

# A Class of Regression Estimator with Cum-Dual Ratio Estimator as Intercept

F. B. Adebola<sup>1</sup>, N. A. Adegoke<sup>1,\*</sup>, Ridwan A. Sanusi<sup>2</sup>

<sup>1</sup>Department of Statistics, Federal University of Technology Akure, Nigeria

<sup>2</sup>Department of Mathematics and Statistics, King Fahd University of Petroleum and Minerals, Saudi Arabia

**Abstract** This paper examines a class of regression estimator with cum-dual ratio estimator as intercept for estimating the mean of the study variable Y using auxiliary variable X. We obtained the bias and the mean square error of the proposed estimator, also, the asymptotically optimum estimator (AOE) was obtained along with its mean square error. Theoretical and numerical validation of the proposed estimator were done to show its superiority over the usual simple random sampling estimator and ratio estimator, product estimator, cum-dual ratio and product estimator. It was found that the estimator while performed better than other competing estimators, performed in almost the same way as the usual regression estimator when compared with the usual simple random estimator for estimating the average sleeping hours of undergraduate students of the department of statistics, Federal University of Technology Akure, Nigeria.

**Keywords** Difference estimator, Auxiliary variable, Cum-Dual ratio estimator, Bias, Mean square error, Efficiency

## 1. Introduction

Ratio estimation has gained relevance in estimation theory because of its improved precision in estimating the population parameters. It has been widely applied in agriculture to estimate the mean yield of crops in a certain area and in forestry, to estimate with high precision, the mean number of trees or crops in a forest or plantation. Other areas of relevance include economics and Population studies to estimate the ratio of income to family size. Utilizing Information from high resolution satellite data, [1] examined the possibilities of different forms of auxiliary information derived from remote sensing data in two-phase sampling design and suggested an appropriate estimator that would be of broad interest and applications by proposing a new class of regression-cum estimators in two-phase sampling. He found it to be more efficient than the usual regression and ratio estimators. A class of product-cum-dual to product estimators was proposed by [2] for estimating finite population mean of the study variate.

The use of auxiliary information at the estimation stage to increase the efficiency of the study variable was proposed by [3]. He used supplementary information on an auxiliary variable X positively correlated with Y to develop the ratio estimator to estimate the population mean or total of the study variable Y. The ratio-estimator is always more

efficient than the normal SRS when the relationship between the study variable Y and the auxiliary variable X is linear through the origin, and Y is proportional to X [4]. Product estimator was proposed by [5]. [6] suggested the use of ratio estimator  $\bar{y}_p$  when  $\frac{\rho C_y}{C_x} > \frac{1}{2}$  and unbiased estimator  $\bar{y}$  when  $-\frac{1}{2} \leq \rho \frac{C_y}{C_x} \leq \frac{1}{2}$ , where  $C_y, C_x$  and  $\rho$  are coefficient of variation of y, coefficient of variation of x and correlation between y and x respectively.

A lots of work have been done using auxiliary information. A ratio-cum-dual to ratio estimator was proposed for finite population mean. It was shown that the proposed estimator is more efficient than the simple mean estimator, usual ratio estimator and dual to ratio estimator under certain given conditions [7]. [8] proposed a modified ratio-cum-product estimator of finite population mean of the study variate Y using known correlation coefficient between two auxiliary characters X1 and X2, while [9] proposed a ratio-cum-product estimator of finite population mean using information on coefficient of variation and coefficient of kurtosis of auxiliary variate and showed that the proposed estimator is more efficient than the sample mean estimator, usual ratio and product estimators and estimators proposed by [10] under certain given conditions. Moreover, two exponential ratio estimators of population mean in simple random sampling without replacement were shown to be more efficient than the regression estimator and some existing estimators under review based on their biases, mean squared errors and also by using analytical and numerical results (at optimal conditions) for comparison [11]. Also, [12] suggested a ratio-cum product estimator of a finite

\* Corresponding author:

nurudeen.adeboke@yahoo.com (N. A. Adegoke)

Published online at <http://journal.sapub.org/ijps>

Copyright © 2015 Scientific & Academic Publishing. All Rights Reserved

population mean using information on the coefficient of variation and the coefficient of kurtosis of auxiliary variate in stratified random sampling.

Suppose that simple random sampling without replacement (SRSWOR) of  $n$  units is drawn from a population of  $N$  units to estimate the population mean

$$\bar{Y} = \frac{1}{N} \sum_{i=1}^N y_i \text{ of the study variable } Y. \text{ All the sample units}$$

are observed for the variables  $Y$  and  $X$ . Let  $(y_i, x_i)$  where  $i = 1, 2, 3, \dots, n$  denotes the set of the observation for the study variable  $Y$  and  $X$ . Let the sample means  $(\bar{x}, \bar{y})$  be unbiased of the population means of the auxiliary variable  $\bar{X}$  and study variable  $\bar{Y}$  based on the  $n$  observations.

The usual ratio estimator of  $\bar{Y}$  is given as  $\bar{y}_R = \frac{\bar{y}}{\bar{x}} \bar{X}$  and the usual regression estimator is given as  $\bar{y}_{reg} = \bar{y} + \hat{\beta}(\bar{X} - \bar{x})$ , where  $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ ,  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$

and  $\hat{\beta} = \frac{s_{xy}}{s_x^2}$  is the estimate slope of regression coefficient of  $Y$  and  $X$ . [13] obtained dual to ratio-cum estimator given as  $\bar{y}_{dR} = \bar{y} \frac{\bar{x}^*}{\bar{X}}$ , where  $\bar{x}^*$  is the un-sampled auxiliary variable in  $X$  given as  $\bar{x}^* = \frac{N\bar{X} - n\bar{x}}{N - n}$ . The use of auxiliary information in sample surveys was

extensively discussed in well-known classical text books such as [14], [15], [16], [17] and [18] among others. Recent developments in ratio and product methods of estimation along with their variety of modified forms are lucidly described in detail by [19].

