European Journal of Operational Research 249 (2016) 657–666

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Stochastics and Statistics

Asymptotic behaviors of stochastic reserving: Aggregate versus individual models[☆]

Jinlong Huang[a], Xianyi Wu[a,*], Xian Zhou[b]

[a] Department of Statistics and Actuarial Science, East China Normal University, China
[b] Department of Applied Finance and Actuarial Studies, Macquarie University, Australia

ARTICLE INFO

Article history:
Received 8 November 2014
Accepted 21 September 2015
Available online 3 October 2015

Keywords:
Risk management
Stochastic reserving
Individual data model
Aggregate data model
Asymptotic variance

ABSTRACT

In this paper, we investigate the asymptotic behaviors of the loss reservings computed by individual data method and its aggregate data versions by Chain-Ladder (CL) and Bornhuetter–Ferguson (BF) algorithms. It is shown that all deviations of the three reservings from the individual loss reserve (the projection of the outstanding liability on the individual data) converge weakly to a zero-mean normal distribution at the $\sqrt{n}$ rate. The analytical forms of the asymptotic variances are derived and compared by both analytical and numerical examples. The results show that the individual method has the smallest asymptotic variance, followed by the BF algorithm, and the CL algorithm has the largest asymptotic variance.

1. Introduction

Predicting future losses (outstanding liabilities) and thus the reserves required to cover them, referred to as *loss reserving*, is a crucial and indispensable task of risk management in the insurance and financial industries. It has been increasingly recognized that loss reservings based directly on individual data (raw data, or micro-level data, see, e.g. Antonio and Plat, 2014) are generally more accurate than those computed from aggregate data (macro-level data), at the cost of higher operation/management expenditure on data collection, storage and computations. This is because the former makes better use of the information included in the data, whereas the latter generally aggregates data without a clear purpose in risk management and rational justification in statistical theory. In other words, the use of aggregate data allows a substantial reduction in operation loadings, but sacrifices the accuracy of loss reserving. To date, however, it is unclear how much accuracy has been sacrificed in loss reserving using the aggregate data versus the more informative individual data. On the other hand, the main reason for aggregating data in old days was the lack of facilities and the high costs to storing, transporting and retrieving information from the raw data, which is no longer a big issue with the fast development of modern technology. As noted ten years ago by England and Verrall (2002) and Taylor and Campbell (2002), lots of useful information about the claims data remains unused if data are aggregated. There has been a limited literature on the efforts to provide evidence for the advantage of loss reservings using individual data. The examples include:

- Antonio and Plat (2014), Pigeon, Antonio, and Denuit (2013; 2014) conducted empirical studies based on a set of common real-life micro-level data from a European insurance company, with parameterized versions (so that the maximum likelihood procedure can be applied to estimate the unknown parameters) of Norberg's Position Dependent Marked Poisson Process model (in continuous time, cf. Norberg, 1993; 1999). It was shown that the micro-level approaches provided closer prediction of the future liabilities for a subset of the data that have been settled so that the liabilities are in fact observed at the assumed reserve evaluating date (referred to as out-of-sample prediction).
- Huang, Qiu, and Wu (2015), under a special micro-level data set with only once payment for every claim at its settlement, proved the almost sure convergence of their loss reserving at a rate of order $o(n)$ theoretically. They further showed that the individual loss reserve (the projection of the outstanding liability on the individual data) outperforms the aggregate loss reserve (the projection of the outstanding liability on the aggregate data) in mean squared deviation from the individual loss reserves, and the individual loss reserving outperforms (with smaller bias and variance than) such traditional aggregate loss reservings as the Chain Ladder (CL) and Bornhuetter–Ferguson (BF) methods via Monte Carlo simulations.

For an outstanding liability L, a broadly used criterion to compare different loss reserving procedures is the (conditional) mean squared

☆ This work was supported by Shanghai Philosophy and Social Science Foundation under Grant no. 2010BJB004, NSFC under Grant no. 71371074 and the 111 Project under No. B14019.

