AN IMPROVED ESTIMATION PROCEDURE FOR TWO-OCCASION SUCCESSIVE SAMPLING

D. T. Ailobhio[1], J. A. Ikughur[2], S. C. Nwoasu[3], T. Uba[4]

[1]Department of Statistics, University of Lafia, Nasarawa State, Nigeria
[2,3,4]Department of Statistics, Joseph Sarwuan Tarka University, Makurdi, Benue State, Nigeria

Abstract

Surveys are frequently conducted repeatedly throughout the course of the years or seasons in order to track changes in characteristics of interest. Estimation in current occasion is possible by utilizing facts from prior occasion, we have spent time in working on the subject of population mean estimation in two occasion successive sampling. An estimator T is Proposed by using the convex linear combination of tow estimators T_u, based on u units of sample drawn afresh at the current occasion, and T_m, based on m units, which is retained from the previous occasion. The mean square error expressions and bias of the proposed estimator is calculated, and the best replacement strategy for the specified scenario is also explained. An empirical study is carried out to evaluate the estimator's efficiency. The coefficient of variation (CV) and mean square error estimates shows that the suggested estimator is more efficient than the current estimators taken into account in this study.

Keywords: Coefficient of variation, mean square error, estimators, efficiency.

1. Introduction

When it comes to survey sampling, researchers continually look for new methods that will improve the precision and accuracy of estimators of population parameters during both selection and estimation phases. Examples of this estimators include mean, total, variance, proportion, and coefficient of variation. Numerous methods of sampling have been explored during the selection stage to choose study units that will closely resemble the population parameters (Singh et.al.,2019). From the review of Kumar *et.al*, (2010), Singh and Pal (2016), Singh and Pal, (2017), Bandyopadhyay and Singh (2016), and Etuk *et.al*., (2016), it is seen, that using an additional variable that has a strong correlation with the variable of interest at either the selection level, estimation level, or both levels of the survey, with any sampling strategies, has proven to give more accurate and efficient Population parameter estimate Compared to those who solely rely on the research variable under investigation.

Information gathered from baseline and end line (mean or variance) surveys are typical cases of how population features may evolve over time depending on the sample taken for investigation. One well-known technique to calculate the population parameters and measuring the difference in change in research variable in survey is the successive sampling strategy, which is a procedure under multi-phase sampling (Singh and Pal, 2017). This was reported by Jessen (1942), who proposed sampling twice as a method, in order to calculate the average population on each occasion and also shift from one occasion to another (Khare and Habib, 2013).

There are a number of early studies on two occasion sampling; a few of them are (Rao & Graham, (1964); Shabbir *et al*., (2005); Singh and Vishwakarma, (2007); Singh *et al*., (2012); Subramani, (2013)). Similarly, there are several population estimators in two occasion successive sampling that have been developed, these include estimators by, Singh and Homa, (2013); Anieting and Ezugwu, (2013); Singh, and Pal, (2016); Singh and Pal, (2017); Singh and Singh, (2017); Singh *et al*., (2018); Singh and Khalid (2018); Zoramthanga, (2018), and numerous others, who have contributed significantly in this direction. Recent studies by Singh *et al*., (2019); Singh, (2020); Muili *et al*., (2022); Tiwari *et al*., (2023), and many others.

This study suggested estimator which consists of mixture of regression cum exponential ratio estimator for the unmatched section of the sample and a mixture of exponential-ratio cum regression estimator of the matched section of the sample. All cases are useful when the variables are positively correlated. The sampling processes and a few existing estimators are explored in this study.

2. Preliminaries and Notations

Let the following $\pi = (\pi_1, \pi_2, \cdots, \pi_N)$ be a defined population of size N, sampled on two occasions. Let y be the research variable and x and z be the auxiliary variables related to y. take a sample of size n on first occasion, utilizing simple random sampling without replacement. On the second occasion, a sizeable random subsample of size $m = n\lambda$ unit was kept (matched), utilizing simple random sample without replacement from the remaining (N-n) fraction of the population, a new sample of size $u = (n-m) = n\theta$ unit is taken. The selection of sample for the second occasion is made so that the sample size is n as well. Here θ and λ represent the percentage of matched and unmatched units in the sample that satisfy the following condition, $0 \le \lambda < 1$, $0 \le \theta \le 1$, and $\lambda + \theta = 1$.

Throughout the investigation, the following notations are utilized. The population average of the variables x, y, and z are respectively, $\bar{X}, \bar{Y},$ and $\bar{Z}$. The sample average of the various variables of size given in suffices are $\bar{y}_u, \bar{y}_m, \bar{x}_m, \bar{x}_n, \bar{z}_u, \bar{z}_m$, and $\bar{z}_n$. The correlation coefficients between the variables shown in subscript are ρ_{yx}, ρ_{yz}, and ρ_{zx}. The coefficients of variation of the variables x, y, and z are respectively C_x, C_y, and C_z. Where $C_x = \dfrac{S_x}{\bar{X}}$, $C_z = \dfrac{S_z}{\bar{Z}}$, $C_y = \dfrac{S_y}{\bar{Y}}$, $C_{yx} = \rho_{yx} C_y C_x$, $C_{yz} = \rho_{yz} C_y C_z$ $C_{zx} = \rho_{zx} C_x C_z$, $K_{yx} = \rho_{yx} \dfrac{C_y}{C_x}$, $K_{yz} = \rho_{yz} \dfrac{C_y}{C_z}$, $K_{zx} = \rho_{zx} \dfrac{C_z}{C_x}$, $S_y^2 = \frac{1}{N-1} \sum_{i=1}^{N} \left(y_i - \bar{Y}\right)^2$, $S_x^2 = \frac{1}{N-1} \sum_{i=1}^{N} \left(x_i - \bar{X}\right)^2$, $S_z^2 = \frac{1}{N-1} \sum_{i=1}^{N} \left(Z_i - \bar{Z}\right)^2$, $\lambda_u = \dfrac{1}{u} - \dfrac{1}{N}$, $\lambda_m = \dfrac{1}{m} - \dfrac{1}{N}$, $\lambda_n = \dfrac{1}{n} - \dfrac{1}{N}$

