Improved Class of Ratio Estimators for Finite Population Variance

By Audu Ahmed, Adedayo Amos Adewara, Ran Vijay Kumar Singh

Usman Danfodiyo University Sokoto, Nigeria

Abstract- In this paper, we have suggested a class of improved ratio estimators for finite population variance. The proposed class of estimators is obtained by transforming both the sample variances of study and auxiliary variables. The MSE of the proposed estimators have been obtained and the conditions for their efficiency over some existing variance estimators have been established. The present family of finite variance estimator, having obtaining the optimal values of the constants, exhibit significant improvement over the estimators considered in the study. The empirical study is also conducted to corroborate the theoretical results and the results show that the proposed class of estimators is more efficient.

Keywords: efficiency, mean square error, ratio estimator, finite population, variance.

GJSFR-F Classification : MSC 2010: 00A05

Strictly as per the compliance and regulations of:
Improved Class of Ratio Estimators for Finite Population Variance

Audu Ahmed *, Adedayo Amos Adewara * & Ran Vijay Kumar Singh *

Abstract: In this paper, we have suggested a class of improved ratio estimators for finite population variance. The proposed class of estimators is obtained by transforming both the sample variances of study and auxiliary variables. The MSE of the proposed estimators have been obtained and the conditions for their efficiency over some existing variance estimators have been established. The present family of finite variance estimator, having obtaining the optimal values of the constants, exhibit significant improvement over the estimators considered in the study. The empirical study is also conducted to corroborate the theoretical results and the results show that the proposed class of estimators is more efficient.

Keywords: efficiency, mean square error, ratio estimator, finite population, variance.

I. Introduction

The variation of produce or yields in the manufacturing industries and pharmaceutical laboratories are sometime a matter of concern to researchers (Ahmed et al. [2]). The use of supplementary (auxiliary) information, being constant with unit (e.g. population mean, population standard deviation, e.t.c.) or unit free constant (e.g. Coefficient of variation, Kurtosis, e.t.c.), can enhance the efficiency at the estimation stage. In recent past, this concept has been utilized by several authors to improve the efficiency of ratio and product type estimators for estimating population mean as well population variance of study variable.

In this paper, an improved class of ratio estimators for estimating finite population variance has been proposed with objective to produce efficient estimators and their properties have established.

Let $\Omega = \{1, 2, 3 \ldots N\}$ be a population of size $N$ and $Y, X$ be two real valued functions having values $(Y_i, X_i) \in \mathbb{R}^+ > 0$ on the $i^{th}$ unit of $U(1 \leq i \leq N)$. We assume positive correlation $\rho > 0$ between the study variable $Y$ and auxiliary variable $X$. Let $S_y^2$ and $S_x^2$ be the finite population variance of $Y$ and $X$ respectively and $s_y^2$ and $s_x^2$ be respective sample variances based on the random sample of size $n$ drawn without replacement.

Singh et al. [8] defined the general family of estimators for estimating finite population variance $S_y^2$ of the study variable $Y$ as
\[ \eta = s_y^2 \left[ \frac{aS_x^2 + b}{\alpha(as_x^2 - b) + (1 - \alpha)(aS_x^2 - b)} \right] \] (1.1)

where \(a\) and \(b\) are constants based on auxiliary variable \(X\) like coefficient of skewness, kurtosis and correlation coefficient etc. \(\alpha\) is the constant that minimizes the mean square error (MSE) of the estimator. Table 1 shows some members of \(\eta\)-family for different values of \(a\), \(b\) and \(\alpha\). The MSEs/Variance of the estimators in Table 1 are given below:

\[ \text{Var}(\eta_0) = \gamma S_y^2 (\psi_{40} - 1) \] (1.2)

\[
MSE(\eta_i) = \begin{cases} 
S_i^4 \gamma \left[ (\psi_{40} - 1) + (\psi_{04} - 1) - 2(\psi_{22} - 1) \right] & i = 1 \\
S_i^4 \gamma \left[ (\psi_{40} - 1) + h_i^2 (\psi_{04} - 1) - 2h_i (\psi_{22} - 1) \right] & i = 2, 3, 4, 5, 6 
\end{cases}
\] (1.3)

\[ h_2 = \frac{S_x^2}{S_x^2 - C_x}, \quad h_3 = \frac{S_x^2}{S_x^2 - \beta_2(x)}, \quad h_4 = \frac{S_x^2 \beta_2(x)}{S_x^2 \beta_2(x) - C_x}, \quad h_5 = \frac{S_x^2 C_x}{S_x^2 C_x - \beta_2(x)}, \quad h_6 = \frac{S_x^2}{S_x^2 + \beta_2(x)} \]

