Mixture modeling of data with multiple partial right-censoring levels

Semhar Michael[1] · Tatjana Miljkovic[2] · Volodymyr Melnykov[3]

Abstract

In this paper, a new flexible approach to modeling data with multiple partial right censoring points is proposed. This method is based on finite mixture models, flexible tool to model heterogeneity in data. A general framework to accommodate partial censoring is considered. In this setting, it is assumed that a certain portion of data points are censored and the rest are not. This situation occurs in many insurance loss data sets. A novel probability function is proposed to be used as a mixture component and the expectation-maximization algorithm is employed for estimating model parameters. The Bayesian information criterion is used for model selection. Additionally, an approach for the variability assessment of parameter estimates as well as the computation of quantiles commonly known as risk measures is considered. The proposed model is evaluated using a simulation study based on four common probability distribution functions used to model right skewed loss data and applied to a real data set with good results.

Semhar Michael semhar.michael@sdstate.edu

1 Department of Mathematics and Statistics, South Dakota State University, Brookings, SD 57007, USA

2 Department of Statistics, Miami University, Oxford, OH 45056, USA

3 Department of Information Systems, Statistics, and Management Science, University of Alabama, Tuscaloosa, AL 35487, USA

Full article available at:

https://doi.org/10.1007/s11634-020-00391-x