The following definitions apply to the error terms that are used to determine the estimator's bias and mean square error (MSE);

ISSN NUMBER: 1116-249X

$\bar{y}_u = \bar{Y}(1+e_{yu}),\quad \bar{y}_m = \bar{Y}(1+e_{ym}),\quad \bar{x}_n = \bar{X}(1+e_{xn}),\quad \bar{x}_m = \bar{X}(1+e_{xm}),$

$\bar{z}_u = \bar{Z}(1+e_{zu}),\quad \bar{z}_m = \bar{Z}(1+e_{zm}),\quad \bar{z}_n = \bar{Z}(1+e_{zn})$

and the expected values are

$E(e_{yu}) = E(e_{ym}) = E(e_{xn}) = E(e_{xm}) = E(e_{zu}) = E(e_{zm}) = E(e_{zn}) = 0$

$E(e_{yu}^2) = \lambda_u C_y^2,\quad E(e_{ym}^2) = \lambda_m C_y^2\quad E(e_{xn}^2) = \lambda_n C_x^2,\quad E(e_{xm}^2) = \lambda_m C_x^2$

$E(e_{zu}^2) = \lambda_u C_z^2\quad E(e_{zm}^2) = \lambda_m C_z^2\quad E(e_{zn}^2) = \lambda_n C_z^2,\quad E(e_{zn}e_{zm}) = \lambda_n C_z^2$

$E(e_{yu}, e_{zu}) = \lambda_u C_{yz},\quad E(e_{ym}, e_{xn}) = \lambda_n C_{yx},\quad E(e_{ym}, e_{xm}) = \lambda_m C_{yx}\quad E(e_{ym}, e_{zm}) = \lambda_m C_{yz},$
$E(e_{xn}, e_{xm}) = \lambda_n C_x^2,\quad E(e_{xn}, e_{zm}) = \lambda_n C_{xz},\quad E(e_{xm}, e_{zm}) = \lambda_m C_{xz}$

3. A Selection of Some Population Mean Estimators for Two-Occasion Successive Sampling

(i) Shabbir et al., (2005) Estimator

An estimator to estimate the population mean on current occasion in two occasion successive sampling was provided by shabbier et al. (2005), as reported in Tiwari et al., (2023);

$$T_{SAG} = \phi T_{SAG(m)} + (1-\phi)T_{SAG(u)} \tag{1}$$

Where the constant ϕ is unknown and the estimators of $\bar{Y}$ based on the matched and unmatched portion of the samples are $T_{SAG(m)}$, and $T_{SAG(u)}$ respectively. The definition of $T_{SAG(m)}$ is

$$T_{SAG(m)} = w_1\left\{\bar{y}_m + b_1(\bar{x}_n - \bar{x}_m)\frac{\bar{x}_n}{\bar{X}}\right\} + w_2\left\{\bar{y}_m + b_2(\bar{z}_n - \bar{z}_m)\frac{\bar{z}_n}{\bar{Z}}\right\} \tag{2}$$

Where b_1 and b_2 are the regression coefficients of y on x and y on z respectively, for the matched section, and w_1 and w_2 are constants such that $w_1 + w_2 = 1$.

$T_{SAG(u)}$ Is defined as;

$$T_{SAG(u)} = \bar{y}_u + b(\bar{Z} - \bar{z}_u) \tag{3}$$

Where b is the unmatched fraction regression coefficient of y on z. The minimized mean square error (MSE) of T_{SAG} is given as;

$$MSE(T_{SAG}) = \frac{S_y^2}{n}\frac{\delta_0\left[1-\rho_{xz}^2-\theta\left(\rho_{yx}^2+\rho_{yz}^2-2\rho_{yx}\rho_{yz}\rho_{zx}\right)\right]}{\left[\left(\delta_0+\theta\rho_{yz}^2\right)\left(1-\rho_{zx}^2\right)-\theta^2\left(\rho_{yx}^2+\rho_{yz}^2-2\rho_{yx}\rho_{yz}\rho_{zx}\right)\right]} \tag{4}$$

(ii) Singh & Pal, (2016) Estimator

Tiwari et al., (2023), presented an estimator proposed by Singh & Pal, (2016) to calculate population mean $\bar{Y}$ on two occasion successive sample as;

$$T_{SP} = \phi T_{SP(m)} + (1-\phi)T_{SP(u)} \tag{5}$$

ϕ is a constant, and $T_{SP(m)}$, and $T_{SP(u)}$ are defined as

$$T_{SP(m)} = \bar{y}_m \exp\left\{\frac{\alpha_2\left(\bar{Z}-\bar{z}_m\right)}{\left(\bar{Z}+\bar{z}_m\right)}\right\} + b_{yx}\left[\bar{x}_n \exp\left\{\frac{\alpha_3\left(\bar{Z}-\bar{z}_n\right)}{\left(\bar{Z}+\bar{z}_n\right)}\right\} - \bar{x}_m \exp\left\{\frac{\alpha_3\left(\bar{Z}-\bar{z}_m\right)}{\left(\bar{Z}+\bar{z}_m\right)}\right\}\right] \tag{6}$$

Where b_{yx} is the regression coefficient based on the matched fraction, and α_2, and α_3 are scalars.

$$T_{SP(u)} = \bar{y}_u \exp\left\{\frac{\alpha_1\left(\bar{Z}-\bar{z}_u\right)}{\left(\bar{Z}+\bar{z}_n\right)}\right\} \tag{7}$$

Where α_1 is a constant, the minimum MSE of T_{SP} is given has

$$MSE(T_{SP}) = \frac{S_y^2}{n}\frac{\delta_0\left[\delta_0-\theta\left(\rho_{yx}^2+\rho_{yx}^2\rho_{zx}^2-2\rho_{yx}\rho_{yz}\rho_{zx}\right)\right]}{\left[\delta_0-\theta^2\left(\rho_{yx}^2+\rho_{yx}^2\rho_{zx}^2-2\rho_{yx}\rho_{yz}\rho_{zx}\right)\right]} \tag{8}$$

(iii) Tiwari et al., (2023) Estimator

Tiwari et al., (2023), suggested an estimator for determining population mean $\bar{Y}$ on current occasion as

$$T_{KSS} = \phi T_{KSS(m)} + (1-\phi)T_{KSS(u)} \tag{9}$$

Where ϕ is a constant derived from $MSE(T_{KSS})$

The estimator for the matched portion is

$$T_{KSS(m)} = \left[\bar{y}_m + \alpha_1\left(\bar{x}_n - \bar{x}_m\right)\right]\exp\alpha_2\left(\frac{\bar{Z}-\bar{z}_m}{\bar{Z}+\bar{z}_n}\right) \tag{10}$$

Where α_1 and α_2 are constant derived from MSE.

