



SOCIETY OF ACTUARIES

Article from:

ARCH 2014.1 Proceedings

July 31-August 3, 2013

Generalized Linear Models for a Dependent Aggregate Claims Model

Juliana Schulz, Dr. José Garrido

Concordia University, Canada; julianaschulz@live.com

A primary objective of property casualty insurers is to adequately price the risk inherent in their portfolio by estimating the expected value of all future costs associated with the insurance protection provided. The standard approach in the industry is to study both the frequency (the number of claims) and severity (the claim amounts) separately. The product of the expected claim frequency and severity then yields the expected loss cost (or pure premium), which represents the total cost of all claims.

In recent years, several insurers have adopted the use of Generalized Linear Models (GLMs) for modeling both the frequency and the severity components of the claims process. The GLM approach allows for the mean of the response variable to be expressed in terms of a linear combination of covariates via a link function. This method requires that the response variable distribution be a member of the Exponential Family (EF). Generally, the EF distribution structure results in a particular mean-variance relation that allows to further characterize the response variable. Specifically, under the EF, the variance is a function of the mean.

The standard approach in the industry is to develop a GLM for the claim frequency separately from the GLM fitted to the claim severity, and then calculate the pure premium as the product of the expected frequency and the expected severity. This approach inherently assumes independence between the frequency and the severity of the claims process, an assumption that is unrealistic: Both GLMs share common explanatory variables and are fitted to the same policyholder data. An alternative approach is to model the total loss cost directly by means of the Tweedie distribution, which models the aggregate claims as a Compound Poisson-Gamma sum. However, this method also assumes independence between the claim count and claim sizes.

In order to address the dependence between the frequency and the severity in the collective risk model, we propose to use a multivariate modeling approach for correlated responses. In particular, a conditional GLM approach is used to find the expected loss cost in the dependent aggregate claims model. Without any assumption of independence between the frequency and severity components, it is shown that the expected loss cost can be written in terms of the marginal mean claim frequency, a modified marginal mean severity and a correction term. The structure obtained for the mean total claims cost includes the independent model as a special case.

This research provides an alternative approach for establishing insurance premiums which takes into account the dependence between the claim frequency and severity. The structure obtained is simple to implement and allows for a straightforward comparison of the dependent model with the traditionally used independent model. Moreover, using a multivariate GLM approach for correlated responses provides a more accurate representation of the insurance data and allows for more precise insurance premiums.