Modeling multiple runoff tables

E.A. Cator *

V. Lous †

April 15, 2013

1 Introduction

Insurance companies exist by virtue of diversification of risks. This diversification can happen within lines of business but also across different lines of business. In recent years there has been an increase of interest in the latter. This is beneficial for insurance companies as it will lead to more accurate predictions, and consequently more accurate determination of risk margins and capital requirements.

In this paper we propose a method to model two or more possibly dependent runoff tables at the same time. A logical way to do this is by making a large model incorporating all tables, and fit the corresponding parameters (of which there will be a lot) at the same time using maximum likelihood or a Bayes method. For example [1, 3, 4] take this approach. A downside of this approach is that the number of parameters might get out of hand. This would lead to computational issues and a large statistical error.

We propose to model each table separately, estimating its own parameters, and then model the dependence between two tables by a single correlation parameter. This parameter can be estimated using the parameter estimates for the individual tables. This would mean that the marginal predictions for each table would not change, but the joint prediction of the two tables would depend on this correlation parameter. In this bottom up approach the reserving actuary can apply its knowledge regarding different runoff tables at a deep level of granularity. In contrast to trying to translate knowledge of all runoff tables into a single model.

Another advantage of our method is that it does not require an explicit description of the nature of the dependence between two tables, which is the approach taken in [1, 3]. Instead, the dependencies between the cells of a single table, also hold between cells of different tables. For example, when cells of a table belonging to a particular loss period are correlated they will also be correlated to the cells of a different table belonging to that loss period.

*Eric Cator is a professor of Mathematical Statistics and Probability at the Radboud University, Nijmegen, the Netherlands. He is also a partner of the consultancy firm Cator & van Zwet, Statistisch Adviseurs.

†Vincent Lous is actuarial consultant and statistical programmer at Posthuma Partners. Lous@posthuma-partners.nl

This paper is organized as follows. In section 2 the model for multiple runoff tables is explained. It starts with the key idea of this article on how to create a full covariance matrix of a portfolio. Then a model for a single runoff table is presented. Finally, estimation issues are tackled. The model for multiple runoff tables, and for a single runoff table are implemented in the loss reserving software IFM.

2 Modeling multiple runoff tables

2.1 Covariance matrix of a portfolio

Let $X_r \in \mathbb{R}^n$ be the cells of the r-th runoff table in a portfolio with $r = 1, \dots, R$ and $n \in \mathbb{N}$. Assume that the different runoff tables have the same loss and development periods, and those periods are of equal length (we could always enforce this condition by aggregating or splitting up cells where necessary). Each cell contains the loss (paid or incurred) over a certain loss period in a certain booking period, so we do not consider cumulative run-off tables. Furthermore, let $X = (X_1, \dots, X_R) \in \mathbb{R}^{Rn}$ be the vector combining the runoff tables of a portfolio. We consider a normal model for the incremental losses:

$$X \sim \mathcal{N}(\mu, \Sigma) \tag{2.1}$$

The diagonal blocks of Σ are the covariance matrices of the individual runoff tables ($\Sigma_1 \dots \Sigma_R$). The "off-diagonal blocks", the object of interest here, are the covariance matrices between two runoff tables r and s and are denoted $\Sigma_{r,s}$. For each Σ_r there exists a unique symmetric $n \times n$-matrix B_r such that

$$\Sigma_r = B_r^2 \tag{2.2}$$

B_r can be found via eigen decomposition as Σ_r has positive eigenvalues. For any $R \times R$ correlation matrix Γ (i.e., a symmetric positive definite matrix with all 1's one the diagonal), given by

$$\Gamma = \begin{pmatrix} 1 & \rho_{1,2} & \cdots & \rho_{1,R} \\ \rho_{2,1} & 1 & \cdots & \rho_{2,R} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{R,1} & \rho_{R,2} & \cdots & 1 \end{pmatrix}, \tag{2.3}$$

we now define $\Sigma_{r,s}$ as

$$\Sigma_{r,s} = \rho_{r,s} B_r B_s \tag{2.4}$$

The condition on Γ implies that $\rho_{r,s} \in [-1, 1]$, and $\rho_{r,s} = \rho_{s,r}$.

Claim 1. *Any valid choice of Γ will lead to a valid covariance matrix Σ.*

Proof: Since Γ is itself a positive definite matrix, there exists vectors $\gamma_1, \ldots, \gamma_R \in \mathbb{R}^R$ such that

$$\Gamma = \begin{pmatrix} \gamma_1^\top \\ \vdots \\ \gamma_R^\top \end{pmatrix} \begin{pmatrix} \gamma_1 & \cdots & \gamma_R \end{pmatrix} = \begin{pmatrix} \gamma_1^\top \gamma_1 & \gamma_1^\top \gamma_2 & \cdots & \gamma_1^\top \gamma_R \\ \gamma_2^\top \gamma_1 & \gamma_2^\top \gamma_2 & \cdots & \gamma_2^\top \gamma_R \\ \vdots & & \ddots & \vdots \\ \gamma_R^\top \gamma_1 & \cdots & & \gamma_R^\top \gamma_R \end{pmatrix}.$$

Note that since Γ has 1's on the diagonal, $\|\gamma_r\| = 1$ for all $1 \leq r \leq R$. Furthermore, $\rho_{r,s} = \gamma_r^\top \gamma_s$. Now define the $nR \times nR$ matrix

$$B = \begin{pmatrix} \gamma_1^\top \otimes B_1 \\ \vdots \\ \gamma_R^\top \otimes B_R \end{pmatrix}.$$