In this paper, we proposed a class of difference estimator with dual to ratio cum as the slope of the estimator instead of  $\bar{y}$ , also,  $\bar{x}^*$  was used instead of  $\bar{x}$  in the usual regression estimator. The proposed estimator is used to estimate the average sleeping hours of undergraduate students of the department of statistics, Federal University of Technology Akure, Nigeria.

The organization of this article is as follows: In Section 2, we provide the conceptual framework of the proposed class of estimator. We derived its bias, Mean Squared Error (MSE) and the resulting optimum value of the MSE, with their rigorous proofs up to order one. In section 3, we compared the MSE of the proposed estimator  $\bar{y}_{Rd}$  with the MSE of  $\bar{y}$  under Simple Random Sampling Scheme, in Section 4, we provide the numerical validation of the proposed estimator by using data on the ages and hours of sleeping by the undergraduate students of the Department of Statistics Federal University of Technology Akure, Ondo State, Nigeria. Finally, Section 5 provides the conclusion of our findings.

## 2. The Proposed Class of Estimator

For estimating population mean  $\bar{Y}$ , we have proposed a class of difference estimator with dual to Ratio cum as the intercept given as

$$\bar{y}_{Rd}^* = \bar{y} \frac{\bar{x}^*}{\bar{X}} + \alpha(\bar{X} - \bar{x}^*) \quad (1)$$

Where  $\alpha$  is a suitably chosen scalar. The bias and mean square error (MSE) of  $\bar{y}_{Rd}^*$  to the first order approximation is obtained by substituting  $\bar{x}^* = \frac{N\bar{X} - n\bar{x}}{N - n}$  into equation (1), hence, equation (1) becomes,

$$\begin{aligned} \bar{y}_{Rd}^* &= \frac{\bar{y}}{\bar{X}} \left( \frac{N\bar{X} - n\bar{x}}{N - n} \right) + \alpha \left( \bar{X} - \frac{N\bar{X} - n\bar{x}}{N - n} \right) \\ \bar{y}_{Rd}^* &= \frac{\bar{y}}{\bar{X}} \left( \frac{N\bar{X} - n\bar{x}}{N - n} \right) + \alpha \left( \frac{(N - n)\bar{X} - N\bar{X} + n\bar{x}}{N - n} \right) \\ \bar{y}_{Rd}^* &= \frac{\bar{y}}{N - n} \left( \frac{N\bar{X} - n\bar{x}}{\bar{X}} \right) + \alpha \left( \frac{n\bar{x} - n\bar{X}}{N - n} \right) \end{aligned} \quad (2)$$

We write,

$$e_0 = \frac{\bar{y} - \bar{Y}}{\bar{Y}} \text{ and } e_1 = \frac{\bar{x} - \bar{X}}{\bar{X}}$$

This implies that

$$\bar{y} = \bar{Y}(1 + e_0) \text{ and } \bar{x} = \bar{X}(1 + e_1),$$

Respectively. Hence, equation (2) becomes,

$$\begin{aligned} \bar{y}_{Rd}^* &= \frac{\bar{Y}(1 + e_0)}{N - n} \frac{(N\bar{X} - n\bar{X}(1 + e_1))}{\bar{X}} + \frac{\alpha}{N - n} (n\bar{X}(1 + e_1) - n\bar{X}) \\ \bar{y}_{Rd}^* &= \frac{\bar{Y}(1 + e_0)}{N - n} (N - n(1 + e_1)) + \frac{\alpha}{N - n} (ne_1\bar{X}) \end{aligned}$$

$$\begin{aligned}\bar{y}_{Rd}^* &= \frac{\bar{Y}(1 + e_0)}{N - n}((N - n) - ne_1) + \frac{\alpha}{N - n}(ne_1\bar{X}) \\ \bar{y}_{Rd}^* &= \bar{Y}(1 + e_0)\left(\frac{(N - n)}{N - n} - ge_1\right) + \alpha ge_1\bar{X}\end{aligned}$$

Where  $g = \frac{n}{N - n}$

$$\begin{aligned}\bar{y}_{Rd}^* &= \bar{Y}(1 + e_0)(1 - ge_1) + \alpha ge_1\bar{X} \\ \bar{y}_{Rd}^* &= \bar{Y}(1 + e_0) - \bar{Y}ge_1(1 + e_0) + \alpha ge_1\bar{X} \\ \bar{y}_{Rd}^* &= \bar{Y} + \bar{Y}e_0 - \bar{Y}ge_1(1 + e_0) + \alpha ge_1\bar{X} \\ (\bar{y}_{Rd}^* - \bar{Y}) &= \bar{Y}e_0 - \bar{Y}ge_1(1 + e_0) + \alpha ge_1\bar{X}\end{aligned}\quad (3)$$

By taking the expectation of equation (3) we have

$$E(\bar{y}_{Rd}^* - \bar{Y}) = E(\bar{Y}e_0) - \bar{Y}g(E(e_1) + E(e_0e_1)) + \alpha g\bar{X}E(e_1) \quad (4)$$

But,  $E(e_0) = E\left(\frac{\bar{y} - \bar{Y}}{\bar{Y}}\right) = \frac{1}{\bar{Y}}(\bar{Y} - \bar{Y}) = 0$ ,

$$E(e_1) = E\left(\frac{\bar{x} - \bar{X}}{\bar{X}}\right) = \frac{1}{\bar{X}}(\bar{X} - \bar{X}) = 0,$$

and

$$E(e_0e_1) = E\left(\frac{\bar{x} - \bar{X}}{\bar{X}}\right)\left(\frac{\bar{y} - \bar{Y}}{\bar{Y}}\right) = \left(\frac{1 - f}{n}\right)\frac{S_{XY}}{\bar{X}}$$

