Improved Exponential Type Estimators using the Information of Two Auxiliary Variables

Muhammad Noor-ul-Amin, Muhammad Hanif and Cem Kadilar

1COMSATS Institute of Information and Technology, Lahore, Pakistan
2National College of Business Administration and Economics, Lahore, Pakistan
3Hacettepe University, Department of Statistics, Beytepe, Ankara, Turkey

Abstract: In this paper, ratio-cum-ratio and product-cum-product exponential type estimators have been proposed for estimating the finite population mean of the study variable using the information of two auxiliary variables. The expressions for mean square errors and biases of the proposed estimators have been derived. The generalized form of the proposed estimators has also been developed. It is shown that the proposed estimators are more efficient as compared to the sample mean estimator, classical ratio and product estimators, estimators of Singh [10] and estimators of Bahl and Tuteja [1] under specified conditions. Empirical study has also been carried out to demonstrate the performance of proposed estimators.

AMS (1991) subject classification: 62D05

Key words: Ratio-cum-ratio, product-cum-product, exponential type estimators, population mean, auxiliary variables, mean square errors, biases

INTRODUCTION

The use of auxiliary information is familiar to improve the precision of estimator for the finite population mean. Cochran [2] introduced ratio method of estimation for positive correlation between the study and auxiliary variables whereas; Robson [7] and Murthy [8] used product method of estimation for negative correlation between study variable and auxiliary variable. (The aim of this paper is to improve the efficiencies of ratio and product type estimators in the simple random sampling. Therefore, the studies about linear regression type estimators, such as Kadilar and Cingi [5, 6] and Singh and Espejo [9] etc., are beyond the scope of this paper.

Let us consider the finite population S = \{s_1, \ldots, s_N\} of size N. Let \( \bar{Y} \) is the population mean of the study variable Y, where X and Z are the population means of auxiliary variables X and Z respectively, where the variable X is positively and Z is negatively correlated with the study variable, Y. The sample is drawn by the simple random sampling without replacement (SRSWOR) of size n (n < N) from the population and we assume that \( \bar{x} \) be the sample mean of the variable X, where \( \bar{y} \) and \( \bar{z} \) are the sample means of variables, Y and Z, respectively. We also have following assumptions:

\[
\begin{align*}
\varepsilon_Y &= \frac{\bar{Y} - \bar{y}}{\bar{Y}}, \quad \varepsilon_X = \frac{\bar{X} - \bar{x}}{\bar{X}} \\
E(\varepsilon_Y) &= E(\varepsilon_X) = E(\varepsilon_Z) = 0, E(\varepsilon_X^2) = \theta C_X^2, E(\varepsilon_Y^2) = \theta C_Y^2, E(\varepsilon_Z^2) = \theta C_Z^2, \quad (1.1)
\end{align*}
\]

\[
\theta = \frac{1}{n} - \frac{1}{N}, \quad C_Y = \frac{S_Y}{\bar{Y}}, \quad \rho_{xy} = \frac{S_{xy}}{S_X S_Y}, \quad H_1 = \rho_{y} \frac{C_Y}{C_X},
\]

These notations are similar for the other auxiliary variables.

Cochran [2] and Robson [7] suggested the usual ratio and product estimators, respectively, for estimating the population mean as:

\[
\bar{Y}_R = \bar{Y} \left( \frac{\bar{X}}{\bar{X}} \right) \quad (1.2)
\]

and

\[
\bar{Y}_P = \bar{Y} \left( \frac{\bar{Z}}{\bar{Z}} \right) \quad (1.3)
\]

The mean square equations (MSE) of the estimators of (1.2) and (1.3) are

\[
\text{MSE}(\bar{Y}_R) = \bar{Y}^2 \left[ C_X^2 + C_Y^2 \left( 1 - 2H_1 \right) \right] \quad (1.4)
\]

and

\[
\text{MSE}(\bar{Y}_P) = \bar{Y}^2 \left[ C_X^2 + C_Z^2 \left( 1 + 2H_1 \right) \right] \quad (1.5)
\]

respectively.
Bahl and Tuteja [1] suggested exponential ratio-type and product-type estimators, respectively, as:

\[ t_1 = \bar{y} \exp \left( \frac{X - x}{x + x} \right) \]  

(1.6)

and

\[ t_2 = \bar{y} \exp \left( \frac{Z - Z}{Z + Z} \right) \]  

(1.7)

The mean square equations (MSE) of the estimators in (1.6) and (1.7) are as

\[ \text{MSE}(t_1) \equiv \bar{y}^2 \left[ C_x^2 + C_x^2 \left( 1 - 4H_v \right) \right] \]  

(1.8)

and

\[ \text{MSE}(t_2) \equiv \bar{y}^2 \left[ C_y^2 + C_y^2 \left( 1 + 4H_v \right) \right] \]  

(1.9)

respectively.

