The Geometric Chain-Ladder

By D. KUANG
Hiscox, 1 Great St Helen's, London EC3A 6HX, U.K.
kuang.di@gmail.com
B. Nielsen
Nuffield College, Oxford OX1 1NF, U.K.
bent.nielsen@nuffield.ox.ac.uk
and J.P. Nielsen
Cass Business School, City University London, 106 Bunhill Row, London EC1Y 8TZ, U.K.
Jens.Nielsen.1@city.ac.uk
1 July 2013

Summary: The log normal reserving model is considered. The contribution of the paper is to derive explicit expressions for the maximum likelihood estimators. These are expressed in terms of development factors which are geometric averages. The distribution of the estimators is derived. It is shown that the analysis is invariant to traditional measures for exposure.
Keywords: Arithmetic chain-ladder, geometric chain-ladder, canonical parameter, identification problem, maximum likelihood, log-normal model.

1 Introduction

The chain-ladder method is one of the main tools for reserving in general insurance. It evolves around development factors which are ratios of arithmetic averages of claims in the current development period relative to claims in the previous period. We will call this method the *arithmetic chain-ladder*. Kremer (1985), Mack (1991), Mack and Venter (2000), Verrall and England (2000) and recently Kuang, Nielsen and Nielsen (2009) and Taylor (2011) have discussed how the arithmetic chain-ladder technique arises through maximum likelihood estimation in a Poisson model.

A long-standing alternative to the arithmetic chain-ladder is a reserving model involving the same parametrisation but for log-normal distributed variables. Kremer (1982), Verrall (1991, 1994) and Doray (1996) have discussed the log-normal model and noted that maximum likelihood estimation is done by linear regression. Mack (1994) mentions the model and refers to it as the 'log-linear cross-classified claims reserving method'. These papers exploit that the least square theory gives a general matrix expression for the estimator of the parameter vector. However, in contrast to the standard arithmetic chain-ladder it appears that there are no explicit expressions available for estimators of individuals parameters. The main contribution of the paper is to present closed form expressions for estimators for individual parameters. The derivations and the results have similarities to the contribution of Kuang, Nielsen and Nielsen(2009) for the standard arithmetic chain-ladder. Therefore, the exact differences of the arithmetic and the geometric chain-ladders are now fully transparent. As a part of the analysis we derive log development factors which play a similar role to the traditional development factors in the arithmetic chain-ladder model. These log development factors are differences of logs of geometric averages. We will therefore refer to the log-normal reserving model as the *geometric chain-ladder* model. We present two additional results. The first result is a simple expression for the variance of the estimators. The second result is an invariance of the estimators to exposure factors.

Choosing between an arithmetic and a geometric chain-ladder will depend on the data at hand. The two methods will typically give approximately the same point forecasts and differences arises in relation to distribution forecasts. The two main types of data are triangles with counts and amounts of claims, respectively. For count triangles the Poisson-based arithmetic model is often appropriate. For amount triangles the residuals have much larger variances than warranted by the Poisson assumption. In the context of the arithmetic model the problem can be addressed by introducing an over-dispersion parameter. The bootstrap method of England (2002) can take the over-dispersion into account, but as the over-dispersed Poisson model does not have a clear distributional basis simulation is not so easy. A more elaborate approach is to model the count and amount data jointly as in the double chain-ladder of Martínez-Miranda, Nielsen and Verrall (2012). In contrast, the geometric chain-ladder is naturally heavy-tailed so that it is easy to simulate from the model. A more subtle difference is that in the (over-dispersed) arithmetic model the mean-variance ratio is constant accross the cells of the triangle, whereas in the geometric model the mean-standard deviation ratio is constant accross cells. It would seem desirable to develop methods choosing between these models on that basis, but we leave this to future work. While we are not particularly wedded to one distribution over the other, the main point of the paper is to show that a similar set of analytic results and methods are available for the arithmetic and the geometric model.

A number of extensions of the geometric chain-ladder have found their way into the literature and applications. An example is the log normal model of Hertig (1985) in which the variance parameter depends on the development year. Another example is the inclusion of calendar effects as suggested by Zehnwirth (1994) and later analysed by Kuang, Nielsen and Nielsen (2008a,b, 2011). For a recent practical study involving a log-normal model see Rehmann and Klugman (2009).

A consequence of working on the log scale as compared to the original scale of observed claim sizes might be an extra level of bias. While the geometric chain-ladder produces unbiased estimators on the log scale the unbiasedness is lost when moving to the original scale. Verrall (1991) and Doray (1996) discussed how to use Finney's result to get unbiased estimators. For the arithmetic chain-ladder Mack (1991) has shown that the development factors are unbiased in a conditional model taking the observations for the first development year as given; see also England and Verrall (2002) and Taylor (2003, 2011) for further discussion. This unbiasedness property falls away in an unconditional Poisson model. Rather than discussing these bias issues the purpose of our current work is to identify the closed form expressions of the maximum likelihood estimators in the geometric chain-ladder model. By comparing the closed form solutions of the maximum likelihood estimators for arithmetic chain-ladder on the usual scale, see Kuang, Nielsen and Nielsen (2009), and the closed form solutions of the geometric chain-ladder on the log scale of this paper, it becomes visually very clear what the difference is between working on the original scale and the log-scale. While the original scale operates with standard arithmetic means, the log-scale operates with geometrical means. An additional feature of the log-normal model is that sampling distribution for the individual estimators and the development factor are readily available.

2 Comparing the arithmetic and geometric models

In the following we compare the arithmetic model based on a Poisson distributional assumption and the geometric model based on a log normal distributional assumption. The starting point is the well-known development factors for the arithmetic model. Subsequently, the underlying statistical model for the arithmetic model is discussed and compared to the geometric model. Finally, we introduce the new log development factors, which are based on geometric averages rather than arithmetic averages.

2.1 The arithmetic development factors

Consider a standard incremental insurance run-off triangle of dimension k. Each entry is denoted Y_{ij} so that i is the accident year index and j is the development year index. Collectively, we have data $\mathbf{Y} = \{Y_{ij}, \forall (i,j) \in \mathcal{I}\}$, where $\mathcal{I}$ is the triangular index set

$$\mathcal{I} = \{(i,j) : i \text{ and } j \text{ belong to } (1,\dots,k) \text{ with } i+j-1 = 1,\dots,k\}. \tag{2.1}$$

The standard chain-ladder technique evolves around arithmetic averages of incremental claims. That is, cumulative claims are computed as $Z_{ij} = \sum_{\ell=1}^{j} Y_{i\ell}$. From this the development factors are computed as

$$F_j = \frac{\sum_{i=1}^{k+1-j} \sum_{\ell=1}^{j} Y_{i\ell}}{\sum_{i=1}^{k+1-j} \sum_{\ell=1}^{j-1} Y_{i\ell}} = \frac{\sum_{i=1}^{k+1-j} Z_{ij}}{\sum_{i=1}^{k+1-j} Z_{i,j-1}} \qquad \text{for } j = 2, \dots k. \tag{2.2}$$

We will call these the *arithmetic development factors*. Here we follow the usual nomenclature and denote statistics by roman letters, unless they are estimators for some parameters. In that case we use a greek letter for the parameter, for which a decoration with a hat indicates the corresponding maximum likelihood estimator. Decorations with a tilde indicates a statistic serves as a predictor or a forecast.

The arithmetic development factors are used to forecast future values of claims in the lower triangle $\mathcal{J} = \{i, j \leq k : i+j-1 > k\}$ through

$$\widetilde{Y}_{ij}^{a} = Z_{i,k-i+1}(F_j - 1) \prod_{\ell=k+2-i}^{j-1} F_\ell. \tag{2.3}$$

The full reserve for accident year i is forecast by the sum of such forecasts and satisfies

$$\widetilde{Z}_{i,k}^{a} - Z_{i,k-i+1} = \sum_{\ell=k+2-i}^{k} \widetilde{Y}_{i\ell}^{a} = Z_{i,k-i+1}\{(\prod_{\ell=k+2-i}^{k} F_\ell) - 1\}. \tag{2.4}$$

2.2 Statistical models

From a statistical view point it is natural to ask whether the arithmetic chain-ladder forecasts arise through maximum likelihood analysis of some statistical model. Indeed, this is the case in the model where Y_{ij} are independent $\mathsf{Poisson}(\mu_{ij})$-distributed so that

$$\log P(Y_{ij} = x) = \mu_{ij} x - \exp(\mu_{ij}) - \log(x!) \qquad \text{for } x \in \mathbb{N}_0,$$

and where the μ_{ij} is the log expectation which is parametrised as

$$\log\{\mathsf{E}(Y_{ij})\} = \mu_{ij} = \alpha_i + \beta_j + \delta. \tag{2.5}$$

This has been discussed in for instance Kremer (1985), Mack (1991) and Kuang, Nielsen and Nielsen (2009). In the latter paper maximum likelihood estimators are found for the parameters $\Delta\alpha_i = \alpha_i - \alpha_{i-1}$, $\Delta\beta_j = \beta_j - \beta_{j-1}$, and $\mu_{11} = \alpha_1 + \beta_1 + \delta$. It is also shown that the arithmetic development factor F_j is maximum likelihood estimator for the parameter

$$\Phi_j = \frac{\sum_{\ell=1}^{j} \exp(\sum_{h=2}^{\ell} \Delta\beta_h)}{\sum_{\ell=1}^{j-1} \exp(\sum_{h=2}^{\ell} \Delta\beta_h)} \qquad \text{for } j = 2, \ldots, k.$$

Another role for the arithmetic development factor appears in the maximum likelihood estimation of $\Delta\alpha_i = \alpha_i - \alpha_{i-1}$ as a correction term to the row sums $R_i = Z_{i,k+1-i}$:

$$\widehat{\Delta\alpha}_i = \Delta \log R_i + \log F_{k+2-i}. \tag{2.6}$$

An advantage of formulating a statistical model is that it gives a framework for altering the method for reserving. This can be done by changing the parametrisation, the distributional assumption or the independence assumption. As an example, the parametrisation (2.5) could be changed to include additional components such as a calendar effect. This has been explored for instance in Zehnwirth (1994) and Kuang, Nielsen and Nielsen (2008a,b,2011). Another example is to change the underlying distribution from a Poisson distribution to a log normal distribution which is what we will explore here. In each case maximum likelihood estimators can be computed numerically. The purpose of the present paper is to show that the basic log-normal model can be analysed in the same way as the Poisson model, although the resulting formulas are geometric rather than arithmetic in nature.

