{l=2}^{t} S_l = S_1.$$

Suppose it is true for k. Then for $k+1$:

$$1-\sum_{l=k+1}^{t} S_l = 1-\sum_{l=k}^{t} S_l + S_k$$

$$=\left[\prod_{l=k}^{t}\lambda_l\right]^{-1}+\frac{\lambda_k-1}{\prod_{l=k}^{t}\lambda_l}=\left[\prod_{l=k+1}^{t}\lambda_l\right]^{-1}. \tag{4.8}$$

Hence, by induction, the result holds. Substituting this result into (4.7) gives

$$\tilde{S}_j = \frac{\sum_{i=1}^{t-j+1} Z_{ij}}{\sum_{i=1}^{t-j+1} C_{i,t-i+1}\Big/\left(1-\sum_{l=t-i+2}^{t}\tilde{S}_l\right)} \tag{4.9}$$

as required.

Now, since $\quad C_{i,t-i+1} = \sum_{j=1}^{t-i+1} Z_{ij}$

and $\quad \prod_{j=t-i+2}^{t}\tilde{\lambda}_j = \left(1-\sum_{j=t-i+2}^{t}\tilde{S}_j\right)^{-1}$

the estimate of total expected outstanding claims for row i,

$$C_{i,t-i+1}\prod_{j=t-i+2}^{t}\tilde{\lambda}_j$$

can be written as

$$\frac{\sum_{j=1}^{t-i+1} Z_{ij}}{1-\sum_{j=t-i+2}^{t}\tilde{S}_j}.$$

This can in turn be written as $\sum_{j=1}^{t-i+1} Z_{ij}\Big/\sum_{j=1}^{t-i+1}\tilde{S}_j$ (4.10) since $$1-\sum_{j=t-i+2}^{t}\tilde{S}_j = 1-\sum_{j=1}^{t-i+1}\tilde{S}_j .$$

4.2 Lemma

Using the estimation method of Lemma 3.1, an estimate of total expected claims for accident year i, $\hat{U}_i$, is given by

$$\hat{U}_i = \left[\prod_{j=1}^{t-i+1} \frac{Z_{ij}}{w_j}\right]^{\frac{1}{t-i+1}} \cdot \sum_{j=1}^{t} w_j \qquad (4.11)$$

where
$$w_j = \frac{\left[\prod_{i=1}^{t-j+1} Z_{ij}\right]^{\frac{1}{t-j+1}}}{\left[\prod_{i=1}^{t-j+1}\left(\prod_{l=1}^{t-i+1} Z_{il}\right)^{\frac{1}{t-i+1}} \Big/ \left(\prod_{l=1}^{t-i+1} w_l\right)^{\frac{1}{t-i+1}}\right]^{\frac{1}{t-j+1}}} . \qquad (4.12)$$

Further,
$$\hat{U}_i = \frac{\left[\prod_{j=1}^{t-i+1} Z_{ij}\right]^{\frac{1}{t-i+1}}}{\left[\prod_{j=1}^{t-i+1} \hat{S}_j\right]^{\frac{1}{t-i+1}}} . \qquad (4.13)$$

This Lemma can be used to show that the estimates of expected total outstanding claims for each row have similar forms using each method, and can be expected to behave in similar ways. The estimate of U_i is obtained by "hatting" the parameters in the identity

$$U_i = e^{\alpha_i} e^{\mu} \sum_{j=1}^{t} e^{\beta_j}$$

which is derived in the proof of this lemma. The resulting estimate of U_i is not the maximum likelihood estimate, neither is it unbiased, but it does serve the purpose of illustrating the similarity between the chain ladder technique and the two-way analysis of variance.

Proof:

The equations (3.17) to (3.19) imply that

$$e^{\alpha_i} = \frac{U_i}{U_1} \qquad (4.14)$$

$$e^{\beta_j} = \frac{S_j}{S_1} \tag{4.15}$$

and $$e^{\mu} = U_1 S_1 \tag{4.16}$$

Since $$\sum_{j=1}^{t} S_j = 1, \quad S_1 = \left(\sum_{j=1}^{t} e^{\beta_j} \right)^{-1},$$

this, together with equations (4.14) and (4.16) gives

$$U_i = e^{\alpha_i} e^{\mu} \sum_{j=1}^{t} e^{\beta_j} . \tag{4.17}$$

Now let $w_j = e^{\beta_j}$; then equation (3.22) is equivalent to equation (4.12).

The best linear unbiased estimate of $\alpha_i + \mu$ is obtained from equation (3.27). Substituting the estimates of $\alpha_i + \mu$ and β_j into equation (4.17) gives the estimate of U_i in equation (4.11).