Singh [10] suggested the estimators using two auxiliary variables as:

\[ t_3 = \bar{y} \exp \left( \frac{XZ - y}{xZ - y} \right) \]  

(1.10)

and

\[ t_4 = \bar{y} \exp \left( \frac{xZ - y}{xZ - y} \right) \]  

(1.11)

The mean square equations of the estimators in (1.10) and (1.11) are

\[ \text{MSE}(t_3) \equiv \bar{y}^2 \left[ C_x^2 + C_x^2 \left( 1 - 2H_v \right) + C_y^2 \left( 1 - 2H_v + 2H_w \right) \right] \]  

(1.12)

and

\[ \text{MSE}(t_4) \equiv \bar{y}^2 \left[ C_y^2 + C_y^2 \left( 1 + 2H_v \right) + C_x^2 \left( 1 + 2H_v + 2H_w \right) \right] \]  

(1.13)

respectively.

**PROPOSED ESTIMATORS**

Combining the estimators in (1.6) and (1.7), we propose the ratio-cum-ratio and product-cum-product exponential type estimators using two auxiliary variables as:

\[ t_5 = \bar{y} \exp \left( \frac{X - X}{x + x} - \frac{Z - Z}{Z + Z} \right) \]  

(2.2)

respectively. Using the notations, given in (1.1), we can write the estimator in (2.1) as

\[ t_5 = \bar{y} \left[ (1 + e_\gamma) \exp \left( -\frac{e_\gamma}{2} \left( 1 + \frac{e_\gamma}{2} \right) - \frac{e_\gamma}{2} \left( 1 + \frac{e_\gamma}{2} \right) \right) \right] \]  

(2.3)

Expanding the right hand side of (2.3) and neglecting the terms with power two or greater, it is possible to re-write \( t_5 \) as

\[ \bar{y} - \bar{y} \equiv \bar{y} \left[ e_\gamma - \frac{e_\gamma}{2} \right] \]  

(2.4)

Squaring both sides of (2.4) and taking the expectation, we get

\[ \text{MSE}(t_5) \equiv \bar{y}^2 \left[ e_\gamma - \frac{e_\gamma}{2} \right]^2 \]  

(2.5)

In order to derive the bias of \( t_5 \), we again use (2.3) and re-write \( t_5 \) as

\[ t_5 = \bar{y} \left[ (1 + e_\gamma) \exp \left( -\frac{e_\gamma}{4} - \frac{e_\gamma}{2} - \frac{e_\gamma}{2} \right) \right] \]  

(2.7)

or

\[ t_5 - \bar{y} = \bar{y} \left[ e_\gamma + \frac{e_\gamma}{2} + \frac{3e_\gamma}{8} + \frac{3e_\gamma}{8} + \frac{e_\gamma}{2} + \frac{e_\gamma}{2} \right] \]  

(2.8)

Applying the expectation of (2.8) with some simplifications, we get

\[ \text{Bias}(t_5) \equiv \frac{\bar{y}^2}{8} \left[ 3(C_\gamma + C_\gamma) + 2C_H + 4C_\gamma (H_v + H_w) \right] \]  

(2.9)

Likewise, the mean square error and bias of \( t_6 \) may be obtained and are

\[ \text{MSE}(t_6) \equiv \bar{y}^2 \left[ C_x^2 + C_x^2 \left( 1 + 4H_v \right) + C_y^2 \left( 1 + 4H_v + 2H_w \right) \right] \]  

(2.10)
EFFICIENCY COMPARISONS

In the section, the proposed estimators have been compared with the existing estimators for population mean. Reddy [12] have proved that in repetitive surveys \( H_{yx} \) is stable.

The mean square errors comparison of proposed estimators have made in terms of \( H_{yx} \) and \( H_{yz} \). Note that the value of parameters \( (H_{yx} \) and \( H_{yz} \) falls between \( (0, \infty) \) and \( (-\infty,0) \) for positive and negative correlation, respectively.

