Insurance: Mathematics and Economics 59 (2014) 258–278

Contents lists available at ScienceDirect

Insurance: Mathematics and Economics

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

A copula based Bayesian approach for paid–incurred claims models for non-life insurance reserving

Gareth W. Peters[a,b,*], Alice X.D. Dong[c], Robert Kohn[d]

[a] Department of Statistical Science, University College London UCL, London, UK
[b] CSIRO Mathematics, Informatics and Statistics, Sydney, Australia
[c] University of Sydney, School of Mathematics and Statistics, Australia
[d] UNSW Australian School of Business, Australia

ARTICLE INFO

Article history:
Received December 2012
Received in revised form
May 2014
Accepted 26 September 2014
Available online 16 October 2014

Keywords:
Chain ladder
Claims reserving
Adaptive Markov chain Monte Carlo

ABSTRACT

Our article considers the class of recently developed stochastic models that combine claims payments and incurred losses information into a coherent reserving methodology. In particular, we develop a family of hierarchical Bayesian paid–incurred claims models, combining the claims reserving models of Hertig (1985) and Gogol (1993). In the process we extend the independent log-normal model of Merz and Wüthrich (2010) by incorporating different dependence structures using a Data-Augmented mixture Copula paid–incurred claims model.

In this way the paper makes two main contributions: firstly we develop an extended class of model structures for the paid–incurred chain ladder models where we develop precisely the Bayesian formulation of such models; secondly we explain how to develop advanced Markov chain Monte Carlo sampling algorithms to make inference under these copula dependence PIC models accurately and efficiently, making such models accessible to practitioners to explore their suitability in practice. In this regard the focus of the paper should be considered in two parts, firstly development of Bayesian PIC models for general dependence structures with specialised properties relating to conjugacy and consistency of tail dependence across the development years and accident years and between Payment and incurred loss data are developed. The second main contribution is the development of techniques that allow general audiences to efficiently work with such Bayesian models to make inference. The focus of the paper is not so much to illustrate that the PIC paper is a good class of models for a particular data set, the suitability of such PIC type models is discussed in Merz and Wüthrich (2010) and Happ and Wüthrich (2013). Instead we develop generalised model classes for the PIC family of Bayesian models and in addition provide advanced Monte Carlo methods for inference that practitioners may utilise with confidence in their efficiency and validity.

1. Introduction

As discussed in Merz and Wüthrich (2010) the main task of reserving actuaries is to predict ultimate loss ratios and outstanding loss liabilities. That is, in order to ensure the financial security of an insurance company, it is important to predict future claims liabilities and obtain the corresponding prediction intervals which take into account parameter uncertainty. In general such predictions are based on past information that comes from a variety of sources. Under a credibility based framework, the weighting of different data sources and their relative contribution to the estimated reserve can be difficult to determine. Therefore, it is important to consider the development of other unified prediction frameworks for the outstanding loss liabilities. Early attempts at such unified combining methods go back to the Munich chain ladder method introduced by Quarg and Mack (2004) which is one of the first claims reserving approaches in the actuarial literature to unify outstanding loss liability prediction based on both sources of information. This method aims to reduce the gap between the two chain ladder predictions that are based on claims payments and incurred losses data, respectively. It is achieved by adjusting the chain ladder factors with paid–incurred ratios to reduce the gap between the two predictions. The main drawback with the Munich chain ladder method is that it involves several parameter estimates whose precisions are difficult to quantify within a stochastic model framework.

* Corresponding author at: Department of Statistical Science, University College London UCL, London, UK.
E-mail address: gareth.peters@ucl.ac.uk (G.W. Peters).

http://dx.doi.org/10.1016/j.insmatheco.2014.09.011

As a consequence recently there have been new attempts to develop models to combine these two important sources of information, the payment and incurred loss data, into a consistent claims reserving model. The approach proposed in Merz and Wüthrich (2010) is known as a sub-family of the paid–incurred chain ladder (PIC) class of models. As a first instance of such a PIC model, Merz and Wüthrich (2010) introduced a log-normal PIC model and used Bayesian methods to estimate the missing (future) part of the claims reserving triangles based on both payment and loss incurred information. The major advantage of the PIC model structure is that the full predictive distribution of the outstanding loss liabilities can be quantified. One important limitation of the model of Merz and Wüthrich (2010) is that it does not develop the dependence properties of the PIC model that may be applicable or of interest to consider in loss reserving data observed in practice.

The first attempt to address the incorporation of dependence was recently proposed in Happ and Wüthrich (2013) where a restricted class of models was proposed for a very simple dependence structure, see Happ and Wüthrich (2013, Figure 1.1). This model is parsimonious but was limited to only three parameters for the correlations which were not incorporated into a formal Bayesian estimation approach, and instead fixed deterministically *a priori* via some hybrid of Bayesian and frequentist methods.