The unmatched portion of the estimator is defined as

$$T_{KSS(u)} = \bar{y}_u \left(\frac{\bar{Z}}{\bar{z}_u} \right) \exp \alpha \left(\frac{\bar{Z} - \bar{z}_u}{\bar{Z} + \bar{z}_u} \right) \tag{11}$$

Where α is a real scalar. The minimum MSE of T_{KSS} is given has

$$MSE(T_{KSS}) = \frac{S_y^2}{n} \frac{\delta_0 \left[\delta_0 \left(1 - \theta \rho_{zx}^2 \right) - \theta V^2 \right]}{\left[\delta_0 \left(1 - \theta \rho_{zx}^2 \right) - \theta^2 V^2 \right]} \tag{12}$$

Key: $\delta_0 = \left(1 - \rho_{yz}^2 \right)$, $V = (\rho_{yx} - \rho_{xz}\rho_{yz})$, $\theta_0 = \dfrac{\delta_0 \pm \sqrt{\delta_0^2 (1 - \rho_{xz}^2) - \delta_0 V^2}}{V^2 + \delta_0 \rho_{xz}^2}$

4. Proposed Estimator

Taking inspiration from Shabbir et al. (2005), Singh and Pal (2016), and Tiwari et al., (2023), an improved estimator is suggested, for calculating the population mean of the research variable (y) on the current occasion in two occasion successive sampling.

$$T = \phi T_m + (1 - \phi) T_u \tag{13}$$

Where ϕ is a constant derived from mean square error, MSE(T).

The estimator for the unmatched portion T_u is defined as;

$$T_u = \beta_1 \left[\bar{y}_u + b_1 (\bar{Z} - \bar{z}_u) \right] + (1 - \beta_1) \left[\bar{y}_u \exp \left(\frac{\bar{Z} - \bar{z}_u}{\bar{Z} + \bar{z}_u} \right) \right] \tag{14}$$

Where β_1 is a constant derived from MSE(T)

The matched portion estimator T_m is given as

$$T_m = \beta_2 \left[\bar{y}_m \exp \left(\frac{\bar{Z} - \bar{z}_m}{\bar{Z} + \bar{z}_n} \right) \right] + (1 - \beta_2) \left[\bar{y}_m + b_2 (\bar{x}_n - \bar{x}_m) \right] \tag{15}$$

Where β_2 is constant derived from MSE(T)

5. Features of the suggested estimator

Transforming T_u and T_m in form of errors (e_{ij}) terms, where $i = x, y, \&z$,and $j = u, m, n$, to obtain the bias and MSE, is as follows.

ISSN NUMBER: 1116-249X

$$T_u = \beta_1\left[\bar{Y}\left(1+e_{yu}\right)+b_1\left(\bar{Z}-\bar{Z}\left(1+e_{zu}\right)\right)\right]+\left(1-\beta_1\right)\left[\bar{Y}\left(1+e_{yu}\right)\exp\left(\frac{\bar{Z}-\bar{Z}\left(1+e_{zu}\right)}{\bar{Z}+\bar{Z}\left(1+e_{zu}\right)}\right)\right] \tag{16}$$

Now, using Taylor's series approximation, expanding the formula and terminate the terms having (e_{ij})'s degree more than two. Assuming $\left|e_{ij}\right| < 1$, $i = x, y, \&z$; $j = u, m, n$. After simplifying gives,

$$\left(T_u - \bar{Y}\right) = \bar{Y}\left[-\beta_1 R_1 e_{zu} - \frac{1}{2}e_{zu} + \frac{3}{8}e_{zu}^2 + e_{yu} - \frac{1}{2}e_{yu}e_{zu} + \frac{1}{2}\beta_1 e_{zu} - \frac{3}{8}\beta_1 e_{zu}^2 + \frac{1}{2}\beta_1 e_{yu}e_{zu}\right] \tag{17}$$

Where $R_1 = \frac{\bar{Z}}{\bar{Y}}$

The bias to the first degree of approximation can be found by taking the expectation of each side of equation (17);

$$B(T_u) = \bar{Y}\lambda_u\left[\left(1-\beta_1\right)\frac{3}{8}C_z^2 + \left(\beta_1 - 1\right)\frac{1}{2}C_{yz}\right]$$

squaring both sides of equation (8), and eliminating the terms whose degree of (e_{ij})'s is more than two, gives;

$$(T_u - \bar{Y})^2 = \bar{Y}^2\begin{bmatrix} -2\beta_1 R_1 b_1 e_{zu}e_{yu} + \beta_1^2 R_1^2 b_1^2 e_{zu}^2 + \beta_1 R_1 b_1 e_{zu}^2 - \beta_1^2 R_1 b_1 e_{zu}^2 + e_{yu}^2 + \frac{1}{4}\beta_1^2 e_{zu}^2 \\ -\frac{1}{2}\beta_1 e_{zu}^2 - e_{zu}e_{yu} + \beta_1 e_{zu}e_{yu} + \frac{1}{4}e_{zu}^2 \end{bmatrix} \tag{18}$$