It is known that the variance of the sample mean estimator, \( \bar{y} \), under SRSWOR is

\[
V(\bar{y}) = \left( \frac{1}{n} - \frac{1}{N} \right) s_y^2
\]

Then,

\[
\text{MSE}(t_5) < \text{MSE}(\bar{y}) \text{ if } H_{yx} > \left( \frac{1+2H_{ux}}{4} \right) \text{ and } H_{yz} > \frac{1}{4}
\]

\[
\text{MSE}(t_5) < \text{MSE}(\bar{y}_y) \text{ if } H_{yx} > \left( \frac{2H_{ux}+1}{4} \right) \text{ and } H_{yz} < \frac{3}{4}
\]

\[
\text{MSE}(t_5) < \text{MSE}(t_1) \text{ if } H_{yx} > \left( \frac{1+2H_{ux}}{4} \right)
\]

\[
\text{MSE}(t_6) < \text{MSE}(t_5) \text{ if } H_{yx} < \left( \frac{3(2H_{ux}+1)}{4} \right) \text{ and } H_{yz} < \frac{3}{4}
\]

\[
\text{MSE}(t_6) < \text{MSE}(\bar{y}) \text{ if } H_{yx} < \left( \frac{1+2H_{ux}}{4} \right) \text{ and } H_{yz} < \frac{1}{4}
\]

\[
\text{MSE}(t_6) < \text{MSE}(t_2) \text{ if } H_{yx} > \left( \frac{2H_{ux}-3}{4} \right) \text{ and } H_{yz} < \frac{1}{4}
\]

\[
\text{MSE}(t_6) < \text{MSE}(t_4) \text{ if } H_{yx} < \left( \frac{2H_{ux}-3}{4} \right) \text{ and } H_{yz} < \frac{1}{4}
\]

When the conditions (3.2)-(3.5) are satisfied, it is clear that the proposed estimator, \( t_5 \), is more efficient than the sample mean estimator, classical ratio estimator, Singh ratio estimator and Bahl-Tuteja ratio estimator respectively. Similarly, when the conditions (3.6)-(3.9) are satisfied, the proposed estimator, \( t_6 \), is more efficient than the sample mean estimator, classical product estimator, Singh product estimator and Bahl-Tuteja product estimator respectively.

GENERALIZED FORM OF PROPOSED ESTIMATORS

The generalized form of ratio-cum-ratio exponential type estimator is given by

\[
t_{1G} = \bar{y} \exp \left[ \frac{X - \bar{x}}{X + (a-1)\bar{x}} - \frac{Z - \bar{z}}{Z + (b-1)\bar{z}} \right]
\]

and the generalized product-cum-product exponential type estimator is given by

\[
t_{2G} = \bar{y} \exp \left[ \frac{X - \bar{x}}{X + (c-1)\bar{x}} - \frac{Z - \bar{z}}{Z + (d-1)\bar{z}} \right]
\]

where, \( a, b, c \) and \( d \) are the real positive constants

(i) For \( a=b=1 \) in (4.1), the estimator reduces to

\[
t_{1G} = \bar{y} \exp \left[ \frac{X - \bar{x}}{X} - \frac{Z - \bar{z}}{Z} \right]
\]

(ii) For \( c=d=1 \) in (4.2), the estimator reduces to

\[
t_{2G} = \bar{y} \exp \left[ \frac{X - \bar{x}}{X} - \frac{Z - \bar{z}}{Z} \right]
\]

(iii) For \( a=b=2 \) in (4.1), the estimator reduces to

\[
t_{1G} = \bar{y} \exp \left[ \frac{X - \bar{x}}{X} - \frac{Z - \bar{z}}{Z} \right]
\]

(iii) For \( c=d=2 \) in (4.1), the estimator reduces to

\[
t_{2G} = \bar{y} \exp \left[ \frac{X - \bar{x}}{X} - \frac{Z - \bar{z}}{Z} \right] = t_5
\]

\[
t_{1G} = \bar{y} \exp \left[ \frac{X - \bar{x}}{X} - \frac{Z - \bar{z}}{Z} \right]
\]
By similar derivation in Section 2, the MSE of $t_{5G}$ is obtained as

$$\text{MSE}(t_{5G}) \equiv 0 \Theta^2 \begin{bmatrix} C_y^2 + \frac{C_y^2}{a^2} (1 - 2aH_y) \\ + \frac{C_y^2}{b^2} (1 - 2bH_y + \frac{2b}{a} H_y) \end{bmatrix}$$