* Corresponding author. Tel.: +86 21 54345058.
E-mail address: xywu@stat.ecnu.edu.cn (X. Wu).

http://dx.doi.org/10.1016/j.ejor.2015.09.039
0377-2217/

error of prediction (abbreviated as MSEP in the literature)

$$\mathrm{MSEP}_{L|\mathcal{D}ata} = \mathrm{E}[(\widehat{L}-L)^2|\mathcal{D}ata] = \mathrm{Var}(L|\mathcal{D}ata) + (\widehat{L}-\mathrm{E}[L|\mathcal{D}ata])^2, \tag{1.1}$$

where $\widehat{L}$ is a $\mathcal{D}ata$-measurable predictor/estimate (or simply referred to as a *loss reserve*) of L. In Antonio and Plat (2014) and Pigeon et al. (2013, 2014), $\mathrm{MSEP}_{L|\mathcal{D}ata}$ was computed by empirical data where the future liabilities were known from the data. Theoretically, it was estimated by the data (in micro- or macro-level) using for example the bootstrap method. Examples that have been discussed are MSEP of the CL method (Mack, 1993), Benktander model (Mack, 2000), the BF method under generalized linear models (Alai, Merz, & Wüthrich, 2009; 2010) and MSEP of loss reservings under Log-normal/Log-normal model, the exponential dispersion model and other stochastic models (see, e.g., Wüthrich & Merz, 2008). Note that, as a function of the observed data that is free of the reserving procedures, the term $\mathrm{Var}(L|\mathcal{D}ata)$ in (1.1) contributes only a constant to MSEP and thus can be ignored in measuring the deviation of the reserving $\widehat{L}$ from the outstanding liability L, so that one can just address the term $(\widehat{L}-\mathrm{E}[L|\mathcal{D}ata])^2$, for comparing different reserving methods. For simplicity, we refer to $(\widehat{L}-\mathrm{E}[L|\mathcal{D}ata])^2$ as the *mean squared error* (MSE) below. It is obvious that the estimated conditional MSE/MSEP is also a random variable and thus, when more than one loss reserving method are in hand, does not give a clear picture of which one would be better.

Generally, a feasible way is to use unconditional $\mathrm{E}[(\widehat{L}-\mathrm{E}[L|\mathcal{D}ata])^2]$ to comprehensively measure the deviation of the loss reserving from the outstanding liability/loss reserve, as what have done in Huang et al. (2015) by means of Monte Carlo simulations. As argued by Huang et al. (2015), in the situations where micro-level data (denoted by $\mathcal{D}$) are available, one should measure the deviation $(\widehat{L}-\mathrm{E}[L|\mathcal{D}])^2$, rather than $(\widehat{L}-\mathrm{E}[L|\mathcal{AD}])^2$ based on aggregated data, where $\mathcal{AD}$ indicates some aggregated data (macro-level data) produced from the micro-level data.

The other way that can give perfect characteristics of a loss reserving $\widehat{L}$ is to compare the distribution of $\widehat{L}-\mathrm{E}[L|\mathcal{D}]$. Because it is usually impossible to obtain the exact distribution of the deviation in fixed sample size, a usual alternative is to study its asymptotic distribution. This paper is dedicated to derive the asymptotic distributions of loss reservings and compare different reserving procedures. More specifically, we investigate the asymptotic behaviors of the loss reservings based on individual data and their CL and BF versions based on the deduced aggregated data proposed in Huang et al. (2015). The key contributions of this paper are highlighted as follows:

(a) We derive the asymptotically normal distributions of $\widehat{L}-\mathrm{E}[L|\mathcal{D}]$ for individual and aggregate loss reservings $\widehat{L}$. While the aggregate procedures have been used for decades, this paper appears the first to derive the asymptotic distributions of $\widehat{L}-\mathrm{E}[L|\mathcal{D}]$ under both individual and aggregate methods.
(b) Based on the asymptotic distributions, we compare the accuracy of the three loss reservings in terms of their asymptotic variances. While they are all weakly convergent at a rate of order $\sqrt{n}$, the numerical examples show that the asymptotic variance is smallest with individual loss reserving and largest with the CL loss reserving, while the BF reserving takes an intermediate place that is closer to the CL method.

These findings indicate the preference order of the loss reservings: Individual loss reserving is more accurate than its BF version, and the BF procedure is more accurate than its CL version.

The remainder of this paper is organized as follows. In Section 2, after introducing the notations for observations and parameters concerned, data structure, model assumptions and estimates of the unknown parameters proposed in Huang et al. (2015) for later reference, we review the three loss reserving methods: individual data method, CL method and BF method. In Section 3, we give asymptotic distributions of the deviations of the three reservings from the individual loss reserve $\mathrm{E}[L|\mathcal{D}]$ and then make some analytical and numerical comparisons on their asymptotic variances. The technically complicated proofs of the asymptotic distributions are relegated in Section 4 so as to smooth the flow of the text. The final section concludes the paper.

