Insurance: Mathematics and Economics 73 (2017) 41–53

Contents lists available at ScienceDirect

Insurance: Mathematics and Economics

journal homepage: www.elsevier.com/locate/ime

Full Bayesian analysis of claims reserving uncertainty

Gareth W. Peters [a], Rodrigo S. Targino [a], Mario V. Wüthrich [b,c,*]

[a] University College London, Department of Statistical Science, UK
[b] ETH Zurich, RiskLab, Department of Mathematics, Switzerland
[c] Swiss Finance Institute, Switzerland

ARTICLE INFO

Article history:
Received May 2016
Received in revised form September 2016
Accepted 30 December 2016
Available online 5 January 2017

Keywords:
Chain-ladder method
Claims reserving uncertainty
Claims development result
Mack's formula
Merz–Wüthrich's formula
Conditional mean square error of prediction
Run-off uncertainty
Full Bayesian chain-ladder model

ABSTRACT

We revisit the gamma–gamma Bayesian chain-ladder (BCL) model for claims reserving in non-life insurance. This claims reserving model is usually used in an empirical Bayesian way using plug-in estimates for the variance parameters. The advantage of this empirical Bayesian framework is that allows us for closed form solutions. The main purpose of this paper is to develop the full Bayesian case also considering prior distributions for the variance parameters and to study the resulting sensitivities.

1. Introduction

The chain-ladder (CL) algorithm is probably to most popular method to set the reserves for non-life insurance claims. Originally, the CL method was introduced in a purely algorithmic fashion and it was not based on a stochastic model. Stochastic models underpinning the CL algorithm with a statistical model were only developed much later. The two most commonly used stochastic representations are Mack's (1993) distribution-free CL model and the over-dispersed Poisson (ODP) model of Renshaw and Verrall (1998) and England and Verrall (2002). In this paper we study the gamma–gamma Bayesian chain-ladder (BCL) model which provides in its non-informative prior limit another stochastic representation for the CL method. This model was first considered in a claims reserving context by Gisler (2006) and Gisler and Wüthrich (2008). The typical application of the gamma–gamma BCL model was done under fixed (given) variance parameters, using plug-in estimates for these variance parameters, see Example 2.13 in Wüthrich and Merz (2015) for such an empirical Bayesian analysis. Of course, this (partially) contradicts the Bayesian paradigm. In a full Bayesian approach one should also model these variance parameters with prior distributions. The aim of this paper is to study the influence of such a *full* Bayesian modeling approach and compare it to the *empirical* Bayesian modeling approach used in Wüthrich and Merz (2015). In particular, we aim at analyzing the sensitivity of the prediction uncertainty in the choice of these variance parameters. This is crucial in solvency considerations and it improves the crude estimates usually used in practice.

We remark that Bayesian models have become very popular in claims reserving, recently. This popularity is partly explained by the fact that Bayesian models can be solved very efficiently with Markov chain Monte Carlo (MCMC) simulation methods. In this paper we aim at finding a class of Bayesian models that provides an analytical closed form solution for the posterior quantities of interest. This has the advantage of getting a deeper mathematical insight and of accelerating the calculations. For an overview on Bayesian methods in claims reserving and corresponding MCMC applications we refer to the literature from this growing area.

Organization.
In the next section we introduce the gamma–gamma BCL model and we show that we recover the classical CL reserves in its non-informative prior limit. In Section 3 we provide the prediction uncertainty formulas in the long-term view and the short-term view, respectively. In Section 4 we give several real data examples and analyze the resulting sensitivities in the choice of the prior distributions. All proofs are provided in Appendix A.

* Corresponding author.
E-mail address: mario.wuethrich@math.ethz.ch (M.V. Wüthrich).

http://dx.doi.org/10.1016/j.insmatheco.2016.12.007
0167-6687/

2. Gamma–gamma Bayesian chain-ladder model

We introduce the gamma–gamma BCL model in this section. In contrast to Model Assumptions 2.6 in Wüthrich and Merz (2015) we also model the variance parameters in a Bayesian way. We then derive the claims predictors in the non-informative prior limit which turn out to be identical to the classical CL predictors, as seen in Theorem 2.2.

We denote accident years by $1 \leq i \leq I$ and development years by $0 \leq j \leq J$. Throughout, we assume $I > J$. The cumulative claim of accident year i after development year j is denoted by $C_{i,j}$, and $C_{i,J}$ denotes the ultimate claim of accident year i. For more background information on the claims reserving problem and the corresponding notation we refer to Chapter 1 in Wüthrich and Merz (2015). We make the following model assumptions.

