African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 4, Issue 2, 2021 (pp. 41-51)
www.abjournals.org

EFFICIENT ESTIMATOR FOR POPULATION MEAN IN STRATIFIED DOUBLE SAMPLING IN THE PRESENCE OF NONRESPONSE USING ONE AUXILIARY VARIABLE

A.E. Anieting[1], E. I. Enang[2] and C. E. Onwukwe[2]

[1]Department of Statistics, University of Uyo, Uyo, Nigeria

[2]Department of Statistics, University of Calabar, Nigeria

Email: akaninyeneanieting@uniuyo.edu.ng.

Cite this article:

A.E. Anieting, E. I. Enang, C. E. Onwukwe (2021), Efficient Estimator for Population Mean in Stratified Double Sampling in the Presence of Nonresponse Using One Auxiliary Variable. African Journal of Mathematics and Statistics Studies 4(2), 40-50. DOI: 10.52589/AJMSS-YF4V11QV.

Manuscript History

Received: 28 April 2021

Accepted: 18 May 2021

Published: 22 May 2021

ABSTRACT: *A modified form of the population mean estimator suggested by Anieting and Enang (2020) in stratified double sampling in the presence of nonresponse using a single auxiliary variable has been proposed. The Mean Squared Error (MSE) and the bias of the proposed estimator have been given using large sample approximation. The empirical study shows that the MSE of the suggested estimator is more efficient than all other existing estimators in the same scheme. Determination of the optimal values of the first and second phases samples has also been done*

KEYWORDS: Efficiency, Nonresponse, Stratified Double Sampling

MSC: 62D05.

INTRODUCTION

Classical double sampling for stratification is a method that uses two random samples where the second sample is a stratified subsample of the first sample. If the problem of nonresponse is there then this subsample may be divided into classes of respondents and non-respondents (Cochran (1977)). It is resorted to when the auxiliary variable x needed for stratification of the population units is not available prior to sampling and the frequency distribution of x and the stratum weights are not known in advance. The presence of nonresponse distorts parameter estimation by increasing bias in estimates resulting in larger variances. In a well-designed research, handling nonresponse is important and several authors have tried to tackle this problem differently. According to Singer (2006) statisticians have been concerned mainly with imputation and weighting as ways of adjusting for the bias caused by nonresponse while social scientists have tended to focus on measuring, understanding and reducing the nonresponse rates themselves.

Hansen and Hurwitz (1946) suggested a method of sub-sampling of non-respondents so as to modify the non-response in their mail surveys. Khoshnevisan et al. (2007), Chaudhary et al. (2009), Chaudhary and Singh (2013), Chaudhary and Kumar (2015) and Anieting and Enang 2020) have all suggested different estimators in two-phase stratified sampling under nonresponse. In this article, a modified efficient ratio-product estimator in stratified double sampling under nonresponse using one auxiliary variable is been proposed. The characteristics of the estimator suggested have been given.

Sampling Design

Consider a population of size N divided into k strata. Let the size of the i^{th} stratum be N_i (i=1,2,….k) such that $\sum_i^k N_i$ = N. A large first sample of size n_i' is drawn from N_i units by (SRSWOR) scheme for the i^{th} stratum and auxiliary variable $\bar{x}_i'$ is observed to estimate the population mean $\bar{X}$, which is unknown. Secondly, a smaller second phase sample of size n_i is drawn from n_i' unit by SRSWOR. Let Y be the study variable with population mean $\bar{Y}$ $=\sum_i^k p_i\bar{Y}_i$ ($\bar{Y}_i$ being the mean of the i^{th} stratum based on N_i units and $p_i = \frac{N_i}{N}$) and assume that the non-response is observed on the study variable at second phase while the auxiliary variable is free from non-response. Also, assume that there are n_{i1} respondent units and n_{i2} non-respondent units in n_i units at the second phase. Applying Hansen and Hurwitz (1946) technique of sub-sampling of non-respondents, a sub-sample of $h_{i2} = \frac{n_{i2}}{L_i}$, $L_i \geq 1$ unit is selected from the sample of n_{i2} non-respondents and the information is collected from all of them.

Some Existing Estimators

Some existing estimators for population mean in stratified double sampling with single auxiliary variable in the presence of non-response shall be presented.

Khare and Srivastava (1997) Estimator

Khare and Srivastava (1997) proposed two ratio estimators given as

$$T_1^* = \bar{y}_{st}^* \frac{\bar{x}_{st}'}{\bar{x}_{st}}$$

Article DOI: 10.52589/AJMSS-YF4V11QV
DOI URL: https://doi.org/10.52589/AJMSS-YF4V11QV

with mean square error,

$$\text{MSE}\ (T_1^*) = \sum_i^k f_i' p_i^2 S_{Y_i}^2 + \sum_i^k f_i^* p_i^2 \left(S_{Y_i}^2 + R^2 S_{X_i}^2 - 2R\rho_i S_{X_i} S_{Y_i}\right) + \sum_i^k \frac{(L_i - 1)}{n_i} W_{i2} P_i^2 S_{Y_{i2}}^2$$