Taking the expectation of both sides of Equation (18), gives

$$MSE(T_u) = \bar{Y}^2\lambda_u\begin{bmatrix} -2\beta_1 R_1 b_1 C_{yz} + \beta_1^2 R_1^2 b_1^2 C_z^2 + \beta_1 R_1 b_1 C_z^2 - \beta_1^2 R_1 b_1 C_z^2 + C_y^2 + \frac{1}{4}\beta_1^2 C_z^2 \\ -\frac{1}{2}\beta_1 C_z^2 - C_{yz} + \beta_1 C_{yz} + \frac{1}{4}C_z^2 \end{bmatrix} \tag{19}$$

In order to minimize $MSE(T_u)$, Equation (19) was differentiate with respect to β_1, and equate to zero.

$$\frac{d}{d\beta_1}\left(MSE(T_u)\right) = 0$$

It provides an ideal value of β_1;

ISSN NUMBER: 1116-249X

$$\beta_1 = \frac{-C_z^2 + 2C_{yz}}{(2R_1b_1 - 1)C_z^2}$$

Putting the value of β_1, into Equation (19), to obtain the minimized value of $MSE(T_u)$, gives

$$MSE(T_u)_{\min} = \bar{Y}^2\lambda_u\left(\frac{C_y^2C_z^2 - C_{yz}^2}{C_z^2}\right) \tag{20}$$

$$= \bar{Y}^2\lambda_u(1-\rho_{yz}^2)C_y^2$$

Now, express T_m in terms of e_{ij}'s to drive the bias and MSE of T_m

$$T_m = (1-\beta_2)\left[\bar{Y}(1+e_{ym}) + b_2\left(\bar{X}(1+e_{xn}) - \bar{X}(1+e_{xm})\right)\right] + \beta_2\bar{Y}\left[(1+e_{ym})\exp\left(\frac{\bar{Z} - \bar{Z}(1+e_{zm})}{\bar{Z} + \bar{Z}(1+e_{zn})}\right)\right] \tag{21}$$

By utilizing the Taylor's series approximation, and expanding while disregarding terms with e_{ij}'s degree greater than two, gives

$$(T_m - \bar{Y}) = \bar{Y}\begin{bmatrix} e_{ym} + R_2b_2e_{xn} - R_2b_2e_{xm} - \beta_2R_2b_2e_{xn} + \beta_2R_2b_2e_{xm} - \frac{1}{2}\beta_2e_{zm} + \frac{1}{4}\beta_2e_{zm}e_{zn} \\ + \frac{1}{8}\beta_2e_{zm}^2 - \frac{1}{2}\beta_2e_{ym}e_{zm} \end{bmatrix} \tag{22}$$

Where $R_2 = \frac{\bar{X}}{\bar{Y}}$

The bias to the first degree of approximation can be found by taking the expectation of each side of equation (22) which yields;

$$B(T_m) = \bar{Y}\left[\frac{1}{4}\beta_2C_z^2\lambda_n + \frac{1}{8}\beta_2C_z^2\lambda_m - \frac{1}{2}\beta_2C_{yz}\lambda_m\right] \tag{23}$$

Equivalently squaring both sides of equation (22), and eliminating the terms whose degree of (e_{ij})'s is more than two, gives;

ISSN NUMBER: 1116-249X

$$\left(T_m-\bar{Y}\right)^2=\bar{Y}^2\begin{bmatrix}\beta_2^2R_2^2b_2^2e_{xm}^2-2\beta_2R_2^2b_2^2e_{xm}^2+\beta_2^2R_2^2b_2^2e_{xn}^2-2\beta_2R_2^2b_2^2e_{xn}^2-2R_2^2b_2^2e_{xm}e_{xn}-2R_2b_2e_{xm}e_{ym}\\+2R_2b_2e_{xn}e_{ym}+e_{ym}^2+\frac{1}{4}\beta_2^2e_{zm}^2+R_2^2b_2^2e_{xm}^2+R_2^2b_2^2e_{xn}^2-\beta_2R_2b_2e_{xn}e_{zm}-2\beta_2^2R_2^2b_2^2e_{xm}e_{xn}\\+4\beta_2R_2^2b_2^2e_{xm}e_{xn}+2\beta_2R_2b_2e_{xm}e_{ym}-\beta_2^2R_2b_2e_{xm}e_{zm}-2\beta_2R_2b_2e_{xn}e_{ym}+\beta_2^2R_2b_2e_{xn}e_{zm}\\+\beta_2R_2b_2e_{xm}e_{zm}-\beta_2e_{ym}e_{zm}\end{bmatrix}\quad(24)$$

Taking the expectation of both sides of Equation (24), gives

$$MSE(T_m)=\bar{Y}^2\begin{bmatrix}\beta_2^2R_2^2b_2^2C_x^2\lambda_m-2\beta_2R_2^2b_2^2C_x^2\lambda_m+\beta_2^2R_2^2b_2^2C_x^2\lambda_n-2\beta_2R_2^2b_2^2C_x^2\lambda_n-2R_2^2b_2^2C_x^2\lambda_n\\-2R_2b_2C_{yx}\lambda_m+2R_2b_2C_{yx}\lambda_n+C_y^2\lambda_m+\frac{1}{4}\beta_2^2C_z^2\lambda_m+R_2^2b_2^2C_x^2\lambda_m+R_2^2b_2^2C_x^2\lambda_n\\-\beta_2R_2b_2C_{xz}\lambda_n-2\beta_2^2R_2^2b_2^2C_x^2\lambda_n+4\beta_2R_2^2b_2^2C_x^2\lambda_n+2\beta_2R_2b_2C_{yx}\lambda_m\\-\beta_2^2R_2b_2C_{xz}\lambda_m-2\beta_2R_2b_2C_{yx}\lambda_n+\beta_2^2R_2b_2C_{xz}\lambda_n+\beta_2R_2b_2C_{xz}\lambda_m-\beta_2C_{yz}\lambda_m\end{bmatrix}\quad(25)$$