Model Assumptions 2.1 (*Gamma–Gamma BCL Model*).

(a) *Conditionally given parameter vectors* $\boldsymbol{\Theta} = (\Theta_0, \dots, \Theta_{J-1})$ *and* $\boldsymbol{\sigma} = (\sigma_0, \dots, \sigma_{J-1})$, *the cumulative claims* $(C_{i,j})_{0 \leq j \leq J}$ *are independent (in accident year i), and Markov processes (in development year j) with conditional distributions*

$$C_{i,j+1}\big|_{\{C_{i,j}, \boldsymbol{\Theta}, \boldsymbol{\sigma}\}} \sim \Gamma\left(C_{i,j}\sigma_j^{-2}, \Theta_j \sigma_j^{-2}\right),$$

for all $1 \leq i \leq I$ *and* $0 \leq j \leq J-1$.

(b) *The parameter vectors* $\boldsymbol{\Theta}$ *and* $\boldsymbol{\sigma}$ *are independent.*

(c) *The components* Θ_j *of* $\boldsymbol{\Theta}$ *are independent and* $\Gamma(\gamma_j, f_j(\gamma_j - 1))$*-distributed with prior parameters* $f_j > 0$ *and* $\gamma_j > 1$ *for* $0 \leq j \leq J-1$.

(d) *The components* σ_j *of* $\boldsymbol{\sigma}$ *are independent and* π_j*-distributed having support in* $(0, d_j)$ *for given constants* $0 < d_j < \infty$ *for all* $0 \leq j \leq J-1$.

(e) $(\boldsymbol{\Theta}, \boldsymbol{\sigma})$ *and* $C_{1,0}, \dots, C_{I,0}$ *are independent and* $C_{i,0} > 0$, $\mathbb{P}$*-a.s., for all* $1 \leq i \leq I$.

These model assumptions imply that we have the following CL properties

$$\mathbb{E}\left[C_{i,j+1}\middle| C_{i,0}, \dots, C_{i,j}, \boldsymbol{\Theta}, \boldsymbol{\sigma}\right] = \Theta_j^{-1} C_{i,j}, \tag{2.1}$$

$$\operatorname{Var}\left(C_{i,j+1}\middle| C_{i,0}, \dots, C_{i,j}, \boldsymbol{\Theta}, \boldsymbol{\sigma}\right) = \Theta_j^{-2}\sigma_j^2 C_{i,j}. \tag{2.2}$$

Thus, for given parameter vectors $\boldsymbol{\Theta}$ and $\boldsymbol{\sigma}$ we obtain a distributional example of Mack's (1993) distribution-free CL model with CL factors Θ_j^{-1} and variance parameters $\Theta_j^{-2}\sigma_j^2$. Moreover, for $\pi_j(\cdot)$ being single point masses for all $0 \leq j \leq J-1$ we exactly obtain Model Assumptions 2.6 of Wüthrich and Merz (2015) assuming given (known) variance parameters σ_j^2.

The main task in claims reserving is to predict the ultimate claims $C_{i,J}$, given observations

$$\mathcal{D}_t = \{C_{i,j} : i + j \leq t, \ 1 \leq i \leq I, \ 0 \leq j \leq J\}, \quad \text{at time } t \geq I.$$

In complete analogy to the derivations in Section 2.2.1 of Wüthrich and Merz (2015), the application of Bayes' rule provides posterior π for the parameters $(\boldsymbol{\Theta}, \boldsymbol{\sigma})$, conditionally given observations $\mathcal{D}_t$, for $t \geq I$,

$$\pi\left(\boldsymbol{\theta}, \boldsymbol{\sigma}\middle| \mathcal{D}_t\right) \propto \prod_{j=0}^{J-1} \left[\theta_j^{\gamma_j + \sum_{i=1}^{(t-j-1)\wedge I} \frac{C_{i,j}}{\sigma_j^2} - 1} \right.$$
$$\times \exp\left\{ -\theta_j \left[f_j(\gamma_j - 1) + \sum_{i=1}^{(t-j-1)\wedge I} \frac{C_{i,j+1}}{\sigma_j^2} \right] \right\}$$
$$\left. \times \prod_{i=1}^{(t-j-1)\wedge I} \frac{\left(\frac{C_{i,j+1}}{\sigma_j^2}\right)^{\frac{C_{i,j}}{\sigma_j^2}}}{\Gamma\left(\frac{C_{i,j}}{\sigma_j^2}\right)} \pi_j(\sigma_j) \right]. \tag{2.3}$$