Now, equation (4.15) implies that $\hat{S}_j = w_j \Big/ \sum_{l=1}^{t} w_l$ and so equations (4.11) and (4.12) can be written as

$$\hat{U}_i = \frac{\left[\prod_{j=1}^{t-i+1} Z_{ij} \right]^{\frac{1}{t-i+1}}}{\left[\prod_{j=1}^{t-i+1} \hat{S}_j \right]^{\frac{1}{t-i+1}}}$$

and $$\hat{S}_j = \frac{\left[\prod_{i=1}^{t-j+1} Z_{ij} \right]^{\frac{1}{t-j+1}}}{\left[\prod_{i=1}^{t-j+1} \left(\prod_{l=1}^{t-i+1} Z_{il} \right)^{\frac{1}{t-i+1}} \Big/ \left(\prod_{l=1}^{t-i+1} \hat{S}_l \right)^{\frac{1}{t-i+1}} \right]^{\frac{1}{t-j+1}}} . \tag{4.18}$$

Now, if all the geometric means are replaced by arithmetic means in equation (4.18), the recurrence relation for the estimators of S_j becomes the same as that in lemma 4.1. Similarly the estimators of U_i are equivalent if geometric means are replaced by arithmetic means. Thus the two estimation methods, the chain ladder method and the linear model, will produce similar results. The structure of the models is identical and the only difference is the estimation technique. It can be argued that the linear model estimates are best in a statistical sense, but it should be emphasized that in using the linear model instead of the crude chain ladder technique, there are no radical changes.

5. UNBIASED ESTIMATION OF RESERVES AND VARIANCES OF RESERVES

It has been shown that the chain ladder can be considered as a two-way analysis of variance. This linear model, and other linear models, can be used effectively for analyzing claims data and producing estimates of expected total outstanding claims for each year of business. The methods have in common the assumption that the data is lognormally distributed, and the linear models are therefore applied to the logged incremental claims rather than the raw incremental claims data. The problem therefore arises of reversing the log transformation to produce estimates on the original scale. It is this problem which is addressed in this section; in particular the unbiasedness of the estimates is considered. It is important that estimates should be unbiased in order that they are aiming at the correct target and do not yield values which consistently under- or over-estimate. It is also important to consider unbiased estimation of the standard error of the estimates of expected total outstanding claims, in order that some measure of the order of the errors can be attached to the predictions. The procedure for analyzing claims data using loglinear models is illustrated by Figure 1.

The final stage in this procedure – reversing the log transformation – is considered here and unbiased estimates of total outstanding claims are derived. Unbiased estimates of the variances of these estimates are also derived. The theory is applied to claims data (obtained from Weke, 2003) using the analysis of variance linear model and the unbiased estimates are compared with maximum likelihood estimates. In order to make the analysis more easily assimilable, a sample of independently, identically distributed observations is considered first. The theory is then extended to the more general case of independent, but not necessarily identically distributed observations. And, it is the more general theory which is applicable to claims data.

5.1 Unbiased Estimates of Total Outstanding Claims

The purpose of the analysis of the claims data is to produce estimates of the expected total outstanding claims, R_i, for each year of business, and the total outstanding claims, R, for the whole triangle.

An unbiased estimate of R_i is $\hat{R}_i$, where

$$\hat{R}_i = \sum_{j=t-i+2}^{t} \hat{\theta}_{ij} \, . \tag{5.1}$$

and

$$\hat{\theta}_{ij} = \exp\left(\underline{X}_{ij}\hat{\underline{\beta}} + \hat{\sigma}^2/2\right) \tag{5.2}$$

is the maximum likelihood estimate of the expected value of the lognormally distributed data, θ_{ij}, which is related to the mean and variance of the normally distributed data by

$$\theta_{ij} = \exp\left(\underline{X}_{ij}\underline{\beta} + \sigma^2/2\right).$$

The variance of $\hat{R}_i$ can be calculated from

$$Var(\hat{R}_i) = Var\left[\sum_{j=t-i+2}^{t} \hat{\theta}_{ij}\right] = \sum_{j=t-i+2}^{t}\left[Var\left(\hat{\theta}_{ij}\right) + 2\sum_{k=j+1}^{t} Cov\left(\hat{\theta}_{ij}, \hat{\theta}_{ik}\right)\right]. \qquad (5.3)$$

By extending the limits of the summations, the total outstanding claims for the whole triangle can also be obtained.