Where $f_i' = \left(\frac{1}{n_i'} - \frac{1}{N_i}\right)$, $f_i^* = \left(\frac{1}{n_i} - \frac{1}{n_i'}\right)$ and

$$T_2^* = \bar{y}_{st}^* \frac{\bar{x}_{st}}{\bar{x}_{st}'}$$

With mean square error,

$$\text{MSE}\ (T_1^*) = \sum_i^k f_i' p_i^2 S_{Y_i}^2 + \sum_i^k f_i^* p_i^2 \left(S_{Y_i}^2 + R^2 S_{X_i}^2 - 2R\rho_i S_{X_i} S_{Y_i}\right) + \sum_i^k \frac{(L_i - 1)}{n_i} W_{i2} P_i^2 S_{Y_{i2}}^2$$

Where $f_i' = \left(\frac{1}{n_i'} - \frac{1}{N_i}\right)$, $f_i^* = \left(\frac{1}{n_i} - \frac{1}{n_i'}\right)$

Chaudhary et al. (2009) Estimator

Chandhary et. al. (2009) suggested a family of combined-type estimators given as

$$T_C = \bar{y}_{st}^* \left[\frac{a\bar{X} + b}{\alpha(a\bar{X}_{st} + b) + (1-\alpha)(a\bar{X} + b)}\right]^g$$

Where $\bar{X}_{st} = \sum_i^k p_i \bar{x}_i$ and $\bar{X} = \sum_i^k p_i \bar{X}_i$; $\bar{x}_i$ and $\bar{X}_i$ are respectively mean based on n_i units and mean based on N_i in the i^{th} stratum for auxiliary variable

They gave mean square error as

$$\text{MSE}\ (T_C) = \sum_i^k f_i p_i^2 \left[S_{Y_i}^2 + \alpha^2 \lambda^2 g^2 R^2 S_{X_i}^2 - 2\alpha\lambda g R \rho_i S_{X_i} S_{Y_i}\right] + \sum_i^k \frac{(L_i - 1) W_{i2} P_i^2 S_{Y_{i2}}^2}{n_i}$$

Where $f_i = \left(\frac{1}{n_i} - \frac{1}{N_i}\right)$, $\lambda = \frac{a\bar{X}}{a\bar{X} + b}$,

Chaudhary et al. (2012) Estimator

Chaudhary et al. (2012) proposed a combined type estimator given as

$$T_{Fc}(\alpha) = \bar{y}_{st}^* \left[\frac{(A+C)\bar{X} + fB\bar{X}_{st}}{(A+B)\bar{X} + C\bar{X}_{st}}\right]$$

Where A = $(\alpha - 1)(\alpha - 2)$, B = $(\alpha - 1)(\alpha - 4)$, C = $(\alpha - 3)(\alpha - 2)(\alpha - 4)$, $\alpha > 0$, f = $\frac{n_i'}{N_i}$

The mean square error as obtained by them is

$$\text{MSE}\ (T_{Fc}(\alpha)) = \sum_i^k \left(\frac{1}{n_i} - \frac{1}{N_i}\right) P_i^2 \left[S_{Y_i}^2 + \emptyset(\alpha)^2 R S_{X_i}^2 - 2\emptyset(\alpha) R \rho_i S_{X_i} S_{Y_i}\right]$$

Article DOI: 10.52589/AJMSS-YF4V11QV
DOI URL: https://doi.org/10.52589/AJMSS-YF4V11QV

African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 4, Issue 2, 2021 (pp. 41-51)

Where $\emptyset(\alpha)=\frac{C-fB}{A+fB+C}$

Chandhary And Kumar (2015) Estimator

A combined type estimator was proposed by Chandhary and Kumar (2015) as follows:

$$T_c' = \bar{y}_{st}^* \left[\frac{a\bar{x}_{st}'+b}{\alpha(a\bar{x}_{st}+b)+(1-\alpha)(a\bar{x}_{st}'+b)}\right]^g$$

With mean square error given as

MSE (T_c') = $\sum_i^k f_i' p_i^2 S_{y_i}^2 + \sum_i^k f_i^* p_i^2 (S_{y_i}^2 + g^2\lambda^2 R^2 \alpha^2 S_{x_i}^2 - 2g\lambda R\alpha p_i S_{x_i} S_{y_i})$ + $\sum_i^k \left(\frac{L_i-1}{n_i}\right) W_{i2} p_i^2 S_{y_i}^2$

Anieting and Enang (2020) Estimator

A product-ratio estimator was suggested by Anieting and Enang (2020) as follows

$T_{ae}=\ \bar{y}_{st}^* \left(\frac{\bar{x}_{st}'+\varphi}{\bar{x}_{st}-\varphi}\right)$

Where $\varphi = \sum_i^k C_{x_i}$ where C_{x_i} is the coefficient of variation of the auxiliary variable. $\bar{y}_{st}^* = \sum_i^k p_i \bar{y}_i^*$; $\bar{y}_i^* = \frac{n_{i1}\bar{y}_{n_{i1}} + n_{i2}\bar{y}_{h_{i2}}}{n_i}$, $p_i = \frac{N_i}{N}$