Differentiating Equation (25) partially with respect to β_2, and equate to zero gives the expression for β_2 as;

$$\beta_2=\frac{2\left(2R_2^2b_2^2C_x^2(\lambda_m-\lambda_n)+R_2b_2C_{xz}(\lambda_n-\lambda_m)+2R_2b_2C_{yx}(\lambda_n-\lambda_m)+C_{yz}\lambda_m\right)}{4R_2^2b_2^2C_x^2(\lambda_m-\lambda_n)+4R_2b_2C_{xz}(\lambda_n-\lambda_m)+C_z^2\lambda_m}$$

Substituting the value of β_2, into Equation (25) gives,

$$MSE(T_m)=\bar{Y}^2\left(\frac{\begin{pmatrix}-w^2\rho_{xz}^2\lambda_1+4w^2\rho_{xz}\rho_{yz}\lambda_1-4w^2\rho_{yx}^2\lambda_1+4w^2\rho_{yz}\lambda_2+3w^2\lambda_2-2w\ \rho_{xz}\rho_{yz}\lambda_2-4w\ \rho_{yx}\rho_{yz}\lambda_2\\-4w\ \rho_{xz}\lambda_2+2w\ \rho_{yx}^2\lambda_2+\rho_{yz}^2\lambda_m+\lambda_m^2\end{pmatrix}C_y^2}{4w^2\lambda_0-4w\ \rho_{xz}\lambda_0+\lambda_m}\right)$$

$$MSE(T_m)=\frac{S_y^2}{m}\left(\frac{K_{1.1}+K_{1.2}+\rho_{yz}^2\lambda_m^2+\lambda_m^2}{K_{1.3}+\lambda_m}\right)\quad(26)$$

Where

$\lambda_0=\lambda_m-\lambda_n$, $\lambda_1=2\lambda_m\lambda_n-\lambda_m^2-\lambda_n^2$, $\lambda_2=\lambda_m^2-\lambda_m\lambda_n$, $w=R_2b_2$

$$K_{1.1}=\left(4w^2\rho_{xz}\rho_{yz}-w^2\rho_{xz}^2-4w^2\rho_{yx}^2\right)\lambda_1$$

$$K_{1.2}=\left(4w^2\rho_{yz}+3w^2-2w\ \rho_{xz}\rho_{yz}-4w\ \rho_{yx}\rho_{yz}-4w\rho_{xz}+2ww^2\rho_{yx}\right)\lambda_2$$

$$K_{1.3}=\left(4w^2-4w\ \rho_{xz}\right)\lambda_0$$

ISSN NUMBER: 1116-249X

By definition;

$$MSE(T) = (1-\phi)^2 MSE(T_u) + \phi^2 MSE(T_m) \tag{27}$$

$$MSE(T) = (1-\phi)^2 \begin{bmatrix} -2\beta_1 R_1 b_1 e_{zu} e_{yu} + \beta_1^2 R_1^2 b_1^2 e_{zu}^2 + \beta_1 R_1 b_1 e_{zu}^2 - \beta_1^2 R_1 b_1 e_{zu}^2 + e_{yu}^2 \\ +\frac{1}{4}\beta_1^2 e_{zu}^2 - \frac{1}{2}\beta_1 e_{zu}^2 - e_{zu}e_{yu} + \beta_1 e_{zu} e_{yu} + \frac{1}{4} e_{zu}^2 \end{bmatrix} + \phi^2$$

$$\begin{bmatrix} \beta_2^2 R_2^2 b_2^2 e_{xm}^2 - 2\beta_2 R_2^2 b_2^2 e_{xm}^2 + \beta_2^2 R_2^2 b_2^2 e_{xn}^2 - 2\beta_2 R_2^2 b_2^2 e_{xn}^2 - 2R_2^2 b_2^2 e_{xm} e_{xn} - 2R_2 b_2 e_{xm} e_{ym} \\ +2R_2 b_2 e_{xn} e_{ym} + e_{ym}^2 + \frac{1}{4}\beta_2^2 e_{zm}^2 + R_2^2 b_2^2 e_{xm}^2 + R_2^2 b_2^2 e_{xn}^2 - \beta_2 R_2 b_2 e_{xn} e_{zm} - 2\beta_2^2 R_2^2 b_2^2 e_{xm} e_{xn} \\ +4\beta_2 R_2^2 b_2^2 e_{xm} e_{xn} + 2\beta_2 R_2 b_2 e_{xm} e_{ym} - \beta_2^2 R_2 b_2 e_{xm} e_{zm} - 2\beta_2 R_2 b_2 e_{xn} e_{ym} + \beta_2^2 R_2 b_2 e_{xn} e_{zm} \\ +\beta_2 R_2 b_2 e_{xm} e_{zm} - \beta_2 e_{ym} e_{zm} \end{bmatrix} \tag{28}$$

Differentiating Equation (28) with respect to β_1, β_2 and ϕ, and equating to zero, yields;

$$\beta_1 = \frac{-C_z^2 + 2C_{yz}}{(2R_1 b_1 - 1)C_z^2}, \quad \beta_2 = \frac{2\left(2R_2^2 b_2^2 C_x^2 (\lambda_m - \lambda_n) + R_2 b_2 C_{xz}(\lambda_n - \lambda_m) + 2R_2 b_2 C_{yx}(\lambda_n - \lambda_m) + C_{yz}\lambda_m\right)}{4R_2^2 b_2^2 C_x^2(\lambda_m - \lambda_n) + 4R_2 b_2 C_{xz}(\lambda_n - \lambda_m) + C_z^2 \lambda_m}$$

$$\phi = \frac{\lambda_u\left(K_{3.1} + C_y^2 C_z^4 \lambda_m - C_{yz}^2 C_z^2 \lambda_m\right)}{\left(K_{3.2} + K_{3.3} + K_{3.4} + K_{3.5}\right)}$$