Hence, equation (4) becomes,

$$E(\bar{y}_{Rd}^* - \bar{Y}) = -\bar{Y}g\frac{(1 - f)S_{XY}}{n\bar{X}\bar{Y}} = -g\left(\frac{N - n}{Nn}\right)\frac{S_{XY}}{\bar{X}}.$$

Recall,

$$\begin{aligned}g &= \frac{n}{N - n} \\ Bias(\bar{y}_{Rd}^*) &= -\left(\frac{n}{N - n}\right)\left(\frac{N - n}{Nn}\right)\frac{S_{XY}}{\bar{X}} \\ Bias(\bar{y}_{Rd}^*) &= -\frac{S_{XY}}{N\bar{X}}\end{aligned}$$

The Mean Square Error of the estimator given as  $MSE(\bar{y}_{pd})$  is obtained by squaring both sides of equation (3) and taking the expectation. We have

$$\begin{aligned}(\bar{y}_{Rd}^* - \bar{Y})^2 &= \bar{Y}^2e_0^2 + \bar{Y}^2g^2(e_1^2 + e_1^2e_0^2 + 2e_0e_1^2) + \alpha^2g^2\bar{X}^2e_1^2 - 2\bar{Y}^2g(e_0e_1 + e_1e_0^2) \\ &\quad + 2g\alpha\bar{Y}\bar{X}e_0e_1 - 2\alpha g^2\bar{Y}\bar{X}(e_1^2 + e_1^2e_0)\end{aligned}$$

Ignoring the higher powers of error greater than or equal to 3, we have.

$$\begin{aligned}(\bar{y}_{Rd}^* - \bar{Y})^2 &= \bar{Y}^2e_0^2 + \bar{Y}^2g^2e_1^2 + \alpha^2g^2\bar{X}^2e_1^2 - 2\bar{Y}^2ge_0e_1 + 2g\alpha\bar{Y}\bar{X}e_0e_1 - 2\alpha g^2\bar{Y}\bar{X}e_1^2 \\ (\bar{y}_{Rd}^* - \bar{Y})^2 &= \bar{Y}^2e_0^2 + e_1^2g^2(\bar{Y}^2 - 2\alpha\bar{X}\bar{Y} + \alpha^2\bar{X}^2) + 2ge_0e_1(\alpha\bar{Y}\bar{X} - \bar{Y}^2) \\ (\bar{y}_{Rd}^* - \bar{Y})^2 &= \bar{Y}^2e_0^2 + e_1^2g^2(\bar{Y} - \alpha\bar{X})^2 - 2ge_0e_1\bar{Y}(\bar{Y} - \alpha\bar{X})\end{aligned}\quad (5)$$

Take the expectation of (5) we have

$$\begin{aligned}xE(\bar{y}_{Rd}^* - \bar{Y})^2 &= \left(\frac{1 - f}{n}\right)\left(\bar{Y}^2\frac{S_y^2}{\bar{Y}^2} - 2g\bar{Y}\frac{S_{xy}}{\bar{X}\bar{Y}}(\bar{Y} - \alpha\bar{X}) + g^2\frac{S_x^2}{\bar{X}^2}(\bar{Y} - \alpha\bar{X})^2\right) \\ MSE(\bar{y}_{Rd}^*) &= \left(\frac{1 - f}{n}\right)\left(S_y^2 - 2g\bar{Y}\frac{\rho S_x S_y}{\bar{X}\bar{Y}}(\bar{Y} - \alpha\bar{X}) + g^2\frac{S_x^2}{\bar{X}^2}(\bar{Y} - \alpha\bar{X})^2\right) \\ MSE(\bar{y}_{Rd}^*) &= \left(\frac{1 - f}{n}\right)(\bar{Y}^2C_y^2 - 2g\rho\bar{Y}C_xC_y(\bar{Y} - \alpha\bar{X}) + g^2C_x^2(\bar{Y} - \alpha\bar{X})^2)\end{aligned}\quad (6)$$

Where  $C_y = \frac{S_y}{\bar{Y}}$ , and  $C_x = \frac{S_x}{\bar{X}}$ .

The optimum value of the  $MSE(\bar{y}_{pd})$  is given as

$$\begin{aligned}\frac{\partial}{\partial \alpha} MSE(\bar{y}_{Rd}^*) &= \left(\frac{1-f}{n}\right) \frac{\partial}{\partial \alpha} (\bar{Y}^2 C_y^2 - 2g\rho \bar{Y} C_x C_y (\bar{Y} - \alpha \bar{X}) + g^2 C_x^2 (\bar{Y} - \alpha \bar{X})^2) \\ \frac{\partial}{\partial \alpha} MSE(\bar{y}_{Rd}) &= \left(\frac{1-f}{n}\right) \left(2g\bar{Y} \frac{S_{xy}}{\bar{X}\bar{Y}} (\bar{X}) - 2g^2 C_x^2 \bar{X} (\bar{Y} - \alpha \bar{X})\right)\end{aligned}\quad (7)$$

Set equation (7) to zero; we have

$$\begin{aligned}2gS_{xy} - 2g^2 C_x^2 \bar{X} (\bar{Y} - \alpha \bar{X}) &= 0 \\ \alpha &= \frac{\bar{Y}}{\bar{X}} - \frac{\beta}{g} \\ \alpha &= R - \frac{\beta}{g}\end{aligned}$$

