Insurance: Mathematics and Economics 52 (2013) 381–390

Contents lists available at SciVerse ScienceDirect

Insurance: Mathematics and Economics

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

Claims reserving in the hierarchical generalized linear model framework

Patrizia Gigante [a,*], Liviana Picech [a], Luciano Sigalotti [b]

[a] Dipartimento di Scienze Economiche, Aziendali, Matematiche e Statistiche "Bruno de Finetti", University of Trieste, Piazzale Europa 1, 34127 Trieste, Italy

[b] Dipartimento di Scienze Economiche e Statistiche, University of Udine, Via Tomadini 30, 33100 Udine, Italy

ARTICLE INFO

Article history:
Received January 2011
Accepted 9 February 2013

JEL classification:
G22

MSC:
IM42

Keywords:
Claims reserving
Conditional mean square error of prediction
Hierarchical generalized linear models
h-likelihood
Quasi-hierarchical generalized linear models

ABSTRACT

We consider an approach based on the hierarchical generalized linear models and *h*-likelihood estimators for claims reserving in non-life insurance. The hierarchical generalized linear models represent a class of flexible mixture models that extend the generalized linear models and the generalized linear mixed models. The fitting algorithm and the inferential analyses can be obtained by applying standard procedures to one or more generalized linear models, suitably defined. Our study examines how the models can be used to obtain predictors of the claims reserves and to determine their prediction uncertainty.

1. Introduction

For several reasons, the insurance companies are unable to settle some claims immediately. For such claims, the insurers need to set up claims reserves to meet their obligations towards the policyholders.

The evaluation of adequate reserves to face the outstanding claims is one of the main actuarial problems in non-life insurance. The literature is rich of stochastic claims reserving models. For an extensive overview, we refer the reader to England and Verrall (2002) and Wüthrich and Merz (2008). In the recent years, the interest in such models has been growing also in the actuarial practice because of the new regulations on solvency requirements (Solvency II) and on accounting standards (IAS/IFRS), which require to provide, in addition to a best estimate of the outstanding claims, also some measure of the prediction uncertainty.

Among the proposed models, several papers consider mixture models, where the distributions of the incremental payments depend on a vector of unknown risk parameters. Such models allow incorporating external information, such as prior expert knowledge about the model parameters or market information. They can be studied following the Bayesian or the credibility approaches (e.g. de Alba, 2002; England and Verrall, 2006; England et al., 2012; Mack, 2000; Ntzoufras and Dellaportas, 2002; Verrall, 2004; Verrall and England, 2005). In the presence of regression components, they can be tackled by combining the techniques of the Generalized Linear Models (GLMs) with those of the credibility theory (e.g. Nelder and Verrall, 1997; Ohlsson, 2008; Ohlsson and Johansson, 2006) or in the framework of the Generalized Linear Mixed Models (GLMMs) (Antonio and Beirlant, 2007; Antonio et al., 2006).

In this paper, we explore the application of stochastic models of the class of *Hierarchical Generalized Linear Models* (HGLMs) (Lee and Nelder, 1996, 2001; Lee et al., 2006) to the claims reserving problem. The use of such models in actuarial applications has been recommended also by Nelder and Verrall (1997).

The HGLMs are mixture models with regression components, in which the response variables conditioned to the risk parameters as well as the risk parameters themselves follow distributions of the exponential dispersion family. In the conjugate models, in particular, the distributions of the risk parameters are conjugate of those of the response variables.

The HGLMs extend on the one hand the class of GLMs admitting random effects in the linear predictors, in addition to the usual fixed effects. Hence, they introduce a particular dependence structure among the response variables. On the other hand, HGLMs extend the class of GLMMs, which are mixed models having both fixed and random effects, the last ones assumed

* Corresponding author. Tel.: +39 040 5587114; fax: +39 040 5587033.
E-mail addresses: patriziag@econ.units.it (P. Gigante), livianap@econ.units.it (L. Picech), luciano.sigalotti@uniud.it (L. Sigalotti).

0167-6687/$ – see front matter
doi:10.1016/j.insmatheco.2013.02.006

to be normally distributed. In GLMMs the estimation approach consists in maximizing the marginal log-likelihood of the response variables and the computation of the estimates often requires analytically intractable integrals. Many numerical integration or approximate methods have been introduced to overcome the problem, however, they are generally computationally intensive.

For the estimation of the fixed and random effects and inferential analyses in HGLMs, Lee and Nelder (1996) introduced the *hierarchical likelihood* or *h-likelihood*, that avoids the integration out of random effects necessary to derive the marginal likelihood of the response variables. Under appropriate conditions, the method leads to estimators with good statistical properties.

An extension of HGLMs, similar to that leading from GLMs to quasi likelihood models (or quasi-GLMs), allows estimating semi-parametric models in which the distributions of the conditional response variables and/or of the risk parameters are not fully specified. Moreover, the dispersion parameters may have a regression structure.

