REVISTA INVESTIGACION OPERACIONAL VOL. 37 , NO. 2, 163-172, 2016

A GENERALIZED CLASS OF ESTIMATORS FOR THE FINITE POPULATION MEAN WHEN THE STUDY VARIABLE IS QUALITATIVE IN NATURE

Prayas Sharma[1]*, Carlos N. Bouza**, Hemant K.Verma***,Rajesh Singh*** and Jose M. Sautto****
*Department of Decision Sciences, College of Management and Economics Studies, University of Petroleum and Energy Studies, Dehradun, India.
**Department of Mathematics and Computation, University of Havana, Havana, Cuba.
***Department of Statistics, Banaras Hindu University, Varanasi, India.
****Campus Costa Chica, Universidad Autónoma de Guerrero, Mexico

ABSTRACT
This paper suggests a generalized class of estimators for the population mean of a qualitative study variable in simple random sampling using information on an auxiliary variable. Asymptotic expressions of bias and mean square error of the proposed class of estimators have been obtained. Asymptotic optimum estimator has been investigated along with its approximate mean square error. It has been shown that the proposed generalized class of estimators are more efficient than all the estimators considered by Singh et al. (2010) in case of a qualitative study variable. In addition theoretical findings are supported by an empirical study based on real population to show the superiority of the constructed estimators over others.

KEYWORDS: Auxiliary variable, Auxiliary Attribute, Bias, Mean Square Error, Simple random sampling.

MSC: 62D05

RESUMEN
En este paper se sugiere una clase de estimadores generalizados para la media poblacional de una variable de estudio cualitativa. Se han obtenido expresiones asintóticas del sesgo y el error cuadrático medio. Se ha investigado sobre la optimalidad en términos del error cuadrático aproximado obtenido. Se demuestra que los estimadores de la clase los generalizados propuesta son más eficientes que todos lo estimadores considerados por Singh et al. (2010) en el caso de una variable cualitativa. Adicionalmente los resultados teóricos son soportados por un estudio empírico basado en una población real notándose la superioridad de los estimadores construidos sobre los demás.

1. INTRODUCTION

Statisticians are often interested to use auxiliary information in sample surveys at estimation stage in order to improve the precision or accuracy of an estimator of unknown population parameter of interest (see Verma et al. (2015)). In some situations the auxiliary information is not available directly but in the form of an attribute that is auxiliary information is qualitative in nature. When auxiliary information is qualitative in nature then using the point bi-serial correlation between the study variable y and the auxiliary attribute ϕ several authors including Naik and Gupta (1996), Jhajj et al. (2006), Shabbir and Gupta (2007), Singh et al.(2008), Singh et al. (2010), Abd-Elfattah et al. (2010), Singh and Solanki (2012) , Sharma et al. (2013a,2013b) and Adichwal et al.(2015) proposed improved estimators of population parameters of interest under different situations.
All the authors have implicitly assumed that the study variable Y is quantitative whereas the auxiliary variable is qualitative. But there may be practical situations when study variable itself is qualitative in nature. For example, consider U.S. presidential elections. Assume that there are two political parties, Democratic and Republican. The dependent variable here is the vote choice between two political parties. Suppose we let Y=1, if the vote is for a Democratic candidate and Y=0, if the vote is for republican candidate. Some of the variables used in the vote choice are growth rate of GDP, unemployment and inflation rates, whether the candidate is running for re-election, etc. For the present purposes, the important thing is to note that the study variable is a qualitative variable. One can think several other examples where the study variable is qualitative in nature. Thus, a family

[1] prayassharma02@gmail.com

either owns a house or it does not, it has disability insurance or it does not, both husband and wife are in the labour force or only one suppose is, etc. In this paper we propose a generalized class estimators in which study variable is qualitative in nature. (see Gujarati and Sangeetha (2007)).

Consider a finite population U= (U_1, U_2, U_3, ….,U_N) containing N distinct and identifiable units. Let a sample of size n drawn by simple random sampling without replacement (SRSWOR) from a population U to estimate the population mean of qualitative variable. Let ϕ_i and x_i denote the observations on variable ϕ and x respectively for i^{th} unit (i=1,2,3…N). $\phi_i = 1$, if i^{th} unit of population possesses attribute ϕ and $\phi_i = 0$, otherwise. Further let $A = \sum_{i=1}^{N} \phi_i$ and $a = \sum_{i=1}^{n} \phi_i$, denotes the total number of units in the population and sample possessing attribute ϕ respectively, $P = \frac{A}{N}$ and $p = \frac{a}{n}$, denotes the proportion of units in the population and sample, respectively, possessing attribute ϕ.

