Claim Sizes in the Compound Poisson Process from a Bayesian viewpoint

Gamze Özel

Bayesian Young Statisticians Meeting (BAYSM), Milan June, 5-6, 2013
Paper no. 27

Hacettepe University, Department of Statistics, Ankara, Turkey
gamzeozl@hacettepe.edu.tr

Abstract

The typical assumption of independence among claim size distributions is not always satisfied in risk modelling. In this study, the exchangeable claim sizes are considered aggregated claims are obtained via compound Poisson proces. Exchangeability of the claim size is obtained by the conditional independence, using parametric and nonparametric measures for the conditioning distribution. A Bayesian analysis of the proposed model is illustrated with Turkish Earthquake Insurance Claims Data between 2000-2003.

Keywords: Bayesian analysis; compound Poisson process; claims size; earthquake data.

1 Introduction

In risk modelling, when the number of claims follows a Poisson process, the aggregated claims amount can be modelled by the CPP which is denoted by $\{X_t, t \geq 0\}$ and defined as follows:

$$X_t = \sum_{i=1}^{N_t} Y_i \tag{1}$$

where $\{N_t, t \geq 0\}$ is a homogeneous Poisson process with parameter $\lambda > 0$ and Y_i, $i = 1, 2, \ldots$ are independent and identically distributed random variables independent of N_t. In risk theory, $\{N_t, t \geq 0\}$ is the number of claims performed to the company during the time interval $(0, t]$ and Y_i, $i = 1, 2, \ldots$ is the ith claim size. Hence, X_t is the aggregated claims up to time t.

Bayesian methods are very useful in actuarial science since it yields to learn about the whole distributions of quantities rather than just obtain an expected value for each parameter. These methods allow to include many levels of randomness in the analysis through the use of prior distributions for each parameter, which highlights the uncertainty regarding individual distributions or parameters. In addition, the posterior distribution can be updated obtained when new information becomes available. The existing literature on the Bayesian analysis of CPP includes, but is not limited to, works by [3], [5] and [6].

In this study, the CPP is used for the aggregate claim and a variety of loss distributions to model the size of individual claims. Bayesian methodology is used to fit distributions to the claim sizes. This approach assumes that all parameters in the distribution are variables. The results derived by using Bayesian methodology with those from classical statistics are compared to see which had the best fit with the data.

2 Estimation of the claim count and claim amount distribution

In classical risk theory, it is very common to assume a homogeneous Poisson process for the claim arrival process since this assumption simplies the derivation of the total claim amount distribution. Also, a gamma distribution model is frequently assumed to describe the usual right skewed shape of the claim size distributions. In this section, the bayesian estimation of the total claim count and the total claim amount in a future time period are given. A Bayesian estimation of the CPP is obtained by calculating the posterior means of their cumulative distributions, also called their predictive cumulative distribution functions. In order to set ideas, started with CPP model with independent claim sizes. $Y_{ij} \sim f(y|\theta)$ for $i = 1, 2, \ldots, N_{tj}$ Due to the independence of the Poisson processes N_{tj} and the claim sizes Y_{ij}, inference for λ and θ is done separately. The number of events N_{tj} follows a homogeneous Poisson process independent of the expenditures sizes Y_{ij} which are gamma distributed with parameters a and b. We then start by assuming that X_{tj} is a CPP with independent claims, that is, Y_{ij} are all independent for $i = 1, \ldots, n_j$ and $j = 1, \ldots, m$. If the prior knowledge on (a, b) can be represented by $\pi(a) = Ga(a|\alpha_a, \beta_a)$ and $\pi(b) = Ga(b|\alpha_b, \beta_b)$ independently, then the posterior distribution, given the sample, is characterized by the conditional distributions. An illustration based on Turkish Earthquake Insurance Claims Data between $2000 - 2003$ is given. Posterior inference for the insurance data carried out by implementing Gibbs samplers for the CPP model. For sampling from each of the conditional distributions random walk Metropolis-Hastings steps are used. This proposal distribution is centered at the previous value of the chain and has a variation coefficient of one. The Gibbs sampler was run for 100,000 iterations with a burn-in of 200,000 keeping every $10th$ observation after burn in to reduce the autocorrelation in the chain. Finally, a predictive analysis for the aggregated expenditures of a patient in a year is carried out. For that the information about the frequencies of occurrence of the claims is needed, modeled by the Poisson processes and in particular by the intensity λ. Considering that we observed 1342 claims in the year, then the posterior distribution for the intensity of the claims per year, λ, is $Ga(1341; 10.658)$. Therefore, the posterior mean rate is 12.42 claims per person per year. With the posterior distribution for λ and the posterior predictive distribution for the whole sequence of claims, we obtain the posterior predictive distribution for the aggregated claims in a year for one individual.

References

[1] J.M. Bernardo, A.F.M. Smith. Bayesian Theory; (1994); Wiley, Chichester.

[2] U.E. Makov, A.F.M. Smith, Y. Liu. **Bayesian methods in actuarial science**; *The Statistician*; 1996; 45(4); pp. 503-515.

[3] H.H. Panjer, G.E. Willmot. **Compound Poisson models in actuarial risk theory**; *Journal of Econometrics*; 1983; 23; pp. 63-76.

[4] R.V. Hogg, S.A. Klugman. Loss Distributions; (1984); Wiley, Toronto.

[5] J.S. Pai. **Bayesian analysis of compound loss distributions**; *Journal of Econometrics*; 1997; 79; pp. 129-146.

[6] C. Dudley, Bayesian analysis of an aggregate claim model usng various loss distributions. Dissertation thesis for Master of Science in Actuarial Management, 2006, Heriot-Watt University Edinburgh.