An Improvement in Regression Estimator Through Exponential Estimator Using Two Auxiliary Variables

Sachin Malik*, Kanika and Atul

Department of Mathematics, SRM University Delhi NCR, Sonipat-131029, India
E-mail: sachin.malik@srmuniversity.ac.in; kanikasehrawat2814@gmail.com; ak4422882@gmail.com
**Corresponding Author*

Received 29 May 2023; Accepted 21 August 2023;
Publication 16 March 2024

Abstract

For the case of simple random sampling, we are introducing a new regression estimator for the population mean with the supporting values of two auxiliary variables. The results for the mean square error (MSE) of the new form of regression estimator is fined. The mean square error's results have also been checked through numerical illustration. It is observed that our introduced estimator is having less mean square error than the traditional ratio and regression estimator for two auxiliary variables.

Keywords: Mean square error (MSE), auxiliary variables, ratio estimator and regression estimator.

1 Introduction

In statistics, we are interested to know the behavior of the population based on a sample. Sample results cannot be accurate as the population results.

Journal of Reliability and Statistical Studies, Vol. 16, Issue 2 (2023), 329–336.
doi: 10.13052/jrss0974-8024.1627

Every time there is a difference between the results of sample and the results of population. Throughout this paper, we are trying to minimize this mean square error (MSE). For this we have proposed an estimator using auxiliary information for two variables. Our target is to find mean square error of some estimators which is already given in literature, is always more than our introduced estimator.

The auxiliary variable's information is effectively used to increase the efficiency of the estimators of the population mean. In the Cochran [1] and Murthy [2], ratio estimators, product estimators, and regression estimators are used in several conditions.

Abu-Dayyeh et al. [3] suggested an estimator when population means of both the auxiliary variables X and Z are known

$$\mathrm{t}_1 = \overline{\mathrm{y}}\left(\frac{\overline{\mathrm{X}}}{\overline{\mathrm{x}}}\right)^{\alpha_1}\left(\frac{\overline{\mathrm{Z}}}{\overline{\mathrm{z}}}\right)^{\alpha_2} \tag{1}$$

Where α_1 and α_2 are constants.

Muhammad Noor-ul-Amin et al. [4] proposed estimator as

The regression estimator for two auxiliary variables is define as

$$\mathrm{t}_2 = \overline{\mathrm{y}} + \mathrm{b}_1(\overline{\mathrm{X}} - \overline{\mathrm{x}}) + \mathrm{b}_2(\overline{\mathrm{Z}} - \overline{\mathrm{z}}) \tag{2}$$

Kadilar and Cingi [5] is defined a new estimator for auxiliary variables as

$$\mathrm{t}_3 = \overline{\mathrm{y}}\left(\frac{\overline{\mathrm{X}}}{\overline{\mathrm{x}}}\right)^{\alpha_1}\left(\frac{\overline{\mathrm{Z}}}{\overline{\mathrm{z}}}\right)^{\alpha_2} + \mathrm{b}_1(\overline{\mathrm{X}} - \overline{\mathrm{x}}) + \mathrm{b}_2(\overline{\mathrm{Z}} - \overline{\mathrm{z}}) \tag{3}$$

The mean squared error of t_1, t_2, and t_3 after taking first order of approximation is hereby

$$\begin{aligned}\mathrm{MSE}(\mathrm{t}_1) = \overline{\mathrm{Y}}^2\left(\frac{1-\mathrm{f}}{\mathrm{n}}\right)[\mathrm{C}_\mathrm{x}^2 + \mathrm{C}_\mathrm{y}^2 + \mathrm{C}_\mathrm{z}^2 - 2\rho_{\mathrm{xy}}\mathrm{C}_\mathrm{x}\mathrm{C}_\mathrm{y} \\ - 2\rho_{\mathrm{yz}}\mathrm{C}_\mathrm{y}\mathrm{C}_\mathrm{z} + 2\rho_{\mathrm{xz}}\mathrm{C}_\mathrm{x}\mathrm{C}_\mathrm{z}]\end{aligned} \tag{4}$$

