An efficient new class of estimators of population variance using information on auxiliary attribute in sample surveys

Housila P. Singh,* and Surya K. Pal,†‡

Abstract
This paper addresses the problem of estimating the population variance $S^2_Y$ of the study variate $Y$ when information on an auxiliary attribute is available. We have suggested a new class of estimators for the population variance $S^2_Y$ using information on an auxiliary attribute. It has been shown that the proposed class of estimators is more efficient than those estimators considered by [16], [19] and [23]. The theoretical findings are well supported through an empirical study.

Keywords: Study variate, Auxiliary attribute, Population variance, Bias, Mean squared error

2000 AMS Classification: 62D05

Received: 08.09.2015 Accepted: 15.03.2016 Doi: 10.15672/HJMS.201615322373

1. Introduction

In the theory of sample surveys, it is well known that the suitable use of auxiliary information provides more efficient estimators of population parameters such as mean or total and variance of the variable under study. In literature of survey sampling various authors have proposed estimators of population parameters based on information about the population parameters of the auxiliary variable. Sometimes, there exist situations when information is available in the form of auxiliary attribute $\phi$ which is highly correlated with study variable $y$. For example (i) $y$ may be the use of drugs and $\phi$ may be the gender, (ii) $y$ may be production of a crop and $\phi$ may be the particular variety, (iii) $y$ may be the amount of milk produced and $\phi$ may be a particular breed of cow, (iv) $y$ may be the yield of wheat crop and $\phi$ may be the particular variety of wheat; (see, [23], p.64). In these situations by taking the advantage of point bi-serial (see, [6]) correlation between the study variable $y$ and the auxiliary attribute $\phi$, the efficient estimators of

---

*School of Studies in Statistics, Vikram University, Ujjain, M.P., India.
†School of Studies in Statistics, Vikram University, Ujjain, M.P., India, Email: suryakantpal6676@gmail.com
‡Corresponding Author.
population parameters under investigation can be formulated. Various authors including [1], [5], [7], [9], [10], [13], [14], [17], [19] and [20] suggested a number of estimators of the population mean $\bar{Y}$ of the study variable $y$ using information on a single auxiliary attribute $\phi$. [8], [12], [15] and [22] have discussed the problem of estimating the population mean $\bar{Y}$ of the study variable $y$ using information on two auxiliary attributes in simple random sampling. However, in many situations of practical importance, the problem of estimating the finite population variance $S_y^2$ of the study variable $y$ also deserves special attention in presence of auxiliary attribute for instance, see [2], [3], [4], [16], [19] and [23].

Consider a finite population $U = (U_1, U_2, \ldots, U_N)$ of size $N$. Let $y_i$ and $\phi_i$ ($i = 1, 2, \ldots, N$) be the observations on the study variable $y$ and the auxiliary attribute $\phi$ respectively.

We note that $\phi_i = \begin{cases} 1, & \text{if } i^{th} \text{ unit processes attribute } \phi; \\ 0, & \text{otherwise} \end{cases}$

For estimating the population variance $S_y^2 = \frac{1}{N-1} \sum_{i=1}^{N} (y_i - \bar{Y})^2$ with population mean $\bar{Y} = \frac{1}{N} \sum_{i=1}^{N} y_i$, a simple random sample of size $n$ is drawn without replacement from the population $U$. Let $M = \sum_{i=1}^{N} \phi_i$ and $m = \sum_{i=1}^{n} \phi_i$ denote the total number of units in the population and sample respectively possessing the attribute $\phi$. Let the corresponding population and sample proportion be $P = (M/N)$ and $p = (m/n)$ respectively.

In many situations, in addition to the population proportion $P$ of the auxiliary attribute $\phi$, various parameters associated with auxiliary attribute such as population variance $S_\phi^2 = \frac{1}{N-1} \sum_{i=1}^{N} (\phi_i - P)^2$, coefficient of variation $C_p = (S_\phi/P)$, coefficient of skewness $\beta_1(\phi)$, coefficient of kurtosis $\beta_2(\phi)$ and point bi-serial correlation coefficient $\rho_{pb}$ etc. may also be known. Utilizing such information [2], [3], [19] and [23] have developed various estimators of the population variance $S_y^2$ of the study variable $y$ to improve the conventional estimators given by [16]. Let $s_y^2 = \frac{1}{n-1} \sum_{i=1}^{n} (y_i - \bar{y})^2$ and $s_\phi^2 = \frac{1}{n-1} \sum_{i=1}^{n} (\phi_i - p)^2$ be the sample variances corresponding to the population variances $S_y^2$ and $S_\phi^2$ respectively, where $\bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i$.

