International Journal of Current Advanced Research
ISSN: O: 2319-6475, ISSN: P: 2319-6505, Impact Factor: 6.614
Available Online at www.journalijcar.org
Volume 10; Issue 03 (B); March 2021; Page No.24031-24035
DOI: http://dx.doi.org/10.24327/ijcar.2021.24035.4762

Research Article

AN IMPROVED CLASS OF REGRESSION ESTIMATORS FOR FINITE POPULATION MEAN IN THREE-PHASE SAMPLING USING TWO AUXILIARY VARIABLES

Priyaranjan Dash* and Bishnupriya Behera

Department of Statistics, Utkal University, Vani Vihar, Bhubaneswar, Odisha – 751004, INDIA

A R T I C L E I N F O

Article History:
Received 13th December, 2020
Received in revised form 11th January, 2021
Accepted 8th February, 2021
Published online 28th March, 2021

Key words:

Multi-phase sampling, regression estimator, bias, mean square error (MSE)

A B S T R A C T

In this paper, we have suggested two different classes of regression-type estimators in two phase sampling as well as in three phase sampling using SRSWOR scheme at all the phases. We have seen that one of the suggested class of estimator is more efficient than some existing estimators as it has a minimum mean square error in three phase sampling.

INTRODUCTION

The auxiliary information, which is available or can be made available by utilizing a small part of the survey cost, can be effectively exploited at the estimation stage in order to improve the efficiency of the estimator (see Watson, 1937 and Cochran, 1942). Tamhane (1978) discussed the problem of hypothesis testing using the regression estimator in double sampling and also discussed the problem of choosing between the double sample regression estimator and the single sample mean estimator. Srivenkataramana (1978) pointed out a cost-effective method of reducing bias of the regression estimator by using suitable transformation. Rao (1987), Bisht and Sisodia (1990), Naik and Gupta (1991), Updhyaya and Singh (1999) and Singh and Tailor (2005a,b) have remarkable contributions in improving the classical regression estimator. Shukla (1965) for multivariate regression estimator for estimating finite population mean.

Swain (2012), Khare *et. al*. (2013), Singh and Majhi (2014), Khan (2016) proposed different modifications to usual regression estimator under different conditions. When the population mean of the auxiliary variable $\bar{X}$ is not known to us, the use of double sampling procedure was studied by Bose (1943) and Cochran (1963). Dellhouse and Joshi (1984) studied the admissibility property of the classical regression estimator under double sampling scheme. Sahoo *et. al*. (1993) suggested that even if population mean of the auxiliary variable $\bar{X}$ is unknown, we can utilize the information on a cheaply ascertainable variable z, closely related to x but compared to x remotely related to y, is available on all units of the population and suggested some ratio-in-regression and regression-in-regression estimators. They have also revisited some other similar estimators for estimating the finite population mean.

*Corresponding author: **Priyaranjan Dash**
Department of Statistics, Utkal University, Vani Vihar, Bhubaneswar, Odisha – 751004, INDIA

Regression Estimator in Two Phase Sampling

Consider a finite population with N distinct and identifiable units with y and x be the study variable and auxiliary variable taking value y_i and x_i for the i^{th} unit of the population. The classical regression estimator assumes the knowledge of population mean $\bar{X}$ of the auxiliary variable which is not sometimes available to us in advance. In such cases, in order to take the advantage of auxiliary information x, we use double sampling or two-phase sampling method. Here, we select a large preliminary sample S' of size n' from N units of the population by SRSWOR and study only x variable which require a very little cost. This sample is known as the first phase sample. From this selected first phase sample, we select a second phase sample S of size $n, (n < n')$, using SRSWOR scheme and study both y and x. The classical regression estimator for estimating population mean $\bar{Y}$ of y is given by

$$t_1 = \bar{y}_n + b_{yx}(\bar{x}_{n'} - \bar{x}_n) \quad (1)$$

where $\bar{y}_n$, $\bar{x}_n$ are the sample means of y and x respectively and $b_{yx} = \frac{s_{yx}}{s_x{}^2}$ is the sample regression coefficient of y on x calculated on the basis of second phase sample S and $\bar{x}_{n'}$ is the sample mean of x basing on the units of first phase sample S'. The mean square error (MSE) of this estimator upto first order of approximation is given by

$$M_1 = M(t_1) = \theta\left(1-\rho^2{}_{yx}\right){S_y}^2 + \theta_1\rho^2{}_{yx}{S_y}^2 \quad (2)$$

Where $\theta = \frac{1}{n} - \frac{1}{N}$, $\theta_1 = \frac{1}{n'} - \frac{1}{N}$, ${S_y}^2 = \frac{1}{N-1}\sum_{i=1}^{N}(y_i - \bar{Y})^2$ is the population mean square of y and ρ_{yx} is the population correlation coefficient between y and x. The notation used for the variables x and z are similar to those of y variable.