When formulating the log-normal model it is convenient to introduce notation for the (natural) logarithm of the observations, that is $y_{ij} = \log Y_{ij}$, noting that the logarithmic transformation requires all observations to be positive. This is restrictive compared to the arithmetic chain-ladder, which only requires that row and columns are non-negative, see (2.6). Collectively, we have the logarithmic data $\mathbf{y} = \{y_{ij}, \forall (i,j) \in \mathcal{I}\}$. The statistical model now assumes that $y_{ij} = \log Y_{ij}$ are independent $\mathsf{N}(\mu_{ij}, \sigma^2)$-distributed with log density

$$\log \mathsf{f}_{y_{ij}}(x) = -\frac{1}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}(x - \mu_{ij})^2$$

where the expectation of the logarithmic data is parametrised as

$$\mathsf{E}(y_{ij}) = \mathsf{E}\{\log(Y_{ij})\} = \mu_{ij} = \alpha_i + \beta_j + \delta. \tag{2.7}$$

The difference between the two parametrisations (2.5), (2.7) is therefore whether μ_{ij} represents the log of the expectation of Y_{ij} or the expectation of the log of Y_{ij}.

In practice, the choice between the (over-dispersed) Poisson model and the log-normal model can to some extent be guided by the variance-mean ratio, which is constant accross cells in the Poisson model

$$\frac{\mathsf{Var}_{Poisson}(Y_{ij})}{\mathsf{E}_{Poisson}(Y_{ij})} = \frac{\exp(\mu_{ij})}{\exp(\mu_{ij})} = 1, \tag{2.8}$$

whereas the log normal model has constant standard deviation-mean ratio accross cells

$$\frac{\mathsf{Sdv}_{LogNormal}(Y_{ij})}{\mathsf{E}_{LogNormal}(Y_{ij})} = \frac{[\{\exp(\sigma^2) - 1\} \exp(2\mu_{ij} + \sigma^2)]^{1/2}}{\exp(\mu_{ij} + \sigma^2/2)} = \{\exp(\sigma^2) - 1\}^{1/2}. \tag{2.9}$$

As a consequence, for the log normal model, the variance-mean ratio will be larger for cells with large mean than for cells with a low mean. Now, suppose some knowledge of the variance-mean ratio of the underlying individual claims was available. If this is constant across cells a Poisson model should be favoured, whereas if it is declining with declining aggregate mean then a log normal model should be favoured.

2.3 Reserving in the log-normal model

The reserves for the log-normal model can be expressed in terms of the maximum likelihood estimator. These arise from the least squares regression of the logarithmic data on a design matrix derived from (2.7). The generic expression for least squares estimators then applies as pointed out by Kremer (1982), Verrall (1991, 1994) and Doray (1996). The generic expression is not so informative about how the reserve is formed. We will show that the maximum likelihood estimators for the individual parameters of the log-normal model resemble those of the Poisson model. In principle, this could be derived directly from the generic least squares estimator, but it turns out to be convenient to use a proof resembling that for Poisson model given in Kuang, Nielsen and Nielsen (2009).

Theorem 3.1 below shows that the maximum likelihood estimators are conveniently written in terms of expressions of the form

$$f_j = \frac{\sum_{i=1}^{k+1-j} \sum_{\ell=1}^{j} \log Y_{i\ell}}{j(k+1-j)} - \frac{\sum_{i=1}^{k+1-j} \sum_{\ell=1}^{j-1} \log Y_{i\ell}}{(j-1)(k+1-j)} \qquad \text{for } j = 2, \ldots k. \tag{2.10}$$

This resembles the log of the classical arithmetic development factor, $\log F_j$, from (2.2), except that geometric averages are used instead of arithmetic averages. We will call f_j a *log geometric development factor*. The formulas for F_j and f_j both compare an average over a $(k+1-j) \times j$-array with an average of a smaller $(k+1-j) \times (j-1)$-sub-array, but there are some subtle differences. First, f_j corrects for the number of summands, whereas F_j does not. Moreover, the reference point for F_j is 1, that is $\log F_j = 0$, when the jth column of observations consists of zeros so that there is no development. This is a boundary outcome, because Poisson-distributed variates are non-negative. Kuang, Nielsen and Nielsen (2009) discuss this boundary outcome in some detail. The reference point for f_j is correspondingly 0, but this occurs when the geometric average of the jth column equals the geometric average of the sub-array. In the common situation where claims fall with the development year then f_j will be negative.

In parallel with the reserving formulas for the classical Poisson model, the log geometric development factors lead to new explicit formulas for reserving in the log-normal model as explored in §4. In particular, the log future claims with indices i, j in the lower triangle, $i, j \leq k$, $i + j - 1 > k$, are forecasted by

$$\tilde{y}_{ij} = \sum_{\ell=1}^{k+1-i} \log Y_{i\ell} + j f_j + \sum_{\ell=k+2-i}^{j-1} f_\ell. \tag{2.11}$$

The median of the claim on the original scale can be forecast by $\tilde{Y}_{ij} = \exp(\tilde{y}_{ij})$. Verrall (1991) and Doray (1996) discuss how to use Finney's result to get unbiased estimators for both Y_{ij} and for the total outstanding reserve. Simple analytic expressions for these unbiased estimators do not seem to exist.

3 Likelihood analysis

The likelihood for the log normal model for the triangle $\mathbf{Y}$ is analysed. This is the same as a normal likelihood for the log data $\mathbf{y}$. At first the parametrisation is discussed. Then the new expressions for the maximum likelihood estimators are presented.

As mentioned above the statistical model assumes that the log observations $y_{ij} = \log Y_{ij}$ are independent, normal $\mathsf{N}(\mu_{ij}, \sigma^2)$-distributed, where μ_{ij} was given in (2.7). The normality assumption can be checked in various ways. One method would be to use a quantile-quantile plot of the residuals against the fitted normal distribution as mentioned by Rehman and Klugman (2010). Another method would be to consider how much the standardised third and fourth moment of the residuals deviate from the normal values. The latter method is easier to implement, but the methods could be used side by side as they convey different distributional information.

The geometric chain-ladder model based on the log-normal distribution has a parametrisation resembling that of the arithmetic chain-ladder model based on the Poisson distribution. It therefore has the same over-parametrisation problem which can be addressed by using the canonical parametrisation explored in Kuang, Nielsen and Nielsen (2008a,b, 2009). The identification problem is that a constant can be added and subtracted to the αs and the βs, respectively, without changing the overall parameter μ_{ij} in (2.7). Using telescopic sums we can write $\alpha_i = \alpha_1 + \sum_{\ell=2}^{i} \Delta\alpha_\ell$, where $\Delta\alpha_\ell = \alpha_\ell - \alpha_{\ell-1}$ is the growth rate from $\alpha_{\ell-1}$ to α_ℓ. Doing the same for the parameter β_j, noting $\mu_{11} = \alpha_1 + \beta_1 + \delta$, and inserting in equation (2.7) gives

$$\mu_{ij} = \mu_{11} + \sum_{\ell=2}^{i} \Delta\alpha_\ell + \sum_{\ell=2}^{j} \Delta\beta_\ell. \tag{3.12}$$

The parameters μ_{ij} can therefore be expressed in terms of a parameter

$$\xi = (\mu_{11}, \Delta\alpha_2, \ldots, \Delta\alpha_k, \Delta\beta_2, \ldots, \Delta\beta_k)' \in \mathbf{R}^{2k-1}. \tag{3.13}$$

Kuang, Nielsen and Nielsen (2009, Theorem 1) show that the parameter vector ξ gives a unique parametrisation of the parameters μ_{ij}.

In the implementation it is convenient to write μ_{ij} in terms of a design matrix. The triangular arrangement of the data does not lend itself well to computation. The triangular data array therefore needs to be ordered as a vector of dimension $k(k+1)/2$. The ordering does not matter as long as the parameters μ_{ij} are stacked in a vector μ in the same order. The stacked vector μ can then be written in terms of a design matrix X with $n = k(k+1)/2$ rows and $q = 2k-1$ columns. It is convenient to let the rows of the design matrix X be indexed by i, j to reflect how the rows relate to the original triangular arrangement of the data. Therefore, the formula (3.12) implies that $\mu_{ij} = X_{ij}'\xi$ where X_{ij} is a q-vector given by

$$X_{ij} = \{1, 1_{(2\leq i)}, \ldots, 1_{(k\leq i)}, 1_{(2\leq j)}, \ldots, 1_{(k\leq j)}\}', \tag{3.14}$$

where the indicator function $1_{(h\leq \ell)}$ takes the value unity if $h \leq \ell$ and zero otherwise.