In this paper our main aim is to estimate the unknown population variance $S_y^2$ of the study variable $y$ by improving the estimators envisaged earlier using information on auxiliary attribute $\phi$ such as $S_\phi^2, C_p, \beta_1(\phi), \beta_2(\phi)$ and $\rho_{pb}$ etc. The remaining part of the paper is organized as follows: In Section 2, a class of estimators of population variance $S_y^2$ is suggested. In addition, some known and unknown members of the suggested class of estimators are given in Table1 and Table2 respectively. The expressions of asymptotic bias and the mean squared error (MSE) of the suggested class of estimators are obtained in Subsection 2.1. The problem of efficiency comparisons of the suggested class of estimators with some existing estimators has been addressed in Section 3. In Section 4, an empirical study is carried out in support of the present study. We conclude with a brief discussion in Section 5.

2. Proposed class of estimators

Taking motivation from [10] and [18] we propose the following class of estimators of the population variance $S_y^2$ as

$$t = \frac{[w_1 s_y^2 + w_2 (S_\phi^2 - s_\phi^2)]}{\alpha (as_\phi^2 + b) + (1 - \alpha) (aS_\phi^2 + b)} \{\exp \left(\frac{hc(S_y^2 - s_y^2)}{c(S_y^2 + s_y^2) + 2d}\right)\}.$$
where $\alpha$ is a suitable constant, $(a \neq 0, b, c \neq 0, d)$ are either real numbers or the functions of the known parameters of the auxiliary attribute $\phi$ such as $C_P, \beta_1(\phi), \beta_2(\phi), \Delta, \rho_{pb}, b_{pb} = \rho_{pb}(C_P/C_R)$, where $C_R = (S_y/Y)$ and $\Delta = (\beta_2(\phi) - \beta_1(\phi) - 1)$; $(g, h)$ are constants may take values between -1 to 1 (i.e. $-1 \leq (g, h) \leq 1$), and $w_1$ and $w_2$ are the scalars such that mean squared error of the proposed class of estimators $t$ is minimum. The proposed class of estimators $t$ includes various known estimators of the population variance which are listed in Table 1 for various suitable values of $(w_1, w_2, g, h, a, b, c, d)$. Some unknown members of the suggested class of estimators of the population variance $S^2_y$ are tabulated in Table 2.

Table 1. Some known members of the proposed class of estimators $t$.

<table>
<thead>
<tr>
<th>Estimator (Est.)</th>
<th>$w_1$</th>
<th>$w_2$</th>
<th>$g$</th>
<th>$h$</th>
<th>$\alpha$</th>
<th>$a$</th>
<th>$b$</th>
<th>$c$</th>
<th>$d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_1 = s_0^2/\sigma^2$</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$t_2 = s_0^2 + b_0(S_1^2 - s_0^2)$</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$t_3 = s_0^2 \exp(\frac{S_1^2}{S_2^2} - 1)$</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$t_4 = s_0^2 \left(\frac{S_1^2 + b_0(S_1^2 - s_0^2)}{S_2^2}\right)$</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$t_5 = s_0^2 \left(\frac{S_1^2 + b_0(S_1^2 - s_0^2)}{S_2^2}\right)$</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$t_6 = s_0^2 \left(\frac{S_1^2 + b_0(S_1^2 - s_0^2)}{S_2^2}\right)$</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$t_7 = s_0^2 \left(\frac{S_1^2 + b_0(S_1^2 - s_0^2)}{S_2^2}\right)$</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$t_8 =</td>
<td>s_0^2 + b_0(S_1^2 - s_0^2)</td>
<td>/\frac{S_2^2}{\rho_{pb}(C_P/C_R)}$</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$t_9 =</td>
<td>s_0^2 + b_0(S_1^2 - s_0^2)</td>
<td>/\frac{S_2^2}{\rho_{pb}(C_P/C_R)}$</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$t_{10} =</td>
<td>s_0^2 + b_0(S_1^2 - s_0^2)</td>
<td>/\frac{S_2^2}{\rho_{pb}(C_P/C_R)}$</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$t_{11} =</td>
<td>s_0^2 + b_0(S_1^2 - s_0^2)</td>
<td>/\frac{S_2^2}{\rho_{pb}(C_P/C_R)}$</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$t_{12} =</td>
<td>s_0^2 + b_0(S_1^2 - s_0^2)</td>
<td>/\frac{S_2^2}{\rho_{pb}(C_P/C_R)}$</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$t_{13} =</td>
<td>s_0^2 + b_0(S_1^2 - s_0^2)</td>
<td>/\frac{S_2^2}{\rho_{pb}(C_P/C_R)}$</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$t_{14} =</td>
<td>s_0^2 + b_0(S_1^2 - s_0^2)</td>
<td>/\frac{S_2^2}{\rho_{pb}(C_P/C_R)}$</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$t_{15} =</td>
<td>s_0^2 + b_0(S_1^2 - s_0^2)</td>
<td>/\frac{S_2^2}{\rho_{pb}(C_P/C_R)}$</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$t_{16} =</td>
<td>s_0^2 + b_0(S_1^2 - s_0^2)</td>
<td>/\frac{S_2^2}{\rho_{pb}(C_P/C_R)}$</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$t_{17} =</td>
<td>s_0^2 + b_0(S_1^2 - s_0^2)</td>
<td>/\frac{S_2^2}{\rho_{pb}(C_P/C_R)}$</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$t_{18} =</td>
<td>s_0^2 + b_0(S_1^2 - s_0^2)</td>
<td>/\frac{S_2^2}{\rho_{pb}(C_P/C_R)}$</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$t_{19} =</td>
<td>s_0^2 + b_0(S_1^2 - s_0^2)</td>
<td>/\frac{S_2^2}{\rho_{pb}(C_P/C_R)}$</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$t_{20} =</td>
<td>s_0^2 + b_0(S_1^2 - s_0^2)</td>
<td>/\frac{S_2^2}{\rho_{pb}(C_P/C_R)}$</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$t_{21} =</td>
<td>s_0^2 + b_0(S_1^2 - s_0^2)</td>
<td>/\frac{S_2^2}{\rho_{pb}(C_P/C_R)}$</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$t_{22} =</td>
<td>s_0^2 + b_0(S_1^2 - s_0^2)</td>
<td>/\frac{S_2^2}{\rho_{pb}(C_P/C_R)}$</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$t_{23} =</td>
<td>s_0^2 + b_0(S_1^2 - s_0^2)</td>
<td>/\frac{S_2^2}{\rho_{pb}(C_P/C_R)}$</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$t_{24} = t_0 = s_0^2$</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