Where

$$K_{3.1} = \lambda_0\left(4w^2 C_x^2 C_y^2 C_z^2 - 4w^2 C_x^2 C_{yz}^2 - 4wC_{xz} C_y^2 C_z^2 + 4wC_{xz} C_{yz}^2\right)$$

$$K_{3.2} = \lambda_{0.0}\left(w^2 C_{xz}^2 C_z^2 - 4w^2 C_{xz} C_{yx} C_z^2 + 4w^2 C_{yx}^2 C_z^2 + 4wC_{xz} C_{yz}^2\right)$$

$$K_{3.3} = \lambda_1\left(4w^2 C_x^2 C_{yz}^2 - 4wC_{xz} C_{yz}^2 - 4w^2 C_x^2 C_y^2 C_z^2 + 4wC_{xz} C_y^2 C_z^2\right)$$

$$K_{3.4} = \lambda_2\left(4w^2 C_x^2 C_y^2 C_z^2 - 4wC_{xz} C_y^2 C_z^2 + 2C_{yx} C_z^4 - 4wC_{yx} C_{yz} C_z^2 - w^2 C_x^2 C_z^4 - 2wC_{xz} C_{yz} C_z^2 + 4w^2 C_x^2 C_{yz} C_z^2\right)$$

$$K_{3.5} = \lambda_4\left(C_y^2 C_z^4 + C_{yz}^2 C_z^2\right)$$

Substituting the value of β_1, β_2 and ϕ, into equation (16), we get

$$MSE(T) = \frac{S_y^2}{n}\left[\frac{\theta K_{2.1} + K_{2.2} + \theta^2 K_{2.3} + \rho_{yz}^4 \lambda_5 + 8\rho_{yz}^2 \lambda_m \lambda_u + \lambda_m^2 + 4\lambda_m \lambda_u}{K_{2.4} + \theta K_{2.5} - \theta^2 K_{2.6} + \rho_{yz}^2(\lambda_m^2 + \lambda_m \lambda_n) + \lambda_4}\right] \tag{29}$$

ISSN NUMBER: 1116-249X

Where

$$K_{2.1} = \lambda_0\left(4w^2\rho_{xz}\rho_{yx}\rho_{yz}^2 - w^2\rho_{yz}^2\rho_{xz}^2 - 4w^2\rho_{yx}^2\rho_{yz}^2 + w^2\rho_{xz}^2 - 4w^2\rho_{xz}\rho_{yx} + 4w^2\rho_{yx}^2\right)$$

$$K_{2.2} = \lambda_1\left(16w\rho_{xz}\rho_{yz}^4 - 16w^2\rho_{yz}^4 + 32w^2\rho_{yz}^2 - 32w\rho_{xz}\rho_{yz}^2 - 16w^2 + 16w\rho_{xz}\right)$$

$$K_{2.3} = \lambda_2\begin{pmatrix} 2w\rho_{xz}\rho_{yz}^3 - 4w^2\rho_{yz}^3 + 4w\rho_{yx}\rho_{yz}^3 - 3w^2\rho_{yz}^2 + 4w\rho_{xz}\rho_{yz}^2 - 2w\rho_{yx}\rho_{yz}^2 + 4w^2\rho_{yz} \\ -2w\rho_{xz}\rho_{yz} - 4w\rho_{yx}\rho_{yz} + 3w^2 - 4w\rho_{xz} + 2w\rho_{yx} \end{pmatrix}$$

$$K_{2..4} = \lambda_0\left(w^2\rho_{xz}^2 - 4w^2\rho_{xz}\rho_{yx} + 4w^2\rho_{yx}^2\right)$$

$$K_{2.5} = \lambda_1\left(4w^2\rho_{yz}^2 - 4w\rho_{xz}\rho_{yz}^2 - 4w^2 + 4w\rho_{xz}\right)$$

$$K_{2.6} = \lambda_2\left(4w^2\rho_{yz} - 2w\rho_{xz}\rho_{yz} - 4w\rho_{yx}\rho_{yz} + 3w^2 - 4w\rho_{xz} + 2w\rho_{yx}\right)$$

6. Optimum Replacement Policy

To obtain the optimum value of θ (fraction of sample to be drawn afresh on the second occasion) to enhance an efficient estimates of $\bar{Y}$ with maximum precision. MSE(T) will be minimize in equation (29) with respect to θ, which resulted to the quadratic equation as shown below.

$$d_1\theta^2 + 2d_2\theta + d_3 = 0 \tag{30}$$

Where $d_1 = K_{2.4}K_{2.1} - K_{22}K_{2.5}$; $d_2 = K_{2.4}K_{2.3} + K_{22}K_{2.5}$; $d_3 = K_{2.3}K_{2.5} + K_{2.6}K_{2.1}$

Solving equation (30) for θ gives;

$$\hat{\theta} = \frac{-d_2 \pm \sqrt{d_2 - d_1 d_3}}{d_1} \tag{31}$$

If the true values of $\hat{\theta}$ exist. Two real values of $\hat{\theta}$ are attainable for any combination of correlation coefficients ρ_{yx}, ρ_{yz} and ρ_{zx} that meet the requirements of a real solution. It is important to keep in mind that $0 \le \theta \le 1$; when selecting the values of $\hat{\theta}$. If all the values are acceptable, the lowest $\hat{\theta}$ value will be the best option, because it will lower the survey's cost. Equation (32) displays the optimal value of the mean square error of T, which is obtained by substituting the admissible value of $\hat{\theta}$, from equation (31) into equation (29).

$$MSE(T) = \frac{S_y^2}{n}\left[\frac{\theta_0 K_{2.1} + K_{2.2} + \theta_0{}^2 K_{2.3} + \rho_{yz}^4\lambda_5 + 8\rho_{yz}^2\lambda_m\lambda_u + \lambda_m^2 + 4\lambda_m\lambda_u}{K_{2.4} + \theta_0 K_{2.5} - \theta_0{}^2 K_{2.6} + \rho_{yz}^2(\lambda_m^2 + \lambda_m\lambda_n) + \lambda_4}\right] \tag{32}$$

7. Efficiency Comparison

If MSE(i) < MSE(j), then estimator (i) is assumed to be more efficient than estimator (j). The task at hand is to determine the circumstances in which the suggested estimator T performs better than the estimator, T_{SAG}, T_{SP}, and T_{KSS}. From equation 4, 8, 12 and 29. It can be shown that;

