

Semiparametric Prediction of Finite Population Total under Informative Sampling and Nonignorable Nonresponse

Abdulhakeem Eideh

Al-Quds University, Palestine, msabdul@staff.alquds.edu

Sverchkov and Pfeiffermann (2004) studied the semiparametric prediction of finite population totals under informative sampling. In this paper we develop semiparametric prediction of finite population total under informative sampling and nonignorable nonresponse. Prediction of finite population total, requires specification of nonsampled and nonresponse models. Fuller (2009, p282) pointed that “The analysis of data with unplanned nonresponse requires the specification of a model for the nonresponse. Models for nonresponse address two characteristics: the probability of obtaining a response and the distribution of the characteristic. In one model it is assumed that the probability of response can be expressed as a function of auxiliary data. The assumption of a second important model is that the expected value of the unobserved variable is related to observable auxiliary data. In some situations models constructed under the two models lead to the same estimator. Similarly, specifications containing models for both components can be developed.”

For this aim we combines two methodologies used in the model-based survey sampling: the prediction of finite population total T , under informative sampling, and full response, and the prediction of T when the sampling design is noninformative and nonresponse mechanism is nonignorable. One incorporates the dependence of the first order inclusion probabilities on the study variable, while the other incorporates the dependence of the probability of nonresponse on unobserved or missing observations. Accordingly, we use the response distribution and relationships between moments of the superpopulation, sample, sample-complement, response, and non-response distributions, for the prediction of finite population totals, see Eideh (2012, 2016), and Pfeiffermann and Sverchkov(1999). The paper is purely mathematical that focus on the role of informativness of sampling design and informativeness of nonresponse in adjusting various predictors for bias reduction. Further experimentation (simulation and real data problem) with this kind of predictors is therefore highly recommended. I hope that the new mathematical results obtained will encourage further theoretical, empirical and practical research in these directions.

References

- Eideh, A.H. (2016). Estimation of finite population mean and superpopulation parameters when the sampling design is informative and nonresponse mechanism is nonignorable. *Pakistan Journal of Statistics and Operation Research (PJSOR)*. *Pak.j.stat.oper.res. Vol.XII No.3, 2016*, pp467-489.
- Eideh A.H. (2012). Estimation and Prediction under Nonignorable Nonresponse via Response and Nonresponse Distributions. *Journal of Indian Society of Agriculture Statistics*. *66(3)2012*, pp 359-380.
- Fuller W.A. (2009). *Sampling Statistics*. Wiley.
- Pfeiffermann, D. and Sverchkov, M.(1999). Parametric and semi-parametric estimation of regression models fitted to survey data. *Sankhya*, 61, B, pp 166-186.
- Sverchkov, M., and Pfeiffermann, D. (2004). Prediction of finite population totals based on the sample distribution. *Survey Methodology*, 30, pp 79-92.