where $b_0$ and $(\eta, \varphi)$ being constants.
Table 2. Some unknown members of the proposed class of estimators $t$ for $(g, h, a) = (1, 1, 1)$.

<table>
<thead>
<tr>
<th>Estimator</th>
<th>$a$</th>
<th>$b$</th>
<th>$c$</th>
<th>$d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_1^* = [w_1 s_o^2 + w_2 (S_o^2 - s_o^2)] (\frac{S_o^2 - s_o^2}{S_o^2 - s_o^2}) \exp[\frac{S_o^2 - s_o^2}{S_o^2 + s_o^2}]$</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>$t_2^* = [w_1 s_o^2 + w_2 (S_o^2 - s_o^2)] (\frac{S_o^2 + s_o^2}{S_o^2 + s_o^2}) \exp[\frac{S_o^2 - s_o^2}{S_o^2 + s_o^2 + 2s_o^2}]$</td>
<td>1</td>
<td>$C_p$</td>
<td>1</td>
<td>$C_p$</td>
</tr>
<tr>
<td>$t_3^* = [w_1 s_o^2 + w_2 (S_o^2 - s_o^2)] (\frac{S_o^2 + s_o^2}{S_o^2 + s_o^2 + 2s_o^2}) \exp[\frac{S_o^2 - s_o^2}{S_o^2 + s_o^2 + 2s_o^2}]$</td>
<td>1</td>
<td>$\beta_2(\phi)$</td>
<td>1</td>
<td>$\beta_2(\phi)$</td>
</tr>
<tr>
<td>$t_4^* = [w_1 s_o^2 + w_2 (S_o^2 - s_o^2)] (\frac{S_o^2 + s_o^2}{S_o^2 + s_o^2 + 2s_o^2}) \exp[\frac{C_p(S_o^2 - s_o^2)}{C_p(S_o^2 + s_o^2) + 2s_o^2}]$</td>
<td>1</td>
<td>$\beta_2(\phi)$</td>
<td>1</td>
<td>$\beta_2(\phi)$</td>
</tr>
<tr>
<td>$t_5^* = [w_1 s_o^2 + w_2 (S_o^2 - s_o^2)] (\frac{S_o^2 + s_o^2}{S_o^2 + s_o^2 + 2s_o^2}) \exp[\frac{S_o^2 - s_o^2}{S_o^2 + s_o^2 + 2s_o^2}]$</td>
<td>1</td>
<td>$\beta_2(\phi)$</td>
<td>1</td>
<td>$\beta_2(\phi)$</td>
</tr>
<tr>
<td>$t_6^* = [w_1 s_o^2 + w_2 (S_o^2 - s_o^2)] (\frac{S_o^2 + s_o^2}{S_o^2 + s_o^2 + 2s_o^2}) \exp[\frac{S_o^2 - s_o^2}{S_o^2 + s_o^2 + 2s_o^2}]$</td>
<td>1</td>
<td>$C_p$</td>
<td>$\beta_2(\phi)$</td>
<td>$C_p$</td>
</tr>
<tr>
<td>$t_7^* = [w_1 s_o^2 + w_2 (S_o^2 - s_o^2)] (\frac{S_o^2 + s_o^2}{S_o^2 + s_o^2 + 2s_o^2}) \exp[\frac{S_o^2 - s_o^2}{S_o^2 + s_o^2 + 2s_o^2}]$</td>
<td>1</td>
<td>$\beta_2(\phi)$</td>
<td>1</td>
<td>$\beta_2(\phi)$</td>
</tr>
<tr>
<td>$t_8^* = [w_1 s_o^2 + w_2 (S_o^2 - s_o^2)] (\frac{S_o^2 + s_o^2}{S_o^2 + s_o^2 + 2s_o^2}) \exp[\frac{S_o^2 - s_o^2}{S_o^2 + s_o^2 + 2s_o^2}]$</td>
<td>1</td>
<td>$C_p$</td>
<td>$\beta_2(\phi)$</td>
<td>$C_p$</td>
</tr>
<tr>
<td>$t_9^* = [w_1 s_o^2 + w_2 (S_o^2 - s_o^2)] (\frac{S_o^2 + s_o^2}{S_o^2 + s_o^2 + 2s_o^2}) \exp[\frac{S_o^2 - s_o^2}{S_o^2 + s_o^2 + 2s_o^2}]$</td>
<td>1</td>
<td>$\beta_2(\phi)$</td>
<td>1</td>
<td>$\beta_2(\phi)$</td>
</tr>
</tbody>
</table>