Where  $\beta = \frac{S_{xy}}{S_x^2}$  and  $R = \frac{\bar{Y}}{\bar{X}}$

Thus, the resulting OPTIMUM MSE of  $\bar{y}_{pd}^*$  is given as

$$MSE(\bar{y}_{Rd}^{*opt}) = \left(\frac{1-f}{n}\right) \left(S_y^2 - 2g\rho \frac{S_x S_y}{\bar{X}} \left(\bar{Y} - \left(R - \frac{\beta}{g}\right) \bar{X}\right) + g^2 \frac{S_x^2}{\bar{X}^2} \left(\bar{Y} - \left(R - \frac{\beta}{g}\right) \bar{X}\right)^2\right)$$

By substituting  $\alpha$  in equation (6)

$$MSE(\bar{y}_{Rd}^{*opt}) = \left(\frac{1-f}{n}\right) \left(S_y^2 - 2g\rho \frac{S_x S_y}{\bar{X}} \left(\bar{Y} - R\bar{X} + \frac{\beta \bar{X}}{g}\right) + g^2 \frac{S_x^2}{\bar{X}^2} \left(\bar{Y} - R\bar{X} + \frac{\beta \bar{X}}{g}\right)^2\right)$$

Where  $R\bar{X} = \bar{Y}$ ,

$$\begin{aligned}MSE(\bar{y}_{Rd}^{*opt}) &= \left(\frac{1-f}{n}\right) \left(S_y^2 - 2g\rho \frac{S_x S_y}{\bar{X}\bar{Y}} \left(\frac{\beta \bar{X}}{g}\right) + g^2 \frac{S_x^2}{\bar{X}^2} \left(\frac{\beta \bar{X}}{g}\right)^2\right) \\ MSE(\bar{y}_{Rd}^{*opt}) &= \left(\frac{1-f}{n}\right) (S_y^2 - 2S_{xy}\beta + S_x^2\beta^2) \\ MSE(\bar{y}_{Rd}^{*opt}) &= \left(\frac{1-f}{n}\right) \left(S_y^2 - 2S_{xy} \left(\frac{S_{xy}}{S_x^2}\right) + S_x^2 \left(\frac{S_{xy}}{S_x^2}\right)^2\right) \\ MSE(\bar{y}_{Rd}^{*opt}) &= \left(\frac{1-f}{n}\right) \left(S_y^2 - \frac{2(S_{xy})^2}{S_x^2} + \frac{(S_{xy})^2}{S_x^2}\right) \\ MSE(\bar{y}_{Rd}^{*opt}) &= \left(\frac{1-f}{n}\right) \left(S_y^2 - \frac{(S_{xy})^2}{S_x^2}\right) \\ MSE(\bar{y}_{Rd}^{*opt}) &= \left(\frac{1-f}{n}\right) S_y^2 \left(1 - \left(\frac{S_{xy}}{S_x S_y}\right)^2\right)\end{aligned}\quad (7)$$

$MSE(\bar{y}_{Rd}^{*opt}) = \left(\frac{1-f}{n}\right) S_y^2 (1 - \rho^2)$  Equation (7) shows that the MSE  $(\bar{y}_{Rd}^{*opt})$  is the same as the MSE Regression estimator.

#### Remark

The Bias of  $\bar{y}_{Rd}^*$  is the same as Bias of the dual ratio estimator  $\bar{y}_R^*$  and when  $\alpha = 0$ ,  $MSE(\bar{y}_{Rd}^*)$  becomes  $MSE(\bar{y}_R^*)$  of dual to ratio estimator  $\bar{y}_R^*$  proposed by [13].

The bias of  $\bar{y}_R^*$  is given as

$$Bias(\bar{y}_R^*) = -\frac{S_{xy}}{N\bar{X}}$$

The MSE  $(\bar{y}_R^*)$  is given as

$$MSE(\bar{y}_R^*) = \left(\frac{1-f}{n}\right) (\bar{Y}^2 C_y^2 - 2\rho g \bar{Y}^2 C_x C_y + \bar{Y} g^2 C_x^2) \quad (7a)$$

### 3. Efficiency Comparisons

In this section, we compared the MSE of the proposed estimator  $\bar{y}_{Rd}$  with the MSE of  $\bar{y}$  under Simple Random Sampling Scheme given as,

$$MSE(\bar{y}) = \left(\frac{1-f}{n}\right) \bar{Y}^2 C_y^2 \quad (8)$$

From equations (5) and (8), the proposed estimator is better than that the usual estimator  $\bar{y}$  if,  $MSE(\bar{y}_{Rd}^*) < MSE(\bar{y})$ . That is,

$$\begin{aligned} \left(\frac{1-f}{n}\right) (\bar{Y}^2 C_y^2 - 2g\rho\bar{Y}C_xC_y(\bar{Y} - \alpha\bar{X}) + g^2C_x^2(\bar{Y} - \alpha\bar{X})^2) &< \left(\frac{1-f}{n}\right) \bar{Y}^2 C_y^2 \\ -2g\rho\bar{Y}C_xC_y(\bar{Y} - \alpha\bar{X}) + g^2C_x^2(\bar{Y} - \alpha\bar{X})^2 &< 0 \\ (\bar{Y} - \alpha\bar{X}) (-2g\rho\bar{Y}C_xC_y + g^2C_x^2(\bar{Y} - \alpha\bar{X})) &< 0 \end{aligned}$$

This holds if and only if,

$$\text{Case (1) } \bar{Y} - \alpha\bar{X} < 0 \text{ and } -2g\rho\bar{Y}C_xC_y + g^2C_x^2(\bar{Y} - \alpha\bar{X}) > 0$$

Or

$$\text{Case (2) } \bar{Y} - \alpha\bar{X} > 0 \text{ and } -2g\rho\bar{Y}C_xC_y + g^2C_x^2(\bar{Y} - \alpha\bar{X}) < 0$$