$\bar{y}_{n_{i1}}$ and $\bar{y}_{h_{i2}}$ are the means based on n_{i1} respondent units and h_{i2} subsampled non-respondent units respectively for the study variable and $\bar{x}_{st}' = \sum_i^k p_i \bar{x}_i'$, $\bar{x}_{st} = \sum_i^k p_i \bar{x}_i$

with mean squared error given as

MSE(T_{ae})=$\Big[b^2R^2 \ \sum_i^k \left[\left(\frac{1}{n_i}-\frac{1}{N_i}\right) p_i^2 S_{X_i}^2\right] - 2abR^2 \ \sum_i^k \left[\left(\frac{1}{n_i'}-\frac{1}{N_i}\right) p_i^2 S_{X_i}^2\right] + a^2R^2 \ \sum_i^k \left[\left(\frac{1}{n_i'} - \frac{1}{N_i}\right) p_i^2 S_{X_i}^2\right] - 2bR\sum_i^k \left(\frac{1}{n_i}-\frac{1}{N_i}\right) p_i^2 \rho_i S_{X_i} S_{Y_i} + 2aR\sum_i^k \left(\frac{1}{n_i'}-\frac{1}{N_i}\right) p_i^2 \rho_i S_{X_i} S_{Y_i} + \ \sum_i^k \left[\left(\frac{1}{n_i} - \frac{1}{N_i}\right) p_i^2 S_{Y_i}^2 + \frac{(L_i-1)}{n_i} p_i^2 W_{i2} S_{Y_{i2}}^2\right]\Big]$

Where R = $\frac{\bar{Y}}{\bar{X}}$, a = $\frac{\bar{X}}{\bar{X}+\varphi}$ and b = $\frac{\bar{X}}{\bar{X}-\varphi}$

Proposed Estimator and its properties

The proposed estimator for population mean in stratified double sampling in the presence of nonresponse using single auxiliary variable is given as,

$$T_{ae1}=\ \bar{y}_{st}^* \left(\frac{\bar{x}_{st}'-\rho_i}{\bar{x}_{st}-\rho_i}\right) \quad (1)$$

Article DOI: 10.52589/AJMSS-YF4V11QV
DOI URL: https://doi.org/10.52589/AJMSS-YF4V11QV

where ρ_i is the correlation coefficient between y and x.

$\bar{y}^*_{st} = \sum_i^k p_i \bar{y}^*_i$; $\bar{y}^*_i = \frac{n_{i1}\bar{y}_{n_{i1}} + n_{i2}\bar{y}_{h_{i2}}}{n_i}$, $p_i = \frac{N_i}{N}$

$\bar{y}_{n_{i1}}$ and $\bar{y}_{h_{i2}}$ are the means based on n_{i1} respondent units and h_{i2} subsampled non-respondent units respectively for the study variable and $\bar{x}'_{st} = \sum_i^k p_i \bar{x}'_i$, $\bar{x}_{st} = \sum_i^k p_i \bar{x}_i$

The Bias and the Mean Squared Error MSE) of the proposed Estimator.

The Bias of T_{ae1}

To obtain the bias of the proposed estimator, the following terms are defined:

$\bar{y}^*_{st} = \bar{Y}(1+e_0)$, $\bar{x}_{st} = \bar{X}(1+e_1)$, $\bar{x}'_{st} = \bar{X}(1+e_2)$ where the e_is are the relative error terms defined as

$e_0 = \frac{\bar{y}^*_{st} - \bar{Y}}{\bar{Y}}$, $e_1 = \frac{\bar{x}_{st} - \bar{X}}{\bar{X}}$, $e_2 = \frac{\bar{x}'_{st} - \bar{X}}{\bar{X}}$

(1) becomes

$T_{ae1} = \bar{Y}(1+e_0)(1+be_2) \quad (1+be_1)^{-1}$ where $b = \frac{\bar{X}}{\bar{X}-\rho_i}$

$= \bar{Y}(1+e_0)(1+be_2) \quad (1 - be_1 + b^2e_1^2)$

$= \bar{Y}(1+e_0)[1 - be_1 + b^2e_1^2 + be_2 - b^2e_1e_2]$

$= \bar{Y}[1 - be_1 + b^2e_1^2 + be_2 - b^2e_1e_2 + e_0 - be_0e_1 + be_0e_2]$ where the e_is are neglected having power more than two. Thus

Bias (T_{ae1}) = E $(T_{ae1} - \bar{Y})$

$T_{ae1} - \bar{Y} = \bar{Y}[-be_1 + b^2e_1^2 + be_2 - b^2e_1e_2 + e_0 - be_0e_1 + be_0e_2]$

Bias $(T_{ae1}) = \bar{Y}[b^2E(e_1^2) - b^2E(e_1e_2) - bE(e_0e_1) + bE(e_0e_2)]$

Since E $(e_0) = (e_1) = (e_2) = 0$

By substituting for the following expectations

E $(e_0) = (e_1) = (e_2) = 0$

E $(e_0^2) = \frac{1}{\bar{Y}^2} \sum_i^k \left[\left(\frac{1}{n_i} - \frac{1}{N_i}\right) p_i^2 S_{Y_i}^2 + \frac{(L_i - 1)}{n_i} p_i^2 W_{i2} S_{Y_{i2}}^2\right]$

E $(e_1^2) = \frac{1}{\bar{X}^2} \sum_i^k \left[\left(\frac{1}{n_i} - \frac{1}{N_i}\right) p_i^2 S_{X_i}^2\right]$

E $(e_2^2) = \frac{1}{\bar{X}^2} \sum_i^k \left[\left(\frac{1}{n'_i} - \frac{1}{N_i}\right) p_i^2 S_{X_i}^2\right]$ = E (e_1e_2)