We shall consider a case when the Sample Means for the Auxiliary Variables are Known. When the sample means $\bar{x}_n$ and $\bar{x}_{n'}$ are available but the units drawn into the two samples are unknown, Rao (1972) suggested an alternative to regression estimator in double sampling as a class of estimator of the form

$$t_2 = \bar{y}_n + \beta_{yx}(\bar{x} - \bar{x}_n) \quad (3)$$

Where $\bar{x} = a_1\bar{x}_{n'} + a_2\bar{x}_n$ is used as a class of estimators for $\bar{X}$. Here $a_1$and $a_2$are chosen so that the estimator t_3 is unbiased for $\bar{Y}$ and $x_{n'}$ is independent of $\bar{x}_n$ and $\bar{y}_n$. Following Rao (1972), when the sample means $x_{n'}$, $z_{n'}$ for the first phase and $\bar{x}_n, \bar{z}_n$ for the second phase samples of the two auxiliary variables x and z are known and the units drawn into the two samples are not known, we consider, a class of estimators for $\bar{X}$ and $\bar{Z}$ as

$$\bar{x} = a_1\bar{x}_{n'} + a_2\bar{x}_n \ , \bar{z} = b_1\bar{z}_{n'} + b_2\bar{z}_n \quad (4)$$

Where a_1, a_2, b_1, b_2 are four constants so that the class of estimator st_{c1} given by

$$t_{c1} = \bar{y}_n + \beta_{yx.z}(\bar{x} - \bar{x}_n) + \beta_{yz.x}(\bar{z} - \bar{z}_n) \quad (5)$$

will be an unbiased class of estimators for $\bar{Y}$. Since, $\bar{y}_n$ is an unbiased estimator of $\bar{Y}$; $\bar{x}_{n'}$ and $\bar{x}_n$ an unbiased estimator of $\bar{X}$ and $\bar{z}_{n'}$ and $\bar{z}_n$ are unbiased estimator of $\bar{Z}$, so the estimator t_{c1} for $\bar{Y}$ to the first order of approximation. Thus, the estimators $\bar{x}$, $\bar{z}$ and t_3 are unbiased for $\bar{X}$, $\bar{Z}$ and $\bar{Y}$ respectively if $a_1 + a_2 = 1$ and $b_1 + b_2 = 1$. Define regression-type estimator of population mean Y bar (in equation form)

$$t_{c1} = \bar{y}_n + a_1\beta_{yx.z}\left(\bar{x}_{n'} - \bar{x}_n\right) + b_1\beta_{yz.x}\left(\bar{z}_{n'} - \bar{z}_n\right) \quad (6)$$

First of all we are required to find a_1 and b_1 for which $M(t_{c1})$ gives the minimum value. Now,

$$\begin{aligned} M_{c1} = V(t_{c1}) &= M_1E_2(t_3) + E_1M_2(t_3) \\ &= M_1(\bar{y}_{n'}) \\ &+ E_1M_2(\bar{y}_n - a_1\beta_{yx.z}\bar{x}_n \\ &- b_1\beta_{yz.x}\bar{z}_n) \\ &= \theta{S_y}^2 + {a_1}^2{\theta_1}'{\beta_{yx.z}}^2{S_x}^2 \\ &+ {b_1}^2{\theta_1}'{\beta_{yz.x}}^2{S_z}^2 \\ &- 2a_1{\theta_1}'\beta_{yx.z}\rho_{yx}S_yS_x \\ &- 2b_1{\theta_1}'\beta_{yz.x}\rho_{yz}S_yS_z \\ &+ 2a_1b_1{\theta_1}'\beta_{yx.z}\beta_{yz.x}\rho_{xz}S_xS_z \end{aligned} \quad (7)$$