where

\[ s_y^2 = \frac{1}{n-1} \sum_{i=1}^{n} (y_i - \bar{y})^2, \quad s_x^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2, \quad S_y^2 = \frac{1}{N-1} \sum_{i=1}^{N} (Y_i - \overline{Y})^2, \quad S_x^2 = \frac{1}{N-1} \sum_{i=1}^{N} (X_i - \overline{X})^2 \]

\[ \bar{X} = \frac{1}{N} \sum_{i=1}^{N} X_i, \quad \bar{Y} = \frac{1}{N} \sum_{i=1}^{N} Y_i, \quad \bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i, \quad \bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i \]

\[ \gamma = \frac{1}{n}, \quad \psi_{rs} = \frac{\lambda_{rs}}{\lambda_{20} \lambda_{02}^2} \quad \text{and} \quad \lambda_{rs} = \frac{1}{N-1} \sum_{i=1}^{N} (Y_i - \overline{Y})^r (X_i - \overline{X})^s \]

The MSE of \(\eta\) to second order approximation is given below:

\[ MSE(\eta) = \gamma S_y^4 \left[ (\psi_{40} - 1) + \alpha^2 \psi_{04}^2 - 2\alpha \psi_{22} \psi_{04} - 2\alpha \psi_{22} - 1 \right] \] (1.4)

The \(MSE(\eta)\) expression is minimized for the optimum values of \(\alpha\) given by equation (1.5). This is obtained by partial differentiation of equation (1.4) with respect to \(\alpha\).

\[ MSE(\eta)_{\text{min}} = \gamma S_y^4 \left[ (\psi_{40} - 1) - \frac{(\psi_{22} - 1)^2}{(\psi_{04} - 1)} \right] \] (1.5)

Many other researchers including Kadilar and Cingi [6], Yadav and Kadilar [15], Singh and Solanki [10], Gupta and Shabir [3], Subramani and Kumarapandian [13], Singh and Vishwakarma [11], Singh, et al. [9], Sanaullah, et al. [7] and Solanki and Singh [12] have significantly contributed to the improvement of both ratio and product mean & variance estimators in sampling survey.
II. PROPOSED ESTIMATOR

After studying the related finite population variance estimators stated in section 1 and motivated by the work of Yadav and Kadilar [16] and Adewara et al.[1] estimators for population mean in which the former (i.e. Yadav and Kadilar [16]) equals the latter (i.e. Adewara et al.[1]) when \( k_i = k_2 = k_3 = k_4 = k_5 = k_6 = 1 \). Their estimators are defined as

\[
\eta_i^* = k_i \bar{y}^* \left( \frac{\bar{X}}{\bar{x}} \right), \quad \eta_2^* = k_2 \bar{y}^* \left( \frac{\bar{X}}{\bar{x}} \right), \quad \eta_3^* = k_3 \bar{y}^* \left( \frac{\bar{X} + C_x}{\bar{x} + C_x} \right),
\]

\[
\eta_4^* = k_4 \bar{y}^* \left( \frac{\bar{X} + \rho}{\bar{x} + \rho} \right), \quad \eta_5^* = k_5 \bar{y}^* \left( \frac{\bar{X} + \rho}{\bar{x} + \rho} \right), \quad \eta_6^* = k_6 \bar{y}^* \left( \frac{\bar{X} + \rho}{\bar{x} + \rho} \right)
\]

Where \( \bar{x}^* \) and \( \bar{y}^* \) are the respective sample means of the auxiliary and study variables, having the relationship: (1) \( \bar{X} = f \bar{x} + (1-f)\bar{x}^* \) (2) \( \bar{Y} = f \bar{y} + (1-f)\bar{y}^* \) where \( f = \frac{n}{N} \) is finite population correction, \( k_i, i = 1,2,3,4,5,6, \) are real constants.

We proposed the following class of ratio estimator

\[
\tau_i^* = k_i S_y^2 \left( \frac{S_x^2}{S_x^2} \right), \quad \tau_2^* = k_2 S_y^2 \left( \frac{S_x^2-C_x}{S_x^2-C_x} \right), \quad \tau_3^* = k_3 S_y^2 \left( \frac{S_x^2-C_x}{S_x^2-C_x} \right), \quad \tau_4^* = k_4 S_y^2 \left( \frac{S_x^2-C_x}{S_x^2-C_x} \right),
\]

\[
\tau_5^* = k_5 S_y^2 \left( \frac{S_x^2-C_x}{S_x^2-C_x} \right), \quad \tau_6^* = k_6 S_y^2 \left( \frac{S_x^2+C_x}{S_x^2+C_x} \right)
\]

Where \( S_x^2 \) and \( S_y^2 \) are the respective sample finite variances of the auxiliary and study variables, having the relationship: (1) \( S_x^2 = f S_x^2 + (1-f)S_y^2 \) (2) \( S_y^2 = f S_x^2 + (1-f)S_y^2 \) with condition that \( n < \frac{1}{2} N \).