Let us define,

$$e_0 = \frac{(p - P)}{P}, \qquad e_1 = \frac{(\bar{x} - \bar{X})}{\bar{X}},$$

Such that,

$$E(e_i) = 0, (i = 0,1)$$

and

$$E(e_0^2) = fC_\phi^2, \qquad E(e_1^2) = fC_x^2, \qquad E(e_0 e_1) = f\rho C_\phi C_{x.},$$

where,

$$f = \left(\frac{1}{n} - \frac{1}{N}\right), \qquad C_\phi^2 = \frac{S_\phi^2}{\phi^2}, \qquad C_x^2 = \frac{S_x^2}{\bar{X}^2},$$

and ρ is the point bi-serial correlation coefficient between ϕ and x.

The remaining portion of this paper is as follows, in section 2 we have considered some existing estimator along with its biases and mean square errors. In section 3, we have suggested a generalized class of estimators along with its members and studied their properties. Section 4, made some comparison of suggested class with other existing estimators. Empirical studies are carried out in section 5. The numerical results derived from the computation allowed measuring the efficiency of the different alternative estimators. We have ended the paper with the final conclusion.

2. AVAILABLE ESTIMATORS IN LITERATURE WHEN STUDY VARIABLE ITSELF AN ATTRIBUTE

A ratio-type estimator proposed by Singh et al. (2010) for estimating unknown population mean in case of qualitative study variable is

$$t_S = \left(\frac{P}{\bar{x}}\right)\bar{X} \tag{2.1}$$

The bias and MSE of the estimator t_S, developed using Taylor series up to the first order of approximation is given as

$$B(t_S) = f\left(\frac{C_x^2}{2} - \rho C_p C_x\right) \tag{2.2}$$

$$MSE(t_S) = fP^2\left(C_\phi^2 + C_x^2 - 2\rho C_\phi C_x\right) \tag{2.3}$$

Singh et al. (2010) suggested another general class of estimator as,

$$t_{GS} = H(p, u) \tag{2.4}$$

where $u = \frac{\bar{x}}{\bar{X}}$ and $H(p, u)$ is a parametric equation of p and u such that

$$H(p,1) = P, \forall P \tag{2.5}$$

and satisfying following regulations:

(i) Whatever be the sample chosen, the point (p,u) assume values in a bounded closed convex subset R_2 of the two-dimensional real space containing the point (p,1).
(ii) The function H(p,u) is a continuous and bounded in R_2 .
(iii) The first and second order partial derivatives of H(p,u) exist and are continuous as well as bounded in R_2.

where,

$$H_1 = \left.\frac{\partial H}{\partial u}\right|_{p=P,u=1}, \qquad H_2 = \left.\frac{1}{2}\frac{\partial^2 H}{\partial u^2}\right|_{p=P,u=1},$$

$$H_3 = \left.\frac{1}{2}\frac{\partial^2 H}{\partial p \partial u}\right|_{p=P,u=1}, \quad \text{and} \quad H_4 = \left.\frac{1}{2}\frac{\partial^2 H}{\partial p \bar{y}^2}\right|_{p=P,u=1}.$$

The bias and minimum MSE of the estimator t_b are respectively, given by –

$$B(t_{GS}) = f\left(P\rho C_\phi C_x H_3 + C_x^2 H_2 + P^2 C_y^2 H_4\right) \tag{2.6}$$

$$MSE(t_{GS})_{min} = fP^2 C_\phi^2 \left(1-\rho^2\right) \tag{2.7}$$

Singh et al. (2010) proposed a new family of estimator for estimating P, as

$$t_{NS} = \left[q_1 P + q_2\left(\bar{X} - \bar{x}\right)\right]\left[\frac{a\bar{X}+b}{a\bar{x}+b}\right]^\alpha \exp\left[\frac{\left(a\bar{X}+b\right)-\left(a\bar{x}+b\right)}{\left(a\bar{X}+b\right)+\left(a\bar{x}+b\right)}\right]^\beta \tag{2.8}$$