As statistical model we therefore assume that $y_{ij} = \log Y_{ij}$ are independent for i, j varying in the triangular index set $\mathcal{I}$ and normal $\mathsf{N}(\mu_{ij}, \sigma^2)$-distributed, and that $\mu_{ij} = X_{ij}'\xi$ is defined in terms of the canonical parameter $\xi \in \mathbf{R}^{2k-1}$ through (3.12), (3.13), (3.14).

It is worth noting that the geometric chain-ladder as set out here can be extended to include calendar effects as done by Zehnwirth (1994). Recently, Kuang, Nielsen and Nielsen (2008a,b, 2011) have discussed how to construct a unique parametrisation in that situation and how to forecast the outstanding reserve in the context of structural breaks in the calendar effect.

The geometric chain-ladder model is a regression model. The maximum likelihood estimator for the parameter vector ξ is then obtained by least squares regression. Kremer (1982) and Verrall (1991, 1994) give the following generic matrix expression for the estimator, see also Hendry and Nielsen (2007, §8). Stack the data $y_{ij} = \log Y_{ij}$ as a n-vector $\mathbf{y}$. An $n \times q$ design matrix X is constructed by stacking the row vectors X_{ij}' conformably. The least squares estimator of ξ is then $\widehat{\xi} = (X'X)^{-1}X'\mathbf{y}$.

Many computer packages allow the user to enter the data vector $\mathbf{y}$ and the design matrix X and perform a least squares regression. Alternatively, the expression $\widehat{\xi} = (X'X)^{-1}X'\mathbf{y}$ can be coded with little effort. These expressions do, however, not give much information about the estimators for the individual components of ξ. We therefore make a more detailed analysis.

The log likelihood function of the geometric chain-ladder model based on normally distributed logged data is given by

$$\log \mathsf{L}(\xi, \sigma^2; \mathbf{y}) = -\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i,j\in\mathcal{I}}(y_{ij} - \mu_{ij})^2. \tag{3.15}$$

Some further notation is needed to express the maximum likelihood estimators. Introduce geometric averages of row and column sums as

$$r_i = \frac{\sum_{j=1}^{k+1-i} y_{ij}}{k+1-i}, \qquad c_j = \frac{\sum_{i=1}^{k+1-j} y_{ij}}{k+1-j},$$

along with the log geometric column and row development factors

$$f_j = \frac{\sum_{i=1}^{k+1-j}\sum_{\ell=1}^{j} y_{i\ell}}{j(k+1-j)} - \frac{\sum_{i=1}^{k+1-j}\sum_{\ell=1}^{j-1} y_{i\ell}}{(j-1)(k+1-j)} \qquad \text{for } j = 2,\ldots,k, \tag{3.16}$$

$$g_i = \frac{\sum_{j=1}^{k+1-i}\sum_{\ell=1}^{i} y_{\ell j}}{(k+1-i)i} - \frac{\sum_{j=1}^{k+1-i}\sum_{\ell=1}^{i-1} y_{\ell j}}{(k+1-i)(i-1)} \qquad \text{for } i = 2,\ldots,k. \tag{3.17}$$

The maximum likelihood estimator for the canonical parameter can now be found. The proof is left to the Appendix.

Theorem 3.1 *Consider the geometric chain-ladder model over the upper triangle $\mathcal{I}$ with canonical parameter ξ for the mean and variance σ^2. The maximum likelihood estimator for ξ, σ^2 is unique. It is given by*

$$\Delta\widehat{\alpha}_i = \Delta r_i + f_{k+2-i} \qquad \textit{for } i = 2,\ldots,k, \tag{3.18}$$

$$\Delta\widehat{\beta}_j = \Delta c_j + g_{k+2-j} \qquad \textit{for } j = 2,\ldots,k, \tag{3.19}$$

$$\widehat{\mu}_{11} = r_1 - \sum_{\ell=2}^{k} f_\ell = c_1 - \sum_{\ell=2}^{k} g_\ell, \tag{3.20}$$

$$\widehat{\sigma}^2 = \frac{2}{k(k+1)}\sum_{i,j\in\mathcal{I}}(y_{ij} - \widehat{\mu}_{ij})^2, \tag{3.21}$$

where the estimator for the mean-values is $\widehat{\mu}_{ij} = \widehat{\mu}_{11} + \sum_{\ell=2}^{i}\Delta\widehat{\alpha}_\ell + \sum_{\ell=2}^{j}\Delta\widehat{\beta}_\ell$. Moreover, the log geometric development factors, $f_j = \hat{\varphi}_j$ and $g_i = \hat{\gamma}_i$ say, are maximum likelihood estimators for the parameters, for $i, j > 1$,

$$\varphi_j = \frac{1}{j}\sum_{\ell=2}^{j}\sum_{m=2}^{\ell}\Delta\beta_m - \frac{1}{j-1}\sum_{\ell=2}^{j-1}\sum_{m=2}^{\ell}\Delta\beta_m, \tag{3.22}$$

$$\gamma_i = \frac{1}{i}\sum_{\ell=2}^{i}\sum_{m=2}^{\ell}\Delta\alpha_m - \frac{1}{i-1}\sum_{\ell=2}^{i-1}\sum_{m=2}^{\ell}\Delta\alpha_m. \tag{3.23}$$

The estimators for $\Delta\alpha_i$ and $\Delta\beta_j$ have rather simple expressions. The expressions can be viewed as geometric versions of the estimators for standard arithmetic chain-ladder, see (2.6). The geometric estimators can be interpreted as follows. If the claims y_{ij} had been observed for the full square $1 \le i, j \le k$ then the model would amount to a model for two-sided analysis of variance and the maximum likelihood estimators for $\Delta\alpha_i$ and $\Delta\beta_j$ would have been Δr_i and Δc_j. For the triangular data the estimators involve a correction in terms of the development factor.

It is convenient to note how the parameters, $\alpha_i - \alpha_1$ and $\beta_j - \beta_1$ can be expressed in terms of the geometric development parameters γ_i and φ_j.

Theorem 3.2 *It holds $\alpha_j - \alpha_1 = j\gamma_j + \sum_{m=2}^{j-1}\gamma_m$ and $\beta_j - \beta_1 = j\varphi_j + \sum_{m=2}^{j-1}\varphi_m$.*

4 Predicting the reserve

In practice predictions of the reserve are most easily done through the formula $\tilde{y}_{ij} = X'_{ij}\widehat{\xi}$, which can be coded with little effort. Once again, this formula does not reveal much about the structure of the prediction and how the predictions (2.3), (2.4) for the arithmetic chain-ladder are changed when moving to the geometric chain-ladder. Simple expressions for the predictions from the geometric chain-ladder model can be constructed.

As a start consider the predictions for the log values y_{ij}. Combine the expressions in Theorem 3.1 with the expression for μ_{ij} in (3.12) to get, for $j = 1$, the predictor

$$\tilde{y}_{i1} = (r_1 - \sum_{\ell=2}^{k} f_\ell) + (r_i - r_1 + \sum_{\ell=2}^{i} f_{k+2-\ell}) = r_i - \sum_{\ell=2}^{k+1-i} f_\ell \qquad \text{for } j = 1, \tag{4.24}$$

with the convention that the empty sum is zero. In particular $\tilde{y}_{k1} = r_k = y_{k1}$ corresponding to the 'corner solution' for the standard arithmetic chain-ladder. This expression is a geometric version of the corresponding arithmetic prediction of $R_i / \prod_{\ell=2}^{k+1-i} F_\ell$, where $R_i = Z_{i,k+1-i}$, see Kuang, Nielsen and Nielsen (2009, equation 15).

To get an expression for the predictor for $j > 1$ combine this with an expression for $\sum_{\ell=2}^{j} \Delta\beta_\ell = \beta_j - \beta_1$ in Theorem 3.2 and the estimator for φ_j in Theorem 3.1 to get

$$\tilde{y}_{ij} = \widehat{\mu}_{ij} = r_i - \sum_{\ell=2}^{k+1-i} f_\ell + j f_j + \sum_{\ell=2}^{j-1} f_\ell \qquad \text{for } j > 1. \tag{4.25}$$

The connection between this expression and the corresponding arithmetic prediction $R_i(F_j - 1)(\prod_{\ell=2}^{j-1} F_\ell)/(\prod_{\ell=2}^{k+1-i} F_\ell)$ is more tenuous because the term $F_j - 1$ does not easily translate into a geometric version, see Kuang, Nielsen and Nielsen (2009, equation 16).

For the lower triangle the expression (4.25) reduces to

$$\tilde{y}_{ij} = r_i + j f_j + \sum_{\ell=k+2-i}^{j-1} f_\ell \qquad \text{for } k - i + 1 < j \leq k, \tag{4.26}$$

whereas for the upper triangle the prediction is

$$\tilde{y}_{ij} = r_i + j f_j - \sum_{\ell=j}^{k+1-i} f_\ell \qquad \text{for } 1 < j \leq k - i + 1. \tag{4.27}$$

In applications it is of interest to make predictions on the original scale. Taking exponential, $\exp(\tilde{y}_{ij})$, yields the same median as Y_{ij}. The expectation of a single entry on the original scale is $\mathsf{E}(Y_{ij}) = \exp(\mu_{ij} + \sigma^2/2)$. The maximum likelihood estimator of this expectation is found by inserting the estimators for μ_{ij}, σ^2 giving

$$\tilde{Y}_{ij} = \exp(\tilde{y}_{ij} + \widehat{\sigma}^2/2) = \exp\{\widehat{\mu}_{11} + (\widehat{\alpha}_i - \widehat{\alpha}_1) + (\widehat{\beta}_j - \widehat{\beta}_1) + \widehat{\sigma}^2/2\}. \tag{4.28}$$

Likewise the maximum likelihood estimator of the full reserve is

$$\sum_{i=2}^{k}\sum_{j=k+2-i}^{k}\tilde{Y}_{ij} = \exp(\widehat{\sigma}^2/2)\sum_{i=2}^{k}\sum_{j=k+2-i}^{k}\exp(\tilde{y}_{ij}) \quad (4.29)$$

$$= \exp(\widehat{\mu}_{11}+\widehat{\sigma}^2/2)\sum_{i=2}^{k}\sum_{j=k+2-i}^{k}\exp\{(\widehat{\alpha}_i-\widehat{\alpha}_1)+(\widehat{\beta}_j-\widehat{\beta}_1)\}.$$

While these predictions on the original scale are biased, Verrall (1991) and Doray (1996) have discussed how to use Finney's result to get unbiased predictors of Y_{ij} as well as of the full reserve. Simple analytic expressions for these predictors do not seem to exist. An alternative approach would be to use simulation to find the predictive distribution.