The Range of  $\alpha$  under which the proposed estimator  $\bar{y}_{pd}^*$  is more efficient than  $\bar{y}$  is given as,

$$\min\left\{R, R\left(1 - \frac{2\rho C_y}{gC_x}\right)\right\}, \max\left\{R, R\left(1 - \frac{2\rho C_y}{gC_x}\right)\right\}.$$

We also compared the proposed estimator  $\bar{y}_{Rd}^*$  with the usual ratio estimator  $\bar{y}_R$ . The MSE of the  $\bar{y}_R$  is given as

$$MSE(\bar{y}_R) = \left(\frac{1-f}{n}\right) (\bar{Y}^2 C_y^2 - 2\rho\bar{Y}^2 C_x C_y + \bar{Y}^2 C_x^2)$$

It is found that the proposed estimator  $\bar{y}_{Rd}^*$  will be more efficient than the usual ratio estimator  $\bar{y}_R$  if  $MSE(\bar{y}_{Rd}^*) < MSE(\bar{y}_R)$ .

That is,

$$\begin{aligned} \left(\frac{1-f}{n}\right) (\bar{Y}^2 C_y^2 - 2g\rho\bar{Y}C_xC_y(\bar{Y} - \alpha\bar{X}) + g^2C_x^2(\bar{Y} - \alpha\bar{X})^2) &\leq \left(\frac{1-f}{n}\right) (\bar{Y}^2 C_y^2 - 2\rho\bar{Y}^2 C_x C_y + \bar{Y}^2 C_x^2) \\ -2g\rho\bar{Y}C_xC_y(\bar{Y} - \alpha\bar{X}) + 2\rho\bar{Y}^2 C_x C_y + g^2C_x^2(\bar{Y} - \alpha\bar{X})^2 - \bar{Y}^2 C_x^2 &< 0 \\ -2\rho\bar{Y}C_xC_y(g(\bar{Y} - \alpha\bar{X}) - \bar{Y}) + C_x^2((\bar{Y} - \alpha\bar{X})^2 - \bar{Y}^2) &< 0 \\ (g(\bar{Y} - \alpha\bar{X}) - \bar{Y}) (-2\rho\bar{Y}C_xC_y + C_x^2(g(\bar{Y} - \alpha\bar{X}) + \bar{Y})) &< 0 \end{aligned}$$

This holds if the following two conditions are satisfied

$$(1). (g(\bar{Y} - \alpha\bar{X}) - \bar{Y}) < 0 \text{ And } -2\rho\bar{Y}C_xC_y + C_x^2(g(\bar{Y} - \alpha\bar{X}) + \bar{Y}) > 0.$$

Or

$$(2). (g(\bar{Y} - \alpha\bar{X}) - \bar{Y}) > 0 \text{ And } -2\rho\bar{Y}C_xC_y + C_x^2(g(\bar{Y} - \alpha\bar{X}) + \bar{Y}) < 0.$$

This condition holds if  $\alpha > R\left(\frac{2n-N}{n}\right)$  and  $\alpha < R\left(\frac{N}{n} - \frac{2\rho C_y}{gC_x}\right)$  or  $\alpha < R\left(\frac{2n-N}{n}\right)$  and  $\alpha > R\left(\frac{N}{n} - \frac{2\rho C_y}{gC_x}\right)$

$$\min\left\{R\left(\frac{2n-N}{n}\right), R\left(\frac{N}{n} - \frac{2\rho C_y}{gC_x}\right)\right\}, \max\left\{R\left(\frac{2n-N}{n}\right), R\left(\frac{N}{n} - \frac{2\rho C_y}{gC_x}\right)\right\}.$$

We also compared the proposed estimator  $\bar{y}_{Rd}^*$  with the usual product estimator  $\bar{y}_p$ . The MSE of the  $\bar{y}_p$  is given as

$$MSE(\bar{y}_p) = \left(\frac{1-f}{n}\right) (\bar{Y}^2 C_y^2 + 2\rho\bar{Y}^2 C_x C_y + \bar{Y}^2 C_x^2)$$

It is found that the proposed estimator  $\bar{y}_{Rd}^*$  will be more efficient than the usual ratio estimator  $\bar{y}_p$  if  $MSE(\bar{y}_{Rd}^*) < MSE(\bar{y}_p)$ .

That is,

$$\begin{aligned} \left(\frac{1-f}{n}\right) (\bar{Y}^2 C_y^2 - 2g\rho\bar{Y}C_xC_y(\bar{Y} - \alpha\bar{X}) + g^2C_x^2(\bar{Y} - \alpha\bar{X})^2) &\leq \left(\frac{1-f}{n}\right) (\bar{Y}^2 C_y^2 + 2\rho\bar{Y}^2 C_x C_y + \bar{Y}^2 C_x^2) \\ -2g\rho\bar{Y}C_xC_y(\bar{Y} - \alpha\bar{X}) - 2\rho\bar{Y}^2 C_x C_y + g^2C_x^2(\bar{Y} - \alpha\bar{X})^2 - \bar{Y}^2 C_x^2 &< 0 \\ -2\rho\bar{Y}C_xC_y(g(\bar{Y} - \alpha\bar{X}) + \bar{Y}) + C_x^2((\bar{Y} - \alpha\bar{X})^2 - \bar{Y}^2) &< 0 \end{aligned}$$

$$(g(\bar{Y} - \alpha\bar{X}) + \bar{Y}) \left( -2\rho\bar{Y}C_xC_y + C_x^2(g(\bar{Y} - \alpha\bar{X}) - \bar{Y}) \right) < 0$$

This holds if the following two conditions are satisfied

$$(1). (g(\bar{Y} - \alpha\bar{X}) + \bar{Y}) < 0 \text{ and } -2\rho\bar{Y}C_xC_y + C_x^2(g(\bar{Y} - \alpha\bar{X}) - \bar{Y}) > 0.$$