The bias and minimum MSE of the estimator to the first order of approximation, are respectively, given as

$$Bias(t_{NS}) = P(q-1) + f\left[\left(q_2 \bar{X} B + q_1 PA\right)C_x^2 - q_1 PB\rho C_\phi C_x\right] \tag{2.9}$$

$$MSE(t_{NS})_{min} = \left[P^2 - \frac{\Delta_1\Delta_5^2 + \Delta_3\Delta_4^2 - 2\Delta_2\Delta_4\Delta_5}{\Delta_1\Delta_3 - \Delta_2^2}\right] \tag{2.10}$$

where,

$$M_1 = P^2 f\left(C_p^2 + B^2 C_x^2 - 2B\rho C_p C_x\right), \qquad M_2 = \bar{X}^2 f\left(C_x^2\right),$$

$$M_3 = P^2 f\left(AC_x^2 - 2B\rho C_p C_x\right), \qquad M_4 = P\bar{X} f\left(-BC_x^2 + \rho C_p C_x\right),$$

$$M_5 = \bar{X} P f\left(-BC_x^2\right).$$

$$q_1^* = \frac{\Delta_1\Delta_4 - \Delta_2\Delta_5}{\Delta_1\Delta_3 - \Delta_2^2} \quad \text{and} \quad q_2^* = \frac{\Delta_1\Delta_5 - \Delta_2\Delta_4}{\Delta_1\Delta_3 - \Delta_2^2} \tag{2.11}$$

where,

$$\Delta_1 = \left(P^2 + M_1 + 2M_3\right), \ \Delta_2 = \left(-M_4 - M_5\right), \quad \Delta_3 = \left(M_2\right), \quad \Delta_4 = \left(P^2 + M_3\right), \ \Delta_5 = \left(-M_5\right),$$

3. THE SUGGESTED GENERALISED CLASS OF ESTIMATORS

We propose a generalized class of estimators for estimating P for a qualitative variable ϕ , as

$$t_N = \left\{d_1 p\left(\frac{\bar{X}}{\bar{x}}\right)^\alpha \exp\left(\frac{\eta\left(\bar{X}-\bar{x}\right)}{\eta\left(\bar{X}+\bar{x}\right)+2\lambda}\right)\right\} + d_2\bar{x} + \left(1 - d_1 - d_2\right)\bar{X} \tag{3.1}$$

where (d_1, d_2) are suitable constants that can be chosen such that MSE of t_N is minimum, η and λ are either real numbers or the functions of the known parameters of auxiliary variables such as coefficient of variation C_x, skewness $\beta_{1(x)}$, kurtosis $\beta_{2(x)}$ and correlation coefficient ρ (see Sharma and Singh (2015)).
It is to be mentioned that

(i) For (d_1, d_2)=(1,0), the class of estimator t_m reduces to the class of estimator as

$$t_{NP} = \left\{ p\left(\frac{\overline{X}}{\overline{x}}\right)^{\alpha} \exp\left(\frac{\eta(\overline{X}-\overline{x})}{\eta(\overline{X}+\overline{x})+2\lambda}\right)\right\} \tag{3.2}$$

(ii) For (d_1, d_2)=(d_1,0), the class of estimator t_N reduces to the class of estimator as

$$t_{NQ} = \left\{ d_1 p\left(\frac{\overline{X}}{\overline{x}}\right)^{\alpha} \exp\left(\frac{\eta(\overline{X}-\overline{x})}{\eta(\overline{X}+\overline{x})+2\lambda}\right)\right\} \tag{3.3}$$

Set of new estimators originated from (3.1) choosing the suitable values of d_1, d_2, α, η and λ are listed in Table 3.1.

Table 3.1: Set of estimators generated from the class of estimators t_N

Subset of proposed estimator	d_1	d_2	α	η	λ
$t_{N1} = p$ (usual unbiased estimator)	1	0	0	0	1
$t_{N2} = p\left(\frac{\overline{X}}{\overline{x}}\right) = t_S$ (Singh et al. 2010 type)	1	0	1	0	1
$t_{N3} = p\left(\frac{\overline{X}}{\overline{x}}\right)^{\alpha} = \hat{M}_3$ (Srivastava, 1967 type)	1	0	α	0	1
$t_{N4} = p\left(\frac{\overline{x}}{\overline{X}}\right) = M_p$ (Murthy , 1964 type)	1	0	-1	0	1
$t_{N5} = d_1 p\left(\frac{\overline{X}}{\overline{x}}\right)$ (Al and Cingi, 2009 type)	1	0	1	0	1
$t_{N6} = d_1 p\left(\frac{\overline{x}}{\overline{X}}\right)$	w_1	0	-1	0	1
$t_{N7} = d_1 p$ (Al and Cingi, 2009)	w_1	0	0	0	1
$t_{N8} = w_1 p + w_2 \overline{x} + (1 - w_1 - w_2)\overline{X}$	w_1	w_2	0	0	1