Thus $\frac{\partial M_{c1}}{\partial a_1} = 0$ and $\frac{\partial M_{c1}}{\partial a_2} = 0 \Rightarrow a_1 = 1, b_1 = 1$. Therefore the estimator t_{c1} defined in (6) reduces to

$$t_{c1(opt)} = \bar{y}_n + \beta_{yx.z}\left(\bar{x}_{n'} - \bar{x}_n\right) + \beta_{yz.x}\left(\bar{z}_{n'} - \bar{z}_n\right) \quad (8)$$

so that $t_{c1(opt)}$defined in (8) has a minimum MSE of the class with minimum MSE given by

$$M_{c1} = V\left(t_{c1(opt)}\right) = \theta\left(1 - {\rho_{y.xz}}^2\right){S_y}^2 + \theta_1{\rho_{y.xz}}^2{S_y}^2 \quad (9)$$

For large population size i.e., for large N, it can be written as

$$M_{c1} = V\left(t_{c1(opt)}\right) \approx \left(1 - {\rho_{y.xz}}^2\right)\frac{{S_y}^2}{n} + {\rho_{y.xz}}^2\frac{{S_y}^2}{n'} \quad (10)$$

The estimator $t_{c1(opt)}$is much simpler to use in the sense that its variance only depends upon the square of multiple correlation coefficient ${\rho_{y.xz}}^2$ of the study variable y on the two auxiliary variables x and z. If the partial regression coefficients of y on x and z given by $\beta_{yx.z}$ and $\beta_{yz.x}$are unknown, we can use them in place of their estimates $b_{yx.z}$ and $b_{yz.x}$ in equation (8). In this case the class of estimator becomes

$$t_{c1}{}' = \bar{y}_n + b_{yx.z}(\bar{x} - \bar{x}_n) + b_{yz.x}(\bar{z} - \bar{z}_n) \quad (11)$$

To the first order of approximation, its MSE is same as M_{c1}given in equation (10).

Case of Three-Phase Sampling

In this case, we select a large preliminary sample, known as *first phase sample*, of size n' from the population of size N and again a *second phase sample* of size n'' is selected from this first phase sample and at the last a third phase sample of size n is selected from this second phase sample in order to observe the study variable y in order to estimate $\bar{Y}$. We use SRSWOR sampling scheme at each phase of selection of the sample. Mukerjee *et. al.* (1987) have proposed an estimator of the population mean $\bar{Y}$ is given by

$$\begin{aligned} t_3 = \bar{y}_n + b_{yx}\left(\bar{x}_{n'} - \bar{x}_n\right) + b_{yx}b_{xz}\left(\bar{z}_{n''} - \bar{z}_{n'}\right) \\ + b_{yz}\left(\bar{z}_{n''} - \bar{z}_n\right) \end{aligned} \quad (12)$$

whose MSE is given by

$$\begin{aligned} M_3 = M(t_3) = \theta\left(1 - {\rho_{y.xz}}^2\right){S_y}^2 \\ + \theta_1\left(1 - {\rho_{yz}}^2\right){\rho_{yx.z}}^2{S_y}^2 \\ + \theta_2{\rho_{yz}}^2{S_y}^2 \end{aligned} \quad (13)$$

Where $\theta_2 = \frac{1}{n''} - \frac{1}{N}$, $\theta_2 = \frac{1}{n''} - \frac{1}{N}$, ${\theta_2}' = \frac{1}{n''} - \frac{1}{n'} = \theta_2 - \theta_1$, $\theta' = \frac{1}{n} - \frac{1}{n''} = \theta - \theta_2$.

Proposed Generalized Estimator

A more generalized estimator of $\bar{Y}$ can be written as

$$\begin{aligned} t_{ge} = \bar{y}_n + \lambda_1\bar{x}_{n'} + \lambda_2\bar{x}_{n''} + \lambda_3\bar{x}_n + \lambda_4\bar{z}_{n'} + \lambda_5\bar{z}_{n''} \\ + \lambda_6\bar{z}_n \end{aligned} \quad (14)$$