In order to obtain the MSE, we defined \( e_0 = \frac{S_x^2 - S_y^2}{S_y^2} \) and \( e_1 = \frac{S_y^2 - S_x^2}{S_x^2} \) such that

\[
\begin{align*}
E(e_0) &= E(e_1) = 0, \quad E(e_0 e_1) = \gamma(\psi_{40} - 1) \\
E(e_1^2) &= \gamma(\psi_{22} - 1)
\end{align*}
\]

Expressing \( \tau_i^* \), \( i = 1,2,3,4,5,6, \) in terms of \( e_0 \) and \( e_1 \), we have

\[
\tau_i^* = \frac{k_i S_y^2 (1 - ve_0) (1 - ve_1)^{-1}}{k_i S_x^2 (1 - ve_0) (1 - ve_1)^{-1}} \quad i = 1,2,3,4,5,6
\]

now assume that \( |ve_0| < 1 \) and \( |ve_1| < 1 \) so that \( (1 - ve_0)^{-1} \) and \( (1 - ve_1)^{-1} \) are expandable. Expanding the right hand side of (2.2) up to second degree approximation,
subtract $S^2$ from its both sides, square the both sides and taking expectation using the results in equation (2.1), we obtain the MSEs of the proposed estimators as:

$$MSE(i^*_i) = \left\{ \begin{array}{ll} S^4_y \gamma v^2 \left[ k_i^2 (\psi_{04} - 1) + (3k_i^2 - 2k_i)(\psi_{04} - 1) \right] + (k_i - 1)^2, & i = 1 \\ S^4_y \gamma v^2 \left[ -2 \left( 2k_i^2 - k_i \right) (\psi_{22} - 1) \right] + (k_i - 1)^2, & i = 2, 3, 4, 5, 6 \end{array} \right.$$  \hspace{1cm} (2.3)

The $MSE(\eta_i), i = 1, 2, 3, 4, 5, 6$ expressions are minimized for the optimum values of $k$ given by

$$k_i^* = \frac{v^2 \gamma \left[ (\psi_{04} - 1) - (\psi_{22} - 1) \right] + 1}{v^2 \gamma \left[ 3(\psi_{04} - 1) - 4(\psi_{22} - 1) + (\psi_{40} - 1) \right] + 1} = \frac{A_i}{B_i}, \hspace{1cm} i = 1 \hspace{1cm} (2.4)$$

$$k_i^* = \frac{v^2 \gamma \left[ h_i^2 (\psi_{04} - 1) - h_i (\psi_{22} - 1) \right] + 1}{v^2 \gamma \left[ 3h_i^2 (\psi_{04} - 1) - 4h_i (\psi_{22} - 1) + (\psi_{40} - 1) \right] + 1} = \frac{A_i}{B_i}, \hspace{1cm} i = 2, 3, 4, 5, 6 \hspace{1cm} (2.5)$$

where $v = \frac{n}{N-n}$

Replacing $k_i$ by $k_i^*, i = 1, 2, 3, 4, 5, 6$ in equation (2.3), we obtain the minimum MSE as

$$MSE_{min}(\tau_i^*) = \left\{ \begin{array}{ll} S^4_y \left( 1 - \frac{A_i^2}{B_i} \right), & i = 1 \\ S^4_y \left( 1 - \frac{A_i^2}{B_i} \right), & i = 2, 3, 4, 5, 6 \end{array} \right.$$  \hspace{1cm} (2.6)

### III. Efficiency Comparisons

In this section efficiencies of the proposed estimators are compared with efficiencies of some estimators in the literature.

