Bootstrap GLM with Log ink (Poisson) for Over-Dispersed Count Data
Anu Babu

Department of Biostatistics, Christian Medical College, Vellore

Bootstrapping is a resampling method widely used in the medical literature. This work is a review on non-parametric bootstrapping of regression data. The aim of this work is to review generalized linear models and then to apply bootstrapping in generalized linear model to analyse Poisson count data. In this study a secondary data set, namely risk factors.sav is used. The data set incorporate a response variable "number of risk factors observed" and a covariate age. There are total of 441 observations. The mean count of the number of risk factors observed from patient 0-5 is 0.90 with variance is 1.23 and hence there is an over-dispersion exists in the response counts. In the Poisson regression model the age does impact the mean number of risk factors observed from patient 0-5. The mean number of risk factors increases 1.02 times per year. The number of risk factors observed from patient 0-5 is increases as age increases. The standard error of bootstrap estimates is smaller than the SE obtained in the simple method while doing the Poisson regression. The confidence intervals are observed to be very close for covariates Age and intercept. So the application of bootstrapping provides us better understanding and better results. The normal theory and percentile intervals are reasonably similar to each other, but the more trustworthy BCa intervals are somewhat different.