Payam Mokhtarian and Ray Chambers

National Institute for Applied Statistics Research Australia (NIASRA)

University of Wollongong

Outlier Robust Block Bootstrap Fitting of Linear Mixed Models - An Application to Small Area Estimation

Outliers are a well-known problem when fitting models with survey data. Estimators of the model parameters and predictors of population quantities using the fitted model become unstable in presence of outliers in data. The main approaches that have been developed so far for this problem have focused on modifying the parameter estimating equations to make them less sensitive to sample outliers. When linear mixed models are used, this leads to outlier robust versions of the maximum likelihood (ML) and restricted maximum likelihood (REML) estimating equations for the mixed model parameters as well as outlier robust versions of the estimating equations for the random area effects. However, to the best of our knowledge, there has been no attempt to use bootstrap-type outlier robust Monte Carlo methods to tackle this problem. Chambers and Chandra (2013) describe a random effect block bootstrap approach which is robust to failure of the dependence assumptions of the assumed mixed model. We propose an outlier robust extension of this idea that can be used to fit a linear mixed model in the presence of both group level as well as individual level outliers. We describe Monte Carlo simulation results that provide some evidence for our claim that the proposed robust block bootstrap (RREB) method is robust to the influence of outliers and also leads to more reliable mixed model parameter estimates than comparable outlier robust approaches that have been proposed in the literature. An important application of linear mixed models in survey statistics is small area estimation, and we describe how the RREB method can be applied in this context. It is well known that the empirical best linear unbiased predictor (EBLUP) can be unstable when there are outliers in sample data, and Sinha and Rao (2009) have proposed an outlier-robust EBLUP, or REBLUP, based on modifying parameter estimating functions to make them less sensitive to sample outliers. Unfortunately, these modified estimating functions can be numerically unstable, and mean squared error estimation for the REBLUP is not straightforward. Using the RREB approach, an outlier robust predictor of a small area mean comparable to the REBLUP can be easily calculated. Moreover, the fact that the RREB is a bootstrap procedure allows one to easily calculate a stable estimate of the mean squared error of this predictor.