5.2 Prediction Intervals

Having found an unbiased estimate of total outstanding claims, it is now possible to produce a prediction interval for total outstanding claims. The purpose of the analysis so far has been to produce an estimate of total outstanding claims and an estimate of the variance of this estimate. It is often desirable to find a 'safe' value which is unlikely to be exceeded by the total actual claims.

Let $R =$ total outstanding claims for the whole triangle, and

$\hat{R}$ be an unbiased estimate of $E(R)$.

Suppose that a $(1-\alpha)\times 100\%$ upper confidence bound on total claims, R, is required, then it can be found from

$$\hat{R} + Z_{\alpha/2}\cdot\sqrt{Var(R) + Var(\hat{R})} \qquad (5.4)$$

where $\sqrt{Var(R) + Var(\hat{R})}$ is the root mean square error of prediction, and an unbiased estimate is used.

6. NUMERICAL EXAMPLE

This example illustrates and compares the two methods of claims reserving considered in this paper: the chain ladder method and the two-way analysis of variance. For the analysis of variance model, both the unbiased and maximum likelihood estimates of outstanding claims are given. The data used is that from Weke (2003). The estimates of the parameters in the analysis of variance model and their standard errors are shown in Table 1.

Table 1: Estimates of the row and column and their standard errors

	Estimate	*Standard Error*
Overall mean	6.106	0.165
Row parameters	0.194	0.161
	0.149	0.168
	0.153	0.176
	0.299	0.186
	0.412	0.198
	0.508	0.214
	0.673	0.239
	0.495	0.281
	0.602	0.379
Column parameters	0.911	0.161
	0.939	0.168
	0.965	0.176
	0.383	0.186
	-0.005	0.198
	-0.118	0.214
	-0.439	0.239
	-0.054	0.281
	-1.393	0.379

The standard errors are obtained from the estimates of the variance-covariance matrix of the parameter estimates:

$$(XX)^{-1} X\hat{\sigma}^2$$

where $\hat{\sigma}^2$ is the estimate of the residual variance. For the above data $\hat{\sigma}^2 = 0.116$.

Since the data is in the form of a triangle (there are the same number of rows and columns) and the matrix X is based solely on the design matrix, the standard errors are the same for each row and column parameter.

The row parameters are contained within a much smaller range than the column parameters: (0.149, 0.673) compared with (-1.393, 0.965). It is to be expected that the row parameters should be contained within a fairly small range, since the rows are expected to be similar. Any pattern in the row parameters gives an insight into, and depends upon, the particular claims experience. It is thus quite common to observe that the row parameters lie in a small range, but not typical that they follow a trend.

The fitted values for the analysis of variance model and the actual observations are presented in Table 2. These are unbiased estimates and are presented with the actual observations for comparison. In this table, the top entries are the estimates and those underneath are the actual observations.

Table 2: Fitted values and the actual observations

286170	711785	731359	750301	418911	283724	252756	182559	266237	67948
357848	766940	610542	482940	527326	574398	146342	139950	227229	67948
410587	1021245	1049329	1076506	601040	407078	362646	261930	381987	
352118	884021	933894	1183289	445745	320996	527804	266172	425046	
379337	943516	969461	994572	555294	376094	335044	241994		
290507	1001799	926219	1016654	750816	146923	495992	280405		
339233	843767	866971	889425	496588	336334	299624			
310608	1108250	776189	1562400	272482	352053	206286			
378676	941872	967773	992840	554327	375439				
443160	693190	991983	769488	504851	470639				
389421	968599	995234	1021012	570056					
396132	937085	847498	805037	705960					
420963	1047052	1075844	1103710						
440832	847631	1131398	1063269						
457887	1138894	1170213							
359480	1061648	1443370							
396651	986582								
376686	986608								
344014									
344014									

Of most interest to practitioners are the predicted outstanding claims for each year of business, which are the row totals of predicted values. Table 3 presents the maximum likelihood predictions of the outstanding claims in the lower triangle in the top entries, and the unbiased predictions as the underneath entries. The method does not produce any predictions for the first row, and each subsequent row contains one more predicted value. It can be seen that the maximum likelihood estimates are all higher than the unbiased estimates.

The total predicted outstanding claims for each year of business (the row totals of the predicted outstanding claims) are presented in Table 4. There are three estimates given, the maximum likelihood and unbiased estimates from the analysis of variance model, and the chain ladder estimate.

The maximum likelihood estimates differ most significantly from the unbiased estimates in the early and late rows. The estimates for the middle rows are the closest together, which is where the number of observations used in the estimation is the greatest. The maximum likelihood estimate is asymptotically unbiased, and the greater the number of observations used to estimate the parameters, the closer are the two. The chain ladder estimates are sometimes higher and sometimes lower than the analysis of variance estimates.