(i) $MSE(T)_{\min} < MSE(T_{SAG})$, if and only if

$$\partial_0\left(1-\rho_{zx}^2-\theta\left(\rho_{yx}^2+\rho_{yz}^2-2\rho_{yx}\rho_{yz}\rho_{xz}\right)\right)-\varsigma_{2.1}-\varsigma_{2.2}-\varsigma_{2.3}+\rho_{yz}^4\lambda_5-8\rho_{yz}^2\lambda_m\lambda_u-\lambda_m^2+4\lambda_m\lambda_u>0$$

(ii) $MSE(T)_{\min} < MSE(T_{SP})$, if and only if

$$\partial_0\left(\partial_0-\theta\left(\rho_{yx}^2+\rho_{yx}^2\rho_{xz}^2-2\rho_{yx}\rho_{yz}\rho_{xz}\right)\right)-\varsigma_{2.1}-\varsigma_{2.2}-\varsigma_{2.3}+\rho_{yz}^4\lambda_5-8\rho_{yz}^2\lambda_m\lambda_u-\lambda_m^2+4\lambda_m\lambda_u>0$$

(iii) $MSE(T)_{\min} < MSE(T_{KSS})$, if and only if

$$\partial_0\left(\partial_0\left(1-\theta\rho_{xz}^2\right)-\theta V^2\right)-\varsigma_{2.1}-\varsigma_{2.2}-\varsigma_{2.3}+\rho_{yz}^4\lambda_5-8\rho_{yz}^2\lambda_m\lambda_u-\lambda_m^2+4\lambda_m\lambda_u>0$$

Another measure of efficiency includes, the coefficient of variations (CV) of the stated estimators, it is calculated to determine how efficient is an estimator in the study. This is given as

$$CV=\frac{\sqrt{MSE(T_i)}}{T_i}\times 100$$

Estimator with the smallest CV is considered the “best” in the class of all estimators.

8. Empirical Study

Three populations are examined in this study; these populations are synthetic population studied using R-studio. The populations are as follows; population1: has N=100, n=50, u=30, m=20, Population2: has N=500, n=150, u=100, m=50 and Popultion3: has N=1000, n=300, u=198, m=102. Table 1 contains the data statistics for this populations.

ISSN NUMBER: 1116-249X

Table 1: Data Statistics

S/N	N	n	M	u	S_y^2	C_y	C_x	C_z	ρ_{yx}	ρ_{yz}	ρ_{xz}
1	100	50	30	30	11.9281	0.1454	0.0138	0.0053	0.7489	0.677	0.6977
2	500	150	50	100	193.32	0.1291	0.0182	0.0261	0.6127	0.445	0.4756
3	1000	300	97	203	971.29	0.2897	0.0322	0.0392	0.2894	0.154	0.2896

Table 2: Evaluating the estimator's performances using simulated data.

S/N	Population	Estimator	Estimate	MSE(T)	CV
1	N=100, n=50, u= 30, m=20	T_{SAG}	45.3230	0.0199	0.3112
		T_{SP}	44.9146	0.0185	0.3028
		T_{KSS}	44.8837	0.0144	0.2667
		T	44.8018	0.0077	0.1936
2	N=500, n=150, u= 100, m=50	T_{SAG}	42.8997	0.0063	0.1850
		T_{SP}	44.7056	0.0057	0.1689
		T_{KSS}	44.0708	0.0046	0.1596
		T	44.7108	0.0014	0.0837
3	N=1000, n=300, u= 198, m=102	T_{SAG}	84.5013	0.1252	0.4264
		T_{SP}	84.7886	0.1194	0.4075
		T_{KSS}	85.3999	0.1134	0.4036
		T	85.2453	0.0020	0.0525

Table 2 shows that the CV values of the suggested estimators are lower than those of the three existing estimators that are being considered in this study for three population sizes. These indicates that the proposed estimator T is more efficient than the estimators of Shabbir, Azam, and Gupta (2005) (T_{SAG}), Singh and Pal (2016) (T_{SP}), and Kuldeep Kumar Tiwari et al. (2023) (T_{KSS}).

9. Application to Case Study

In this section, we use data from the broadsheet of students results in faculty of sciences, in their first and second semester exams, to show the performances of the proposed estimator. The data comprises of their grade point average (GPA) denoted as (Y), their use of English (X) and General mathematical knowledge (Z). Table 3, shows their estimates of mean, their MSE and their coefficient of variation CV.

ISSN NUMBER: 1116-249X

Table 3: Evaluating the Estimator's Performances, using a Case Study

n	Estimator	Estimate	MSE	CV
n=100; u=82; m=18	$T_{(SAG)}$	2.1308	0.0029	2.527297169
	$T_{(SP)}$	2.8413	0.0024	1.724203529
	$T_{(KSS)}$	2.3021	0.0022	2.037450919
	T	2.7142	0.0002	0.521042503
n=150; u=105; m=45	$T_{(SAG)}$	0.1503	0.2727	97.4430177
	$T_{(SP)}$	2.7705	0.2496	18.03284028
	$T_{(KSS)}$	2.4949	0.2337	19.37653928
	T	2.6843	0.0091	3.553772683
n=200; u=131; m=69	$T_{(SAG)}$	0.7453	0.3045	74.03933805
	$T_{(SP)}$	2.8741	0.2621	17.8127772
	$T_{(KSS)}$	2.2376	0.2469	22.20639658
	T	2.7108	0.0097	3.633192342
n=250; u=135; m=115	$T_{(SAG)}$	0.8127	0.1901	53.64889715
	$T_{(SP)}$	2.7178	0.162	14.80948694
	$T_{(KSS)}$	2.1209	0.1544	18.5269298
	T	2.6449	0.0057	2.854487669

Table 3, shows evaluation metrics of the estimators, when applied to a case study of 100L student performances in their exams. From their MSE and CV the proposed estimator is seen to have the lowest value of MSE and CV at all the sample sizes. Which also agreed with the result obtained from the simulation studies.