$$\begin{aligned}\mathrm{MSE}(\mathrm{t}_2) = \left(\frac{1-\mathrm{f}}{\mathrm{n}}\right)[\overline{\mathrm{Y}}^2\mathrm{C}_\mathrm{y}^2 + \mathrm{B}_1^2\overline{\mathrm{X}}^2\mathrm{C}_\mathrm{x}^2 + \mathrm{B}_2^2\overline{\mathrm{Z}}^2\mathrm{C}_\mathrm{z}^2 + 2\mathrm{B}_1\mathrm{B}_2\overline{\mathrm{X}}\overline{\mathrm{Z}}\rho_{\mathrm{xz}}\mathrm{C}_\mathrm{x}\mathrm{C}_\mathrm{z} \\ - 2\overline{\mathrm{X}}\overline{\mathrm{Y}}B_1\rho_{\mathrm{xy}}\mathrm{C}_\mathrm{x}\mathrm{C}_\mathrm{y} - 2\overline{\mathrm{Y}}\overline{\mathrm{Z}}\mathrm{B}_2\rho_{\mathrm{yz}}\mathrm{C}_\mathrm{y}\mathrm{C}_\mathrm{z}]\end{aligned} \tag{5}$$

$$
\begin{aligned}
\mathrm{MSE}(t_3) = \left(\frac{1-f}{n}\right) & [\alpha_1^2(\overline{Y}^2 C_x^2) + \alpha_2^2(\overline{Y}^2 C_z^2) + 2\alpha_1\alpha_2(\overline{Y}^2 \rho_{xz} C_x C_z) \\
& - 2\alpha_1(\overline{Y}^2 \rho_{xy} C_x C_y - \overline{Y} B_1 \overline{X} C_x^2 - \overline{Y} B_2 \overline{Z} \rho_{xz} C_x C_z) \\
& - 2\alpha_2(\overline{Y}^2 \rho_{yz} C_y C_z - \overline{Y} B_2 \overline{Z} C_z^2 - \overline{Y} B_1 \overline{X} \rho_{xz} C_x C_z) \\
& + (B_1^2 \overline{X}^2 C_x^2 + B_2^2 \overline{Z}^2 C_z^2 + 2B_1 B_2 \overline{X}\overline{Z} \rho_{xz} C_x C_z) \\
& + (\overline{Y}^2 C_y^2 - 2\overline{Y} B_1 \overline{X} \rho_{xy} C_x C_y - 2\overline{Y} B_2 \overline{Z} \rho_{yz} C_y C_z)] \qquad (6)
\end{aligned}
$$

2 The Proposed Estimator

Using (2) and (3), we proposed a new estimator defined as

$$
\begin{aligned}
t_{akm} = \overline{y} \exp\left(\frac{\overline{X} - \overline{x}}{\overline{X} + (\alpha_1 - 1)\overline{x}}\right) \exp\left(\frac{\overline{Z} - \overline{z}}{\overline{Z} + (\alpha_2 - 1)\overline{x}}\right) \\
+ b_1(\overline{X} - \overline{x}) + b_2(\overline{Z} - \overline{z}) \qquad (7)
\end{aligned}
$$

To find the Mean square error (MSE) of t_{akm} up to the first order of approximation, we are using the following notations in literature [4, 6] and [7]

$$
\overline{y} = \overline{Y}(1 + e_0), \overline{x} = \overline{X}(1 + e_1), \overline{z} = \overline{Z}(1 + e_2)
$$

$$
e_0 = \frac{\overline{y}}{\overline{Y}} - 1, e_1 = \frac{\overline{x}}{\overline{X}} - 1, e_2 = \frac{\overline{z}}{\overline{Z}} - 1
$$

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

$$
E(e_0^2) = \left(\frac{1-f}{n}\right) C_y^2, E(e_1^2) = \left(\frac{1-f}{n}\right) C_x^2, E(e_2^2) = \left(\frac{1-f}{n}\right) C_z^2,
$$

$$
E(e_0 e_1) = \left(\frac{1-f}{n}\right) \rho_{xy} C_x C_y, E(e_0 e_2) = \left(\frac{1-f}{n}\right) \rho_{yz} C_y C_z,
$$

$$
E(e_1 e_2) = \left(\frac{1-f}{n}\right) \rho_{xz} C_x C_z
$$

$b_1 = \frac{s_{yx}}{s_x^2}$ and $b_2 = \frac{s_{yz}}{s_z^2}$, where s_{yx} and s_{yz} are the sample covariances between y and x and between z respectively.