Another set of estimators generated from class of estimator t_{NQ} given in (3.3) using suitable values of η and λ are summarized in table 3.2

Table 3.2: Set of estimators generated from the estimator t_{NQ}

Subset of proposed estimator	α	η	λ
$t_{NQ}^{(1)} = \left\{ d_1 p\left(\frac{\overline{X}}{\overline{x}}\right) \exp\left(\frac{(\overline{X}-\overline{x})}{(\overline{X}+\overline{x})+2}\right)\right\}$	1	1	1

$t_{NQ}^{(2)} = \left\{ d_1 p \left(\frac{M_x}{\hat{M}_x} \right) \exp\left(\frac{(\bar{X} - \bar{x})}{(\bar{X} + \bar{x}) + 2\rho} \right) \right\}$	1	1	ρ
$t_{NQ}^{(3)} = \left\{ d_1 p \left(\frac{\bar{X}}{\bar{x}} \right) \exp\left(\frac{(\bar{X} - \bar{x})}{(\bar{X} + \bar{x}) + 2\bar{X}} \right) \right\}$	1	1	$\bar{X}$
$t_{NQ}^{(4)} = \left\{ d_1 p \left(\frac{\bar{X}}{\bar{x}} \right) \exp\left(\frac{(\bar{X} - \bar{x})}{(\bar{X} + \bar{x})} \right) \right\}$	1	1	0
$t_{NQ}^{(5)} = \left\{ w_1 p \left(\frac{\bar{x}}{\bar{X}} \right) \exp\left(\frac{(\bar{X} - \bar{x})}{(\bar{X} + \bar{x})} \right) \right\}$	-1	1	1
$t_{NQ}^{(6)} = \left\{ d_1 p \left(\frac{\bar{X}}{\bar{x}} \right) \exp\left(\frac{\bar{X}(\bar{X} - \bar{x})}{\bar{X}(\bar{X} + \bar{x}) + 2\rho} \right) \right\}$	1	$\bar{X}$	ρ
$t_{NQ}^{(7)} = \left\{ d_1 p \exp\left(\frac{\bar{X}(\bar{X} - \bar{x})}{\bar{X}(\bar{X} + \bar{x}) + 2\rho} \right) \right\}$	0	$\bar{X}$	ρ
$t_{NQ}^{(8)} = \left\{ d_1 p \left(\frac{\bar{X}}{\bar{x}} \right) \exp\left(\frac{\rho(\bar{X} - \bar{x})}{\rho(\bar{X} + \bar{x}) + 2\bar{X}} \right) \right\}$	1	ρ	$\bar{X}$
$t_{NQ}^{(9)} = \left\{ d_1 p \left(\frac{\bar{x}}{\bar{X}} \right) \exp\left(\frac{\rho(\bar{X} - \bar{x})}{\rho(\bar{X} + \bar{x}) + 2\bar{X}} \right) \right\}$	-1	ρ	$\bar{X}$

Expressing (3.1) in terms of e's, we have

$$t_N = d_1 P(1+e_0)(1+e_1)^{-\alpha} \exp\left\{-ke_1(1+ke_1)^{-1}\right\} + d_2\bar{X}(1+e_1) + (1-d_1-d_2)\bar{X}$$

where, $k = \frac{\eta\bar{X}}{2(\eta\bar{X}+\lambda)}$. (3.4)

Up to the first order of approximation we have,

$$(t_N - P) = \left[(d_1 - 1)b + d_2 P\left\{e_0 - ae_1 + de_1^2 - ae_0e_1\right\} + d_2\bar{X}e_1\right] \tag{3.5}$$

where $a = (\alpha + k)$, $b = (P - \bar{X})$ and $d = \left\{\frac{3}{2}k^2 + \alpha k + \frac{\alpha(\alpha+1)}{2}\right\}$

from equation (3.5), we have

$$(t_N - P)^2 = \left[(1-2d_1)b^2 + d_1^2\left\{b^2 + P^2\left(e_0^2 + a^2e_1^2 - 2ae_0e_1\right)\right\} + d_2^2\bar{X}^2e_1^2 + 2d_1d_2P\bar{X}\left(e_0e_1 - ae_1^2\right)\right]$$