Thus, the HGLMs and their extensions represent a class of flexible models, that allow us to choose the distributions of the conditional responses and of the risk parameters in a wide class of distributions, to take account of fixed effects and random effects in the linear predictors, to consider different link functions, to model structured dispersion components. Thanks to the h-likelihood approach, the algorithm for fitting the models can be reduced to the fitting of one or more GLMs or quasi GLMs. Moreover, the model-checking techniques developed for GLMs can be applied to this wide class of models.

Within claims reserving, such models can be used both when the data are aggregated claim figures, provided by the so-called run-off triangles, and when the data consist of individual claim figures. Moreover, through the parameters of the distributions of the random effects, we can take account of some external or initial information that are then updated according to the available portfolio data.

In this paper, we consider the class of HGLMs and their extensions in the estimation and prediction problems connected with claims reserving, assuming that the data are summarized in a run-off triangle of incremental payments. In addition to predictions or best estimates of the reserves, we also provide measures of their quality. As for the last problem, in order to evaluate the variability of the reserve estimates we consider the conditional mean square error of prediction. In this way we can take account of the variability due to the pure randomness as well as of the uncertainty in the estimation of the parameters.

By exploiting some results in Lee and Ha (2010) and Lee and Nelder (1996), we develop approximate formulae for the prediction error, that can be easily calculated once the parameter estimates are available.

Some numerical results are supplied for illustrative and comparison purposes.

The rest of the paper is structured as follows. Section 2 presents the problem setting. In Section 3 we provide a quite extensive overview of the HGLMs, since they are non-standard in claims reserving. We describe the model assumptions of the conjugate models and of their extensions, HGLMs and quasi-HGLM. We describe the estimation procedure based on the h-likelihood for HGLMs and on the double extended quasi-likelihood for quasi-HGLMs. In Section 4 we discuss the prediction problem in claims reserving and give approximate formulae, which can be easily calculated, in order to determine the prediction uncertainty. In Section 5 a numerical example is provided. The data are taken from Tables 2.4–2.5 in Wüthrich and Merz (2008). The results are compared to those in Alai et al. (2009) and in England et al. (2012).

Table 1
Run-off table.

Origin year	Development year				
	0	…	$t-i$	…	t
0	Y_{00}		…		Y_{0t}
⋮	⋮				
i	Y_{i0}	…	$Y_{i,t-i}$	…	Y_{it}
⋮	⋮				
t	Y_{t0}		…		Y_{tt}

2. Notations

As usual in claims reserving, we assume that the data of a portfolio consist of a run-off triangle of values $y_{ij},\ i,j = 0,\dots,t$, $i + j \leq t$, where i refers to the origin year (accident year, underwriting year, ...), j to the development year and t denotes the most recent origin year assumed to be equal to the latest delay year (we do not consider tail factors). Moreover, we assume that the data refer to an incremental quantity, e.g. incremental payments, incremental payments standardized with respect to some exposure measure, number of payments or number of reported claims in the cell (i,j), ... For illustrative purposes, we assume that y_{ij} denotes the incremental payments in cell (i,j).

Within stochastic models, in connection with the above data set, we introduce the random vector $\left\{Y_{ij} \mid i,j = 0,\dots,t\right\}$ represented in Table 1, where Y_{ij} denotes the random payment in development year j, for claims with origin year i.

The upper triangle,

$$\mathcal{D}_t = \left\{Y_{ij} \mid i+j \leq t\right\},$$

relates to the run-off triangle of observations. The lower triangle,

$$\left\{Y_{ij} \mid i+j > t\right\},$$

needs to be predicted. We are often interested in some function of the lower triangle. In particular, we are concerned with the *outstanding loss liabilities* or *reserve for origin year i*,

$$R_i = \sum_{j=t-i+1}^{t} Y_{ij}, \quad i = 1,\dots,t,$$

and with the *total outstanding loss liabilities* or *total reserve*,

$$R = \sum_{i,j:i+j>t} Y_{ij}.$$

In the following, for the process $\left\{Y_{ij} \mid i,j = 0,\dots,t\right\}$, we consider mixture models of the class of the hierarchical generalized linear models and their extensions.

3. Hierarchical generalized linear models

In this section, we provide an overview of the *Hierarchical Generalized Linear Models* (HGLMs), starting from the class of the *conjugate* ones (Lee and Nelder, 1996, 2001; Lee et al., 2006). Our description of these models is adapted to the context of claims reserving, when the data are summarized in a run-off triangle of payments as described in the previous section.

Let $Y_{ij},\ i,j = 0,\dots,t$, be the response variables (incremental payments), $U_i,\ i = 0,\dots,t$, unobservable risk parameters related to the origin years, $\boldsymbol{U} = (U_0,\dots,U_t)^T$, $\boldsymbol{x}_{ij}$ a vector of covariates for Y_{ij}. The covariate vectors could include any observable feature affecting the distribution of the response variables. In our case the available information for any payment is given by the origin year, the development year and the payment year. Among these elements we select the covariates, that can be either numeric or categorical.