2.1. Bias and MSE of the proposed class of estimators $t$. To study the large sample properties of the suggested class of estimators $t$ in Eq. (2.1) of the population variance $S_o^2$ of the study variable $y$, we define

$s_o^2 = S_o^2(1 + e_0)$ and $S_o^2 = S_o^2(1 + e_1)$.

such that

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

and to the first degree of approximation, ignoring finite population correction ($fpc$) term (i.e. $(1 - f) \cong 1, f = (n/N) \cong 0$ ), we have

$E(e_0^2) = (1/n)(\lambda_{40} - 1), E(e_1^2) = (1/n)(\lambda_{41} - 1)$

and

$E(e_0e_1) = (1/n)(\lambda_{22} - 1),$

where $\lambda_{rs} = \frac{\mu_{rs}}{(\mu_{rs}^2)(\mu_{rs}^2)}$, $\mu_{rs} = \frac{1}{N - 1} \sum_{i=1}^{N} (y_i - \bar{Y})(x_i - \bar{X})^2$;

$(r, s)$ being non-negative integers.
Expressing Eq. (2.1) in terms of $e$’s we have

\[
(2.2) \quad t = S^2_w[w_1(1 + e_0) - w_2 r e_1](1 + \alpha \tau e_1)^{-\theta} \exp\left\{ -\frac{h \tau_0 e_1}{2} \left( 1 + \frac{\tau_0 e_1}{2} \right)^{-1} \right\},
\]

where \( \tau = \frac{a_s S^2_w}{(\alpha S^2_w + d)} \), \( \tau_0 = \frac{e_s S^2_w}{(\alpha S^2_w + d)} \), and \( r = \frac{S^2_w}{S^2_w} \).

We assume that \( |\alpha \tau e_1| < 1 \) so that \( (1 + \alpha \tau e_1)^{-\theta} \) is expandable.

Expanding the right hand side of Eq. (2.2) we have

\[
t = S^2_w[w_1(1 + e_0) - w_2 r e_1](1 - \alpha \gamma e_1 + \frac{\alpha (\gamma + 1)}{2} \alpha^2 \tau^2 e_1^2 - \ldots) \left( 1 - \frac{h \tau_0 e_1}{2} \left( 1 + \frac{\tau_0 e_1}{2} \right)^{-1} \right)
+ \frac{\alpha^2 \tau^2 e_1^2}{8} \left( 1 + \frac{\tau_0 e_1}{2} \right)^{-2} - \ldots
\]

\[
= S^2_w[w_1(1 + e_0) - w_2 r e_1](1 - \alpha \gamma e_1 + \frac{\alpha (\gamma + 1)}{2} \alpha^2 \tau^2 e_1^2 - \ldots) \left( 1 - \frac{h \tau_0 e_1}{2} \left( 1 + \frac{\tau_0 e_1}{2} \right)^{-1} \right)
+ \frac{\alpha^2 \tau^2 e_1^2}{8} \left( 1 + \frac{\tau_0 e_1}{2} \right)^{-2} - \ldots
\]