Taking expectations both sides, we get the MSE of the estimator t_N to the first order of approximation as

$$MSE(t_N) = \left[(1-2w_1)b^2 + d_1^2M + d_2^2N + 2d_1d_2O\right] \tag{3.7}$$

where,

$$M = b^2 + P^2f\left(C_\phi^2 + a^2C_x^2 - 2a\rho C_\phi C_x\right),$$

$$N = \bar{X}^2fC_x^2,$$

$$O = P\bar{X}f\left(\rho C_\phi - aC_x\right)C_x.$$

The optimum values of d_1 and d_2 are obtained by minimizing (3.7) and is given by

$$d_1^* = \frac{b^2N}{(MN - O^2)} \quad \text{And} \quad d_2^* = \frac{-b^2O}{(MN - O^2)} \tag{3.8}$$

Substituting the optimal values of d_1 and d_2 in equation (3.7) we obtain the minimum MSE of the estimator t_N as

$$MSE_{min}(t_N) = b^2\left[1 - \frac{b^2 N}{\left(MN - O^2\right)}\right] \tag{3.9}$$

Putting the values of M, N, O and b and simplifying, we get the minimum MSE of estimator t_N as

$$MSE_{min}(t_N) = \left[\frac{P^2(1-R)^2 fC_\phi^2\left(1-\rho^2\right)}{\left[(1-R)^2 + fC_\phi^2\left(1-\rho^2\right)\right]}\right] \tag{3.10}$$

where $R = \frac{\overline{X}}{P}$ and P is defined earlier.

Similarly, the minimum MSE of the class of estimators t_{NQ} is given by

$$MSE_{min}(t_{NQ}) = P^2\left[\frac{\left(fC_\phi^2 + a^2\gamma C_x^2 - 2af\rho C_\phi C_x\right)}{\left(1 + fC_\phi^2 + a^2 fC_x^2 - 2af\rho C_\phi C_x\right)}\right] \tag{3.11}$$

4. EFFICIENCY COMPARISONS

From equations (2.3) and (2.7) we have

$$MSE(t_S) \geq MSE_{min}(t_{GS}) = fP^2\left(C_\phi^2 + C_x^2 - 2\rho C_\phi C_x\right) \geq P^2 fC_\phi^2\left(1-\rho^2\right)$$

Or $\rho^2 C_\phi^2 + C_x^2 - 2\rho C_\phi C_x \geq 0$ (4.1)

From equations (2.7) and (3.10) we have

$$MSE_{min}(t_{GS}) \geq MSE_{min}(t_N) = P^2 fC_\phi^2\left(1-\rho^2\right) \geq \left[\frac{P^2(1-R)^2 fC_\phi^2\left(1-\rho^2\right)}{\left[(1-R)^2 + fC_\phi^2\left(1-\rho^2\right)\right]}\right]$$

Or $(1-R)^2 + fC_\phi^2\left(1-\rho^2\right) \geq (1-R)^2$ (4.2)

The condition given in (4.2) shows always true.

From equation (2.10) and (3.10) we have

$$MSE_{min}(t_{NS}) \geq MSE_{min}(t_N)$$

If, $\left[P^2 - \frac{\Delta_1\Delta_5^2 + \Delta_3\Delta_4^2 - 2\Delta_2\Delta_4\Delta_5}{\Delta_1\Delta_3 - \Delta_2^2}\right] \geq \left[\frac{P^2(1-R)^2 fC_\phi^2\left(1-\rho^2\right)}{\left[(1-R)^2 + fC_\phi^2\left(1-\rho^2\right)\right]}\right]$ (4.3)

It follows from (4.1), (4.2) and (4.3), that the proposed class of estimators t_N is better than the ratio estimator t_S, general class of estimator t_{GS} and the family of estimators t_{NS}, due to Singh et al. (2010) under certain conditions .