Or

$$(2). (g(\bar{Y} - \alpha\bar{X}) + \bar{Y}) > 0 \text{ and } -2\rho\bar{Y}C_xC_y + C_x^2(g(\bar{Y} - \alpha\bar{X}) - \bar{Y}) < 0.$$

This condition holds if  $\alpha > R\left(\frac{N}{n}\right)$  and  $\alpha < R\left(\frac{2n-N}{n} - \frac{2\rho C_y}{gC_x}\right)$  or  $\alpha < R\left(\frac{N}{n}\right)$  and  $\alpha > R\left(\frac{2n-N}{n} - \frac{2\rho C_y}{gC_x}\right)$

$$\min\left\{R\left(\frac{N}{n}\right), R\left(\frac{2n-N}{n} - \frac{2\rho C_y}{gC_x}\right)\right\}, \max\left\{R\left(\frac{N}{n}\right), R\left(\frac{2n-N}{n} - \frac{2\rho C_y}{gC_x}\right)\right\}.$$

We also compared the MSE of the proposed estimator with MSE of dual product estimator  $\bar{y}_p^*$  proposed by [20]. The MSE ( $\bar{y}_p^*$ ) proposed by [20] (1980) is given as

$$MSE(\bar{y}_p^*) = \left(\frac{1-f}{n}\right) (\bar{Y}^2C_y^2 + 2g\rho\bar{Y}^2C_xC_y + \bar{Y}^2g^2C_x^2)$$

It is found that the proposed estimator  $\bar{y}_{pd}^*$  will be more efficient than that of [20] estimator  $\bar{y}_p^*$  if  $MSE(\bar{y}_{pd}^*) < MSE(\bar{y}_p^*)$ . That is

$$\left(\frac{1-f}{n}\right) (\bar{Y}^2C_y^2 - 2g\rho\bar{Y}C_xC_y(\bar{Y} - \alpha\bar{X}) + g^2C_x^2(\bar{Y} - \alpha\bar{X})^2) \leq \left(\frac{1-f}{n}\right) (\bar{Y}^2C_y^2 + 2g\rho\bar{Y}^2C_xC_y + \bar{Y}^2g^2C_x^2)$$

That is,

$$\begin{aligned} -2g\rho\bar{Y}C_xC_y(\bar{Y} - \alpha\bar{X}) + g^2C_x^2(\bar{Y} - \alpha\bar{X})^2 &< 2g\rho\bar{Y}^2C_xC_y + \bar{Y}^2g^2C_x^2 \\ -2g\rho\bar{Y}C_xC_y[(\bar{Y} - \alpha\bar{X}) + \bar{Y}] + g^2C_x^2[(\bar{Y} - \alpha\bar{X})^2 - \bar{Y}^2] &< 0 \\ ((\bar{Y} - \alpha\bar{X}) + \bar{Y})[-2g\rho\bar{Y}C_xC_y + g^2C_x^2[(\bar{Y} - \alpha\bar{X}) - \bar{Y}]] &< 0 \\ (2\bar{Y} - \alpha\bar{X})[-2g\rho\bar{Y}C_xC_y - g^2C_x^2[\alpha\bar{X}]] &< 0 \end{aligned}$$

This holds if,

1.  $2\bar{Y} - \alpha\bar{X} < 0$  and  $-2g\rho\bar{Y}C_xC_y - g^2C_x^2[\alpha\bar{X}] > 0$
2.  $2\bar{Y} - \alpha\bar{X} > 0$  and  $-2g\rho\bar{Y}C_xC_y - g^2C_x^2[\alpha\bar{X}] < 0$

This condition holds if  $2R > \alpha$  and  $\frac{-2R\rho C_y}{gC_x} < \alpha$  or  $2R < \alpha$  and  $\frac{-2R\rho C_y}{gC_x} > \alpha$

The range of  $\alpha$  under which the proposed estimator  $\bar{y}_{pd}^*$  is more efficient than  $\bar{y}_p^*$  is

$$\min\left\{2R, \frac{-2R\rho C_y}{gC_x}\right\}, \max\left\{2R, \frac{-2R\rho C_y}{gC_x}\right\}$$

Lastly, we compared MSE of the proposed estimator  $\bar{y}_{rd}^*$  with that of dual to ratio estimator  $\bar{y}_R^*$  proposed [13] given in equation (7a). The proposed estimator will be more efficient than  $\bar{y}_R^*$  if

$MSE(\bar{y}_{rd}^*) < MSE(\bar{y}_R^*)$ . That is,

$$\begin{aligned} \left(\frac{1-f}{n}\right) (\bar{Y}^2C_y^2 - 2g\rho\bar{Y}C_xC_y(\bar{Y} - \alpha\bar{X}) + g^2C_x^2(\bar{Y} - \alpha\bar{X})^2) &\leq \left(\frac{1-f}{n}\right) (\bar{Y}^2C_y^2 - 2g\rho\bar{Y}^2C_xC_y + \bar{Y}^2g^2C_x^2) \\ -2g\rho\bar{Y}C_xC_y[(\bar{Y} - \alpha\bar{X}) - \bar{Y}] + g^2C_x^2[(\bar{Y} - \alpha\bar{X})^2 - \bar{Y}^2] &< 0 \\ ((\bar{Y} - \alpha\bar{X}) - \bar{Y})[-2g\rho\bar{Y}C_xC_y + g^2C_x^2[(\bar{Y} - \alpha\bar{X}) + \bar{Y}]] &< 0 \\ -\alpha\bar{X}[-2g\rho\bar{Y}C_xC_y + g^2C_x^2(2\bar{Y} - \alpha\bar{X})] &< 0 \end{aligned}$$