5 Distribution of estimator

Regression theory shows that $\widehat{\xi}=(X'X)^{-1}X'y$ is normally $\mathsf{N}\{\xi,\sigma^2(X'X)^{-1}\}$ distributed, whereas $\sigma^2=\sigma^2 n^{-1}y'\{I_n-X(X'X)^{-1}X'\}y$ is independent thereof and χ^2_{df}/n-distributed with $n=k(k+1)/2$ and $df=n-(2k-1)=(k-2)(k-1)/2$. It turns out that simple expressions can be derived for the diagonal elements of the matrix $(X'X)^{-1}$ and hence for the variances of the components of $\widehat{\xi}$. Thereby some analytic insight can be gained in the estimation uncertainty for the geometric chain-ladder model. Indeed a clear pattern emerges for the uncertainty of the estimators.

It should be noted that in a stochastic reserving exercise the forecast errors need to be considered. If the estimation error is to be taken into account in the forecast error calculation, this is easily computed directly from $(X'X)^{-1}$.

Theorem 5.3 *The estimators have variances, for $\ell+1=2,\dots k$,*

$$\mathsf{var}(\widehat{\mu}_{11}) = \frac{2\sigma^2}{k}\left(\frac{k-1}{k}+\frac{1}{k+1}\sum_{\ell=1}^{k}\frac{1}{\ell}\right), \quad (5.30)$$

$$\mathsf{var}(\Delta\widehat{\alpha}_{\ell+1}) = \mathsf{var}(\Delta\widehat{\beta}_{\ell+1}) \quad (5.31)$$

$$= \sigma^2\left\{\frac{1}{k-\ell}+\frac{1}{k+1-\ell}+\frac{1}{(k-\ell)(k+1-\ell)\ell}\right\}, \quad (5.32)$$

$$\mathsf{var}(f_{\ell+1}) = \mathsf{var}(g_{\ell+1}) = \frac{\sigma^2}{(k-\ell)\ell(\ell+1)}. \quad (5.33)$$

The proof is given in the Appendix. It can be extended to considering the variance of $\widehat{\mu}_{ij}$, but that result does not appear to be particularly illuminating. The result would of course show that $\widehat{\mu}_{1k}$ and $\widehat{\mu}_{k1}$ have variance σ^2 reflecting that the 'corner solution' properties $\tilde{y}_{1k}=y_{1k}$ and $\tilde{y}_{k1}=y_{k1}$.

A first implication arises for large triangles, that is for large k. In that case $\sum_{\ell=1}^{k}\ell^{-1}$ is of order $\log k$. As a consequence the variance of $\widehat{\mu}_{11}$ will have leading term $2\sigma^2/k$. In other words the variance of $\widehat{\mu}_{11}$ is more or less the innovation variance σ^2 divided by the number of observations.

The expression for $\mathsf{var}(\Delta\widehat{\alpha}_\ell)$ is increasing in ℓ, which is proved in detail in the appendix. It follows that the accident parameters $\Delta\alpha_\ell$ and corresponding development parameters $\Delta\beta_\ell$ are more imprecisely estimated nearer the corners of the reserve triangle. This is exactly what one would have expected. In particular for large triangles, so k is large, then $\mathsf{var}(\Delta\widehat{\alpha}_2)$ and $\mathsf{var}(\Delta\widehat{\alpha}_k)$ have leading terms $2\sigma^2/k$ and $(3/2)\sigma^2$. Thus the quality of the estimators for $\Delta\alpha_k$ and $\Delta\beta_k$ remains relatively poor even in large triangles.

6 Exposure

Often information about exposure is available, that is information about the size of the portfolio for a given accident year. Typically claim observations are divided by the exposure factor before the analysis. The resulting reserves are then scaled up with the exposure factor, as in Taylor and Ashe (1983), Renshaw (1989), Wright (1990), Verrall (1991, 1994) and Zehnwirth (1994, §12.3). The next theorem shows that for the geometric chain-ladder the same reserve would actually be achieved if the exposure were ignored. This result contrasts with the usual arithmetic chain-ladder where correction for exposure alters the reserve. In particular, when the geometric development factors are the same when computed from the original data and from exposure corrected data. This is not the case for the usual arithmetic development factors.

The proof of the result exploits the logarithmic transformation. That is, on the original scale the claims numbers Y_{ij} are corrected for the exposure H_i by the scaling Y_{ij}/H_i. On the logarithmic scale $y_{ij} = \log Y_{ij}$ this corresponds to the linear translation into $\log(Y_{ij}/H_i) = y_{ij} - \log H_i$. In the log normal model this corresponds to translating the accident year effect α_i into $\alpha_i - \log H_i$. The linearity properties of least squares estimation then imply that the geometric chain-ladder delivers the same reserve when exposure is ignored altogether and when exposure is taken out before estimation and then put back in again. The details of the proof are in the Appendix.

Theorem 6.4 *Consider a run-off triangle $\mathbf{Y} = \{Y_{ij}, \forall i, j \in \mathcal{I}\}$ with exposure factors H_i, $i = 1, \ldots, k$. The following two calculations result in the same prediction of Y_{ij} for $1 \le i, j \le k$.*

1. *Apply the geometric chain-ladder to $\mathbf{Y}$. That is, log transform the data $\mathbf{Y}$, estimate and predict as outlined in §3, §4.*

2. *Construct $\mathbf{Y}_H = \{Y_{ij}/H_i, \forall i, j \in \mathcal{I}\}$ which is the triangle of claims divided by exposure. Apply the geometric chain-ladder. That is, log transform the data $\mathbf{Y}_H$, estimate and predict as outlined in §3, §4. Multiply the reserve by H_i.*

The variance estimator $\widehat{\sigma}^2$ is the same when computed from the log transformations of $\mathbf{Y}$ and of $\mathbf{Y}_H$.

The proof given in the appendix exploits a general feature of least squares regression. It can be applied also for extensions of the model to situations where closed form estimators are not readily available, such as when including a calendar effect, see also

Kuang, Nielsen and Nielsen (2011). As an alternative proof the closed form expressions found above can be used. By inspection of the expression (2.10) for the log column development factor f_j this is seen to be invariant to exposure. Inserting this in the formula (2.11) shows that the forecast for the exposure corrected data is $\tilde{y}_{ij}^H = \tilde{y}_{ij} - \log H_i$. Adding in the log exposure gives the forecast based on the original scale. Likewise, the residuals from the exposure corrected data are of the form $(y_{ij} - \log H_i) - \tilde{y}_{ij}^H = y_{ij} - \tilde{y}_{ij}$ which results in the same variance estimator for both data sets. Therefore the invariance to exposure extends to distribution forecasts because the variance estimator is invariant to the exposure.

7 Empirical illustration

To illustrate the results consider a run-off triangle from motor third party liability. The data originate from the general insurer RSA and were also used in for instance Miranda-Martínez, Nielsen, Nielsen and Verrall (2011). At first the proposed estimators and predictions are computed and compared with the corresponding estimators and predictions from the usual arithmetic chain-ladder. Subsequently, distribution forecasts are presented.

7.1 Motor data and model fit

i vs. j	1	2	3	4	5	6	7	8	9	10
1	451288	339519	333371	144988	93243	45511	25217	20406	31482	1729
	(-0.2)	(-1.0)	(1.9)	(0.3)	(-0.1)	(-0.4)	(-1.4)	(-0.3)	(1.2)	(0)
2	448627	512882	168467	130674	56044	33397	56071	26522	14346	*1560*
	(0.1)	(0.9)	(-0.2)	(0.3)	(-1.6)	(-1.1)	(1.9)	(1.0)	(-1.2)	**1.08**
3	693574	497737	202272	120753	125046	37154	27608	17864	*21916*	*1694*
	(1.4)	(0.5)	(0.2)	(-0.3)	(1.1)	(-1.1)	(-1.0)	(-0.7)	**1.25**	**1.20**
4	652043	546406	244474	200896	106802	106753	63688	*31799*	*32050*	*2477*
	(-0.2)	(-0.6)	(-0.5)	(0.2)	(-0.9)	(1.4)	(0.6)	**0.86**	**0.97**	**0.93**
5	566082	503970	217838	145181	165519	91313	*48907*	*29002*	*29231*	*2259*
	(-0.4)	(-0.5)	(-0.6)	(-0.7)	(1.0)	(1.2)	**0.95**	**0.85**	**0.96**	**0.92**
6	606606	562543	227374	153551	132743	*62736*	*46382*	*27505*	*27722*	*2143*
	(0.1)	(0.0)	(-0.3)	(-0.3)	(0.4)	**1.10**	**1.04**	**0.93**	**1.05**	**1.01**
7	536976	472525	154205	150564	*96137*	*51088*	*37771*	*22398*	*22575*	*1745*
	(0.4)	(0.2)	(-0.9)	(0.4)	**1.07**	**1.14**	**1.07**	**0.97**	**1.08**	**1.04**
8	554833	590880	300964	*176229*	*126009*	*66963*	*49507*	*29358*	*29589*	*2287*
	(-0.5)	(-0.0)	(0.5)	**0.96**	**1.00**	**1.07**	**1.01**	**0.91**	**1.01**	**0.98**
9	537238	701111	*261229*	*176067*	*125893*	*66902*	*49461*	*29331*	*29562*	*2285*
	(-0.6)	(0.6)	**1.03**	**1.00**	**1.05**	**1.11**	**1.05**	**0.94**	**1.06**	**1.02**
10	684944	*637250*	*281207*	*189532*	*135521*	*72018*	*53244*	*31574*	*31823*	*2460*
	(0)	**1.01**	**1.02**	**1.00**	**1.04**	**1.11**	**1.04**	**0.94**	**1.05**	**1.01**

Table 7.1: chain-ladder analysis of motor data. Roman figures are data. Figures in parentheses are standardised residuals. Italic figures are forecasts from a median forecast from a geometric chain-ladder. Bold figures is the ratio of an arithmetic chain-ladder and the median forecast from a geometric chain-ladder.