From this we see that the posteriors of (Θ_j, σ_j) are independent for different development periods $0 \leq j \leq J-1$. A non-informative prior limit for Θ_j corresponds here to letting $\gamma_j \to 1$ (and the terms $f_j(\gamma_j - 1)$ will vanish in (2.3)). Therefore, we refer as the non-informative prior limit when the (component-wise) limits $\boldsymbol{\gamma} \to \mathbf{1}$ are taken, where we set $\boldsymbol{\gamma} = (\gamma_0, \dots, \gamma_{J-1})$ and $\mathbf{1} = (1, \dots, 1)$. We have the following theorem for the claims prediction under this non-informative prior limit (the proof is given in the Appendix).

Theorem 2.2 (*CL Predictor*). *Under Model Assumptions 2.1 and for* $t \geq I \geq i > t - n \geq t - J$, *the Bayesian predictor for* $C_{i,n}$ *in its non-informative prior limit is given by*

$$\lim_{\boldsymbol{\gamma} \to \mathbf{1}} \mathbb{E}\left[C_{i,n}\middle| \mathcal{D}_t\right] = C_{i,t-i} \prod_{j=t-i}^{n-1} \widehat{f}_j^{CL(t)} \stackrel{\text{def.}}{=} \widehat{C}_{i,n}^{CL(t)},$$

with CL factors $\widehat{f}_j^{CL(t)}$ *defined by*

$$\widehat{f}_j^{CL(t)} = \frac{\sum_{\ell=1}^{(t-j-1)\wedge I} C_{\ell,j+1}}{\sum_{\ell=1}^{(t-j-1)\wedge I} C_{\ell,j}}.$$

Theorem 2.2 states that in the non-informative prior limit $\boldsymbol{\gamma} \to \mathbf{1}$ we exactly obtain the classical CL predictor, see Mack (1993). Thus, we have found another stochastic representation that underpins the CL algorithm with a statistical model. Therefore, we (may) use this model in its non-informative prior limit to analyze the prediction uncertainty of the CL algorithm. In contrast to Remark 2.11 of Wüthrich and Merz (2015) we now obtain this result in the full Bayesian framework, also considering prior distributions for the standard deviation parameters $\boldsymbol{\sigma}$.

3. Prediction uncertainty formulas

3.1. Long-term prediction uncertainty formula

The main purpose of this paper is to analyze the influence of the standard deviation parameters $\boldsymbol{\sigma}$ on the ultimate claim prediction uncertainty, where in contrast to Chapter 2 of Wüthrich and Merz (2015), these standard deviation parameters $\boldsymbol{\sigma}$ are also modeled with prior distributions. We analyze the prediction uncertainty at time $t \geq I$ in terms of the conditional mean square error of prediction (MSEP) given by

$$\begin{aligned}
\operatorname{msep}_{C_{i,J}|\mathcal{D}_t}\left(\mathbb{E}\left[C_{i,J}\middle| \mathcal{D}_t\right]\right) &= \mathbb{E}\left[\left(C_{i,J} - \mathbb{E}\left[C_{i,J}\middle| \mathcal{D}_t\right]\right)^2\middle| \mathcal{D}_t\right] \\
&= \operatorname{Var}\left(C_{i,J}\middle| \mathcal{D}_t\right) \\
&= \operatorname{Var}\left(\mathbb{E}\left[C_{i,J}\middle| \mathcal{D}_t, \boldsymbol{\sigma}\right]\middle| \mathcal{D}_t\right) \\
&\quad + \mathbb{E}\left[\operatorname{Var}\left(C_{i,J}\middle| \mathcal{D}_t, \boldsymbol{\sigma}\right)\middle| \mathcal{D}_t\right].
\end{aligned} \tag{3.1}$$

We aim at calculating this conditional MSEP in the gamma–gamma BCL model which provides, in its non-informative prior limit $\boldsymbol{\gamma} \to \mathbf{1}$, an uncertainty estimate for the CL algorithm, that is, we aim at calculating the limit

$$\widehat{\operatorname{msep}}_{C_{i,J}|\mathcal{D}_t}\left(\widehat{C}_{i,J}^{CL(t)}\right) \stackrel{\text{def.}}{=} \lim_{\boldsymbol{\gamma} \to \mathbf{1}} \operatorname{msep}_{C_{i,J}|\mathcal{D}_t}\left(\mathbb{E}\left[C_{i,J}\middle| \mathcal{D}_t\right]\right).$$