\[
= S^2_w[w_1(1 + e_0) - w_2 r e_1](1 - \alpha \gamma e_1 + \frac{\alpha (\gamma + 1)}{2} \alpha^2 \tau^2 e_1^2 - \ldots) \left( 1 - \frac{h \tau_0 e_1}{2} \left( 1 + \frac{\tau_0 e_1}{2} \right)^{-1} \right)
+ \frac{\alpha^2 \tau^2 e_1^2}{8} \left( 1 + \frac{\tau_0 e_1}{2} \right)^{-2} - \ldots
\]

\[
. \quad S^2_w[w_1(1 + e_0) - w_2 r e_1](1 - \alpha \gamma e_1 + \frac{\alpha (\gamma + 1)}{2} \alpha^2 \tau^2 e_1^2 - \ldots) \left( 1 - \frac{h \tau_0 e_1}{2} \left( 1 + \frac{\tau_0 e_1}{2} \right)^{-1} \right)
+ \frac{\alpha^2 \tau^2 e_1^2}{8} \left( 1 + \frac{\tau_0 e_1}{2} \right)^{-2} - \ldots
\]

Now, multiplying out and neglecting terms of $e$’s having power greater than two we have

\[
t \approx S^2_w \left[ w_1 \left( 1 - (\alpha \gamma + \frac{h \tau_0}{2}) e_1 + e_0 - (\alpha \gamma + \frac{h \tau_0}{2}) e_0 e_1 + (2(\alpha \gamma + \frac{h \tau_0}{2}))^2 + (2\alpha^2 \tau^2 + h \tau_0)^2 \right) \right.
- \left. w_2 r \left\{ e_1 - \left( \alpha \gamma + \frac{h \tau_0}{2} \right) e_1^2 \right\} \right]
\]

or

\[
(2.3) \quad (t - S^2_w) \approx S^2_w \left[ w_1 \left( 1 + e_0 - \theta e_1 - \theta e_0 e_1 + (2\theta^2 + 2\alpha^2 \tau^2 + h \tau_0^2) \left( \frac{e_1^2}{4} \right) \right) + w_2 r \left( \theta e_1^2 - e_1 - 1 \right) \right],
\]

where \( \theta = \left( \alpha \gamma + \frac{h \tau_0}{2} \right) \).

Taking expectation of both sides of Eq. (2.3) we get the bias of $t$ to the first degree of approximation as

\[
(2.4) \quad B(t) = S^2_w \left[ w_1 \left( 1 + \frac{1}{n} \left( 2\theta^2 + 2\alpha^2 \tau^2 + h \tau_0^2 \right) \left( \lambda_0^4 - 1 - \theta \lambda_2 - 1 \right) \right) + w_2 r \frac{\theta}{n} \left( \lambda_0^4 - 1 \right) \right].
\]

Squaring both sides of Eq. (2.3) and neglecting terms of $e$’s having power greater than two, we have

\[
(t - S^2_w)^2 \approx S^2_w \left[ w_1 \left( 1 + w_2 \left( 1 + e_0 - 2 \theta e_1 + e_1 \right) - 4 \theta e_0 e_1 + (2\theta^2 + 2\alpha^2 \tau^2 + h \tau_0^2) \right) e_1^2 \right] + w_2^2 r^2 e_1^2 + 2 w_1 w_2 r t
\]

\[
(2.5) \quad (2\theta e_1^2 - e_1 - e_0 e_1) - 2 w_2 r (\theta e_1^2 - e_1) - 2 w_1 \left( 1 + e_0 - \theta e_1 - \theta e_0 e_1 + (2\theta^2 + 2\alpha^2 \tau^2 + h \tau_0^2) \left( \frac{e_1^2}{4} \right) \right)
\]

Taking expectation of both sides of Eq. (2.5) we get the MSE of the estimator $t$ to the first degree of approximation as

\[
(2.6) \quad MSE(t) = S^2_w \left[ w_1^2 A_1 + w_2^2 A_2 + w_1 w_2 A_3 - 2 w_1 A_4 - 2 w_2 A_5 \right],
\]

where

\[
A_1 = \left[ 1 + \frac{1}{n} \left\{ (\lambda_0^4 - 1 - 4 \theta \lambda_2 - 1) \right\} + (2\theta^2 + 2\alpha^2 \tau^2 + h \tau_0^2) \right] \left( \lambda_0^4 - 1 \right),
A_2 = \frac{\theta}{n} \left( \lambda_0^4 - 1 \right),
A_3 = \frac{\theta}{n} \left( \lambda_0^4 - 1 - \lambda_2 - 1 \right),
A_4 = \left[ 1 + \frac{1}{n} \left\{ (2\theta^2 + 2\alpha^2 \tau^2 + h \tau_0^2) \right\} \lambda_0^4 - 1 - \theta \lambda_2 - 1 \right],
A_5 = \frac{\theta}{n} \left( \lambda_0^4 - 1 \right).
\]

