Loss modelling with mixtures of Erlang distributions

Roel Verbelen[1], Lan Gong[2], Katrien Antonio[1,3], Andrei Badescu[2], and Sheldon Lin[2]

[1]LStat, Faculty of Economics and Business, KU Leuven, Belgium.

[2]Department of Statistics, University of Toronto, Canada.

[3]Faculty of Economics and Business, University of Amsterdam, The Netherlands.

Modeling data on claim sizes is crucial when pricing insurance products. Such loss models require on the one hand the flexibility of nonparametric density estimation techniques to describe the insurance losses and on the other hand the feasibility to analytically quantify the risk. Mixtures of Erlang distributions with a common scale are very versatile as they are dense in the space of distributions on $\mathbb{R}^+$(Tijms (1994, p. 163)). At the same time, it is possible to work analytically with this kind of distributions. Closed-form expressions of quantities of interest, such as the Value-at-Risk (VaR) and the Tail-Value-at-Risk (TVaR), can be derived as well as appealing closure properties (Lee and Lin (2010), Willmot and Lin (2011) and Klugman et al. (2012)). In particular, using these distributions in aggregate loss models leads to an analytical form of the corresponding aggregate loss distribution which avoids the need for simulations to evaluate the model.

In actuarial science, claim severity data is often censored and/or truncated due to policy modifications such as deductibles and policy limits. Lee and Lin (2010) formulate a calibration technique based on the EM algorithm for fitting mixtures of Erlangs with a common scale parameter to complete data. Here, we construct an adjusted EM algorithm which is able to deal with censored and truncated data, inspired by McLachlan and Peel (2001) and Lee and Scott (2012). Using the developed R program, we demonstrate the approximation strength of mixtures of Erlangs and model e.g. the left truncated Secura Re data from Beirlant et al. (2004), and use the mixtures of Erlangs approach to price an excess-of-loss reinsurance contract.

References

Beirlant, J., Goegebeur, Y., Segers, J., Teugels, J., De Waal, D., and Ferro, C. (2004). *Statistics of Extremes: Theory and Applications*. Wiley Series in Probability and Statistics. Wiley.

Klugman, S. A., Panjer, H. H., and Willmot, G. E. (2012). *Loss models: from data to decisions*, volume 715. Wiley.

Lee, G. and Scott, C. (2012). EM algorithms for multivariate Gaussian mixture models with truncated and censored data. *Computational Statistics & Data Analysis*, 56(9):2816 – 2829.

Lee, S. C. and Lin, X. S. (2010). Modeling and evaluating insurance losses via mixtures of Erlang distributions. *North American Actuarial Journal*, 14(1):107.

McLachlan, G. and Peel, D. (2001). *Finite mixture models*. Wiley.

Tijms, H. C. (1994). *Stochastic models: an algorithmic approach*. Wiley.

Willmot, G. E. and Lin, X. S. (2011). Risk modelling with the mixed Erlang distribution. *Applied Stochastic Models in Business and Industry*, 27(1):2–16.