Remark 4.1: Estimator Based on optimum values

Putting the optimum values of w_1^* and w_2^* in the equation (3.1) we get the optimum estimator as:

$$t'_m = \left\{d_1^* p\left(\frac{\overline{X}}{\overline{x}}\right)^\alpha \exp\left(\frac{\eta\left(\overline{X} - \overline{x}\right)}{\eta\left(\overline{X} + \overline{x}\right) + 2\lambda}\right)\right\} + w_2^*\overline{x} + \left(1 - d_1^* - d_2^*\right)\overline{X} \tag{4.4}$$

If the experimenter is not able to specify the value precisely, then it may be desirable to estimate the optimum values from the samples, therefore the values of w_1^* and w_2^* are given as:

$$d_1^* = \frac{\hat{b}^2\hat{N}}{\hat{M}\hat{N} - \hat{O}^2} \text{ and } d_2^* = \frac{\hat{b}^2\hat{O}}{\hat{M}\hat{N} - \hat{O}^2}$$

where $M = \hat{b}^2 + P^2\gamma\left(\hat{C}_y^2 + \hat{a}^2\hat{C}_x^2 - 2a\hat{\rho}\hat{C}_y\hat{C}_x\right),$

$N = \overline{X}\gamma\hat{C}_x^2, \ \hat{\rho}_c = 4\left(4\hat{p}_{11} - 1\right)$

$O = P\overline{X}\gamma\left(\hat{\rho}\hat{C}_\phi - a\hat{C}_x\right)\hat{C}_x., \ \hat{b} = \left(P - \overline{X}\right), \quad \hat{a} = \left(\alpha + \hat{k}\right)$ and $\hat{k} = \dfrac{\eta\hat{M}_x}{2\left(\eta\hat{M}_x + \lambda\right)}$

Here, we have assumed that the population median of auxiliary variable x is known, therefore $\hat{M}_x$ can also be remain as M_x .

Expressing (4.4) in terms of e's, we have

$$t'_m = w_1^* P(1+e_0)(1+e_1)^{-\alpha} \exp\left\{-\hat{k}e_1\left(1+\hat{k}e_1\right)^{-1}\right\} + w_2^*\overline{X}(1+e_1) + \left(1 - w_1^* - w_2^*\right)\overline{X}$$

Proceeding as above, we get the minimum MSE of the estimator t'_m given as:

$$MSE_{\min}(t'_m) = \left[\frac{\hat{M}_y^2\left(1-\hat{R}\right)^2\hat{\gamma}\hat{C}_y^2\left(1-\hat{\rho}_c^2\right)}{\left[(1-\hat{R})^2 + \hat{\gamma}\hat{C}_y^2\left(1-\hat{\rho}_c^2\right)\right]}\right] \tag{4.5}$$

5. EMPIRICAL STUDY

Data Statistics: To illustrate the efficiency of proposed generalized class of estimators in the application, we consider the following population data set.

The data used for empirical study has been taken from Gujrati and Sangeetha (2007) -pg, 601. And using raw data we have calculated the following values. Where,

y: Home ownership.

x: Income (in thousands of dollars

The values of the required parameters are:

N=40, n=11, $P = 0.525$, $\overline{X} = 14.4$, $C_\phi = 0.963$, $C_x = 0.308$ $\rho = 0.897$, R= 27.42,

$\lambda_{12} = -0.118$, $\lambda_{04} = 1.75$, $\lambda_{03} = 0.963$

Table 5.1: Variances / MSEs/minimum MSEs of different Estimators

Estimators	MSE	PRE
$V(p)$	0,061122	100
$MSE(t_S)$	0,32271	189,3812
$MSE_{\min}(t_{GS})$	0,01190	511,7912
$MSE_{\min}(t_{NS})$	0,01171	518,9214
$MSE_{\min}(t_N)$	**0,00329**	**1856,8818**
$MSE_{\min}(t_{N1})$	0,01682	362,8112
$MSE_{\min}(t_{N2})$	0,00881	687,2571
$MSE_{\min}(t_{N3})$	0,01191	511,7912
$MSE_{\min}(t_{N4})$	0,02801	216,3089
$MSE_{\min}(t_{N5})$	0,00881	687,2763
$MSE_{\min}(t_{N6})$	0,02821	216,3019
$MSE_{\min}(t_{N7})$	0,01681	362,8229