E $(e_0e_1) = \frac{1}{\bar{X}\bar{Y}} \sum_i^k \left(\frac{1}{n_i} - \frac{1}{N_i}\right) p_i^2 \rho_i S_{X_i} S_{Y_i}$

Article DOI: 10.52589/AJMSS-YF4V11QV
DOI URL: https://doi.org/10.52589/AJMSS-YF4V11QV

African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 4, Issue 2, 2021 (pp. 41-51) www.abjournals.org

E ($e_0 e_2$) = $\frac{1}{\bar{X}\bar{Y}} \sum_i^k \left(\frac{1}{n_i'} - \frac{1}{N_i}\right) p_i^2 \rho_i S_{X_i} S_{Y_i}$

Bias(T_{ae1})=$\bar{Y} \left[(\frac{\bar{X}}{\bar{X}-\rho_i})^2 \frac{1}{\bar{X}^2} \sum_i^k \left[\left(\frac{1}{n_i} - \frac{1}{N_i}\right) p_i^2 S_{X_i}^2\right] - \left(\frac{\bar{X}}{\bar{X}-\rho_i}\right)^2 \frac{1}{\bar{X}^2} \sum_i^k \left[\left(\frac{1}{n_i'} - \frac{1}{N_i}\right) p_i^2 S_{X_i}^2\right] - (\frac{\bar{X}}{\bar{X}-\rho_i}) \frac{1}{\bar{Y}\bar{X}} \sum_i^k \left(\frac{1}{n_i} - \frac{1}{N_i}\right) p_i^2 \rho_i S_{X_i} S_{Y_i} + (\frac{\bar{X}}{\bar{X}-\rho_i}) \frac{1}{\bar{Y}\bar{X}} \sum_i^k \left(\frac{1}{n_i'} - \frac{1}{N_i}\right) p_i^2 \rho_i S_{X_i} S_{Y_i} \right]$ (2)

The Mean Square Error (MSE) of T_{ae1}

The mean squared error of T_{ae1} is defined as

MSE (T_{ae1}) = E$(T_{ae1} - \bar{Y})^2$

=E$(\bar{Y}[-be_1 + b^2 e_1^2 + be_2 - b^2 e_1 e_2 + e_0 - be_0 e_1 + be_0 e_2])^2$

=E ($\bar{Y}$ $[-be_1 + be_2 + e_0]$ * $\bar{Y}$ $[-be_1 + be_2 + e_0]$) since the e_is are neglected having power more than two.

= E [$\bar{Y}^2(b^2 e_1^2 - 2b^2 e_1 e_2 - 2\, be_0 e_1 + b^2 e_2^2 + 2\, be_0 e_2 + e_0^2)$]

=$\bar{Y}^2[b^2 E(e_1^2) - b^2 E(e_1 e_2) - 2\, bE(e_0 e_1) + 2\, bE(e_0 e_2) + E(e_0^2)]$ since E $(e_1 e_2)$ = E (e_2^2)

By substituting for the expectations from section (4.1.1.)

MSE (T_{ae1}) = $\bar{Y}^2[b^2 E(e_1^2) - b^2 E(e_1 e_2) - 2\, bE(e_0 e_1) + 2\, bE(e_0 e_2) + E(e_0^2)]$

=$\bar{Y}^2 \left[b^2 \frac{1}{\bar{X}^2} \sum_i^k \left[\left(\frac{1}{n_i} - \frac{1}{N_i}\right) p_i^2 S_{X_i}^2\right] - b^2 \frac{1}{\bar{X}^2} \sum_i^k \left[\left(\frac{1}{n_i'} - \frac{1}{N_i}\right) p_i^2 S_{X_i}^2\right] - 2b \frac{1}{\bar{Y}\bar{X}} \sum_i^k \left(\frac{1}{n_i} - \frac{1}{N_i}\right) p_i^2 \rho_i S_{X_i} S_{Y_i} + 2b \frac{1}{\bar{Y}\bar{X}} \sum_i^k \left(\frac{1}{n_i'} - \frac{1}{N_i}\right) p_i^2 \rho_i S_{X_i} S_{Y_i} + \frac{1}{\bar{Y}^2} \sum_i^k \left[\left(\frac{1}{n_i} - \frac{1}{N_i}\right) p_i^2 S_{Y_i}^2 + \frac{(L_i - 1)}{n_i} p_i^2 W_{i2} S_{Y_{i2}}^2\right]\right]$

MSE(T_{ae1})=$\left[(\frac{\bar{X}}{\bar{X}-\rho_i})^2 R^2 \sum_i^k \left[\left(\frac{1}{n_i} - \frac{1}{N_i}\right) p_i^2 S_{X_i}^2\right] - \left(\frac{\bar{X}}{\bar{X}-\rho_i}\right)^2 R^2 \sum_i^k \left[\left(\frac{1}{n_i'} - \frac{1}{N_i}\right) p_i^2 S_{X_i}^2\right] - 2(\frac{\bar{X}}{\bar{X}-\rho_i}) R \sum_i^k \left(\frac{1}{n_i} - \frac{1}{N_i}\right) p_i^2 \rho_i S_{X_i} S_{Y_i} + 2(\frac{\bar{X}}{\bar{X}-\rho_i}) R \sum_i^k \left(\frac{1}{n_i'} - \frac{1}{N_i}\right) p_i^2 \rho_i S_{X_i} S_{Y_i} + \sum_i^k \left[\left(\frac{1}{n_i} - \frac{1}{N_i}\right) p_i^2 S_{Y_i}^2 + \frac{(L_i - 1)}{n_i} p_i^2 W_{i2} S_{Y_{i2}}^2\right]\right]$ where R = $\frac{\bar{Y}}{\bar{X}}$ (3)

The Optimum Values Of n_i, n_i' And L_i for T_{ae1}

The cost function analysis adopted by Chaudhary and Kumar (2015) is used in the cost function analysis of this estimator.