Differentiating Eq. (2.6) partially with respect to $w_1$ and $w_2$ and equating to zero, we get
Thus, we established the following theorem.

**Theorem 1:** To the first degree of approximation, $MSE(t) \geq S_0^2(1 - A)$ with equality holding if

\[
\begin{align*}
  w_1 &= w_{10}, \\
  w_2 &= w_{20}.
\end{align*}
\]

**Special Case:** For $w_1 = 1$ and $w_2 = d$ in Eq. (2.1) we get the class of estimators for $S_0^2$ as

\[
(2.10) \quad t(1) = \left[ s_y^2 + d(S_0^2 - s_0^2) \right] \left[ \frac{aS_0^2 + b}{\alpha(aS_0^2 + b) + (1 - \alpha)(aS_0^2 + b)} \right]^{\frac{\alpha}{\alpha - \beta}} \exp \left[ \frac{hc(S_0^2 - s_0^2)}{c(S_0^2 + s_0^2) + 2d} \right]
\]

Putting $w_1 = 1$ and $w_2 = d$ in Eq. (2.4) and (2.6), we get the bias and MSE of $t(1)$ to the first degree of approximation, respectively as

\[
(2.11) \quad B(t(1)) = \frac{S_0^2}{n} \left[ \left( 2b^2 + 2a^2r^2 + h\tau_0 \right) \left( \frac{\lambda_{04} - 1}{4} \right) - \theta(\lambda_{22} - 1) + dr\theta(\lambda_{04} - 1) \right]
\]

and

\[
(2.12) \quad MSE(t_1) = S_0^2[1 + A_1 - 2A_4 + d^2A_2^2 - 2d(A_5 - A_3)].
\]

The $MSE(t_1)$ at Eq. (12) is minimum when $d = \frac{(A_3 - A_4)}{A_2}$

\[
(2.13) \quad = \left[ \frac{\rho^*}{\theta} \right] \sqrt{\left( \frac{\lambda_{04} - 1}{\lambda_{04} - 1} \right) - 1}) = d_0 \text{(say)}
\]

where

\[
\rho^* = \frac{(A_3 - A_4)}{(\lambda_{04} - 1)(\lambda_{04} - 1)}.
\]

Thus, the resulting minimum MSE of $t(1)$ is given by

\[
(2.14) \quad MSE(t_1) = S_0^2[1 + A_1 - 2A_4 - \frac{(A_5 - A_3)^2}{A_2}],
\]

\[
(2.15) \quad = \frac{S_0^2}{n} \left[ \left( \frac{\lambda_{04} - 1}{(1 - \rho^2)^2} \right) \right].
\]

The minimum $MSE$ of $t(1)$ Eq. (2.14) (or (2.15)) is equal to the minimum $MSE$ of the difference-type estimator $t_2 = [s_y^2 + bo(S_0^2 - s_0^2)]$

\[
(2.16) \quad = t_d \text{(say)}
\]

which is due to [19].

Now, we state the following corollary. **Corollary 1:** To the first degree of approximation,
Thus from Eq.(2.9) and (2.14) (or (3.4)), we have

\[ t \text{ estimator whose minimum MSE to the first degree of approximation is equal to the difference} \]

\[ \text{MSE}(t) \geq \frac{S_0^4[(\lambda_{40} - 1)(1 - \rho^2)]}{\lambda_3^2} \]

with equality holding if

\[ d = \left( \frac{S_0^2}{\lambda_3} \right) \sqrt{\frac{(\lambda_{40} - 1)}{(\lambda_{30} - 1)}} - 1 \].