Using above notations, we get

$$\begin{aligned}
t_{akm} = \overline{Y}(1+e_0)\exp\left[\frac{\overline{X}-\overline{X}-\overline{X}e_1}{\overline{X}(1+(\alpha_1-1)(1+e_1))}\right] \\
\times\left[\frac{\overline{Z}-\overline{Z}-Ze_2}{\overline{Z}(1+(\alpha_2-1)(1+e_2))}\right] \\
+ b_1(\overline{X}-\overline{X}-\overline{X}e_1) + b_1(\overline{X}-\overline{X}-\overline{Z}e_2) \\
= \overline{Y}(1+e_0)\exp\left[\frac{-e_1}{1+(\alpha_1-1)(1+e_1)}\right]\left[\frac{-e_2}{1+(\alpha_2-1)(1+e_2)}\right] \\
-(b_1\overline{X}e_1 + b_2\overline{Z}e_2) \qquad (8)
\end{aligned}$$

After solving Equation (8), we get

$$\begin{aligned}
t_{akm} = \overline{Y}\bigg(1 - \frac{e_1}{\alpha_1} + k_1e_1^2 + \frac{e_1^2}{2\alpha_1^2} - \frac{e_2}{\alpha_2} + \frac{e_1e_2}{\alpha_1\alpha_2} + k_2e_2^2 + \frac{e_2^2}{2\alpha_2^2} \\
+ e_0 - \frac{e_0e_1}{\alpha_1} - \frac{e_0e_2}{\alpha_2}\bigg) - (b_1\overline{X}e_1 + b_2\overline{Z}e_2) \qquad (9)
\end{aligned}$$

Let's assume $k_1 = \frac{\alpha_1-1}{\alpha_1^2}, k_2 = \frac{\alpha_2-1}{\alpha_2^2}$.
After avoiding the higher power of e's in (9), we have

$$(t_{akm} - \overline{Y}) = \left[\overline{Y}(e_0 - \frac{e_1}{\alpha_1} - \frac{e_2}{\alpha_2}) - (b_1\overline{X}e_1 + b_2\overline{Z}e_2)\right] \qquad (10)$$

To find the MSE taking square of both sides of (10), we get

$$(t_{akm} - \overline{Y})^2 = \left[\overline{Y}(e_0 - \frac{e_1}{\alpha_1} - \frac{e_2}{\alpha_2}) - (b_1\overline{X}e_1 + b_2\overline{Z}e_2)\right]^2 \qquad (11)$$

Taking expectations on both the sides of (11), we have

$$\begin{aligned}
E(t_{akm} - \overline{Y})^2 = E\bigg[\frac{1}{\alpha_1^2}(\overline{Y}^2e_1^2) + \frac{1}{\alpha_2^2}(\overline{Y}^2e_2^2) + \frac{2}{\alpha_1\alpha_2}(\overline{Y}^2e_1e_2) \\
- \frac{2}{\alpha_1}(\overline{Y}^2e_0e_1 - b_1\overline{XY}e_1^2 - b_2\overline{YZ}e_1e_2) \\
- \frac{2}{\alpha_2}(\overline{Y}^2e_0e_2 - b_1\overline{XY}e_1e_2 - b_2\overline{YZ}e_2^2)
\end{aligned}$$

$$+\left(\mathrm{b}_1^2\overline{\mathrm{X}}^2\mathrm{e}_1^2+\mathrm{b}_2^2\overline{\mathrm{Z}}^2\mathrm{e}_2^2+2\mathrm{b}_1\mathrm{b}_2\overline{\mathrm{X}}\overline{\mathrm{Z}}\mathrm{e}_1\mathrm{e}_2+\overline{\mathrm{Y}}^2\mathrm{e}_0^2\right.$$

$$\left.-2\mathrm{b}_1\overline{\mathrm{X}}\overline{\mathrm{Y}}\mathrm{e}_0\mathrm{e}_1-2\mathrm{b}_2\overline{\mathrm{Y}}\overline{\mathrm{Z}}\mathrm{e}_0\mathrm{e}_2)\right]$$

$$\mathrm{MSE}(\mathrm{t}_{\mathrm{akm}})=\frac{1}{\alpha_1^2}\mathrm{A}_1+\frac{1}{\alpha_2^2}\mathrm{A}_2+\frac{2}{\alpha_1\alpha_2}\mathrm{A}_3-\frac{2}{\alpha_1}\mathrm{A}_4-\frac{2}{\alpha_2}\mathrm{A}_5+\mathrm{A}_6 \tag{12}$$

where,

$$\mathrm{A}_1=\mathrm{E}(\overline{\mathrm{Y}}^2\mathrm{e}_1^2)=\overline{\mathrm{Y}}^2\mathrm{f}\mathrm{C}_\mathrm{x}^2$$

$$\mathrm{A}_2=\mathrm{E}(\overline{\mathrm{Y}}^2\mathrm{e}_2^2)=\overline{\mathrm{Y}}^2\mathrm{f}\mathrm{C}_\mathrm{z}^2$$