Let c_i' be the cost per unit associated with the first phase sample of size n_i' and c_{i0} be the unit cost of first attempt on study variable with second phase sample of size n_i . Let c_{i1} and c_{i2} be respectively the cost per unit of enumerating the n_{i1} respondent units and h_{i2} non-respondent units. Then the total cost for the i^{th} stratum shall be given as

C_i= $c_i' n_i'$ + $c_{i0} n_i$+ $c_{i1} n_{i1}$ + $c_{i2} h_{i2}$ ∀i = 1, 2, ..., k. (4)

Article DOI: 10.52589/AJMSS-YF4V11QV
DOI URL: https://doi.org/10.52589/AJMSS-YF4V11QV

the expected cost per stratum is given as

$$\mathrm{E}\,(C_i) = c_i' n_i' + n_i\,(c_{i0} + c_{i1}W_{i1} + c_{i2}\frac{W_{i2}}{L_i})\ . \tag{5}$$

Thus the total cost over all the strata is represented as

$$C_0 = \sum_i^k \mathrm{E}\,(C_i)$$

$$= \sum_i^k \left[c_i' n_i' + n_i\left(c_{i0} + c_{i1}W_{i1} + c_{i2}\frac{W_{i2}}{L_i}\right)\right] \tag{6}$$

Given a cost function as defined as (6), the optimization function is

$$\phi = \mathrm{MSE}\,(T_{ae1}) + \lambda C_0 \tag{7}$$

where λ is Lagrange's multiplier

$$= \left[b^2R^2 \sum_i^k \left[\left(\frac{1}{n_i} - \frac{1}{N_i}\right) p_i^2 S_{X_i}^2\right] - b^2R^2 \sum_i^k \left[\left(\frac{1}{n_i'} - \frac{1}{N_i}\right) p_i^2 S_{X_i}^2\right] - 2bR\sum_i^k \left(\frac{1}{n_i} - \frac{1}{N_i}\right) p_i^2 \rho_i S_{X_i} S_{Y_i} + 2bR\sum_i^k \left(\frac{1}{n_i'} - \frac{1}{N_i}\right) p_i^2 \rho_i S_{X_i} S_{Y_i} + \sum_i^k \left[\left(\frac{1}{n_i} - \frac{1}{N_i}\right) p_i^2 S_{Y_i}^2 + \frac{(L_i-1)}{n_i} p_i^2 W_{i2} S_{Y_{i2}}^2\right]\right] + \lambda \sum_i^k \left[c_i' n_i' + n_i\left(c_{i0} + c_{i1}W_{i1} + c_{i2}\frac{W_{i2}}{L_i}\right)\right]$$

In order to get the optimum values of n_i, n_i' and L_i, we differentiate ϕ with respect to n_i, n_i' and L_i respectively and equate the derivatives to zero. Thus, we have for stratum i

$$\frac{\partial\phi}{\partial n_i} = \frac{-P_i^2}{n_i^2}\left[S_{Y_i}^2 + b^2R^2S_{X_i}^2 - 2bR\rho_i S_{Y_i} S_{X_i}\right] - \frac{(L_i-1)W_{i2}S_{Y_{i2}}^2 P_i^2}{n_i^2} + \lambda\left(c_{i0} + c_{i1}W_{i1} + c_{i2}\frac{W_{i2}}{L_i}\right) = 0 \tag{8}$$

$$\frac{\partial\phi}{\partial n_i'} = \frac{P_i^2}{n_i'^2}\left[b^2R^2S_{X_i}^2 - 2bR\rho_i S_{Y_i} S_{X_i}\right] + \lambda c_i' = 0 \tag{9}$$

$$\frac{\partial\phi}{\partial L_i} = \frac{P_i^2}{n_i}\left[W_{i2}S_{Y_{i2}}^2\right] - \lambda n_i\, c_{i2}\frac{W_{i2}}{L_i^2} = 0 \tag{10}$$

From (8)

$$\lambda\left(c_{i0} + c_{i1}W_{i1} + c_{i2}\frac{W_{i2}}{L_i}\right) = \frac{P_i^2}{n_i^2}\left[S_{Y_i}^2 + b^2R^2S_{X_i}^2 - 2bR\rho_i S_{Y_i} S_{X_i} + (L_i-1)W_{i2}S_{Y_{i2}}^2\right]$$

$$n_i = \frac{P_i\sqrt{S_{Y_i}^2 + b^2R^2S_{X_i}^2 - 2bR\rho_i S_{Y_i} S_{X_i} + (L_i-1)W_{i2}S_{Y_{i2}}^2}}{\sqrt{\lambda\left(c_{i0} + c_{i1}W_{i1} + c_{i2}\frac{W_{i2}}{L_i}\right)}} \tag{11}$$