The $\tau_i^*$ family of estimators of the population variance is more efficient than $\eta_0$ if,

$$MSE_{min}(\tau_i^*) < Var(\eta_0), \hspace{1cm} i = 1, 2, 3, 4, 5, 6$$

$$\left\{ \begin{array}{ll} 1 - \frac{A_i^2}{B_i} < \gamma \left[ (\psi_{04} - 1) \right], & i = 1 \\ 1 - \frac{A_i^2}{B_i} < \gamma \left[ (\psi_{04} - 1) \right], & i = 2, 3, 4, 5, 6 \end{array} \right.$$  \hspace{1cm} (3.1)

The $\tau_i^*$ family of estimators of the population variance is more efficient than $\eta_i$ family if,
\[ MSE_{\min}(\tau^*_i) < MSE(\eta_i) \quad i = 1, 2, 3, 4, 5, 6 \]

\[
\begin{align*}
\left( 1 - \frac{A_i^2}{B_1} \right) &< \gamma \left[ (\psi_{40} - 1) + (\psi_{04} - 1) - 2(\psi_{22} - 1) \right] & i = 1 \\
\left( 1 - \frac{A_i^2}{B_2} \right) &< \gamma \left[ (\psi_{40} - 1) + h_i^2(\psi_{04} - 1) - 2h_i(\psi_{22} - 1) \right] & i = 2, 3, 4, 5, 6
\end{align*}
\]

(3.2)

The \( \tau^*_i - \text{family} \) of estimators of the population variance is more efficient than \( \eta \) if,

\[ MSE_{\min}(\tau^*_i) < MSE_{\min}(\eta) \quad i = 1, 2, 3, 4, 5, 6 \]

\[
\begin{align*}
\left( 1 - \frac{A_i^2}{B_1} \right) &< \gamma \left[ (\psi_{40} - 1) - (\psi_{04} - 1) \right] & i = 1 \\
\left( 1 - \frac{A_i^2}{B_2} \right) &< \gamma \left[ (\psi_{40} - 1) - (\psi_{04} - 1) \right] & i = 2, 3, 4, 5, 6
\end{align*}
\]

(3.3)

When conditions (3.1), (3.2) and (3.3) are satisfied, we can conclude that the family of proposed class of estimators is more efficient than the \( \eta - \text{family} \) estimator.

IV. Empirical Study

In order to investigate the merits of the proposed estimators, we have considered the following two real populations given as:

Data 1: Subramani and Kumarapandiyan [13]

\[ N = 49, \, n = 20, \, \bar{Y} = 116.1633, \, \bar{X} = 98.6765, \, \rho = 0.6904, \, S_y = 98.8286 \]

\[ S_x = 102.9709, \, C_y = 0.8508, \, C_x = 1.0435, \, \psi_{40} = 4.9245, \, \psi_{04} = 5.9878, \, \psi_{22} = 4.6977 \]

Data 2: Subramani and Kumarapandiyan [13]

\[ N = 80, \, n = 20, \, \bar{Y} = 51.8264, \, \bar{X} = 11.2646, \, \rho = 0.9413, \, S_y = 18.3569 \]

\[ S_x = 8.4563, \, C_y = 0.3542, \, C_x = 0.7507, \, \psi_{40} = 2.2667, \, \psi_{04} = 2.8664, \, \psi_{22} = 2.2209 \]

The numerical demonstration to justify the appropriateness of the suggested class of variance estimator has been conducted using the two data sets. Table 2 and Table 3 show the mean square errors of proposed estimators and that of some existing estimators and their percentage relative efficiencies of different estimators with respect to \( s_y^2 \) respectively.

V. Conclusion

From section 3, the theoretical conditions obtained for the efficiencies of the proposed estimators supported by the numerical illustration in section 4 given in the Table 2 and 3. From the result, we infer that the suggested variance estimators produce a better estimate of finite population variance than the existing estimators in the sense of having higher percentage relative efficiency which implies lesser mean square error.
REFERENCES  Références Referencias

### Table 1: Some Member of \( \eta \)-family for different values of \( a, b \) and \( \alpha \)

<table>
<thead>
<tr>
<th>Estimator</th>
<th>( a )</th>
<th>( b )</th>
<th>( \alpha )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \eta_0 = s_y^2 )</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sample variance</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \eta_1 = s_y^2 \left( \frac{S_x^2}{s_y^2} \right) )</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Isaki [4]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \eta_2 = s_y^2 \left( \frac{S_x^2 - C_x}{s_y^2 - C_x} \right) )</td>
<td>1</td>
<td>( C_x )</td>
<td>1</td>
</tr>
<tr>
<td>Kadilar and Cingi [5]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \eta_3 = s_y^2 \left( \frac{S_x^2 - \beta_2(x)}{s_y^2 - \beta_2(x)} \right) )</td>
<td>1</td>
<td>( \beta_2(x) )</td>
<td>1</td>
</tr>
<tr>
<td>( \eta_4 = s_y^2 \left( \frac{S_x^2 \beta_2(x) - C_x}{s_y^2 \beta_2(x) - C_x} \right) )</td>
<td>( \beta_2(x) )</td>
<td>( C_x )</td>
<td>1</td>
</tr>
<tr>
<td>( \eta_5 = s_y^2 \left( \frac{S_x^2 C_x - \beta_2(x)}{s_y^2 C_x - \beta_2(x)} \right) )</td>
<td>( C_x )</td>
<td>( \beta_2(x) )</td>
<td>1</td>
</tr>
<tr>
<td>( \eta_6 = s_y^2 \left( \frac{S_x^2 + \beta_2(x)}{s_y^2 + \beta_2(x)} \right) )</td>
<td>1</td>
<td>( -\beta_2(x) )</td>
<td>1</td>
</tr>
</tbody>
</table>