$$\mathrm{A}_3=\mathrm{E}(\overline{\mathrm{Y}}^2\mathrm{e}_1\mathrm{e}_2)=\overline{\mathrm{Y}}^2\mathrm{f}\rho_{\mathrm{xz}}\mathrm{C}_\mathrm{x}\mathrm{C}_\mathrm{z}$$

$$\begin{aligned}\mathrm{A}_4&=\mathrm{E}(\overline{\mathrm{Y}}^2\mathrm{e}_0\mathrm{e}_1-\overline{\mathrm{Y}}\mathrm{b}_1\overline{\mathrm{X}}\mathrm{e}_1^2-\overline{\mathrm{Y}}\mathrm{b}_2\overline{\mathrm{Z}}\mathrm{e}_1\mathrm{e}_2)\\&=\overline{\mathrm{Y}}^2\mathrm{f}\rho_{\mathrm{xy}}\mathrm{C}_\mathrm{x}\mathrm{C}_\mathrm{y}-\overline{\mathrm{Y}}\mathrm{b}_1\overline{\mathrm{X}}\mathrm{f}\mathrm{C}_\mathrm{x}^2-\overline{\mathrm{Y}}\mathrm{b}_2\mathrm{f}\overline{\mathrm{Z}}\rho_{\mathrm{xz}}\mathrm{C}_\mathrm{x}\mathrm{C}_\mathrm{z}\end{aligned}$$

$$\begin{aligned}\mathrm{A}_5&=\mathrm{E}(\overline{\mathrm{Y}}^2\mathrm{e}_0\mathrm{e}_2-\overline{\mathrm{Y}}\mathrm{b}_2\overline{\mathrm{Z}}\mathrm{e}_2^2-\overline{\mathrm{Y}}\mathrm{b}_1\overline{\mathrm{X}}\mathrm{e}_1\mathrm{e}_2)\\&=\overline{\mathrm{Y}}^2\mathrm{f}\rho_{\mathrm{yz}}\mathrm{C}_\mathrm{y}\mathrm{C}_\mathrm{z}-\overline{\mathrm{Y}}\mathrm{b}_2\overline{\mathrm{Z}}\mathrm{f}\mathrm{C}_\mathrm{z}^2-\overline{\mathrm{Y}}\mathrm{b}_1\overline{\mathrm{X}}\mathrm{f}\rho_{\mathrm{xz}}\mathrm{C}_\mathrm{x}\mathrm{C}_\mathrm{z}\end{aligned}$$

$$\begin{aligned}\mathrm{A}_6&=\mathrm{E}(\mathrm{b}_1^2\overline{\mathrm{X}}^2\mathrm{e}_1^2+\mathrm{b}_2^2\overline{\mathrm{Z}}^2\mathrm{e}_2^2+2\mathrm{b}_1\mathrm{b}_2\overline{\mathrm{X}}\overline{\mathrm{Z}}\mathrm{e}_1\mathrm{e}_2\\&\quad+\overline{\mathrm{Y}}^2\mathrm{e}_0^2-2\overline{\mathrm{Y}}\mathrm{b}_1\overline{\mathrm{X}}\mathrm{e}_0\mathrm{e}_1-2\overline{\mathrm{Y}}\mathrm{b}_2\overline{\mathrm{Z}}\mathrm{e}_0\mathrm{e}_2)\\&=\mathrm{b}_1^2\overline{\mathrm{X}}^2\mathrm{f}\mathrm{C}_\mathrm{x}^2+\mathrm{b}_2^2\overline{\mathrm{Z}}^2\mathrm{f}\mathrm{C}_\mathrm{z}^2+2\mathrm{b}_1\mathrm{b}_2\overline{\mathrm{X}}\overline{\mathrm{Z}}\mathrm{f}\rho_{\mathrm{xz}}\mathrm{C}_\mathrm{x}\mathrm{C}_\mathrm{z}\\&\quad+\overline{\mathrm{Y}}^2\mathrm{f}\mathrm{C}_\mathrm{y}^2-2\overline{\mathrm{Y}}\mathrm{b}_1\overline{\mathrm{X}}\mathrm{f}\rho_{\mathrm{xy}}\mathrm{C}_\mathrm{x}\mathrm{C}_\mathrm{y}-2\overline{\mathrm{Y}}\mathrm{b}_2\overline{\mathrm{Z}}\mathrm{f}\rho_{\mathrm{yz}}\mathrm{C}_\mathrm{y}\mathrm{C}_\mathrm{z}\end{aligned}$$