From (9)

$$\lambda c_i' = \frac{P_i^2}{n_i'^2}\left[2bR\rho_i S_{Y_i} S_{X_i} - b^2R^2S_{X_i}^2\right]$$

$$n_i' = \frac{P_i\sqrt{Q}}{\sqrt{\lambda\, c_i'}} \tag{12}$$

Article DOI: 10.52589/AJMSS-YF4V11QV
DOI URL: https://doi.org/10.52589/AJMSS-YF4V11QV

where Q = $2bR\rho_i S_{Y_i} S_{X_i} - b^2R^2S_{X_i}^2$

From (10)

$$\lambda = \frac{L_i^2 P_i^2 S_{Y_{i2}}^2}{n_i^2 c_{i2}} \text{ so that } \sqrt{\lambda} = \frac{L_i P_i S_{Y_{i2}}}{n_i\sqrt{C_{i2}}} \tag{13}$$

putting the value of $\sqrt{\lambda}$ in (13) into (11), we get

$$n_i = \frac{P_i\sqrt{S_{Y_i}^2 + b^2R^2S_{X_i}^2 - 2bR\rho_i S_{Y_i}S_{X_i} + (L_i-1)W_{i2}S_{Y_{i2}}^2}}{\frac{L_iP_iS_{Y_{i2}}}{n_i\sqrt{C_{i2}}}\sqrt{\left(c_{i0} + c_{i1}W_{i1} + c_{i2}\frac{W_{i2}}{L_i}\right)}}$$

$$L_i S_{Y_{i2}}\sqrt{\left(c_{i0} + c_{i1}W_{i1} + c_{i2}\frac{W_{i2}}{L_i}\right)} = \sqrt{(S_{Y_i}^2 + b^2R^2S_{X_i}^2 - 2bR\rho_i S_{Y_i}S_{X_i} + (L_i-1)W_{i2}S_{Y_{i2}}^2)\, c_{i2}}$$

Squaring both sides, we have

$$L_i^2\, S_{Y_{i2}}^2\left(c_{i0} + c_{i1}W_{i1} + c_{i2}\frac{W_{i2}}{L_i}\right) = (S_{Y_i}^2 + b^2R^2S_{X_i}^2 - 2bR\rho_i S_{Y_i}S_{X_i} + (L_i-1)W_{i2}S_{Y_{i2}}^2)\, c_{i2}$$

$$L_i^2\, S_{Y_{i2}}^2 c_{i0} + L_i^2\, S_{Y_{i2}}^2 c_{i1}W_{i1} + S_{Y_{i2}}^2 c_{i2}L_iW_{i2} - S_{Y_{i2}}^2 c_{i2}L_iW_{i2} = (S_{Y_i}^2 + b^2R^2S_{X_i}^2 - 2bR\rho_i S_{Y_i}S_{X_i} - W_{i2}S_{Y_{i2}}^2)\, c_{i2}$$

$$L_i^2\, S_{Y_{i2}}^2(c_{i0} + c_{i1}W_{i1}) = (S_{Y_i}^2 + b^2R^2S_{X_i}^2 - 2bR\rho_i S_{Y_i}S_{X_i} - W_{i2}S_{Y_{i2}}^2)\, c_{i2}$$

$$L_{i(opt)} = \frac{\sqrt{c_{i2}}B_i}{S_{Y_{i2}}A_i} \tag{14}$$

where $A_i = \sqrt{c_{i0} + c_{i1}W_{i1}}$;

$$B_i = \sqrt{S_{Y_i}^2 + b^2R^2S_{X_i}^2 - 2bR\rho_i S_{Y_i}S_{X_i} - W_{i2}S_{Y_{i2}}^2}$$

On substituting the value of $L_{i(opt)}$ from (14) into (11), we can express n_i as

$$n_i = \frac{P_i\sqrt{S_{Y_i}^2 + b^2R^2S_{X_i}^2 - 2bR\rho_i S_{Y_i}S_{X_i} + \left(\frac{\sqrt{c_{i2}}B_i}{S_{Y_{i2}}A_i} - 1\right)W_{i2}S_{Y_{i2}}^2}}{\sqrt{\lambda\left(c_{i0} + c_{i1}W_{i1} + c_{i2}\frac{W_{i2}}{\frac{\sqrt{c_{i2}}B_i}{S_{Y_{i2}}A_i}}\right)}}$$

$$n_i = \frac{P_i\sqrt{S_{Y_i}^2 + b^2R^2S_{X_i}^2 - 2bR\rho_i S_{Y_i}S_{X_i} - W_{i2}S_{Y_{i2}}^2 + \left(\frac{\sqrt{c_{i2}}B_iS_{Y_{i2}}W_{i2}}{A_i}\right)}}{\sqrt{\lambda\left(c_{i0} + c_{i1}W_{i1} + \sqrt{c_{i2}}\frac{W_{i2}S_{Y_{i2}}A_i}{B_i}\right)}}$$

Article DOI: 10.52589/AJMSS-YF4V11QV
DOI URL: https://doi.org/10.52589/AJMSS-YF4V11QV