2. Models and methods

2.1. Model formulation and parameter estimation

This section first reviews the data structure, distribution assumptions, and the parameter estimation developed in Huang et al. (2015). Some notations are also introduced in this section.

Following Huang et al. (2015), all claims are organized in accident years $i = 0, 1, \ldots, I$ and loss reserving is evaluated at the end of the latest accident year. The data are structured with the following items:

(a) In every accident year i, there are n_i policies so that $n := \sum_{i=0}^{I} n_i$ is the total number of policies exposed to all $I+1$ accident years, where " := " reads "defined as".
(b) For the kth policy of accident year i, referred to as policy (i, k), denote by $e_{ik} \in [0, 1]$ its exposure, M_{ik} the number of its claims, and (i, k, l) its lth claim event, $l = 1, 2, \ldots, M_{ik}$ so that, the quantities $e_i := \sum_{k=1}^{n_i} e_{ik}$ and $N_i := \sum_{k=1}^{n_i} M_{ik}$ are respectively the total exposure and number of claims occurring in accident year i. Also denote $e_{(r)} = \sum_{i=0}^{I-r} e_i$, $r = 0, 1, \ldots, J_1$.
(c) For each claim event (i, k, l), there exists a maximum reporting delay J_1 and a maximum settlement delay J_2 satisfying $I = J = J_1 + J_2$, where J indicates the maximum development year, and a claim is only paid once at the end of its claim settlement year. Denote by R_{ikl} its reporting delay, T_{ikl} settlement delay and Y_{ikl} claim amount.

Next presented are the technical assumptions regarding the joint distribution of the data.

Assumption 2.1.

(a) The claims data $\left\{M_{ik}; \{R_{ikl}, T_{ikl}, Y_{ikl}\}_{l=1}^{\infty}\right\}$ are mutually independent over policies (i, k) for $k = 1, 2, \ldots, n_i$ and $i = 0, 1, \ldots, I$.
(b) For every policy (i, k),
- *the claims number $M_{ik} \sim$ Poisson(λe_{ik}), independent of the sequence of random triplets $\{(R_{ikl}, T_{ikl}, Y_{ikl}), l = 1, 2, \ldots\}$, where λ is the unknown intensity of claims numbers from a risk with unit exposure, and*
- *the triplets $(R_{ikl}, T_{ikl}, Y_{ikl}), l = 1, 2\ldots,$ are independent and identically distributed (i.i.d.) as a representative (R, T, Y), whose joint distribution is determined by the unknown parameters*

$$p_r = \Pr(R = r), \quad q_{rt} = \Pr(T = t|R = r) \quad \text{and}$$
$$\begin{pmatrix}\mu_{rt}\\ \nu_{rt}\end{pmatrix} = \mathrm{E}\left[\begin{pmatrix}Y\\ Y^2\end{pmatrix}\middle| R = r, T = t\right],$$
$$r = 0, 1, \ldots, J_1, \; t = 0, 1, \ldots, J_2.$$

Write $\lambda_r = \lambda p_r$, $\mu_r' = (\mu_{r0}, \mu_{r1}, \ldots, \mu_{rJ_2})$, $\mathbf{q}_r' = (q_{r0}, q_{r1}, \ldots, q_{rJ_2})$ and

$$c_{rt} = \frac{q_{rt}}{\sum_{l=t}^{J_2} q_{rl}} = \Pr(T = t|R = r, T \geq t), \quad t = 0, 1, \ldots, J_2,$$
$$r = 0, 1, \ldots, J_1.$$

(c) The limits $n_i/n \to \kappa_i \in (0, 1)$ and $\sqrt{n}(e_i/n - \breve{e}_i) \to 0$ hold for every $i = 0, \ldots, I$ as the total number of individuals $n \to \infty$. Accordingly, denote $\breve{e}_{(r)} = \sum_{i=0}^{I-r} \breve{e}_i$, $r = 0, 1, \ldots, J_1$.

For any vector $\mathbf{x}$, use diag($\mathbf{x}$) to indicate the diagonal matrix generated by the components of $\mathbf{x}$, so that if we have two column vectors