Analysis of Longitudinal Data under Complex Surveys

Abdulhakeem A.H. Eideh
Department of Mathematics, Al-Quds University, Abu-Dies, Palestine
msabdul@science.alquds.edu

Data collected by sample surveys, and in particular by longitudinal surveys, are used extensively to make inferences on assumed population models. Often, survey design features (clustering, stratification, unequal probability selection, etc.) are ignored and the longitudinal sample data are then analyzed using classical methods based on simple random sampling. This approach can, however, lead to erroneous inference because of sample selection bias implied by informative sampling - the sample selection probabilities depend on the values of the model outcome variable (or the model outcome variable is correlated with design variables not included in the model). The predominant methods of analysis for longitudinal data have long been based on three approaches: random effects model, general linear model, and time series model. The purpose of this paper is to account for informative sampling in fitting these three models for longitudinal survey data. The idea behind the proposed approach is to extract the model holding for the sample data as a function of the model in the population and the first order inclusion probabilities, and then fit the sample model using maximum likelihood, pseudo maximum likelihood and estimating equations methods. In order to assess the performance of the estimators obtained, we designed and executed a simulation study. The simulation study carried out shows that the sample maximum likelihood estimation method produces better estimators, in the sense of smaller bias and smaller relative root mean square error, than pseudo maximum likelihood method.

Key Words: Informative sampling, longitudinal survey data, random effect model, time series model.