$MSE_{min}(t_{N8})$	**0,00329**	**1856,8818**
$MSE_{min}(t^{1}_{NQ})$	0,00636	960,8345
$MSE_{min}(t^{2}_{NQ})$	0,00631	963,0277
$MSE_{min}(t^{3}_{NQ})$	0,00744	820,9345
$MSE_{min}(t^{4}_{NQ})$	0,00621	983,6847
$MSE_{min}(t^{5}_{NQ})$	0,02211	276,3287
$MSE_{min}(t^{6}_{NQ})$	0,00622	982,1553
$MSE_{min}(t^{7}_{NQ})$	0,01245	490,7537
$MSE_{min}(t^{8}_{NQ})$	0,00151	812,9560
$MSE_{min}(t^{9}_{NQ})$	0,02521	242,0966

Table 5.1 exhibits variance, mean square errors and percent relative efficiencies of the existing estimators p, t_S, t_{GS} , t_{NS} and proposed generalised class of estimators along with its different members. Analysing table 5.1 we conclude that the estimators based on auxiliary information are more efficient than the one which does not use the auxiliary information as p .The members t^{1}_{NQ}, t^{2}_{NQ}, t^{4}_{NQ} and t^{6}_{NQ} of the class of estimators t_{NQ} ,obtained from generalized class of estimators t_N , are almost equally efficient but more than the usual unbiased estimator p , usual ratio estimator t_S , class of estimators t_{GS} and family of estimators t_{NS} . Among the proposed class of estimators t_N and t^{j}_{NQ} (j=1,2,...9) the performance of the estimator t_N , which is equal efficient to the estimator t_{N8} , is best in the sense of having the least MSE (maximum PRE) followed by the estimator t^{4}_{NQ} and t^{6}_{NQ} which utilize the information of population mean $\bar{X}$ and correlation coefficient ρ .

We evaluated the proposal derived in this paper using a population of persons with Dengue. The physician diagnosed whether they were affected by the disease or not. The correct classification was the qualitative variable. The study variable was the number of days of permanence in a hospital (X). The values of the required parameters are:

$N = 152, n = 15, \bar{X} = 7.32 \quad C_{\phi} = 0.259, C_x = 0.114, R = 18{,}92, \lambda_{12} = 0{,}0195, \lambda_{04} = 1{,}44, \lambda \quad = ,806.$

The results are given in the Table 5.2. The conclusion derived from them are qualitatively similar to those derived from Table 5.1.

Table 5.2: Variances / MSEs/minimum MSEs of different Estimators. A real life case

Estimators	MSE	PRE
V(p)	10,60	1,00000
$MSE(t_S)$	7,93	0,74811
$MSE_{min}(t_{NS})$	7,46	0,00704
$MSE_{min}(t_N)$	1,21	0,11415
$MSE_{min}(t_{N1})$	8,45	0,79717
$MSE_{min}(t_{N2})$	7,40	0,69811
$MSE_{min}(t_{N3})$	4,63	0,00437

MSE_{min} (t_{N4})	4,81	0,00454
MSE_{min} (t_{N5})	4,96	0,00468
MSE_{min} (t_{N6})	5,75	0,54245
MSE_{min} (t_{N7})	8,32	0,78491
MSE_{min} (t_{N8})	1,21	0,00114
MSE_{min} (t^1_{NQ})	3,41	0,00322
MSE_{min} (t^2_{NQ})	3,31	0,00312
MSE_{min} (t^3_{NQ})	3,44	0,00325
MSE_{min} (t^4_{NQ})	2,49	0,00235
MSE_{min} (t^5_{NQ})	5,23	0,00493
MSE_{min} (t^6_{NQ})	4,44	0,41887
MSE_{min} (t^7_{NQ})	8,38	0,79057
MSE_{min} (t^8_{NQ})	5,49	0,00518
MSE_{min} (t^9_{NQ})	8,42	0,00794

6. CONCLUSIONS

In this article we have suggested a generalized class of estimators for the population mean when study variable is qualitative in nature using auxiliary information in simple random sampling without replacement (SRSWOR). In addition, some known estimators of population mean when study variable is an attribute such as usual unbiased estimator p, ratio estimator due to Singh et al.(2010) , Srivastava (1967) type estimator , Murthy (1964) type estimator, Al and Chingi (2009) and Singh and Solanki (2013) type estimators are found to be members of the proposed generalized class of estimators. Some new members are also generated from the proposed generalized class of estimators using auxiliary information. We have determined the biases and mean square errors of the proposed class of estimators up to the first order of approximation. The proposed generalized class of estimators are advantageous in the sense that the properties of the estimators, which are members of the proposed class of estimators, can be easily obtained from the properties of the proposed generalized class. Thus the study unifies properties of several estimators for population mean when study variable is qualitative in nature. In theoretical and empirical efficiency comparisons, it has been shown that the proposed generalized class of estimators t_N are more efficient than the estimators considered here.