The expected value of the estimator t_{ge} is given by

$$\begin{aligned} E\left(t_{ge}\right) = E_1E_2E_3\left(t_{ge}\right) \\ &= E_1E_2\left[\bar{y}_{n''} + \lambda_1\bar{x}_{n'} + \lambda_2\bar{x}_{n''} \right. \\ &\left. + \lambda_3\bar{x}_{n''} + \lambda_4\bar{z}_{n'} + \lambda_5\bar{z}_{n''} + \lambda_6\bar{z}_{n''}\right] \\ &= E_1\left[\bar{y}_{n'} + \lambda_1\bar{x}_{n'} + \lambda_2\bar{x}_{n'} + \lambda_3\bar{x}_{n'} \right. \\ &\left. + \lambda_4\bar{z}_{n'} + \lambda_5\bar{z}_{n'} + \lambda_6\bar{z}_{n'}\right] \\ &= E_1\left[\bar{y}_{n'} + (\lambda_1 + \lambda_2 + \lambda_3)\bar{x}_{n'} \right. \\ &\left. + (\lambda_4 + \lambda_5 + \lambda_6)\bar{z}_{n'}\right] \end{aligned} \quad (15)$$

The estimator t_{ge} will be an unbiased estimator of $\bar{Y}$, iff $\lambda_1 + \lambda_2 + \lambda_3 = 0$ and $\lambda_4 + \lambda_5 + \lambda_6 = 0$. Hence the regression estimatort_{ge}takes the form

$$t_{ge} = \bar{y}_n + \lambda_2(\bar{x}_{n''} - \bar{x}_{n'}) + \lambda_3(\bar{x}_n - \bar{x}_{n'}) + \lambda_5(\bar{z}_{n''} - \bar{z}_{n'}) + \lambda_6(\bar{z}_n - \bar{z}_{n'}) \quad (16)$$

Now, the variance of this estimator is given by

$$\mathrm{Var}(t_{ge}) = V_1E_2E_3(t_{ge}) + E_1V_2E_3(t_{ge}) + E_1E_2V_3(t_{ge}) \quad (17)$$

$$V_1E_2E_3(t_{ge}) = Var(\bar{y}_{n'}) = \theta_1 S_y{}^2 \quad (18)$$

$$E_1V_2E_3(t_{ge}) = \theta_2{}'[S_y{}^2 + (\lambda_2+\lambda_3)^2 S_x{}^2 + (\lambda_5+\lambda_6)^2 S_z{}^2 + 2(\lambda_2+\lambda_3)\rho_{yx}S_yS_x + 2(\lambda_5+\lambda_6)\rho_{yz}S_yS_z + 2(\lambda_2+\lambda_3)(\lambda_5+\lambda_6)\rho_{xz}S_xS_z] \quad (19)$$

$$E_1E_2V_3(t_{ge}) = \theta'[S_y{}^2 + \lambda_3{}^2S_x{}^2 + \lambda_6{}^2S_z{}^2 + 2\lambda_3\rho_{yx}S_yS_x + 2\lambda_6\rho_{yz}S_yS_z + 2\lambda_3\lambda_6\rho_{xz}S_xS_z] \quad (20)$$

Thus, by using equations (18,19 and 20) in equation (17), we get the variance of the estimator t_{ge}as

$$\mathrm{Var}(t_{ge}) = \theta_1 S_y{}^2 + \theta_2{}'[(\lambda_2+\lambda_3)^2S_x{}^2 + (\lambda_5+\lambda_6)^2S_z{}^2 + 2(\lambda_2+\lambda_3)\rho_{yx}S_yS_x + 2(\lambda_5+\lambda_6)\rho_{yz}S_yS_z + 2(\lambda_2+\lambda_3)(\lambda_5+\lambda_6)\rho_{xz}S_xS_z] + \theta'[\lambda_3{}^2S_x{}^2 + \lambda_6{}^2S_z{}^2 + 2\lambda_3\rho_{yx}S_yS_x + 2\lambda_6\rho_{yz}S_yS_z + 2\lambda_3\lambda_6\rho_{xz}S_xS_z] \quad (21)$$

