Development of an Efficient Methodology for Analyzing Correlated Survival and Count Responses

A thesis submitted for the Degree of Doctor of Philosophy

A.A. Sunethra
Faculty of Science
University of Colombo
October 2018

Abstract

This research was aspirated for developing an improved analytical solution for analyzing data with correlated survival and count response variables. A bivariate model/joint model which simultaneously model these two responses was developed using two approaches named as shared random effects models and joint random effects models where both belongs to the family of random effects models. The use of random effects models for joint model development doesn't impose any restrictions on the choice of the marginal model for the two responses where this study deployed a parametric survival model with lognormal distribution for the survival response and a Poisson generalized linear model for the count response. As the choice of a statistical model depends on the design of the data, two commonly found designs of simple randomized data and cluster randomized data was considered and joint models were developed for these two designs. For each type of the joint model, for each design of the data, separate simulation studies were carried out to test the functionality of the joint model and was compared critically with the fit of separate univariate models of each response variable. The results of the simulation studies established the appropriateness of the proposed joint models for analyzing correlated survival and count responses. The use of nested random effects models for joint model development, developing a joint model for both simple randomized and cluster randomized data, developing joint models for both positively and negatively correlated responses and developing joint models consisting of fully parametric marginal models serves as the methodological contributions of this study.

Then, the proposed joint model of this study was used for analyzing data from a randomized control trial on Epilepsy where the two responses were the timing of seizures and count of seizures. The performance of the joint model was overwhelming in the analysis of this data where model diagnostics and validations were very satisfactory in the joint model compared to that of the univariate models. It was identified that the use of the proposed joint model provides better analysis of the data in terms of identifying and quantifying the risk factors for each response variable and predicting the survival probabilities of the Epilepsy patients. Therefore, the practical importance of the joint model was demonstrated with the analysis of the Epilepsy data.