Table 3: Maximum likelihood and unbiased predictions of outstanding claims

								101269
								96238
							357398	93599
							350362	88841
						217465	319835	83761
						215218	313105	79394
					335047	243001	357392	93597
					332848	240075	349268	88564
				386433	345088	250283	368102	96402
				384305	342028	246696	358900	91006
			617309	418743	373941	271209	398880	104462
			613257	415031	369373	266419	387593	98281
		1206369	674243	457364	408430	296223	435668	114097
		1193906	666126	450811	401216	289387	421005	106752
	1026594	1053911	589034	399564	356813	258787	380610	99678
	1006382	1031734	575643	389575	346716	250077	363813	92248
888831	913640	937951	524224	355600	317554	230313	338732	88710
844677	867203	889047	496032	335695	298762	215487	313486	79483

Table 4: Total predicted outstanding claims

	Analysis of Variance		*Chain Ladder*
Row	*Maximum Likelihood*	*Unbiased*	
2	101269	96238	94630
3	450997	439203	464668
4	621061	607717	702101
5	1029037	1010755	965576
6	1446307	1422934	1412202
7	2184544	2149953	2176089
8	3592393	4520202	3897142
9	4164990	4056189	4289473
10	4595556	4339873	4618035

The total predicted outstanding claims are:

Analysis of Variance	*Maximum Likelihood*	18186154
	Unbiased	17652064
Chain Ladder		18619916

Table 5 below shows the unbiased estimates of the total outstanding claims for each year of business, the standard errors of these estimates and the root mean square error of prediction. This table can be used in setting safe reserves, and gives an idea of the likely variation of outstanding claims.

Table 5: Unbiased estimates, standard errors and root MSE for each year

Unbiased Estimate	*Standard Error*	*Mean Square Error of prediction*
96238	35105	47202
439203	108804	163217
607717	127616	182847
1010755	195739	269224
1422934	273082	357593
2149953	429669	538533
3529202	775256	942851
4056189	1052049	1197009
4339873	1534943	1631306

The unbiased estimate of total outstanding claims is 17652064 and the root mean square error of prediction is 2759258. Thus a 95% upper bound on total outstanding claims is

$$17652064 + 1.645 \times 2759258 = 22191043.$$

This is a 'safe' reserve for this triangle according to the chain ladder linear model using unbiased estimation.

7. CONCLUSION

It can be seen that the maximum likelihood estimates differ most significantly from the unbiased estimates in the early and late rows. The estimates for the middle rows are the closest together, which is where the number of observations used in the estimation is the greatest. The maximum likelihood estimate is asymptotically unbiased, and the greater the number of observations used to estimate the parameters, the closer are the two. The chain ladder estimates are sometimes higher and sometimes lower than the analysis of variance estimates. There is nothing significant that can be inferred from the differences. This confirms that the crude chain ladder method is a reasonable 'rough and ready' method for calculating outstanding claims, although the more proper method, statistically, is the analysis of variance method (using unbiased estimation).

In conclusion, some practical aspects of claims reserving have been considered. These are the stability of the estimation and predictions, the use of the predictions, their standard errors and the 'safe' reserves in practice. The connection between the linear model and the chain ladder technique has been outlined.

REFERENCES

Baker, R. J. & Nelder, J. A. (1978). *The GLIM System: Release 3*, Numerical Algorithms Group, Oxford.

Bradu, D. & Mundlak, Y. (1970). Estimation in Lognormal Linear Models. *JASA*, Vol. **6**, 198-211.

Finney, D. J. (1942). On the Distribution of a Variate whose Logarithm is Normally Distributed. *JRSS Suppl.*, **7**, 155-161.

Kremer, E. (1982). IBNR Claims and the Two-Way Model of ANOVA. *Scand. Act. J.*, Vol. **1**, 47-55.

Renshaw, A. E. (1989). Chain Ladder and Interactive Modelling: Claims Reserving and GLIM. *Journal of the Institute of Actuaries*, Vol. **116**, Part 3, 559-587.

Verrall, R. J. (1989). A State Space Representation of the Chain Linear Model. *Journal of the Institute of Actuaries*, Vol. **116**, Part 3, 589-609.

Weke, P. G. O. (2003). Estimating IBNR Claims Reserves using Gamma Model and GLIM. *The Nigerian Journal of Risk and Insurance*, Vol. **4**, No. 1, 1-11.

Weke, P. G. O. & Mureithi, A. T. (2006). Deterministic Claims Reserving in Short-Term Insurance Contracts. *The East African Journal of Statistics*, Vol. **1**, No. 2, 198-213.

r