Now, in order to find the optimum values of λ_i 's (i = $2,3,5,6$), that give the minimum variance of t_{ge}, we differentiate equation (21) with respect toλ_i 's (i = $2,3,5,6$) partially and equating to zero, we get

$$(\lambda_2+\lambda_3)S_x{}^2 + (\lambda_5+\lambda_6)\rho_{xz}S_xS_z = -\rho_{yx}S_yS_x \quad (22)$$

$$\lambda_3S_x{}^2 + \lambda_6\rho_{xz}S_xS_z = -\rho_{yx}S_yS_x \quad (23)$$

$$(\lambda_5+\lambda_6)S_z{}^2 + (\lambda_2+\lambda_3)\rho_{xz}S_xS_z = -\rho_{yz}S_yS_z \quad (24)$$

$$\lambda_6S_z{}^2 + \lambda_3\rho_{xz}S_xS_z = -\rho_{yz}S_yS_z \quad (25)$$

Solving these equations (22,23,24) and (25), we get

$$\lambda_2 = 0 = \lambda_5 \,, \lambda_3 = -\beta_{yx.z} \,, \lambda_6 = -\beta_{yz.x} \quad (26)$$

Hence the unbiased minimum variance regression estimator of $\bar{Y}$ is given by

$$t_{ge(opt.)} = \bar{y}_n + \beta_{yx.z}\,(\bar{x}_{n''} - \bar{x}_n) + \beta_{yz.x}\,(\bar{x}_{n''} - \bar{z}_n) \quad (27)$$

and the minimum variance is given by

$$Var(t_{ge(opt.)}) = \theta\,(1 - \rho_{y.xz}{}^2)S_y{}^2 + \theta_1\rho_{y.xz}{}^2S_y{}^2 \quad (28)$$

For large N, we can have

$$Var(t_{ge(opt.)}) \approx (1 - \rho_{y.xz}{}^2)\frac{S_y{}^2}{n} + \rho_{y.xz}{}^2\frac{S_y{}^2}{n'} \quad (29)$$

which depends only upon the first and third phase sample sizes n and n' but independent of the sample size of second phase sample n''. The optimum variance of the estimator t_{ge} does not increase with addition of another phase for selecting the sample.

Remark 1. At this stage, it is a point to be emphasized that the estimator t_{ge} defined in equation (16) depends upon the sample values at all the three phases i.e., $\bar{x}_{n'}$, $\bar{z}_{n'}$as sample means for the first phase sample of n' units; sample means $\bar{x}_{n''}$, $\bar{z}_{n''}$ of second phase sample of n'' units and sample means $\bar{y}_n, \bar{x}_n, \bar{z}_n$of third phase sample of n units.

Table 1 Description of the Population and their Sources

Pop No.	Source	y	x	z
1	Rawlings *et. al*. (1998), p.98.	Age of Patients	Average number of days survival of 10 control patients.	Number of days survival after date of untreatability
2	Rawlings *et. al*. (1998), p.124.	Oxygen uptake rate (ml per kg body weight per minute)	Time to run 1.5 miles (minutes)	Heart rate while running
3	Rawlings *et. al*. (1998), p.163.	Aerial biomass of the substrate	pH-Value of the substrate	Zinc content of the substrate
4	Gujarati (2004), p.235-236.	Quantity of roses sold, dozens	Average wholesale price of roses, $/dozen	Average wholesale price of carnations, $/dozen
5	Gujarati (2004), p.237-238.	Defense budget-outlay for a year, $ billions	GNP for same year , $ billions	U.S. military sales/assistance in same year, $ billions
6	Gujarati (2004), p.238-239.	Per capita consumption of chickens, lb	Real disposable income per capita, $	Real retail price of chicken per lb, ¢.
7	Gujarati (2004), p.238-239.	Billions of Drachmas at constant 1970 prices	Thousands of workers per year.	Amount Invested
8	Gujarati (2004), p.238-239.	Billions of Drachmas at constant 1970 prices	Thousands of workers per year	Capital to Labor Ratio
9	Chaterjee and Hadi (2006), p.55-56.	Scores in the Final	Scores in Second Preliminary	Scores in First Preliminary
10	Chaterjee and Hadi (1988), p.207-208.	infant deaths per 1,000 live births	Number of inhabitants per physician	Gross national product per capita, 1957 U.S. dollars