African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 4, Issue 2, 2021 (pp. 41-51)

$$n_i = \frac{P_i\sqrt{B_i^2+\frac{\sqrt{c_{i2}}B_iW_{i2}S_{Y_{i2}}}{A_i}}}{\sqrt{\lambda}\sqrt{A_i^2+\frac{\sqrt{c_{i2}}A_iW_{i2}S_{Y_{i2}}}{B_i}}} \tag{15}$$

To obtain the value of $\sqrt{\lambda}$ in terms of total cost C_0, we put the value of n_i, n_i' and $L_{i(opt)}$ from (15), (12) and (14) into (6), thus we get

$$C_0 = \sum_i^k \left[c_i'n_i' + n_i\left(c_{i0} + c_{i1}W_{i1} + c_{i2}\frac{W_{i2}}{L_i}\right)\right]$$

$$= \sum_i^k \left[c_i' \frac{P_i\sqrt{Q}}{\sqrt{\lambda c_i'}} + \frac{P_i\sqrt{B_i^2+\frac{\sqrt{c_{i2}}B_iW_{i2}S_{Y_{i2}}}{A_i}}}{\sqrt{\lambda}\sqrt{A_i^2+\frac{\sqrt{c_{i2}}A_iW_{i2}S_{Y_{i2}}}{B_i}}}\left(c_{i0} + c_{i1}W_{i1} + c_{i2}\frac{W_{i2}}{\frac{\sqrt{c_{i2}}B_i}{S_{Y_{i2}}A_i}}\right)\right]$$

$$= \sum_i^k \left[\sqrt{c_i'}\frac{P_i\sqrt{Q}}{\sqrt{\lambda}} + \frac{P_i\sqrt{B_i^2+\frac{\sqrt{c_{i2}}B_iW_{i2}S_{Y_{i2}}}{A_i}}}{\sqrt{\lambda}\sqrt{A_i^2+\frac{\sqrt{c_{i2}}A_iW_{i2}S_{Y_{i2}}}{B_i}}}\left(c_{i0} + c_{i1}W_{i1} + \frac{\sqrt{c_{i2}}A_iW_{i2}S_{Y_{i2}}}{B_i}\right)\right]$$

$$= \sum_i^k \left[\sqrt{c_i'}\frac{P_i\sqrt{Q}}{\sqrt{\lambda}} + \frac{P_i\sqrt{B_i^2+\frac{\sqrt{c_{i2}}B_iW_{i2}S_{Y_{i2}}}{A_i}}}{\sqrt{\lambda}\sqrt{A_i^2+\frac{\sqrt{c_{i2}}A_iW_{i2}S_{Y_{i2}}}{B_i}}}\left(A_i^2 + \frac{\sqrt{c_{i2}}A_iW_{i2}S_{Y_{i2}}}{B_i}\right)\right]$$

$$\sqrt{\lambda} = \frac{1}{C_0}\sum_i^k \left[\sqrt{c_i'}\,P_i\sqrt{Q} + P_i\sqrt{B_i^2 + \frac{\sqrt{c_{i2}}B_iW_{i2}S_{Y_{i2}}}{A_i}}\sqrt{A_i^2 + \frac{\sqrt{c_{i2}}A_iW_{i2}S_{Y_{i2}}}{B_i}}\right]$$

$$= \frac{1}{C_0}\sum_i^k \left[\sqrt{c_i'}\,P_i\sqrt{Q} + P_i\sqrt{\left(B_i^2 + \frac{\sqrt{c_{i2}}B_iW_{i2}S_{Y_{i2}}}{A_i}\right)\left(A_i^2 + \frac{\sqrt{c_{i2}}A_iW_{i2}S_{Y_{i2}}}{B_i}\right)}\right]$$

$$= \frac{1}{C_0}\sum_i^k \left[\sqrt{c_i'}\,P_i\sqrt{Q} + P_i\sqrt{\frac{B_i^3A_i^3+2\sqrt{c_{i2}}W_{i2}B_i^2A_i^2S_{Y_{i2}}+c_{i2}B_iA_iW_{i2}^2S_{Y_{i2}}^2}{B_iA_i}}\right]$$

$$= \frac{1}{C_0}\sum_i^k \left[\sqrt{c_i'}\,P_i\sqrt{Q} + P_i\sqrt{B_i^2A_i^2 + 2\sqrt{c_{i2}}B_iA_iW_{i2}S_{Y_{i2}} + c_{i2}W_{i2}^2S_{Y_{i2}}^2}\right]$$

$$= \frac{1}{C_0}\sum_i^k \left[\sqrt{c_i'}\,P_i\sqrt{Q} + P_i\sqrt{\left(B_iA_i + \sqrt{c_{i2}}W_{i2}S_{Y_{i2}}\right)^2}\right]$$

$$\sqrt{\lambda} = \frac{1}{C_0}\sum_i^k\left[P_i\sqrt{c_i'}\sqrt{Q} + P_i\left(B_iA_i + \sqrt{c_{i2}}W_{i2}S_{Y_{i2}}\right)\right] \tag{16}$$

 Article DOI: 10.52589/AJMSS-YF4V11QV
DOI URL: https://doi.org/10.52589/AJMSS-YF4V11QV