Table 7.1 reports the motor data triangle shown with roman figures. The table also reports, in parantheses, standardised residuals of the form $(y_{ij} - \widehat{\mu}_{ij})/s$, where s^2 is the degrees of freedom corrected residual variance. These residuals can be calculated either using the standard least squares formula or from new explicit formulas of the individual estimators given in Theorem 3.1. A detailed discussion of the estimates follows in §7.2.

A consequence of Theorem 3.1 is that the corners, Y_{k1} and Y_{1k}, of the run-off triangle have perfect fit, just as for the usual geometric chain-ladder. The corresponding residuals are therefore equal to zero. Indeed, for the geometric chain-ladder the formula (4.24) gives the prediction

$$\tilde{Y}_{k1} = \exp(\tilde{y}_{k1}) = \exp(\widehat{\mu}_{11} + \sum_{i=2}^{k} \Delta\widehat{\alpha}_i) = \exp(r_k) = Y_{k1}.$$

For the corresponding formula for the arithmetic chain-ladder, see for instance Kuang, Nielsen and Nielsen (2009, equation 15).

A formal cumulant based test for normality does not indicate any deviation from normality. The residuals in Table 7.1 do, however, indicate one potential model mis-specification in that all the large standardised residuals, those larger than 1 in absolute value, say, are to be found for the first 5 accident years. At the same time the largest outstanding liabilities are likely to be found for the recent accident years. If it is the case that the variance drops then forecasts confidence bands are therefore likely to be too wide. This effect could be investigated by the approach of Hertig (1985), but as this would leave the presented analytic framework we will not pursue this here.

7.2 Estimates and predictions

	Arithmetic CL				Geometric CL			
j	$\widehat{\Delta\alpha}$	$\widehat{\Delta\beta}$	F	G	$\widehat{\Delta\alpha}$	$\widehat{\Delta\beta}$	f	g
2	-0.03	-0.07	1.937	1.97	-0.10	-0.07	-0.04	-0.051
3	0.19	-0.80	1.217	1.60	0.08	-0.82	-0.28	0.010
4	0.12	-0.42	1.117	1.42	0.38	-0.39	-0.24	0.100
5	-0.10	-0.29	1.078	1.27	-0.09	-0.34	-0.21	0.042
6	0.04	-0.57	1.041	1.22	-0.05	-0.63	-0.25	0.019
7	-0.17	-0.36	1.027	1.15	-0.21	-0.30	-0.22	-0.016
8	0.21	-0.63	1.014	1.16	0.27	-0.52	-0.23	0.022
9	0.04	0.12	1.016	1.15	-0.00	0.01	-0.18	0.017
10	0.07	-2.60	1.001	1.14	0.07	-2.56	-0.40	0.021
	$\widehat{\mu}_{11} = 13.071$				$\widehat{\mu}_{11} = 13.085$			
					$\widehat{\sigma}^2 = 0.049$, $s^2 = 0.075$			

Table 7.2: Chain-ladder analysis of motor data. Columns 2-5 use a standard arithmetic chain-ladder. Columns 6-9 use a geometric chain-ladder.

Table 7.2 presents the estimates for the geometric chain-ladder and for the standard arithmetic chain-ladder. These are computed using the new Theorem 3.1 and the corresponding Theorem 3 of Kuang, Nielsen and Nielsen (2009). The estimates for $\Delta\alpha$ and $\Delta\beta$ could of course also be computed using a (generalized) linear model regression package, with a design matrix constructed from formula (3.14).

The estimates for $\Delta\alpha$ and $\Delta\beta$ are very similar for the arithmetic and the geometric chain-ladder. The arithmetic and geometric nature of the calculations of for instance the $\widehat{\Delta\alpha_i}$s as seen from (2.6) and (3.18) do, however, give some minor differences, that in turn translate into different predictions. There does not seem to be a general rule for ordering of the estimators coming from the formulas in (2.6) and (3.18).

Table 7.1 also reports predictions from the geometric and the arithmetic chain-ladders. Figures in italic are median predictions from the geometric chain-ladder, $\tilde{Y}_{ij} = \exp(\tilde{\mu}_{ij})$ say. Figures in bold are ratios of predictions from the geometric and the arithmetic chain-ladder, $\tilde{Y}^a_{ij}/\tilde{Y}_{ij} = \exp(\tilde{\mu}^a_{ij} - \tilde{\mu}_{ij})$ say. Thus, these ratios are larger than one when the arithmetic chain-ladder has the larger forecast. For instance, for rows 2, 3 these ratios are larger than one. This relates to differences in the estimates of $\Delta\alpha$ for the two methods as reported in Table 7.2. Indeed partial sums of the type $\widehat{\alpha}_i - \widehat{\alpha}_1 = \sum_{\ell=2}^{i} \Delta\alpha_\ell$ are larger for the arithmetic chain-ladder for $i = 2, 3$. For rows 4, 5 the situation is the other way around.

Table 7.2 also gives information on the development factors. Actuaries are used to manipulating the development factors F_ℓ for the arithmetic chain-ladder. It is less obvious how the arithmetic development factors f_ℓ should be manipulated. Their interpretation arises through (3.23) showing that, for instance, f_j estimates a weighted average of $\Delta\beta_1, \ldots, \Delta\beta_j$. However, the development factors for the two models have one thing in common, which is that they are the appropriate corrections to differences of row sums, such as $r_i - r_{i-1}$, through formulas like (3.18).

The final piece of information in Table 7.2 is the variance estimate σ^2 for the geometric chain-ladder. Here, $\widehat{\sigma}^2$ is the maximum likelihood estimator, found by dividing the sum of squared residuals with the number of cells, $n = k(k+1)/2 = 55$, while s^2 is the degrees of freedom corrected estimator, found by dividing by $df = (k-2)(k-1)/2 = 36$. The variance estimate is used if a prediction of the mean of Y_{ij} is preferred to a median prediction. Formula (4.28) shows that this would involve multiplication of the predictions indicated with italic in Table 7.1 by $\exp(\widehat{\sigma}^2/2) \approx 1 + \widehat{\sigma}^2/2$. This means adding 3.8% or 2.5% to the predictions depending on the use of a degrees of freedom correction or not. This will give a corresponding reduction in the relative predictions or the arithmetic and geometric chain-ladder. Note, that the anchoring at the corner points Y_{k1} and Y_{1k} will now disappear.

7.3 Distribution forecasts

In practice it is desirable to supplement the point forecasts presented in Table 7.1 with distribution forecasts. This can be done through a relatively simple simulation exercise.

For a single entry in the lower triangle the distribution forecast can be done analytically due to the log normality underpinning the model. The process error and the estimation error are independent normal and therefore, following, Doray (1996), the forecast error of the log point forecast $\tilde{\mu}_{ij} = X_{ij}\widehat{\xi}$ is

$$y_{ij} - \tilde{\mu}_{ij} \overset{\mathsf{D}}{=} \mathsf{N}[0, \sigma^2\{1 + X'_{ij}(X'X)^{-1}X_{ij}\}],$$

noting that X is the design matrix for the observations in the upper triangle. Inserting the degrees of freedom corrected variance estimator s^2 results in errors that follow a t_{df}-distribution scaled by $\{1+X_{ij}'(X'X)^{-1}X_{ij}\}^{1/2}$. When converting to the original scale by $Y_{ij} = \exp(y_{ij})$ the analytical simplicity is lost, although explicit formulas can be found as discussed by Verrall (1991) and Doray (1996).

For a distribution forecast for the overall reserve, possibly broken down by calendar year, we propose a simulation study of the following type.

Algorithm 1 *Simulation of reserve.*

1. *draw independent, standard normal process errors $\varepsilon_{p,i,j,r}$ for each i, j;*
2. *draw estimation error $\varepsilon_{\xi,r}$ from a $\mathsf{N}\{0, (X'X)^{-1}\}$ distribution;*
3. *draw squared scale error $\varepsilon^2_{s,r}$ from $n\chi^2_{df}/df$ distribution;*
4. *compute total reserve $R_r = \sum_{i,j} \exp\{X_{ij}'\widehat{\xi} + \widehat{\sigma}\varepsilon_{s,r}(\varepsilon_{\xi,r} + \varepsilon_{p,i,j,r})\}$;*
5. *repeat steps 1–5 for $r = 1, \ldots, Rep$ and get the empirical distribution.*

Table 7.3 shows the outcome from simulating the forecast distribution. The results are broken down by accident year, development year, and calendar year by varying the summation in item 4 of the algorithm. Columns 5-7 show the pure process error, setting the estimation error to zero in item 2 and the squared scale error to unity in item 3 of the algorithm. Columns 8-10 show the process error combined with the estimation error, while the squared scale error is set to unity. Inclusion of the squared scale error only had a minor effect. In both cases the maximum likelihood estimator for the variance is used.

