

Claim Count Prediction with Over Dispersed Outcome and Mismeasured Covariates in Actuarial Practice*

Minkun Kim

ADAPT Centre, Dublin City University, Ireland,

Marija Bezbradica

ADAPT Centre, Dublin City University, Ireland,

Martin Crane

ADAPT Centre, Dublin City University, Ireland,

Correspondence should be addressed to: Minkun Kim; minkun.kim4@mail.dcu.ie

* Presented at the 40th IBIMA International Conference, 23-24 November 2022, Seville, Spain

Copyright © 2022. Minkun Kim, Marija Bezbradica and Martin Crane

Abstract

The problem of overdispersed claim counts and mismeasured covariates is common in insurance. On the one hand, the presence of overdispersion in the count data violates the homogeneity assumption, and on the other hand, measurement errors in covariates highlight the model risk issue in actuarial practice. The consequence can be inaccurate premium pricing which would negatively affect business competitiveness. Our goal is to address these two modelling problems simultaneously by capturing the unobservable correlations between observations that arise from overdispersed outcome and mismeasured covariate in actuarial process. To this end, we establish novel connections between the count-based generalized linear mixed model (GLMM) and a popular error-correction tool for non-linear modelling - Simulation Extrapolation (SIMEX). We consider a modelling framework based on the hierarchical Bayesian paradigm. To our knowledge, the approach of combining a hierarchical Bayes with SIMEX has not previously been discussed in the literature. We demonstrate the applicability of our approach on the workplace absenteeism data. Our results indicate that the hierarchical Bayesian GLMM incorporated with the SIMEX outperforms naïve GLMM / SIMEX in terms of goodness of fit.

Keywords: Count-based regression, Hierarchical Bayes, Negative Binomial GLMM, SIMEX