Abstract

It is no longer uncommon these days to find the need in actuarial practice to model claim counts from multiple types of coverage, such as the ratemaking process for bundled insurance contracts. Since different types of claims are conceivably correlated with each other, the multivariate count regression models that emphasize the dependency among claim types are more helpful for inference and prediction purposes. Motivated by the characteristics of an insurance dataset, we investigate alternative approaches to constructing multivariate count models based on the negative binomial distribution. A classical approach to induce correlation is to employ common shock variables. However, this formulation relies on the NB-I distribution which is restrictive for dispersion modeling. To address these issues, we consider two different methods of modeling multivariate claim counts using copulas. The first one works with the discrete count data directly using a mixture of max-id copulas that allows for flexible pair-wise association as well as tail and global dependence. The second one employs elliptical copulas to join continuumized data while preserving the dependence structure of the original counts. The empirical analysis examines a portfolio of auto insurance policies from a Singapore insurer where claim frequency of three types of claims (third party property damage, own damage, and third party bodily injury) are considered. The results demonstrate the superiority of the copula-based approaches over the common shock model. Finally, we implemented the various models in loss predictive applications.

Keywords: Negative binomial, Insurance claim count, Copula, Jitter, Multivariate model