3. Efficiency comparisons

To compare the proposed class of estimators \( t \) with usual unbiased estimator \( s_0^2 \), ratio-type estimator \( t_1 = s_0^2(S_0^2/s_0^2) \), exponential ratio-type estimator \( t_2 \) and difference type estimator \( t_3 = s_0^2 + \phi_0(S_0^2 - s_0^2) = t_d \) (say), we write the MSEs of \( s_0^2 \) and \( t_1 \), and the minimum mean squared error of \( t_2 \) to the first degree of approximation (ignoring finite population correction (fpc) term) respectively as

\[ \text{MSE}(s_0^2) = \left( \frac{S_0^4}{n} \right)(\lambda_{40} - 1), \]

\[ \text{MSE}(t_1) = \left( \frac{S_0^4}{n} \right)(\lambda_{40} - 1 + (\lambda_{30} - 1) - 2(\lambda_{22} - 1)), \]

\[ \text{MSE}(t_2) = \left( \frac{S_0^4}{n} \right)(\lambda_{40} - 1 + (1/4)(\lambda_{30} - 1) - (\lambda_{22} - 1)), \]

\[ \min.\text{MSE}(t_2 \text{ or } t_d) = \left( \frac{S_0^4}{n} \right)[(\lambda_{40} - 1) - (\lambda_{22} - 1)^2/(\lambda_{30} - 1)]. \]

From Eq. (3.1) and (3.4), we have

\[ \text{MSE}(s_0^2) - \min.\text{MSE}(t_d) = \left( \frac{S_0^4}{n} \right)(\lambda_{22} - 1)^2/(\lambda_{30} - 1), \]

\[ \text{MSE}(t_1) - \min.\text{MSE}(t_d) = \left( \frac{S_0^4}{n} \right)(\lambda_{30} - \lambda_{22})^2/(\lambda_{30} - 1). \]

\[ \text{MSE}(t_3) - \min.\text{MSE}(t_d) = \left( \frac{S_0^4}{n} \right)(\lambda_{30} - 2\lambda_{22} + 1)^2/(\lambda_{30} - 1). \]

It is clear that the expressions Eq.(3.5), (3.6) and (3.7) are positive. Therefore, the difference-type estimator \( t_d \) is more efficient than the usual unbiased estimator \( s_0^2 \), ratio-type estimator \( t_1 \) and exponential ratio-type estimator \( t_3 \).

Now, we compare the proposed class of estimators \( t \) with the proposed estimator \( t_{(1)} \) whose minimum MSE to the first degree of approximation is equal to the difference estimator \( t_d \).

Thus from Eq.(2.9) and (2.14) (or (3.4)), we have

\[ \min.\text{MSE}(t_{(1)}) \text{ or } t_d = \frac{S_0^4[A_2(A_1 - A_4) - A_3(A_3 - A_5)]^2}{A_2(A_1A_2 - A_5^2)} \]

which is always positive. It follows from Eq.(3.5) to (3.8) that the proposed class of estimators \( t' \) is better than the usual unbiased estimator \( s_0^2 \), ratio-type estimator \( t_1 \), exponential ratio-type estimator \( t_3 \), the proposed estimator \( t_{(1)} \) and its subclasses of estimators \( t_i \) (i - 3 to 23) listed in Table 1 and the difference estimator \( t_d \) (ort 2). [19] suggested another class of estimators for the population variance \( S_0^2 \) as

\[ t_M = s_0^2[m_1 + m_2(s_0^2 - s_0^2)]\exp[-\frac{h_c(S_0^2 - s_0^2)}{c(S_0^2 + s_0^2) + 2d}], \]

where \( c, d \) and \( h \) are same as defined earlier and \( (m_1 \text{ and } m_2) \) are suitable chosen constants such that MSE of \( t_M \) is minimum.

To the first degree of approximation (ignoring fpc term), the MSE of \( t_M \) is given by

\[ \text{MSE}(t_M) = S_0^4[1 + m_1^2R_1 + m_2^2R_2 + 2m_1m_2R_3 - 2m_1R_4 - 2m_2R_5] \]
which is minimized for

\[
(3.11) \begin{cases}
    m_1 = \frac{(R_2 R_4 - R_2 R_3)}{(R_1 R_2 - R_3^2)}, \\
    m_2 = \frac{(R_1 R_6 - R_3 R_4)}{(R_1 R_2 - R_3^2)},
\end{cases}
\]

where

\[
R_1 = \left[ 1 + \frac{1}{n} \left\{ (\lambda_0 - 1) - 2h \tau_0 (\lambda_22 - 1) + \frac{h(h+1)}{4} \tau_0^2 (\lambda_04 - 1) \right\} \right],
\]

\[
R_2 = \frac{s_0^4}{\pi} (\lambda_04 - 1),
\]

\[
R_3 = \frac{s_0^4}{\pi} [h \tau_0 (\lambda_04 - 1) - 2(\lambda_22 - 1)],
\]

\[
R_4 = \left[ 1 + \frac{\lambda \tau_0}{\pi} \right] \frac{\frac{h(\lambda+2)}{4} (\lambda_04 - 1) - (\lambda_22 - 1)}{\lambda_04 - 1} - (\lambda_22 - 1)],
\]

\[
R_5 = \frac{s_0^4}{\pi} [: h \tau_0 (\lambda_04 - 1) - (\lambda_22 - 1)],
\]

\[
\tau_0 = \frac{\phi_2}{\phi_2^2 + \rho_2^2}.
\]