For reference, it would be useful to present distribution forecasts for the arithmetic chain-ladder. It is not so clear how to do this because of the underlying Poisson model is over-dispersed. As a compromise we use the bootstrap method of England (2002). columns 2-4 show the outcome from the implementation in R Development Core Team (2006) package Chain Ladder 0.1.5-0 by Gesmann, Zhang, and Murphy (2011). This method has a tendency to generate negative reserves at the longest developments even though the underlying model does not permit this.

When comparing the arithmetic based bootstrap values with the geometric based simulations it is useful to recall that the Poisson distribution has constant variance-mean ratio whereas the log normal distribution has constant standard deviation-mean ratio, see (2.8) and (2.9). Accordingly, we observe a tendency to smaller variance for the bootstrap values than for the simulated log normal values for sums that include the lower left corner of the lower triangle and vice versa for the right hand part of the lower triangle.

For this triangle an additional variance issue is observed. In §7.1 it was suggested that the variance is possibly lower for the second half of accident years than for the first half of accident years. If this could be confirmed by the substantial context it would seem sensible to shrink the forecast distribution towards the pure process error. Indeed, Algorithm 1 could easily be adapted by adjusting the value of $\widehat{\sigma}$. An alternative solution would be to alter the model so as to allow different variances for the two blocks of accident years along the lines of Hertig (1985). The resulting likelihood could still be estimated numerically.

Year	Bootstrap			LogN: process			LogN: full		
	1%	50%	99%	1%	50%	99%	1%	50%	99%
				Accident years					
2	-17	0.38	31	0.9	1.6	2.6	0.7	1.6	3.4
3	-6.3	25	103	15	24	38	13	24	44
4	5.5	56	157	47	67	96	42	68	110
5	28	97	219	82	111	151	73	112	171
6	73	169	318	129	169	223	114	170	255
7	125	245	419	182	236	310	159	237	360
8	290	471	719	381	488	637	324	491	752
9	508	756	1091	594	754	970	472	757	1212
10	996	1440	2145	1132	1456	1926	797	1463	2743
				Development years					
2	398	631	977	380	636	1066	296	636	1386
3	349	550	819	380	549	795	325	554	947
4	328	529	795	406	551	743	355	555	875
5	297	496	772	381	493	639	332	498	756
6	170	345	596	258	326	412	224	329	487
7	112	284	547	235	291	360	199	294	432
8	24	174	431	168	205	250	139	207	309
9	9.1	221	600	190	229	276	147	231	363
10	-91	7.7	193	16	19	23	11	19	35
				Calendar years					
11	1023	1347	1763	1040	1355	1818	951	1367	2143
12	524	750	1040	591	756	985	544	764	1134
13	307	484	718	377	487	641	344	493	749
14	176	314	501	237	309	412	212	313	489
15	81	180	330	135	179	238	120	181	284
16	35	110	239	85	116	158	74	117	192
17	6.2	59	163	45	64	92	38	65	115
18	-5.3	31	119	21	34	56	16	34	77
19	-23	0.011	40	1.5	2.5	4.1	1.0	2.5	6.2
				Total Reserve					
	2563	3303	4239	2916	3324	3848	2514	3362	4785

Table 7.3: Distribution forecasts by accident year, development year, calendar year and for full reserve: Column 2-4 use the Bootstrap method of England (2001), see also England and Verrall (2002), as implemented by Gesmann (2009) using a seed of unity and 39999 bootstrap replications; Column 5-10 use simulation of log Normal distribution based on 39999 replications. Column 5-7 have process error only; Column 8-10 have process error combined with estimation error for mean.

8 Conclusion

The geometric chain-ladder is often used in practice, although not as commonly as the arithmetic chain-ladder. Both methods can be implemented as generalized linear models. However, one of the attractions of the arithmetic chain-ladder is the analytic expressions for estimators and forecasts in terms of development factors. The main contribution of this paper is to present similar analytic expressions for the geometric chain-ladder. The results and the derivation resemble the contribution of Kuang, Nielsen and Nielsen (2009) for the standard arithmetic case. Therefore, the exact difference of the poisson model and the lognormal model are now fully transparent. Additionally, some analytic expressions were derived for the distribution of the estimators and it was shown how inclusion of exposure affects the geometric chain-ladder.

The point forecasts are broadly similar for the two chain-ladder models although minor difference appears in the way they weigh the information from different accident years. When it comes to distribution forecasting there are some methodological differences. The arithmetic chain-ladder is essentially based on a Poisson distribution. In practice, one often finds a considerable over-dispersion. As a consequence one does not really have a model for the variation in the data. This means that it is hard to test the distributional validity of the model and distribution forecasting is made on shaky ground. In contrast, the log normal model underlying the geometric chain-ladder is much more reasonable than a Poisson distribution, so one can directly test the distributional validity and make distribution forecasts. One consequence is that it is easier to think about improving the geometric chain-ladder if the residuals suggest a systematic deviation from the model. One such example was briefly discussed in the empirical illustration. Thus, the geometric chain-ladder may have an edge from a viewpoint of distribution forecasting.

The substantive interpretation of the geometric chain-ladder is less clear. The reserving triangle is formed by aggregation over a large number of policies. While the aggregated distribution may be well approximated by a log normal distribution it is not obvious how a log normal distribution could arise through aggregation. An over-dispersed Poisson distribution can, however, be motivated in terms of a compound Poisson aggregation mechanism. When faced with additional data, for instance from a triangle of counts of reported claims, it may be easier to generalise a micro foundation of this type, see for instance Martínez-Miranda, Nielsen, Nielsen and Verrall (2011), Martínez-Miranda, Nielsen and Verrall (2012), Martínez-Miranda, Nielsen and Wüthrich (2012).

Acknowledgements

The second author gratefully acknowledges financial support from the Institute of Economic Modelling. The paper was presented at the International Cramér Symposium on Insurance Mathematics in Stockholm in June 2013 with discussion by Mario Wüthrich.

A Appendix: Proofs of Theorems

Proof of Theorem 3.1. It is convenient to introduce row and column sums

$$\breve{r}_i = \sum_{j=1}^{k+1-i} y_{ij} = (k+1-i)r_i, \qquad \breve{c}_j = \sum_{i=1}^{k+1-j} y_{ij} = (k+1-j)c_j. \tag{A.34}$$

The likelihood (3.15) can then be rewritten as

$$\log \mathsf{L}(\xi, \sigma^2; y) = -\frac{k(k+1)}{4} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i,j \in \mathcal{I}} \mu_{ij}^2$$
$$- \frac{1}{2\sigma^2} \sum_{i,j \in \mathcal{I}} y_{ij}^2 + \frac{\mu_{11}}{\sigma^2} \sum_{i,j \in \mathcal{I}} y_{ij} + \frac{1}{\sigma^2} \sum_{i=2}^{k} (\alpha_i - \alpha_1) \breve{r}_i + \frac{1}{\sigma^2} \sum_{j=2}^{k} (\beta_j - \beta_1) \breve{c}_j.$$

Thus, the model is a full exponential family with minimal sufficient statistic

$$T = (\sum_{i,j \in \mathcal{I}} y_{ij}^2, \sum_{i,j \in \mathcal{I}} y_{ij}, \breve{r}_2, \ldots, \breve{r}_k, \breve{c}_2, \ldots, \breve{c}_k).$$

For a full exponential family the maximum likelihood estimator is unique if and only if the natural statistic is interior to its convex support (Barndorff-Nielsen, 1978, Theorem 9.13). For a regression model this happens with probability one when $\dim(y) = k(k+1)/2$ is larger than $\dim(\xi) = 2k - 1$, that is for $k \geq 3$.