Table 2 Value of Different Population Parameters

P. No.	N	$S_y{}^2$	ρ_{yx}	ρ_{yz}	ρ_{xz}	$\rho_{yx.z}{}^2$	$\rho_{y.xz}{}^2$
1	17	124.191	−0.568	−0.226	−0.049	−0.595	0.355
2	31	28.379	−0.862	0.314	−0.398	−0.846	0.761
3	45	435703.709	0.774	−0.722	−0.624	0.598	0.608
4	16	4173090.4	−0.784	0.472	−0.023	−0.878	0.771
5	20	839.708	0.952	0.897	0.801	0.882	0.921
6	23	54.360	0.947	0.932	0.840	0.835	0.911
7	27	1382.729	0.947	0.955	0.989	0.062	0.979
8	27	1382.729	0.947	0.997	0.943	0.295	0.898
9	22	124.338	0.927	0.884	0.896	0.652	0.886
10	49	1263.437	0.568	−0.484	−0.53	0.42	0.408

Table 3 Sample Sizes and MSE of Different Estimators under Two Phase Sampling

P.No.	n'	n	M_1	M_2	M_3	M_{C1}	M_{C2}
1	10.	3	24.757	24.847	24.926	23.857	23.801
2	20	5	1.596	1.583	1.605	2.196	1.519
3	25	6	29856.589	36639.687	29925.653	50893.841	29364.510
4	10	3	531472.377	552878.464	555845.7	379972.28	379816.144
5	15	3	34.915	29.004	35.111	175.548	31.78
6	16	4	2.083	1.528	2.097	9.134	1.943
7	18	5	46.073	19.149	48.164	225.318	29.764
8	18	5	46.073	23.552	46.090	223.396	45.943
9	12	3	9.06	5.546	9.184	33.181	8.243
10	30	8	94.746	100.424	96.132	117.454	84.92

Table 4 Sample Sizes and Variance of Different Estimators under Three Phase Sampling

P.No.	n'	n''	n	$M_3 = M_{C2}$	$M_{ge} = M_{C1}$
1	10	5	3	23.833	23.801
2	20	10	5	1.743	1.519
3	25	12	6	36724.064	29364.510
4	10	6	3	380014.767	379816.144
5	15	8	3	63.223	31.780
6	16	10	4	3.381	1.943
7	18	11	5	77.627	29.764
8	18	11	5	89.378	45.943
9	12	6	3	16.556	8.243
10	30	20	8	90.835	84.92

But, under the optimal conditions, the estimator $\mathrm{t_{ge}}$becomes the MVU estimator $\mathrm{t_{ge(opt.)}}$defined in equation (27), which variance given in equation (7) is independent of the units of the second phase sample of size $\mathrm{n''}$but depends only upon the sizes $\mathrm{n'}$and n of the first and third phase samples respectively.

Remark 2. The variance of the estimator $\mathrm{t_{ge}}$for three phase sampling is same as the corresponding estimator $\mathrm{t_{c1}}$case of two phase sampling, when both are in their optimum cases i.e., $\mathrm{Var}(\mathrm{t_{ge(opt.)}})$in three phase sampling is equal to $\mathrm{Var}(\mathrm{t_{c1(opt.)}})$in two phase sampling given in Section 2.

Simulation Study

The performance of proposed classes of estimators t_{c1}, are studied along with some competitive estimators. We have considered 10 different natural populations available from different text books for the comparison between these estimators. Table 1 gives the description of these populations and Table 2 gives the values of the population size (N), population mean square of y values $S_y{}^2$and the simple correlation coefficients between y, x *and* zi. e.; $\rho_{yx}, \rho_{xz}, \rho_{yz}$,partial correlation coefficient between y and $x, i.e.;$ $\rho_{yx.z}$and the multiple correlation coefficient of y on x and z, i.e.; $\rho_{y.xz}$. Table 3 gives the first phase and second phase sample sizes (n' *and* n),mean square error of these estimators. The proposed class of estimator t_{c2} in its optimum case i.e., $t_{c2(opt.)}$is equivalent to the estimator t_3 proposed by Mukerjee *et. al*. (1987) as both the estimators are unbiased and have same mean square error (MSE). The proposed class of estimator t_{ge}in its optimum case dominates the class t_{c1}.The proposed class of estimator t_{ge} in its optimum case, i.e., the estimator$t_{ge(opt.)}$ is independent of the number of units in the second phase and its size n''.

CONCLUSION

In the present paper, we have discussed some regression type estimators using two auxiliary variables under both two phase and three phase sampling schemes using SRSWOR at all the phases. We have found that the estimator suggested Mukerjee *et. al*. (1987), having same mean square error upto first order of approximation, performs better under two phase sampling among several competitive estimators. The suggested classes t_{c1} and t_{c2} have same mean square errors. The proposed class of estimator t_{ge} not only dominates all these estimators (t_3, t_{c2}) but also it becomes independent of the second phase sample size n'' as its variance does not increase with addition of second phase after first phase for selecting the third phase sample.