10. Conclusion

This work presents an alternative estimator, for determining population mean in two occasion successive sampling. The estimates of mean, MSE, and CV expression of the suggested estimator is given. Also, the theoretical conditions under which the proposed estimator will outperforms the existing estimators studied in this work have been determined. An empirical investigation has been conducted using an artificial population. Three other current estimators that have been published by Shabbir, Azam, and Gupta (2005), Singh and Pal (2016), and Kuldeep Kumar Tiwari et al. (2023) are compared with the suggested estimator, the suggested estimator consistently performs better than the other given estimators considered in the study, judging by the mean square error (MSE) and Coefficient of variation (CV). Based on this empirical justifications, the suggested estimator is recommended for further research and practical application to real-world problems in sampling in two occasions especially when the correlation among the variables are positive.

References

Anieting, A. E., & Ezugwu, V. O. (2013). Estimation of the Current Occasion Parameters Using Successive Sampling Approach. *America Journal of Engineering Research (AJER)*, *2*(09), 204–207.

Bandyopadhyay, A., and Singh, G.N. (2016). Estimation of ratio of population means in two occasion successive sampling. communications in statistics theory and methods, 45(14), 4136-4155.

Etuk, S.I., Enang, E.I., and Ekpenyong, E,J. (2016). A modified class of ratio estimators for population mean. communication in statistics theory and methods, 45(4), 837-849.

Jessen, R. J. (1942). Statistical investigation of a sample survey for obtaining farm facts. *Research Bulletin*.

Khare, B.B. and Habib, U.R. (2013). Improved chain type estimators for population mean using two auxiliary variables and double sampling scheme. International Journal of Advanced Statistics and Probability, Vol.1(3) 82-87

Kumar, S. (2016). Improved estimation of population mean in presence of nonresponse and measurement error. Journal of statistical theory and practice, 10(4), 707-720.

Kumar, S., Bhougal, S., & Joorel, J. P. S. (2010). Improved Class of Estimators Using Multi-Auxiliary Information in Successive Sampling. *Pakistan Journal of Statistics and Operation Research*, *6*(2), 117. https://doi.org/10.18187/pjsor.v6i2.160

Muili, J. O., Aliero, T., Audu, A., and Yunusa, A. (2022). *On The Efficiency of Ratio Estimators of Finite Population Mean using Auxiliary Information*. *Oriental Journal of Physical Sciences February*. https://doi.org/10.13005/OJPS06.01-02.03

Rao, J. N. K., and Graham, J. E. (1964). Rotation Designs for Sampling on Repeated Occasions. *Australian and New Zealand Journal of Statistics*, *59*(306), 37–41.

Shabbir, J., Azam, M., and Sciences, A. (2005). Estimation of Current Population Mean in Successive Sampling. *Proc. ICCS-VIII, Lahore, Pakistan*, *13*(December), 77–85.

Singh, H. P. and Pal, S. K. (2016). An efficient effective rotation pattern in successive sampling over two. *Communication in Statistics*, *45*(17), 5017–5027.

Singh, G. N., and Homa, F. (2013). Effective Rotation Patterns in Successive Sampling Over Two Occasions. *Journal of Statistics Theory and Practice*, *7*(1), 37–41. https://doi.org/10.1080/15598608.2012.755484

Singh, G. N., and Khalid, M. (2018). Efficient class of estimators for finite population mean using auxiliary information in two-occasion successive sampling. *Journal of Modern Applied Statistical Methods*, *17*(2). https://doi.org/10.22237/jmasm/1554138114

Singh, G. N., Kumar, A., and Vishwakarma, G. K. (2018). Development of chain-type exponential estimators for population variance in two-phase sampling design in presence of random non-response. *Statistics in Transition New Series*, *19*(4), 575–596. https://doi.org/10.21307/stattrans-2018-031

Singh, G. N., Majhi, D., and Prasad, S. (2012). Estimation of population mean in two-phase successive sampling. *2012 1st International Conference on Recent Advances in Information Technology, RAIT-2012*, 861–866. https://doi.org/10.1109/RAIT.2012.6194538

Singh, G. N., and Singh, A. K. (2017). An Efficient Estimation Procedure in Two-Occasion Successive Sampling. *Journal Statistical Theory and Application, 16*(1), 26–37.

Singh, H. P., and Pal, S. K. (2017). A Review of Chain-Type Estimators and a New Chain-Type Multivariate Regression Approach for Population Mean Under Two-Occasion Successive Sampling. *Proceedings of the National Academy of Sciences, India Section A: Physical Sciences, December*. https://doi.org/10.1007/s40010-016-0313-x

Singh, M. K. G. N. (2020). General Estimation Procedures of Population Mean in Two-Occasion Successive Sampling under Random. *Proceedings of the National Academy of Sciences, India Section A: Physical Sciences*. https://doi.org/10.1007/s40010-020-00713-w

Singh, H. P., and Pal, S. K. (2017). A Modified Procedure for Estimating the Population Mean in Two-occasion Successive Samplings. Afrika Statistika, Vol. 13 (2), 2017, pages 1347 - 1365.

Singh, & Vishwakarma. (2007). A general class of estimators in successive sampling. *Metron*, *65*(2), 201–227.

Singh, G.N., Pandey, A.K., Singh, C., and Suman, S. (2019). An improved alternative estimation procedure for current population mean in presences of positively and negatively correlated auxiliary variables in two occasion rotation patterns. communication in statistics theory and methods, 49(13), 3094-3125. https://dio.org/10.1080/03610926.2019.1586932

Subramani, J. (2013). Generalized Modified Ratio Estimator for Estimation of Finite Population Mean. Journal of Modern Applied Statistical Methods, Vol.12(2) dOI: 10.22237/jmasm/1383278760 http://digitalcommons.wayne.edu/jmasm/vol12/iss2/7

Tiwari, K. K., Bhougal, S., & Kumar, S. (2023). On Efficient Estimation in Successive Sampling Over Two Occasions. *Thailand Statistician, 21*(2), 305–313.

Zoramthanga, R. (2018). Estimation of current population mean using two-occasion successive sampling with one auxiliary variable Sci Vis 18 (2), 74—77 (2018)

ISSN NUMBER: 1116-249X