Statistics and Applications {ISSN 2452-7395 (online)}
Volume 15 Nos. 1&2, 2017 (New Series), pp 27-36

A Generalized Mixture Estimator Of The Mean Of A Sensitive Variable In The Presence Of Non-Sensitive Auxiliary Information

Sat Gupta[1], Tanja Zatezalo[1] & Javid Shabbir[2]
[1]*Department of Mathematics and Statistics, University of North Carolina at Greensboro, Greensboro, North Carolina, USA*
[2] *Department of Statistics, Quaid-i-Azam University, Islambad, Pakistan*

Received: December 12, 2016; Revised: January 15, 2017; Accepted: January 25, 2017

Abstract

Gupta et al. (2012) proposed a generalized regression-cum-ratio estimator and Koyuncu et al. (2014) proposed a generalized exponential estimator for the mean of the sensitive variable utilizing a non sensitive auxiliary variable. We propose a new generalized mixture estimator for estimating the population mean of a sensitive study variable. The expressions for Bias and Mean Square Error are derived up to the first order of approximation. Numer-ical examples show that the proposed generalized mixture estimator performs better than many of the existing estimators.

Keywords: Generalized regression-cum-ratio estimator, Generalized exponential estimator, Generalized mixture Estimator, Population mean, Auxiliary information

1 Introduction

Randomized response technique (RRT) can be used to estimate the mean of a sensitive variable Y where direct observation on Y is subject to bias. We assume a non sensitive auxiliary variable X is available and can be observed directly. Sousa et al. (2010) introduced a ratio type estimator and Gupta et al. (2012) proposed a regression and generalized regression-cum-ratio estimators based on RRT models to deal with this situation. Following Bahl & Tuteja (1991), Koyuncu et al. (2014) also proposed a generalized exponential type estimator to improve the efficiency of the mean estimator based on RRT models.

In this paper we propose an ordinary exponential ratipo type estimator and two generalized mixture estimators where the RRT estimators of the mean of Y are further improved by using information from an auxiliary variable X. Expressions for the Bias and Mean Square Error are derived up to the first order of approximation. We will use the following notations.

Let Y be the sensitive study variable which cannot be observed directly. Let X be a non sensitive auxiliary variable which has a positive correlation with Y, and let S be a scrambling variable. Assume that S is independent of Y and X. The respondent is asked to report a scrambled response for Y given by $Z = Y + S$, but is asked to provide the true response for X. Let a random sample of size n be drawn without replacement from a finite population $U = (U_1, U_2, \ldots, U_N)$. For ith population dement, let y_i and x_i respectively be the values of the study variable Y and auxiliary variable X. Let $\overline{Y} = E(Y), \overline{X} = E(X)$ and $\overline{Z} = E(Z)$ be the population means for Y, X and Z respectively. We assume that the population mean $\overline{X}$ and the population variance S_x^2 of the auxiliary variable are known. Also, assume that population mean and the population variance for the scrambling variable S are known and given as $\overline{S} = E(S) = 0$

Corresponding author: Sat Gupta
E-mail: sngupta@uncg.edu

of the commonly known RRT estimators. For the proposed estimators all the percent relative efficiencies are greater 100 indicating that all these estimators are better than the RRT ordinary mean estimator. We also note that both of the proposed generalized mixture estimators are more efficient than the other estimators considered here. Furthemore, the choice $\alpha = 2$ works better than $\alpha = 1$. We may note that at a theoretical level, one may be tempted to optimize α. Our goal though was to have a general family of estimators where many of the existing estimators become special cases of the proposed estimator with specific choice of α. For example, with $\alpha = 0$ our generalized mixture estimator II becomes combination of the regression and exponential ratio type estimators. For $\alpha = 1$, it involves the ratio term also. For $\alpha = -1$, it involves the product term.

References

Gupta S., Shabbir J. and Sehra S. (2010). Mean and sensitivity estimation in optional randomized response model. *Journal of Statistical Planning and Inference,* **140 (10)**, 2870-2874

Bahl,S. ,Tuteja, R.K. (1991). Ratio and product type exponential estimators. *Information and Optimization Sciences,* **12(1)**, 159-163

Gupta S., Shabbir J., Sousa R. and Corte-Real P. (2012). Estimation of the mean of a sensitive variable in the presence of auxiliary information. *Communications in Statistics - Theory and Methods,* **41 (13-14)**, 2394-2404

Koyuncu, N., Gupta S. and Sousa R. (2014). Exponential type estimators of the mean of a sensitive variable in the presence of non sensitive auxiliary information. *Communications in Statistics- Simulation and Computation*, **43**, 1583-1594

Sousa R., Shabbir J., Real P.C. and Gupta S. (2010). Ratio estimation of the mean of a sensitive variable in the presence of auxiliary information. *Journal of Statistical Theory and Practice,* **4 (3)**, 495-507.

Sukhatme, P. , Sukhatme, B. and Sukhatme, S. (1970). *Sampling theory of surveys with applications.* Third edition, Iowa state university press