Acknowledgments: This paper was benefited by the support of some authors from a project PNCB on Mathematical modelling for Environment and Campus Costa Chica UAGro.

RECEIVED: JUNE, 2015
REVISED: OCTOBER , 2015

REFERENCES

[1] ABD-ELFATTAH, A.M. EL-SHERPIENY, E.A. MOHAMED, S.M. and ABDOU, O. F. (2010): Improvement in estimating the population mean in simple random sampling using information on auxiliary attribute. **Applied Mathematics and Computations,** 215, 4198-4202.

[2] ADICHWAL, N.K., SHARMA, P., VERMA, H.K. and SINGH, R. (2015): Generalized class of estimators for population variance using auxiliary attribute, **International Journal of Applied and Computational Mathematics,** 1(3), DOI: 10.1007/s40819-015-0073-3.

[3] AL, S. and CINGI, H. (2009): New estimators for the population median in simple random sampling. **Tenth Islamic Countries Conference on Statistical Sciences**, held in New Cairo, Egypt.
[4] GUJARAT, D. N., SANGEETHA, (2007): Basic Econometrics. Tata McGraw – Hill, USA.
[5] JHAJJ, H. S., SHARMA, M. K. and GROVER, L. K.,(2006): A family of estimators of population mean using information on auxiliary attribute. Pak. J. Statist., 22, 43-50.
[6] NAIK,V.D. and GUPTA, P.C., (1996): A note on estimation of mean with known population proportion of an auxiliary character. Jour. Ind. Soc. Agr. Stat., 48,151-158.
[7] MURTHY, M. N. (1964): Product method of estimation. **Sankhya** 26,294–307.
[8] SHARMA, P. SANAULLAH, A., VERMA, H. and SINGH, R. (2013a): Some Exponential Ratio-Product Estimators using information on Auxiliary Attributes under Second Order Approximation. **International Journal of Statistics and Economics,** 12, 57-66.
[9] SHARMA, P., SINGH, R. and JONG, MIN-KIM. (2013b): Study of Some Improved Ratio Type Estimators Using information on Auxiliary Attributes under Second Order Approximation. **Journal of Scientific Research,** 57, 138-146.
[10] SHARMA, P. and SINGH, R. (2015): Generalized class of estimators for population median using auxiliary information. **Hacettepe Journal of Mathematics and Statistics**, 44, 443-453.
[11] SHABBIR, J. and GUPTA, S.,(2007): On estimating the finite population mean with known population proportion of an auxiliary variable, **Pakistan Journal of Statistics** 23, 1–9.
[12] SINGH, H.P. and SOLANKI, R. S.,(2012): Improved estimation of population mean in simple random sampling using information on auxiliary attribute. **Applied Mathematics and Computation**. 218, 7798–7812.
[13] SINGH, H.P. and SOLANKI, R. S., (2013): Some Classes of estimators for the Population Median Using Auxiliary Information. **Com. in Stat**. 42, 4222-4238.
[14] SINGH, R., CHAUHAN, P., SAWAN, N. and SMARANDACHe,F. (2008): Ratio estimators in simple random sampling using information on auxiliary attribute. **Pak. J. Stat. Oper. Res**. 4, 47-53.
[15] SINGH, R**.,** KUMAR, M. and SMARANDACHE, F., (2010): Ratio estimators in simple random sampling when study variable is an attribute**. World Applied Sciences Journal** 11, 586-589.
[16] SRIVASTAVA, S. K. (1967): An estimator using auxiliary information in sample surveys. **Calcutta Statist. Assoc. Bull**. 6,121–132.
[17] UPADHYAYA, L. N. and SINGH, H. P., (1999) Use of transformed auxiliary variable in estimating the finite population mean. **Biom. Jour**., 41, 627-636.
[18] VERMA H.K., SHARMA, P. and SINGH, R. (2015): Some Families of Estimators Using Two auxiliary variables in stratified random sampling. **Revista Investigacion Operacional**, 36, 140-150.