Differentiating (12) partially with respect to α_1 and α_2, we get

$$\frac{\mathrm{A}_1}{\alpha_1}+\frac{\mathrm{A}_3}{\alpha_2}=\mathrm{A}_4 \tag{13}$$

$$\frac{\mathrm{A}_3}{\alpha_1}+\frac{\mathrm{A}_2}{\alpha_2}=\mathrm{A}_5 \tag{14}$$

On solving Equations (13) and (14), we get

$$\alpha_1=\frac{\mathrm{A}_1\mathrm{A}_2-\mathrm{A}_3^2}{\mathrm{A}_2\mathrm{A}_4-\mathrm{A}_3\mathrm{A}_5}\quad\text{and}\quad\alpha_2=\frac{\mathrm{A}_3^2-\mathrm{A}_1\mathrm{A}_2}{\mathrm{A}_3\mathrm{A}_4-\mathrm{A}_1\mathrm{A}_5}$$

Table 1 Data statistics

Population I	Population II
$N = 34$	$N = 25$
$n = 20$	$n = 10$
$\overline{X} = 208.88$	$\overline{X} = 14.3$
$\overline{Y} = 856.41$	$\overline{Y} = 75.28$
$\overline{Z} = 199.44$	$\overline{Z} = 6.82$
$S_x = 150.22$	$S_x = 3.17$
$S_y = 733.14$	$S_y = 17.27$
$S_z = 150.22$	$S_z = 1.53$
$\rho_{yx} = 0.45$	$\rho_{yx} = 0.99$
$\rho_{yz} = 0.45$	$\rho_{yz} = 0.89$
$\rho_{xz} = 0.98$	$\rho_{xz} = 0.92$
$B_1 = 2.19$	$B_1 = 2.60$

Table 2 Results MSE values of different estimators

Estimators	MSE (Data Set I)	MSE (Data Set II)
t_1	26344.84	17.44
t_2	10976.42	15.19
t_3	8967.45	2.35
t_{akm}	**8802.54**	**0.30**

3 Numerical Illustration

The performance of the proposed improve regression estimator through exponential estimator are assessed with two different data sets. From Singh and Chaudhary [8] data set I is taken and data set II is taken from the Cingi and Kadilar [9]. In first data set area under wheat (1974) is our study variable, area under wheat (1971) is first auxiliary variable and area under wheat (1973) is the second auxiliary variable. For the second data set the population mean of the height of the fish is our study variable, the population mean of the length of the head is first auxiliary variable and the population mean of the length of the fin is the second auxiliary variable.

4 Conclusion

In the present paper, we have introduced an improvement in regression estimator through exponential estimator for finding the study variable's population mean using available information of two auxiliary variables. Taking results given in Table 2, we can have an idea that the introduced estimator $\mathrm{t}_{\mathrm{akm}}$ is performing better than other estimators in literature.

References

[1] Cochran WG. Sampling Techniques. New York: John Wiley & Sons; 1977.
[2] Murthy MN. Sampling Theory and Methods. Calcutta: Statistical Publishing Society; 1967.
[3] Akingbade T., Okafor F. (2019): A class of Ratio-Type Estimator Using Two Auxiliary Variables for Estimating the Population Mean with Sum Known Population Parameters, Pakistan Journal of statistics and operation research, Vol. XV, No. 2, 329–340.
[4] M. Amin, M. Hanif and C. Kadilar (2014): Improved Exponential Type Estimators using the Information of Two Auxiliary Variables, Middle-East Journal of Scientific Research 19(12): 1711–1715.
[5] C. Kadilar, H. Cingi (2005): A new estimator using two auxiliary variables, Applied Mathematics and Computation 162, 901–908.
[6] Subramani J. (2018): Two parameter modified ratio estimators with two auxiliary variables for the estimation of finite population mean, Biometrics & Biostatistics International Journal, Volume 7 Issue 6 559–568.
[7] Manoj K. Chaudhary and Amit Kumar (2020): An Improvement in Estimation of Population Mean using Two Auxiliary Variables and Two-Phase Sampling Scheme under Non-Response, Journal of Reliability and Statistical Studies, Vol. 13, 349–362.
[8] Singh D, Chaudhary FS. Theory and analysis of sample survey designs. New Age International Publisher. 1986.
[9] Kadilar C, Cingi H. Advances in Sampling Theory – Ratio Method of Estimation. Bentham Science Publishers. 2009.