Substituting the value of $\sqrt{\lambda}$ from (16) into (15) and (12), we respectively get the optimum values of n_i and n_i' as

$$n_{i(opt)} = \frac{P_i C_0 \sqrt{B_i^2 + \frac{\sqrt{c_{i2}} B_i W_{i2} S_{Y_{i2}}}{A_i}}}{\sqrt{A_i^2 + \frac{\sqrt{c_{i2}} A_i W_{i2} S_{Y_{i2}}}{B_i}} \sum_i^k \left[P_i \sqrt{c_i'}\sqrt{Q} + P_i (B_i A_i + \sqrt{c_{i2}} W_{i2} S_{Y_{i2}}) \right]}$$

$$n'_{i(opt)} = \frac{P_i C_0 \sqrt{Q}}{\sqrt{c_i'} \sum_i^k \left[P_i \sqrt{c_i'}\sqrt{Q} + P_i (B_i A_i + \sqrt{c_{i2}} W_{i2} S_{Y_{i2}}) \right]}$$

Empirical Study

Using the data set used by Chaudhary and Kumar 2015) shown on table 1 below

Table 1

Stratum no	N_i	n_i'	n_i	$\bar{Y}_i$	$\bar{X}_i$	$S_{Y_i}^2$	$S_{X_i}^2$	ρ_i	$S_{Y_{i2}}^2$
1	73	65	26	40.85	39.56	6369.1	6624.44	0.999	618.88
2	70	25	10	27.57	27.57	1051.07	1147.01	0.998	240.91
3	97	48	19	25.44	25.44	2014.97	2205.4	0.999	265.52
4	44	11	5	20.36	20.36	538.47	485.27	0.997	83.69

Table 2. Below shows the Mean Squared Error (MSE) and Percent Relative Efficiency (PRE) of the proposed estimator with respect to $\bar{y}_{st}^*$ for the different choices W_{i2} and L_i

Table 2

W_{i2}	L_i	$V(\bar{y}_{st}^*)$	MSE(T_{ae1})	MSE(T_c')	MSE(T_{ae})	PRE(T_{ae})	PRE(T_{ae1})
0.1	2	34.42	4.38	6.28	4.66	738.6	785.8
	2.5	34.67	4.64	6.54	4.92	704.7	747.2
	3	4.92	4.89	6.79	5.17	675.4	714.1
0.2	2	34.92	4.89	6.79	5.26	663.9	714.1
	2.5	35.43	5.64	7.3	5.66	625.97	628.2
	3	35.94	6.12	7.8	6.18	581.6	587.3
0.3	2	35.43	5.40	7.3	5.66	625.97	656.1
	2.5	36.19	6.43	8.06	6.44	561.96	562.8
	3	36.95	7.12	8.82	7.22	511.8	518.9

Article DOI: 10.52589/AJMSS-YF4V11QV
DOI URL: https://doi.org/10.52589/AJMSS-YF4V11QV

CONCLUSION

The optimum values of first phase sample size n_i', second phase sample size n_i and inverse sampling rate L_i for the different strata using the proposed estimator have been determined under the cost survey. From Table 2 it is seen that the proposed estimator is more efficient than other existing estimators, hence, the proposed estimator is recommended for use in practice.

REFERENCES

Anieting, A. E. and Enang, E. I. (2020) Two-Phase Stratified Sampling Estimator for Population Mean in the Presence of Nonresponse Using One Auxiliary Variable. Math. J. Interdiscip. Sci.Vol. 8, No.2.pp.49–56

Cochran, W. G. (1977). *Sampling Techniques*. New York: Wiley and SonsKhare, B. B. & Srivastava, S. (1997).Transformed ratio type estimators for the population mean in the presence of Nonresponse. *Comminication in Statistics – Theory and Methods*, 26 (7),1779-1791.

M. H. Hansen, W. N. Hurwitz (1946). The problem of non-response in sample surveys. Journal of The American Statistical Association, , no. 41, pp. 517–529.

M. K. Chaudhary and A. Kumar(2015) Use of double sample scheme in estimating the mean of Stratified population Under nonresponse. STATISTICA, anno, Lxxv,n.4

M. K. Chaudhary, R. Singh, R. K. Shukla, M. Kumar (2009). A family of estimators for estimating population mean in stratified sampling under non- response. Pakistan Journal of Statistics and Operation Research, 5, no. 1, pp. 47–54.

M. K. Chaudhary, V. K. Singh (2013). Estimating the population mean in two- stage sampling with equal size clusters under non-response using auxiliary characteristic. Mathematical Journal of Interdisciplinary Sciences, 2, no. 1, pp. 43–55.

M. K. Chaudhary, V. K. Singh, R. K. Shukla (2012). Combined-type fam- ily of estimators of population mean in stratified random sampling under non-response. Journal of Reliability and Statistical Studies (JRSS), 5, no. 2, pp. 133–142.

M. Khoshnevisan, R. Singh, P. Chauhan, N. Sawanand, F. Smarandache (2007). A general family of estimators for estimating population mean using known value of some population parameter(s). Far East Journal of Theoretical Statistics, 22, pp. 181–191.

Singer, E (2006) Introduction: Nonresponse Bias in Household Survey. *The Public Opinion Quarterly*. 70(5) 637-645.

Article DOI: 10.52589/AJMSS-YF4V11QV
DOI URL: https://doi.org/10.52589/AJMSS-YF4V11QV