Since the exponential family is regular the $2k$ likelihood equations are $T = \mathsf{E}T$ (Barndorff-Nielsen, 1978, Corollary 9.6). Since $\sum_{i,j \in \mathcal{I}} y_{ij} = \sum_{i=1}^{k} \breve{r}_i = \sum_{j=1}^{k} \breve{c}_j$ this in turn implies the equations

$$\breve{r}_i = \mathsf{E}\breve{r}_i, \qquad \breve{c}_j = \mathsf{E}\breve{c}_j, \qquad \text{for } i, j = 1, \ldots k. \tag{A.35}$$

Estimating the development parameters. The log development factor f_j as defined in (2.10) can be written in terms of the minimal sufficient statistics $\breve{r}_i$, $\breve{c}_j$ as follows. The numerator of, for instance, the first term in f_j is the sum over the first j columns of the first $k+1-j$ rows of $y_{i\ell} = \log Y_{i\ell}$. This is the same as the sum over the first $k+1-j$ rows of $y_{i\ell}$ minus the sum over the last columns sums indexed $j+1, \ldots, k$, that is, using the definitions in (A.34),

$$\sum_{i=1}^{k+1-j} \sum_{\ell=1}^{j} y_{i\ell} = \sum_{i=1}^{k+1-j} \sum_{\ell=1}^{k+1-i} y_{i\ell} - \sum_{\ell=j+1}^{k} \sum_{i=1}^{k+1-\ell} y_{i\ell} = \sum_{i=1}^{k+1-j} \breve{r}_i - \sum_{\ell=j+1}^{k} \breve{c}_j.$$

A similar manipulation of the numerator of the second term in f_j gives

$$f_j = \frac{\sum_{i=1}^{k+1-j} \breve{r}_i - \sum_{\ell=j+1}^{k} \breve{c}_\ell}{j(k+1-j)} - \frac{\sum_{i=1}^{k+1-j} \breve{r}_i - \sum_{\ell=j}^{k} \breve{c}_\ell}{(j-1)(k+1-j)}.$$

The likelihood equations (A.35) therefore imply that f_j is maximum likelihood estimator for the parameter

$$\varphi_j = \frac{\sum_{i=1}^{k+1-j} \mathsf{E}\breve{r}_i - \sum_{\ell=j+1}^{k} \mathsf{E}\breve{c}_\ell}{j(k+1-j)} - \frac{\sum_{i=1}^{k+1-j} \mathsf{E}\breve{r}_i - \sum_{\ell=j}^{k} \mathsf{E}\breve{c}_\ell}{(j-1)(k+1-j)}$$

A similar manipulation to that done for f_j then shows that

$$\varphi_j = \frac{\sum_{i=1}^{k+1-j} \sum_{\ell=1}^{j} \mathsf{E}y_{i\ell}}{j(k+1-j)} - \frac{\sum_{i=1}^{k+1-j} \sum_{\ell=1}^{j-1} \mathsf{E}y_{i\ell}}{(j-1)(k+1-j)}.$$

Insert that

$$\mathsf{E}y_{i\ell} = \mu_{i\ell} = \mu_{11} + \sum_{m=2}^{i} \Delta\alpha_m + \sum_{m=2}^{\ell} \Delta\beta_m, \tag{A.36}$$

noting that the μ_{11} and $\Delta\alpha_m$ terms cancel so

$$\varphi_j = \frac{1}{j} \sum_{\ell=2}^{j} \sum_{m=2}^{\ell} \Delta\beta_m - \frac{1}{j-1} \sum_{\ell=2}^{j-1} \sum_{m=2}^{\ell} \Delta\beta_m. \tag{A.37}$$

The expressions for g_i and γ_i arise similarly.

Estimating the difference parameters. The argument proceeds in a similar fashion as above. Due to equation (A.34) then Δr_i is a function of the minimal sufficient statistic. Indeed $\Delta r_i = r_i - r_{i-1} = \breve{r}_i/(k+1-i) - \breve{r}_{i-1}/(k+2-i)$. Thus the equations (A.35) imply that $\Delta r_i = \Delta \mathsf{E} r_i$. Note that

$$\Delta \mathsf{E} r_i = \frac{1}{k+1-i} \sum_{\ell=1}^{k+1-i} \mathsf{E} y_{i\ell} - \frac{1}{k+2-i} \sum_{\ell=1}^{k+2-i} \mathsf{E} y_{i\ell},$$

and insert the expression for $\mathsf{E}y_{i\ell}$ in (A.36) to get

$$\begin{aligned} \Delta \mathsf{E} r_i &= \frac{1}{k+1-i} \sum_{j=1}^{k+1-i} (\mu_{11} + \sum_{\ell=2}^{i} \Delta\alpha_\ell + \sum_{\ell=2}^{j} \Delta\beta_\ell) \\ &\quad - \frac{1}{k+2-i} \sum_{j=1}^{k+2-i} (\mu_{11} + \sum_{\ell=2}^{i-1} \Delta\alpha_\ell + \sum_{\ell=2}^{j} \Delta\beta_\ell). \end{aligned}$$

The terms involving μ_{11}, $\Delta\alpha_2, \ldots, \Delta\alpha_{i-1}$ cancel so this reduces to

$$\Delta \mathsf{E} r_i = \Delta\alpha_i + \frac{1}{k+1-i} \sum_{j=2}^{k+1-i} \sum_{\ell=2}^{j} \Delta\beta_\ell - \frac{1}{k+2-i} \sum_{j=2}^{k+2-i} \sum_{\ell=2}^{j} \Delta\beta_\ell.$$

Using the expression in (A.37) this reduces to

$$\Delta \mathsf{E} r_i = \Delta\alpha_i - \varphi_{k+2-i}.$$

Equating $\mathsf{E}r_i = r_i$ and $\varphi_{k+2-i} = f_{k+2-i}$ implies the desired expression for the maximum likelihood estimator for $\Delta\alpha_i$. The $\Delta\beta_j$ terms are dealt with similarly.

Estimating the level. The first expression for $\hat{\mu}_{11}$ arises from the likelihood equation for $r_k = y_{k1}$. This is

$$r_k = \mathsf{E} r_k = \mu_{11} + \sum_{\ell=2}^{k} \Delta\alpha_\ell.$$

Insert, the estimators for $\Delta\alpha_\ell$ to get the desired expression. The second expression arises in a similar way from the likelihood equation $c_k = \mathsf{E} c_k$.

Proof of Theorem 3.2. The two expressions are proved in the same way. Consider therefore the second expression linking βs and φs. Recall the expression for φ_j in (3.23). Multiply by j and replace $\ell - 1$ by $\sum_{m=2}^{\ell} 1$ to get

$$j\varphi_j = \frac{1}{j-1} \sum_{\ell=2}^{j} \Delta\beta_\ell \sum_{m=2}^{\ell} 1.$$

Interchange sums to get

$$\begin{aligned} j\varphi_j = \frac{1}{j-1} \sum_{m=2}^{j} \sum_{\ell=m}^{j} \Delta\beta_\ell &= \frac{1}{j-1} \sum_{m=2}^{j} \{(\beta_j - \beta_1) - (\beta_{m-1} - \beta_1)\}. \\ &= (\beta_j - \beta_1) - \frac{1}{j-1} \sum_{m=2}^{j-1} (\beta_m - \beta_1). \end{aligned}$$

Formulated in terms of matrices:

$$\begin{pmatrix} 2 & 0 & \cdots & 0 \\ 0 & 3 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & k \end{pmatrix} \begin{pmatrix} \varphi_2 \\ \varphi_3 \\ \vdots \\ \varphi_k \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ -\frac{1}{2} & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ -\frac{1}{k-1} & \cdots & -\frac{1}{k-1} & 1 \end{pmatrix} \begin{pmatrix} \beta_2 - \beta_1 \\ \beta_3 - \beta_1 \\ \vdots \\ \beta_k - \beta_1 \end{pmatrix}$$

Inverting the matrix on the right hand side gives

$$\begin{pmatrix} \beta_2 - \beta_1 \\ \beta_3 - \beta_1 \\ \vdots \\ \beta_k - \beta_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ \frac{1}{2} & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ \frac{1}{2} & \cdots & \frac{1}{k-1} & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 & \cdots & 0 \\ 0 & 3 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & k \end{pmatrix} \begin{pmatrix} \varphi_2 \\ \varphi_3 \\ \vdots \\ \varphi_k \end{pmatrix}.$$

Multiplying the two matrices on the right gives the desired result.

Proof of Theorem 5.3. *Derivation of* $\mathsf{var}(\widehat{\mu}_{11})$. The variance is given in terms of the top left element of $(X'X)^{-1}$, that is the element $e_1'(X'X)^{-1}e_1$ where $e_1 = (1, 0)'$ is a unit vector of length $2k - 1$. Now, the unit vector e_1 can also be written as $e_1 = X'\check{\mathbf{y}}_{11}$ where $\check{\mathbf{y}}_{11} = (1, 0)'$ is a unit vector of length $k(k + 1)/2$. Therefore it holds that

$$e_1'(X'X)^{-1}e_1 = e_1'(X'X)^{-1}X'\check{\mathbf{y}}_{11}.$$

Here $(X'X)^{-1}X'\check{\mathbf{y}}_{11}$ is the least squares estimator from regressing the auxiliary data vector $\check{\mathbf{y}}_{11}$ on X. Closed form expressions for the elements of this least squares estimator were found in Theorem 3.1. In particular, inserting the auxiliary data $\check{\mathbf{y}}_{11}$ instead of the original data $\mathbf{y}$ gives development factors

$$f_j = \frac{1}{j(k + 1 - j)} - \frac{1}{(j - 1)(k + 1 - j)},$$

and hence, through the expression for $\widehat{\mu}_{11}$, that

$$\begin{aligned} e_1'(X'X)^{-1}e_1 &= e_1'(X'X)^{-1}X'\check{\mathbf{y}}_{11} \\ &= \frac{1}{k} - \sum_{\ell=2}^{k}\left\{\frac{1}{\ell(k+1-\ell)} - \frac{1}{(\ell-1)(k+1-\ell)}\right\}. \end{aligned}$$

To simplify the sum note that for instance

$$\frac{1}{\ell(k+1-\ell)} = \frac{1}{k+1}\left(\frac{1}{\ell} + \frac{1}{k+1-\ell}\right). \tag{A.38}$$

Introducing the function $C_k = \sum_{\ell=1}^{k}(1/\ell)$ then

$$e_1'(X'X)^{-1}e_1 = \frac{1}{k} - \frac{1}{k+1}(2C_k - 1 - \frac{1}{k}) + \frac{2}{k}(C_k - \frac{1}{k}).$$

This in turn reduces to $2(k-1)/k^2 + 2C_k/\{k(k+1)\}$ as desired.