(4.7)

The optimal values of $a$ and $b$ are found by

$$a^* = \frac{1}{H_y - H_{x'}}, \quad b^* = \frac{1 - \rho_{xy}^2}{H_y - H_{x'} H_{y'}}$$

(4.8)

when the optimal values in (4.8) are replaced with $a$ and $b$ in (4.7), respectively, the minimum MSE of $t_{5G}$, denoted by $t_{5Gmin}$, is obtained by

$$\text{MSE}(t_{5Gmin}) \equiv 0 \Theta^2 \begin{bmatrix} C_y^2 - C_y^2 (H_{y'}^2 - (KH_n)^2) \\ + C_y^2 K \left( 2H_{y'} (H_{y} - KH_n) - 2H_{y'} + K \right) \end{bmatrix}$$

(4.9)

Note that, $\text{MSE}(t_{5Gmin}) < \text{MSE}(t_5)$ if

$$K^2 < \left( \frac{1 - 4H_{y'}}{4H_{y'}^2} \right)$$

and

$$H_{y'} > \frac{2(4KH_n - 1)H_{y'} + 4K^2 (1 - 2\rho_{xy}^2) - 1}{4H_{y'}(2K - 1)}$$

(4.10)

when the condition (4.10) is satisfied, the generalized form of the proposed ratio-cum-ratio estimator is more efficient than the proposed ratio-cum-ratio estimator. Similarly,

$$\text{MSE}(t_{5G}) \equiv 0 \Theta^2 \begin{bmatrix} C_y^2 + \frac{C_y^2}{c^2} (1 + 2cH_y) \\ + \frac{C_y^2}{d} (1 + 2dH_y + \frac{2d}{c} H_y) \end{bmatrix}$$

(4.11)

The optimal values of $c$ and $d$ are

$$c^* = \frac{1}{H_y K - H_{y'}}$$

$$d^* = \frac{1 - \rho_{xy}^2}{H_y H_{xy} - H_{y'}}$$

(4.12)

when the optimal values in (4.12) are replaced with $c$ and $d$ in (4.11), respectively, the minimum MSE of $t_{6G}$, denoted by $t_{6Gmin}$, is obtained by

$$\text{MSE}(t_{6Gmin}) \equiv 0 \Theta^2 \begin{bmatrix} C_y^2 - C_y^2 (H_{y'}^2 - (KH_n)^2) \\ + C_y^2 K \left( 2H_{y'} (H_{y} - KH_n) + 2H_{y'} - K \right) \end{bmatrix}$$

(4.13)

Note that, $\text{MSE}(t_{6Gmin}) < \text{MSE}(t_6)$ if

$$K^2 < \left( \frac{1 + 4H_{y'} (1 + \rho_{xy})}{4H_{y'}^2} \right)$$

and

$$H_{y'} > \frac{2(1 - 4KH_n)H_{y'} - 4K^2 (1 - 2\rho_{xy}^2) - 1}{4(1 - 2K)}$$

(4.14)

when the condition (4.14) is satisfied, the generalized form of the proposed product-cum-product estimator is more efficient than the proposed ratio-cum-ratio estimator.

**EMPIRICAL STUDY**

In order to examine the performance of proposed estimators, we take the real population data. The percent relative efficiencies of $\bar{y}_b, \bar{y}_r, t_1, t_2, t_3, t_4, t_5, t_6, t_{5G}$ and $t_{6G}$ based on the sample mean estimator, $\bar{y}$, are presented in Table 1. The description of populations in this table is as follows:

**Population I:** Source: Cochran [3]

- **Y:** Number of “placebo” children.
- **X:** Number of paralytic polio cases in the placebo group.
- **Z:** Number of paralytic polio cases in the “not inoculated” group.

The required parameters of the population are:

- $\bar{Y} = 4.92$, $C_y^2 = 1.0248$, $\rho_{xy} = 0.7326$, $N = 34$
- $\bar{X} = 2.59$, $C_x^2 = 1.5175$, $\rho_{xy} = 0.6430$, $n = 15$
- $\bar{Z} = 2.91$, $C_z^2 = 1.1492$, $\rho_{xy} = 0.6837$. 