This holds if

1.  $-\alpha\bar{X} < 0$  and  $-2g\rho\bar{Y}C_xC_y + g^2C_x^2(2\bar{Y} - \alpha\bar{X}) < 0$
2.  $-\alpha\bar{X} > 0$  and  $-2g\rho\bar{Y}C_xC_y + g^2C_x^2(2\bar{Y} - \alpha\bar{X}) > 0$

This condition holds if  $\alpha > 0$  and  $\alpha < 2R \left(1 - \frac{\rho C_y}{g C_x}\right)$  or  $\alpha < 0$  and  $\alpha > 2R \left(1 - \frac{\rho C_y}{g C_x}\right)$

Therefore, the range of  $\alpha$  under which the proposed estimator  $\bar{y}_{Rd}^*$  is more efficient than dual ratio estimator

$$\min \left\{ 0, 2R \left(1 - \frac{\rho C_y}{g C_x}\right) \right\}, \max \left\{ 0, 2R \left(1 - \frac{\rho C_y}{g C_x}\right) \right\}$$

Thus it seems from the above that the proposed estimator  $\bar{y}_{Rd}^*$  may be made better than the usual estimator, ratio estimator, product estimator, dual to product estimator  $\bar{y}_p^*$  and the dual to ratio estimator  $\bar{y}_R^*$ , if the given conditions are satisfied.

#### Comparison of 'AOE' to $\bar{y}_{Rd}^{OPT}$

$\bar{y}_{Rd}^{OPT}$  is more efficient than the other existing estimators  $\bar{y}$ , the ratio estimator  $\bar{y}_R$ , the product estimator  $\bar{y}_p$ , the dual to ratio estimator  $\bar{y}_R^*$  and the dual to product estimator  $\bar{y}_p^*$  since

$$MSE(\bar{y}) - MSE(\bar{y}_{Rd}^{OPT}) = \left(\frac{1-f}{n}\right) (\bar{Y}^2 \rho^2 C_y^2) > 0$$

$$MSE(\bar{y}_R) - MSE(\bar{y}_{Rd}^{OPT}) = \left(\frac{1-f}{n}\right) \left( \bar{Y}^2 C_x^2 \left(1 - \frac{\rho C_y}{C_x}\right)^2 \right) > 0$$

$$MSE(\bar{y}_p) - MSE(\bar{y}_{Rd}^{OPT}) = \left(\frac{1-f}{n}\right) \left( \bar{Y}^2 C_x^2 \left(1 + \frac{\rho C_y}{C_x}\right)^2 \right) > 0$$

$$MSE(\bar{y}_R^*) - MSE(\bar{y}_{Rd}^{OPT}) = \left(\frac{1-f}{n}\right) \left( \bar{Y}^2 C_x^2 \left(\frac{\rho C_y}{C_x} - g\right)^2 \right) > 0$$

$$MSE(\bar{y}_p^*) - MSE(\bar{y}_{Rd}^{OPT}) = \left(\frac{1-f}{n}\right) \left( \bar{Y}^2 C_x^2 \left(\frac{\rho C_y}{C_x} + g\right)^2 \right) > 0$$

Hence, we conclude that the proposed class of estimator  $\bar{y}_{Rd}^*$  is more efficient than other estimator in case of its optimality.

## 4. Numerical Validation

To illustrate the efficiency of the proposed estimator over the other estimators  $\bar{y}$ ,  $\bar{y}_R$ ,  $\bar{y}_p$ ,  $\bar{y}_R^*$  and  $\bar{y}_p^*$ . Data on the ages and hours of sleeping by the undergraduate students of the Department of Statistics Federal University of Technology Akure, Ondo State, Nigeria. A sample of 150 out of 461 students of the department was obtained using simple random sampling without replacement. The information on the age of the students was used as auxiliary information to increase the precision of the estimate of the average sleeping hours. The estimate of the average hours of sleeping of the students were obtained and also the 95% confidence intervals of the average hours of sleeping were obtained for the proposed estimator and the other estimators. Table 1., gives the estimates of the average sleeping hours and the 95% confidence Interval. As shown in Table 1.0, the proposed estimator performed better than the other estimators, the width of the confidence interval of the proposed estimator is smallest than the other competing estimators.

**Table 1.** Average Sleeping Hours and 95% confidence intervals for Different Estimators for the undergraduate Students of Department of Statistics, Federals University of Technology Akure. Nigeria

ESTIMATOR	Average Sleeping Hours	LCL	UCL	WIDTH
$\bar{y}$	6.08	5.930386531	6.229613469	0.299226939
$\bar{y}_R$	6.210472103	6.042844235	6.378099971	0.335255737
$\bar{y}_p$	5.952268908	5.778821411	6.125716404	0.346894993
$\bar{y}_R^*$	6.141606636	5.988421023	6.294792249	0.306371226
$\bar{y}_p^*$	6.019011342	5.862732122	6.175290562	0.31255844
$\bar{y}_{RD}^*$	<b>6.194882566</b>	<b>6.050452623</b>	<b>6.33931251</b>	<b>0.288859888</b>

The proposed estimator performed the same way as the regression estimator when compared with the usual simple random sampling. The average Sleeping Hours and 95% confidence intervals for the proposed estimator and the regression estimator is given below, the two estimators have the same width.