Derivation of $\mathsf{var}(\Delta\widehat{\alpha}_i)$. The variance is given in terms of the ith diagonal element of $(X'X)^{-1}$. Let $e_i = (0_{i-1}, 1, 0_{2k-i-1})'$ be the ith unit vector of length $2k-1$. Then $e_i = X_{i1} - X_{i-1,1} = X'(\check{\mathbf{y}}_{i1} - \check{\mathbf{y}}_{i-1,1})$, where $\check{\mathbf{y}}_{ij}$ takes the value unity for element (i,j) and zero otherwise. Thus, evaluate the expression for $\Delta\widehat{\alpha}_i$ in Theorem 3.1 at the data point $\check{\mathbf{y}}_{i1} - \check{\mathbf{y}}_{i-1,1}$. Now, $\Delta r_i = 1/(k+1-i)$ and $f_{k+2-i} = 0$ when evaluated at $\check{\mathbf{y}}_{i1}$, whereas $\Delta r_i = -1/(k+2-i)$ and $f_{k+2-i} = -1/\{(k+2-i)(i-1)\} - (-1)/\{(k+1-i)(i-1)\}$ when evaluated at $\check{\mathbf{y}}_{i-1,1}$. In combination it holds

$$\begin{aligned} e_i'(X'X)^{-1}e_i = \frac{1}{k+1-i} - \frac{-1}{k+2-i} \\ - \frac{-1}{(k+2-i)(i-1)} - \frac{1}{(k+1-i)(i-1)}, \end{aligned} \tag{A.39}$$

which reduces to the desired expression.

Derivation of $\mathsf{var}(f_j)$. Use the expression (2.10). This shows that f_j is the differences of two averages of terms $\log Y_{i\ell}$ that are independent normal with variance σ^2. The variances of the two terms are therefore σ^2 divided by the number of elements. To find the covariance write the numerator of the first term as the numerator of the second term plus a component that is independent of the second term. In combination it follows that

$$\mathsf{var}(f_j) = \frac{\sigma^2}{(k+1-j)j} + \frac{\sigma^2}{(k+1-j)(j-1)} - \frac{2\sigma^2}{(k+1-j)j},$$

which reduces to the desired expression.

Derivation of $\mathsf{var}(\widehat{\Delta\beta}_i)$ *and* $\mathsf{var}(g_j)$. Similar arguments as for $\mathsf{var}(\widehat{\Delta\alpha}_i)$ and $\mathsf{var}(f_j)$.

Proof of Theorem 6.4. The geometric chain-ladder is estimated by least squares regression of the logged data $\mathbf{y}$ on the design matrix X giving the estimator $\widehat{\xi}$. The data corrected for log exposure, $h_i = \log H_i$, is given by $\mathbf{y}^\dagger = \mathbf{y} - X\xi^\dagger$ where $\xi^\dagger = (h_1, \Delta h_2, \ldots, \Delta h_k, 0, \ldots, 0)'$. Due to the linear nature of the least squares estimator the regression of $\mathbf{y}^\dagger$ on X gives $\widehat{\xi} - \xi^\dagger$. Due to the formula (3.12) the forecast of y_{ij} based on the exposure corrected data is $X'_{ij}(\widehat{\xi}-\xi^\dagger) = X'_{ij}\widehat{\xi} - h_i = \widehat{\mu}_{ij} - h_i$. When adding h_i this gives the same forecast as for the original data.

For the variance estimator note $\mathbf{y} = \mathbf{y}^\dagger + X\xi^\dagger$ so $\mathbf{y}$ and $\mathbf{y}^\dagger$ yield the same residuals when regressing on X. The variance estimators are then identical.

Proof of $\mathsf{var}(\Delta\widehat{\alpha}_\ell)$ is increasing in $\ell = 2, \ldots, k$ for each k. Let $\sigma^2 = 1$ without loss of generality. It has to be argued that

$$d_\ell = \mathsf{var}(\Delta\widehat{\alpha}_{\ell+1}) - \mathsf{var}(\Delta\widehat{\alpha}_\ell)$$

is positive for $\ell = 2, \ldots, k-1$. Insert the expression for the variances to get

$$d_\ell = \frac{1}{k-\ell} + \frac{1}{k+1-\ell} + \frac{1}{(k-\ell)(k+1-\ell)\ell} - \frac{1}{k+1-\ell} - \frac{1}{k+2-\ell} - \frac{1}{(k+1-\ell)(k+2-\ell)(\ell-1)}$$

Note that the $1/(k+1-\ell)$-terms cancel. Multiply d_ℓ by the positive term $(k-\ell)(k+1-\ell)(k+2-\ell)$ to get

$$\begin{aligned} d_{\ell,k} &= (k-\ell)(k+1-\ell)(k+2-\ell)d_\ell \\ &= (k+2-\ell)\{(k+1-\ell)+\frac{1}{\ell}\} - (k-\ell)\{(k+1-\ell)+\frac{1}{\ell-1}\}. \end{aligned}$$

This expression reduces to

$$\begin{aligned} d_{\ell,k} &= 2(k+1-\ell) + \frac{(k+2-\ell)(\ell-1)-(k-\ell)\ell}{\ell(\ell-1)} \\ &= (k+1-\ell)\{2 - \frac{1}{\ell(\ell-1)}\} + \frac{2\ell-1}{\ell(\ell-1)}, \end{aligned}$$

which is positive for $\ell = 2, \ldots, k-1$.

References

Barndorff-Nielsen, O.E. (1978) *Information and Exponential Families*. New York: Wiley.

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

England, P.D. (2002) Addendum to "Analytic and bootstrap estimates of prediction errors in claims reserving". *Insurance: Mathematics and Economics* 31, 461–466.

England, P.D. & Verrall, R.J. (2002) Stochastic claims reserving in general insurance. *British Actuarial Journal* 8, 519–44.

Gesmann, M., Zhang, W. and Murphy, D. (2011) The R-package Chain Ladder version 0.1.5-0.

Hendry, D.F. & Nielsen, B. (2007) *Econometric Modeling* Princeton, NJ: Princeton University Press.

Hertig, J. (1985) A statistical approach to IBNR-reserves in marine reinsurance. *ASTIN Bulletin* 15, 171–183.

Kremer, E. (1982) IBNR-Claims and the Two-way model of ANOVA. *Scandinavian Actuarial Journal*, 47–55.

Kremer, E. (1985) *Einführung in die Versicherungsmathematik*. Göttingen: Vandenhoek & Ruprecht.

Kuang, D., Nielsen, B. & Nielsen, J.P. (2008a) Identification of the age-period-cohort model and the extended chain-ladder model. *Biometrika* 95, 979–986.

Kuang, D., Nielsen, B. & Nielsen, J.P. (2008b) Forecasting with the age-period-cohort model and the extended chain-ladder model. *Biometrika* 95, 987–991.

Kuang, D., Nielsen, B. & Nielsen, J.P. (2009) Chain-ladder as maximum likelihood revisited. *Annals of Actuarial Science* 4, 105–121.

Kuang, D., Nielsen, B. & Nielsen, J.P. (2011) Forecasting in an extended chain-ladder-type model. *Journal of Risk and Insurance* 78, 345–359.

Mack, T. (1991) A simple parametric model for rating automobile insurance or estimating IBNR claims reserves. *ASTIN Bulletin* 21, 93–109.

Mack, T. (1994) Which stochastic model is underlying the chain-ladder method? *Insurance: Mathematics and Economics* 15, 133–138.

Mack, T. and Venter, G. (2000) A comparison of stochastic models that reproduce chain-ladder reserve estimates. *Insurance: Mathematics and Economics* 26, 101–107.

Martínez-Miranda, M.D., Nielsen, B., Nielsen, J.P. and Verrall, R. (2011) Cash flow simulation for a model of outstanding liabilities based on claim amounts and claim numbers. *ASTIN Bulletin* 41, 107–129.

Martínez-Miranda, M.D., Nielsen, J.P. and Verrall, R. (2012) Double chain-ladder *Astin Bulletin* 42, 59–76.

Martínez-Miranda M.D., Nielsen, J.P. and Wúthrich, M. (2012) Statistical modelling and forecasting of outstanding liabilities in non-life insurance. *Statistics and Operations Research Transactions* 36, 195–218.

R Development Core Team (2006) *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Rehmann, Z. and Klugman, S. (2010). Quantifying uncertainty in reserve estimates. *Variance* 4, 30–46.

Renshaw, A.E. (1989) Chain ladder and interactive modeling. *Journal of the Institute of Actuaries* 116, 559–587.

Taylor, G.C.(2003). Chain ladder bias. *ASTIN Bulletin* 33, 313–330.

Taylor. G. (2011). Maximum likelihood and estimation efficiency of the chain-ladder. *ASTIN Bulletin* 41, 131 - 155.

Taylor, G.C. and Ashe, F.R. (1983) Second moments of estimates of outstanding claims *Journal of Econometrics* 23, 37–61.

Verrall, R.J. (1991) On the estimation of reserves from log-linear models. *Insurance: Mathematics and Economics* 10, 75–80.

Verrall, R.J. (1994) Statistical methods for the chain-ladder technique. *Casualty Actuarial Society Forum, Spring 1994*, 393–446.

Verrall, R.J. & England, P.D. (2000) Comments on "A comparison of stochastic models that reproduce chain-ladder reserve estimates", by Mack and Venter. *Insurance: Mathematics and Economics* 26, 109–111.

Wright, T.S. (1990) A stochastic method for claims reserving in general insurance. *Journal of the Institute of Actuaries* 117, 677–731.

Zehnwirth, B. (1994). Probabilistic development factor models with applications to loss reserve variability, prediction intervals, and risk based capital. *Casualty Actuarial Society Forum* Spring 1994, 447–605.