Upadhya and Singh [14]

### Table 2: MSE of different estimators

<table>
<thead>
<tr>
<th>Estimator</th>
<th>Data 1</th>
<th>Data 2</th>
<th>Estimator</th>
<th>Data 1</th>
<th>Data 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \eta_0 )</td>
<td>187190</td>
<td>7191.859</td>
<td>( \eta_{opt} )</td>
<td>5643700</td>
<td>2657.427</td>
</tr>
</tbody>
</table>

\( \eta \)-family

(Singh et al. [7] estimators)

<table>
<thead>
<tr>
<th>Estimator</th>
<th>Data 1</th>
<th>Data 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \eta_1 )</td>
<td>723531</td>
<td>3924.948</td>
</tr>
<tr>
<td>( \eta_2 )</td>
<td>723652</td>
<td>4003.906</td>
</tr>
<tr>
<td>( \eta_3 )</td>
<td>724227</td>
<td>4249.508</td>
</tr>
<tr>
<td>( \eta_4 )</td>
<td>723551</td>
<td>3952.035</td>
</tr>
<tr>
<td>( \eta_5 )</td>
<td>724198</td>
<td>4372.12</td>
</tr>
<tr>
<td>( \eta_6 )</td>
<td>722837</td>
<td>3658.1</td>
</tr>
</tbody>
</table>

\( \tau^* \)-family

(NEWLY PROPOSED ESTIMATORS)

<table>
<thead>
<tr>
<th>Estimator</th>
<th>Data 1</th>
<th>Data 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \tau_1^* )</td>
<td>3053941</td>
<td>429.9105</td>
</tr>
<tr>
<td>( \tau_2^* )</td>
<td>3054289</td>
<td>438.3024</td>
</tr>
<tr>
<td>( \tau_3^* )</td>
<td>3055940</td>
<td>464.3576</td>
</tr>
<tr>
<td>( \tau_4^* )</td>
<td>3053999</td>
<td>432.7903</td>
</tr>
<tr>
<td>( \tau_5^* )</td>
<td>3055857</td>
<td>477.3385</td>
</tr>
<tr>
<td>( \tau_6^* )</td>
<td>3051947</td>
<td>401.5034</td>
</tr>
</tbody>
</table>

Ref

Table 3: PRE of different estimators with respect to $s^2_r$

<table>
<thead>
<tr>
<th>Estimator</th>
<th>Data 1</th>
<th>Data 2</th>
<th>Estimator</th>
<th>Data 1</th>
<th>Data 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\eta_0$</td>
<td>100.00</td>
<td>100.00</td>
<td>$\eta_{opt}$</td>
<td>331.6813</td>
<td>270.6324</td>
</tr>
<tr>
<td>$\eta$ - family (Singh et al. [7] estimators)</td>
<td></td>
<td></td>
<td>$\tau^*$ - family (Newly proposed estimators)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\eta_1$</td>
<td>258.7184</td>
<td>183.2345</td>
<td>$\tau^*_1$</td>
<td>612.949</td>
<td>1672.873</td>
</tr>
<tr>
<td>$\eta_2$</td>
<td>258.6751</td>
<td>179.6211</td>
<td>$\tau^*_2$</td>
<td>612.8791</td>
<td>1640.844</td>
</tr>
<tr>
<td>$\eta_3$</td>
<td>258.4697</td>
<td>169.2398</td>
<td>$\tau^*_3$</td>
<td>612.5479</td>
<td>1548.776</td>
</tr>
<tr>
<td>$\eta_4$</td>
<td>258.7112</td>
<td>181.9786</td>
<td>$\tau^*_4$</td>
<td>612.9373</td>
<td>1661.742</td>
</tr>
<tr>
<td>$\eta_5$</td>
<td>258.4801</td>
<td>164.4934</td>
<td>$\tau^*_5$</td>
<td>612.5646</td>
<td>1506.658</td>
</tr>
<tr>
<td>$\eta_6$</td>
<td>258.9668</td>
<td>196.5956</td>
<td>$\tau^*_6$</td>
<td>613.3494</td>
<td>1791.233</td>
</tr>
</tbody>
</table>