Biostatistics and Biometrics
Open Access Journal
ISSN: 2573-2633

Juniper Publishers
key to the Researchers

Mini Review
Volume 3 Issue 2 - October 2017
DOI: 10.19080/BBOAJ.2017.03.555607

Biostat Biometrics Open Acc J

An Improved Estimation Procedure of the Mean of a Sensitive Variable Using Auxiliary Information

Tanveer AT[1] and Housila PS[2]

[1]*Department of Computer Science and Engineering, Islamic University of Science and Technology, India*

[2]*School of Studies in Statistics, Vikram University Ujjain, India*

Submission: August 22, 2017; **Published:** October 02, 2017

***Corresponding author:** Tanveer A Tarray, Department of Computer Science and Engineering, Islamic University of Science and Technology, India, Email: tanveerstat@gmail.com

Abstract

This paper proposes new ratio and regression estimators for the mean of sensitive variable utilizing information on a non – sensitive auxiliary variable. Expressions for the Biases and mean square errors of the suggested estimators correct up to first order of approximation are derived. It has been shown that the suggested new ratio and regression estimators are better than conventional unbiased estimators which do not utilize the auxiliary information, Sousa et al. [1] ratio estimator and Gupta et al. [2] regression estimator under a very realistic condition. In support of the present study we have also given the numerical illustrations.

Keywords: Ratio estimator; Regression estimator; Randomized response technique; Mean Square error; Bias; Auxiliary variable

Introduction

Let Y be the variable under study, a sensitive variable which can't be observed directly. Let X is a non – sensitive auxiliary variable which is strongly correlated with Y. Let S be a scrambling variable independent of the study variable Y and the auxiliary variable X. The usual additive model used for gathering information on quantitative sensitive variable is due to Himmelfarb & Edgell [3]. Their model allows the interviewee to hide personal information using a scrambling variable to their response. The respondent is asked to report a scrambled response for the study variable Y (based on additive model) given by Za =Y+S, but is asked to provided a true response for the auxiliary variable X [1].

Hussain [4] have discussed the use of subtracting scrambling. Thus following Hussain [4], the respondent is asked to report a scrambled response for the study variable Y (based on subtractive model) given by Zs = Y-S, but is asked to provide a true response for the auxiliary variable X. It is interesting to mention that the proposed model generalizes both usual additive and subtractive models. Gjestvang & Singh [5] have pointed out that "the practical application of an additive model is much easier than the multiplicative model, that is, respondents may like to add two numbers rather than doing painstaking work of multiplying two numbers or dividing two numbers: thus the improvement of the additive model has its own importance in the literature". Looking at the form the additive model, subtractive model and above arguments due to Gjestvang & Singh [5] we have introduced a new model (which is additive in nature)

$$Z_\varphi = Y + \varphi S$$

where φ is a known scalar such that $-1 \le \varphi \le 1$.

Thus keeping the proposed model $Z_\varphi = Y + \varphi S$ in view, the respondent is asked to report a scrambled response for Y given but is asked to provide a true response for X. Let a simple random sample of size n be drawn without replacement from a finite population U=(U1,U2,...UN). For the ith unit (i=1,2,... ,N), let and respectively be the values of the study variable Y and the auxiliary variable X. Further, let $\bar{y} = (1/n)\sum y_i$, $\bar{x} = (1/n)\sum_{i=1}^{n} x_i$, $\bar{z}_\varphi = (1/n)\sum_{i=1}^{n} z_{\varphi i}$ be the sample means and $\bar{Y} = E(Y$, $\bar{X} = E(X)$, $\bar{Z}_\varphi = E(Z_\varphi)$ be the population mean for Y, X and Z_φ respectively. We assume that the population mean $\bar{X}$ of the auxiliary variable X is known and $\bar{S} = E(S) = 0$

. Thus, $E(Z_\varphi) = E(Y)$. We also define

$$\bar{x} = \bar{X}(1 + e_x) \text{ and } \bar{z}_\varphi = \bar{Z}_\varphi(1 + e_{z_\varphi})$$

such that $E(e_x) = E(e_{z_\varphi}) = 0$, and

$$E(e_x^2) = \frac{(1-f)}{n} C_x^2 \ , \ E(e_{z_\varphi}^2) = \frac{(1-f)}{n} C_{z_\varphi}^2 ,$$