We then have

$$\begin{aligned} BB^\top &= \begin{pmatrix} \gamma_1^\top \otimes B_1 \\ \vdots \\ \gamma_R^\top \otimes B_R \end{pmatrix} \begin{pmatrix} \gamma_1 \otimes B_1 & \cdots & \gamma_R \otimes B_R \end{pmatrix} \\ &= \begin{pmatrix} (\gamma_1^\top \gamma_1) B_1^2 & (\gamma_1^\top \gamma_2) B_1 B_2 & \cdots & (\gamma_1^\top \gamma_R) B_1 B_R \\ (\gamma_2^\top \gamma_1) B_2 B_1 & (\gamma_2^\top \gamma_2) B_2^2 & \cdots & (\gamma_2^\top \gamma_R) B_2 B_R \\ \vdots & & \ddots & \vdots \\ (\gamma_R^\top \gamma_1) B_R B_1 & \cdots & & (\gamma_R^\top \gamma_R) B_R^2 \end{pmatrix} \\ &= \Sigma. \end{aligned}$$

This proves that Σ is indeed a symmetric positive definite matrix. □

2.2 Modeling a single runoff table

In this section a model for a single runoff table is proposed. This model is explained in more detail in [2]. However, the model for multiple runoff tables can also be used in combination with different models for individual tables.

We define l and k as loss period and settlement period respectively. Furthermore $X_{lk}^{(1)}$ and $X_{lk}^{(2)}$ are incremental incurred and incrementel paid cells of a runoff table. The following independent auxilliary variables are introduced.

$$Z_{lk}^{(1)} \sim \mathcal{N}\left(\mu_{lk}^{(1)}, V_{lk}^{(1)}\right)$$
$$Z_{lk}^{(2)} \sim \mathcal{N}\left(\mu_{lk}^{(2)}, V_{lk}^{(2)}\right)$$

When the event $R = \left\{\sum_k Z_{lk}^{(1)} = \sum_k Z_{lk}^{(2)} \quad (\forall l)\right\}$ occurs, the sum of changes in case reserves for each loss period equals zero; this corresponds to the fact that eventually, the total incurred loss equals the total paid loss. $X^{(1)}$, and $X^{(2)}$ are assumed to have the following distribution.

$$
\begin{aligned}
X^{(1)} &\sim Z^{(1)} \,|\, R \\
X^{(2)} &\sim Z^{(2)} \,|\, R
\end{aligned}
\tag{2.5}
$$

This is again a multivariate normally distributed, the parameters of which are simple functions of the parameters of $Z^{(1)}$ and $Z^{(2)}$. The data belonging to $X = (X^{(1)}, X^{(2)})$ is given by SX, with S a matrix consisting of zeros and ones that can also be used to aggregate observations. In the following μ_r and Σ_r are the estimated parameters of table r.

2.3 Estimation

The idea is to first estimate the marginal expectation and covariance matrix for all individual run-off tables, and then to use maximum likelihood to determine the correlation matrix Γ, keeping the marginal distributions fixed. There are $R(R-1)/2$ unknown parameters in Γ which we need to estimate. When estimating these parameters we need to work under the nonlinear constraint that Γ is positive definite. To enforce this condition we use the LU decomposition of Γ as in the proof of Claim 2.1.

We construct the vectors γ_r as follows

$$
\begin{aligned}
\gamma_1 &= (1, 0, \ldots, 0) \\
\gamma_2 &= (sin(a_{21}), cos(a_{21}), 0 \ldots, 0) \\
\gamma_3 &= (sin(a_{31})sin(a_{32}), sin(a_{31})cos(a_{32}), cos(a_{31}), 0 \ldots, 0) \\
\gamma_4 &= (sin(a_{41})sin(a_{42})sin(a_{43}), sin(a_{41})sin(a_{42})cos(a_{43}), sin(a_{41})cos(a_{42}), cos(a_{41}), 0 \ldots, 0) \\
&\vdots
\end{aligned}
$$

In this way $\|\gamma_r\| = 1$. Note that this exactly corresponds to $R(R-1)/2$ free parameters to be estimated. Let S be a matrix of zeros and ones such that $W = SX$ gives a vector of observed elements of the portfolio. Expectation and covariance are given by $\mathbb{E}(W) = S\mu$ and $\text{Cov}(W) = S\Sigma S^\top$ respectively. We can now maximize the log-likelihood

$$
-\frac{1}{2}(W - \mathbb{E}W)^\top \text{Cov}(W)^{-1}(W - \mathbb{E}W) - \frac{1}{2}\log(\det(\text{Cov}(W)))
$$

over all choices of $a_{ij} \in [-1, 1]$ $(i > j)$. Standard maximum likelihood theory also provides us with a way of determining the estimation uncertainty of the estimated correlations, simply by determining the second derivative of the likelihood at the maximum. This could then be taken into account when calculating predictions for future payments.

3 Conclusion

In this article we proposed a method for modelling multiple runoff tables. This method entails a bottom up approach where first the individual tables are modeled. In the second step the correlations between tables are modeled, where the parameters of the individual tables are taken as given.

References

[1] P. de Jong. Modeling dependence between loss triangles. *North American Actuarial Journal*, 16(1):74–86, 2012.

[2] B. Posthuma, E. A. Cator, W. Veerkamp, and E. van Zwet. Combined analysis of paid and incurred losses. *Casualty Actuarial Society E-forum Fall 2008*, 2008.

[3] P. Shi, S. Basu, and G. G. Meyers. A bayesian log-normal model for multivariate loss reserving. *North American Actuarial Journal*, 16(1):29–51, 2012.

[4] P. Shi and E. W. Frees. Dependent loss reserving using copulas. *Astin Bulletin*, 41(2):449–486, 2011.