**Table 2.** Average Sleeping Hours and 95% confidence intervals for the proposed estimators and regression estimators for the undergraduate Students of Department of Statistics, Federal University of Technology Akure, Nigeria

ESTIMATOR	Average Sleeping Hours	LCL	UCL	WIDTH
$\bar{y}_{RD}^*$	6.194882566	6.050452623	6.33931251	0.288859888
$\bar{y}_{REG}^*$	6.089652737	5.945222793	6.234082681	0.288859888

To examine the gain in the efficiency of the proposed estimator  $\bar{y}_{RD}^*$  over the estimator  $\bar{y}$ ,  $\bar{y}_R$ ,  $\bar{y}_p$ ,  $\bar{y}_R^*$  and  $\bar{y}_p^*$ , we obtained the percentage relative efficiency of different estimator of  $\bar{Y}$  with respect to the usual unbiased estimator  $\bar{y}$  in Table 2. The proposed estimator  $\bar{y}_{RD}^*$  performed better than the other estimators  $\bar{y}$ ,  $\bar{y}_R$ ,  $\bar{y}_p$ ,  $\bar{y}_R^*$  and  $\bar{y}_p^*$  and performed exactly the same way as regression estimator.

**Table 3.** The percentage relative efficiency of different estimator of  $\bar{Y}$  with respect to the usual unbiased estimator  $\bar{y}$ 

ESTIMATOR	PERCENTAGE RELATIVE EFFICIENCY
$\bar{y}$	100
$\bar{y}_R$	79.66158486
$\bar{y}_p$	74.40554745
$\bar{y}_R^*$	95.39056726
$\bar{y}_p^*$	91.65136111
$\bar{y}_{REG}^*$	107.3067159
$\bar{y}_{RD}^*$	107.3067159

## 5. Conclusions

We have proposed a class of regression estimator with cum-dual ratio estimator as intercept for estimating the mean of the study variable  $Y$  using auxiliary variable  $X$  as in equation (1) and obtained 'AOE' for the proposed estimator. Theoretically, we have demonstrated that proposed estimator is always more efficient than other under the effective ranges of  $\alpha$  and its optimum values.

Table 1. shows that the proposed estimator performed better than the other estimators as the width of the confidence interval of the proposed estimator is smallest than the other competing estimators. The percentage relative efficiency of different estimator of  $\bar{Y}$  with respects to the usual unbiased estimator  $\bar{y}$  in Table 2. shows that the proposed estimator  $\bar{y}_{RD}^*$  performed better than the other estimators  $\bar{y}$ ,  $\bar{y}_R$ ,  $\bar{y}_p$ ,  $\bar{y}_R^*$  and  $\bar{y}_p^*$  and performed exactly the same way as regression estimator. Hence, it is preferred to use the proposed class of estimator in practice.

## REFERENCES

- [1] B. K. Handique, "A Class of Regression-Cum-Ratio Estimators in Two-Phase Sampling for Utilizing Information From High Resolution Satellite Data," *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. 1-4, pp. 71-76, Jul. 2012.
- [2] S. Choudhury and B. K. Singh, "A Class of Product-cum-dual to Product Estimators of the Population Mean in Survey Sampling Using Auxiliary Information," *Asian J Math Stat*, vol. 6, 2012.
- [3] W. G. Cochran, "The estimation of the yields of cereal experiments by sampling for the ratio of grain to total produce," *J. Agric. Sci.*, vol. 30, no. 02, pp. 262-275, 1940.
- [4] S. Choudhury and B. K. Singh, "An efficient class of dual to product-cum-dual to ratio estimators of finite population mean in sample surveys," *Glob. J. Sci. Front. Res.*, vol. 12, no. 3-F, 2012.
- [5] D. S. Robson, "Applications of multivariate polykeys to the theory of unbiased ratio-type estimation," *J. Am. Stat. Assoc.*, vol. 52, no. 280, pp. 511-522, 1957.
- [6] M. N. Murthy, "Product method of estimation," *Sankhyā Indian J. Stat. Ser. A*, pp. 69-74, 1964.
- [7] B. Sharma and R. Tailor, "A new ratio-cum-dual to ratio estimator of finite population mean in simple random sampling," *Glob. J. Sci. Front. Res.*, vol. 10, no. 1, 2010.
- [8] H. P. Singh, "Estimation of finite population mean using known correlation coefficient between auxiliary characters," *Statistica*, vol. 65, no. 4, pp. 407-418, 2005.
- [9] R. Tailor and B. K. Sharma, "A modified ratio-cum-product estimator of finite population mean using known coefficient of variation and coefficient of kurtosis," *Stat. Transition-new Ser.*, vol. 10, p. 1, 2009.
- [10] L. N. Upadhyaya and H. P. Singh, "Use of transformed

- auxiliary variable in estimating the finite population mean,” *Biometrical J.*, vol. 41, no. 5, pp. 627–636, 1999.
- [11] E. J. Ekpenyong and E. I. Enang, “Efficient Exponential Ratio Estimator for Estimating the Population Mean in Simple Random Sampling.”
  - [12] R. Tailor, B. Sharma, and J.-M. Kim, “A generalized ratio-cum-product estimator of finite population mean in stratified random sampling,” *Commun Korea Stat Soc*, vol. 18, no. 1, pp. 111–118, 2011.
  - [13] T. Srivenkataramana, “A dual to ratio estimator in sample surveys,” *Biometrika*, vol. 67, no. 1, pp. 199–204, 1980.
  - [14] W. G. Cochran, “Sampling techniques,” 1977.
  - [15] P. V Sukhatme and B. V Sukhatme, “Sampling theory of surveys with applications,” 1970.
  - [16] P. V Sukhatme, B. V Sukhatme, S. Sukhatme, and C. Asok, “Sampling theory of surveys with applications,” 1984.
  - [17] M. N. Murthy, “Sampling theory and methods,” *Sampl. theory methods*, 1967.
  - [18] F. Yates, “Sampling methods for censuses and surveys,” *Griffin Books Stat.*, 1960.
  - [19] S. Singh, *Advanced Sampling Theory With Applications: How Michael Selected Amy*, vol. 2. Springer Science & Business Media, 2003.
  - [20] S. Bandyopadhyay, “Improved ratio and product estimators,” *Sankhya Ser. C*, vol. 42, no. 2, pp. 45–49, 1980.