References

Bisht, K. K. S., Sisodia, B. V. S. (1990). Efficient estimators of mean of finite populations with known coefficient of variation. *Journal of Indian Society of Agricultural Statistics*, 42(1), pp. 131-139.

Bose, C. (1943). Note on the sampling error in the method of double sampling. Sankhya: *The Indian Journal of Statistics,* 6, pp. 329-330.

Cochran, W. G. (1963). *Sampling Techniques*, John Wiley and Sons, New York

Cochran W. G. (1942). Sampling Theory When the Sampling-Units are of Unequal Sizes, *Journal of the American Statistical Association*, 37(218), pp.199-212.

Dellhouse, D. R. and Joshi, V.M. (1984). On the Admissibility of the Regression Estimator, *Journal of the Royal Statistical Society. Series B (Methodological)*, 46(2), pp. 268-269.

Khan, M. (2016). A ratio chain-type exponential estimator for finite population mean using double sampling. *Springer Plus* 5, pp. 1–9.

Khare, B.B., Srivastava, U., Kumar, K. (2013). A generalized chain ratio in regression estimator for population mean using two auxiliary characters in sample survey, *Jour Sci Res Banaras Hindu Univ Varanasi,*57, pp. 147–153.

Mukerjee, R., Rao, T. J. and Vijayan, K. (1987) Regression type estimators using multiple auxiliary information, *Australian Journal of Statistics*, 29(3), pp. 244–254.

Naik, V. D., Gupta, P. C., 1991. A general class of estimators for estimating population mean using auxiliary information. *Metrika*, 38, pp. 11-17.

Rao, P. S. R. S. (1972). On Two Phase Regression Estimator, *Sankhya: The Indian Journal of Statistics*, Series A, 34(4), pp. 473-476.

Rao, P.S.R.S., (1987). Ratio and regression estimates with sub sampling the non-respondents. Paper presented at a special contributed session of the *International Statistical Association Meeting, Sept*., 2-16, Tokyo, Japan.

Sahoo, J., Sahoo, L. N., Mohanty, S. (1993). A regression approach to estimation in two phase sampling using two auxiliary variables. *Current Science*, 65(1), pp. 73-75.

Shukla, G. K. (1965). Multivariate regression estimate, *Journal of Indian Statistical Association*, 3, pp. 202–211.

Singh, G.N., Majhi, D. (2014). Some chain-type exponential estimators of population mean in two-phase sampling. *Statistical Transitions*, 15(2), pp. 221–230.

Singh, H. P., Tailor, R. (2005). Estimation of finite population mean with known coefficient of variation of an auxiliary character. *Statistica*, LXV(3), pp. 407-418.

Singh, H. P., Tailor, R.(2005). Estimation of finite population mean using known correlation coefficient between auxiliary characters. *Statistica*, LXV(3), pp. 301- 313.

Srivenkataramana, T. (1978).Change of Origin and Scale in Ratio and Difference Methods of Estimation in Sampling, *The Canadian Journal of Statistics*, 6, pp. 79-86.

Swain, A.K.P.C. (2012). On Classes of Modified Ratio type and Regresion-cum-Ratio type estimators in Sample Surveys using Two Auxiliary Variables, *Statistics in Transition-new series*, 13(3), pp. 473—494.

Tamhane, A. C. (1978). Inference Based on Regression Estimator in Double Sampling, *Biometrika*, 65(2), pp. 419-427.

Upadhyaya, L. N., Singh, H. P. (1999). An estimator for population variance that utilizes the kurtosis of an auxiliary variable in sample surveys. *Vikram Mathematical Journal*, 19, pp. 14-17.

Watson, D. J. (1937). The estimation of leaf areas. *Journal of Agricultural Science,* 27, pp. 474.

How to cite this article:

Priyaranjan Dash and Bishnupriya Behera (2021) 'An Improved Class of Regression Estimators for Finite Population Mean in Three-Phase Sampling using Two Auxiliary Variables', *International Journal of Current Advanced Research,* 10(03), pp. 24031-24035. DOI: http://dx.doi.org/10.24327/ijcar.2021.24035.4762

****