

Estimation Methods for Incomplete Insurance Data

Noah Gblonyah, Seth Okyere, Michael Throolin

3/25/2022

Introduction

The growth of the insurance industry is fueled by societal expectations for protection against a variety of risks associated with unfavorable random events that have a large economic impact. Insurance is a process that involves the payment of a premium in exchange for an equitable manner of offsetting the risk of a potential future loss. The basic idea is to set up a fund to which insured members contribute specified sums of premium for specific loss levels. When the random events that policyholders are protected against occur giving rise to claims then claims are settled from the fund.

The maximum likelihood (MLE) technique is commonly used by insurance firms to estimate claim distribution parameters. Maximum likelihood techniques are typically applied to complete data sets where we have exact values for all of the data points; however, data is rarely perfect in real life. When modeling the underlying loss variable, insurance contracts have coverage adjustments that must be taken into account. Typically, loss control methods like deductibles, policy limitations, and coinsurance are implemented to reduce undesirable policyholder behavioral effects such as adverse selection.

Estimation by method of moments and maximum likelihood are often easy to do, but these estimators tend to perform poorly, mainly because they use a few features of the data rather than the entire set of observations. It is important to use as much information as possible when the population has a heavy tail as in insurance data. This project discusses some considerations for properly handling truncated and censored data for modeling of insurance data. The methods to be discussed in this study are Maximum Likelihood Estimations and the Expectation Maximization Algorithm.

References

- Hartley, H. O. 1958. "Maximum Likelihood Estimation from Incomplete Data." *Biometrics* 14: 174.
- Klugman, Stuart et al. 2004. *Loss Models : From Data to Decisions, 2nd Edition*. Hoboken, N.J: Wiley Interscience.
- Truxillo, Catherine. 2005. "Maximum Likelihood Parameter Estimation with Incomplete Data." In *Proceedings of the Thirtieth Annual SAS (r) Users Group International Conference*, 111–30. Citeseer.