Our article extends the proposed Bayesian PIC models to capture a more flexible range of dependence structures that includes as special cases the model classes of Merz and Wüthrich (2010) and Happ and Wüthrich (2013). We note that in our framework we also show how to extend the model of Happ and Wüthrich (2013) into a complete Bayesian estimation framework rather than a hybrid frequentist and Bayesian approach that they considered. We aim to significantly enhance the class of dependence model structures one may consider in the PIC setting whilst sticking to a complete and formal Bayesian formulation of the problem. Achieving this is non-trivial in both the development of the Bayesian model structures and the estimation under such models. We note that in some cases we will show that this can also be done in a parsimonious manner under mixture Archimedean copula structures. This will be particularly relevant if only vague *a-priori* information is known about the PIC model parameters and the loss payment and incurred loss triangles are small.

In general one may consider three forms of dependence, namely, dependence within payment data, within incurred loss data, and even between payment and incurred loss data. In general it will be up to practitioners as to which of these forms of dependence they find most practically relevant in practice. The intention of this paper is not to advocate that dependence is always present, instead we provide a general class of PIC Bayesian models that can accommodate a wide array of dependence structures as well as a suitable estimation framework based on MCMC methods that practitioners can utilise to explore aspects of their given data accurately.

It is important to recognise that in this paper we are not attempting to argue for or against particular model structures or dependence features in the PIC family of models for application in practice. Instead we simply demonstrate carefully how to develop generalised classes of such models that are complete in a Bayesian sense and preserve certain important statistical assumptions such as consistency of the tail dependence assumptions throughout the development and accident year structures in the likelihood, even in the presence of unobserved components as arises in the triangle structures of the payment and incurred loss triangles.

2. Contributions: extending dependence structures for PIC Bayesian models

The main focus of this paper is to develop new flexible classes of copula dependent PIC Bayesian models with appropriate non-trivial inference procedures that will allow practitioners and actuaries to explore in their given applications this class of models in order to test their worth in practical settings. To be able to utilise efficiently general dependence structures in PIC Bayesian models in practice we introduce to the actuarial literature the data augmentation method which is an auxiliary variable framework that is not previously utilised in the actuarial literature. Given these extended Bayesian PIC models we also provide the appropriate MCMC methods to handle these new PIC Bayesian models, since standard MCMC methods will not be adequate to efficiently work with such models in practice. This is an important contribution to introduce to the actuarial literature modern classes of adaptive MCMC and auxiliary variable methods recently developed in statistics that actuaries can then explore for their given applications. In particular we make two main contributions to the literature for this class of PIC models.

Contribution 1: We extend the class of dependence structures available for the PIC reserving models in the process making the inference procedure into a consistent and coherent fully Bayesian formulation. This is unlike previous proposed approaches in Happ and Wüthrich (2013) which have only very simple and restrictive correlation based dependence structures which were highly constrained in the form of dependence between the development and accident years of the payments and incurred claims triangles and which did not allow for the possibility of interesting features such as tail dependence. In addition we note that we achieve this also in more sophisticated dependence structures in some cases without incurring an increase in the number of parameters relative to the simplified model of Happ and Wüthrich (2013). In other cases since our approach produces a complete and consistent fully Bayesian formulation we may utilise prior beliefs to inform the *a-priori* belief in a particular form of dependence which can therefore still be estimated sensibly in applications where the size of the loss reserving triangle is not large. Then since we are developing complete Bayesian models, practitioners can utilise standard model selection methods to select appropriate dependence structures for their data and application.

There are two technical difficulties we address when extending the dependence structures utilised in the PIC reserving models within a formal complete Bayesian framework. The first occurs when working with general linear dependence structures in the PIC model likelihood, as encoded by covariance-correlation structures between rows (columns) of the payment or incurred or both loss data triangles. The challenge involves being able to specify and evaluate in closed form "point-wise" the PIC Bayesian posterior model (up to proportionality). This is challenging as we need to ensure the prior and posterior admit appropriate restrictions on the support upon which they are defined to produce densities for different structures of positive definite symmetric covariance matrices (with constraints). The second problem occurs when other forms of dependence are considered such as those obtained when considering copula dependence models not purely obtained from a covariance-correlation matrix structure.

We solve the first challenge through development of a generic PIC Bayesian model with special specification of a class of matrix-variate Inverse-Wishart priors defined over the space of positive definite matrices that will admit conditional posterior conjugacy structures under our family of PIC Bayesian models. The second problem is addressed by developing models for general copula dependence structures in the likelihood that will produce closed form