1714
Table 1: Percent relative efficiencies of different estimators for the population mean with respect to the sample mean

<table>
<thead>
<tr>
<th></th>
<th>Populations</th>
<th>I</th>
<th>II</th>
<th>III</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimators</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{Y}$</td>
<td>$\bar{y}$</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>$\gamma_r$</td>
<td>$\gamma_r$</td>
<td>143.30</td>
<td>155.53</td>
<td>*</td>
</tr>
<tr>
<td>$\gamma_y$</td>
<td>$\gamma_y$</td>
<td>*</td>
<td>*</td>
<td>92.85</td>
</tr>
<tr>
<td>$t_1$</td>
<td>$t_1$</td>
<td>87.16</td>
<td>86.77</td>
<td>*</td>
</tr>
<tr>
<td>$t_2$</td>
<td>$t_2$</td>
<td>*</td>
<td>*</td>
<td>100.59</td>
</tr>
<tr>
<td>$t_3$</td>
<td>$t_3$</td>
<td>45.06</td>
<td>50.53</td>
<td>*</td>
</tr>
<tr>
<td>$t_4$</td>
<td>$t_4$</td>
<td>*</td>
<td>*</td>
<td>37.63</td>
</tr>
<tr>
<td>$t_5$</td>
<td>$t_5$</td>
<td>192.80</td>
<td>242.35</td>
<td>*</td>
</tr>
<tr>
<td>$t_6$</td>
<td>$t_6$</td>
<td>*</td>
<td>*</td>
<td>278.02</td>
</tr>
<tr>
<td>$t_{5G}$</td>
<td>$t_{5G}$</td>
<td>226.88</td>
<td>255.86</td>
<td>*</td>
</tr>
<tr>
<td>$t_{6G}$</td>
<td>$t_{6G}$</td>
<td>*</td>
<td>*</td>
<td>615.74</td>
</tr>
</tbody>
</table>

*: Data not applicable

Table 2: Optimal values of a, b, c and d

<table>
<thead>
<tr>
<th></th>
<th>Populations</th>
<th>I</th>
<th>II</th>
<th>III</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Constants</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a^*$</td>
<td>$a^*$</td>
<td>1.83</td>
<td>1.77</td>
<td>-</td>
</tr>
<tr>
<td>$b^*$</td>
<td>$b^*$</td>
<td>10.93</td>
<td>12.53</td>
<td>-</td>
</tr>
<tr>
<td>$c^*$</td>
<td>$c^*$</td>
<td>-</td>
<td>-</td>
<td>14.41</td>
</tr>
<tr>
<td>$d^*$</td>
<td>$d^*$</td>
<td>-</td>
<td>-</td>
<td>1.08</td>
</tr>
</tbody>
</table>

**Population II:** Source: Sukhatme and Chand [11]

X: Bushels of apples harvested in 1964.
Z: Bushels of apples harvested in 1959

The required parameters of the population are:

\[ \overline{Y} = 0.103182 \times 10^4, \quad C_2 = 2.55280, \quad \rho_{yx} = 0.93, \quad N = 200 \]
\[ \overline{X} = 0.293458 \times 10^4, \quad C_2 = 4.0250, \quad \rho_{zx} = 0.77, \quad n = 30 \]
\[ \overline{Z} = 0.365149 \times 10^4, \quad C_2 = 2.09379, \quad \rho_{zx} = 0.84 \]

**Population III:** Source: Gujarati [4]

Y: Average miles per gallon
X: Top speed, miles per hour
Z: Cubic feet of cab space

The required parameters of the population are:

\[ \overline{Y} = 33.83457, \quad C_2 = 0.088324, \quad \rho_{yx} = -0.69079, \quad N = 81 \]
\[ \overline{X} = 112.4568, \quad C_2 = 0.015765, \quad \rho_{zx} = -0.36831, \quad n = 20 \]
\[ \overline{Z} = 98.76543, \quad C_2 = 0.050987, \quad \rho_{zx} = -0.04265 \]

The optimal values of a, b and c, d in Table 2 are computed using (4.8) and (4.12), respectively. While obtaining the percent relative efficiencies of the estimators in Table 1, the MSE values of the estimators are computed using (3.1), (1.4), (1.5), (1.8), (1.9), (1.12), (1.13), (2.6), (2.10), (4.9) and (4.13). From Table 1, it is clearly observed that the suggested estimators are more efficient than sample mean estimator, classical ratio and product estimators, Bahl-Tuteja estimators and Singh estimators. It is also shown that the performances of $t_{5G}$ and $t_{6G}$, based on optimal values presented in Table 2, are better than all of the other estimators.

**REFERENCES**