Putting Eq. (3.11) in (3.10), we get the minimum MSE of \( t_M \) is given by

\[
(3.12) \quad \min.MSE(t_M) = S_0^4 (1 - R),
\]

where

\[
R = \frac{(R_2 R_4 - R_2 R_3)}{(R_1 R_2 - R_3^2)}.
\]

From Eq. (2.9) and (3.12), we have

\[
(3.13) \quad \min.MSE(t_M) - \min.MSE(t) = S_0^4 (A - R),
\]

which is positive if

\[
(3.14) \quad R < A,
\]

where \( A \) and \( R \) respectively defined for Eq. (2.9) and (3.12).

Thus, the proposed class of estimators \( t \) is more efficient than the class of estimators \( t_M \) due to [19] as long as the condition Eq. (3.14) is satisfied. Note that if \( R > A \) then the above conclusion will be just the opposite and this is also possible.

4. Empirical Study

To judge the merits of the proposed class of estimators \( t \) over other existing estimators, we have computed the percent relative efficiencies (PREs) of the members of the suggested class of estimators with respect to the usual unbiased estimator \( s_0^2 \) by using the formula:

\[
(4.1) \quad PRE(t, s_0^2) = \frac{MSE(s_0^2)}{\min.MSE(t)} \times 100.
\]

For this we use the same data used by [19] given in ([21], p. 256). The description of the study variable, auxiliary attribute \( \phi \) and required parameters are given below:

\[
y: \quad \text{is the number of villages in the circle,}
\]

\[
\phi: \quad \text{represents a circle consisting more than five villages;}
\]

\[
N = 89, n = 23, S_0^2 = 4.074, S_0^2 = -0.110, \quad C_y = -0.601, \quad C_p = 2.678, \quad \rho_p = -0.766,
\]

\[
\beta_2(\phi) = -6.162, \quad \lambda_22 = -3.996, \quad p_c = -3.811, \quad \lambda_04 = -6.162.
\]

We have computed the PREs of the existing estimators and various members of the proposed class of estimators \( t \) with respect to the usual unbiased estimator \( s_0^2 \) and findings are given in Table 3 and Table 4.

We note that the empirical studies carried out in [19] suggest that the performance of the subclass of estimators:

\[
(4.2) \quad t_M = s_0^2 [m_1 + m_2 (S_0^2 - s_0^2)] \exp \left( \frac{(S_0^2 - s_0^2)}{(S_0^2 + s_0^2)} \right)
\]
Table 3. The PREs of various existing estimators of the population variance $s_y^2$ with respect to the usual unbiased estimator $s_y^2$.

<table>
<thead>
<tr>
<th>Estimator</th>
<th>PRE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_0$</td>
<td>100.000</td>
</tr>
<tr>
<td>$t_1$</td>
<td>141.898</td>
</tr>
<tr>
<td>$t_3$</td>
<td>254.270</td>
</tr>
<tr>
<td>$t_4$</td>
<td>108.834</td>
</tr>
<tr>
<td>$t_5$</td>
<td>103.823</td>
</tr>
<tr>
<td>$t_6$</td>
<td>155.191</td>
</tr>
<tr>
<td>$t_7$</td>
<td>110.306</td>
</tr>
<tr>
<td>$t_8$</td>
<td>91.351</td>
</tr>
<tr>
<td>$t_9$</td>
<td>96.214</td>
</tr>
<tr>
<td>$t_{10}$</td>
<td>51.726</td>
</tr>
<tr>
<td>$t_{11}$</td>
<td>89.956</td>
</tr>
<tr>
<td>$t_{12}$</td>
<td>262.187</td>
</tr>
<tr>
<td>$t_{13}$</td>
<td>284.570</td>
</tr>
</tbody>
</table>

Table 4. The PREs of the members of the proposed class of estimators $t$ of the population variance $s_y^2$ with respect to the usual unbiased estimator $s_y^2$.

<table>
<thead>
<tr>
<th>Estimator</th>
<th>PRE</th>
<th>RANK</th>
<th>Estimator</th>
<th>PRE</th>
<th>RANK</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_{1}^*$</td>
<td>356.62</td>
<td>XIX</td>
<td>$t_{11}^*$</td>
<td>13011.54</td>
<td>I</td>
</tr>
<tr>
<td>$t_{2}^*$</td>
<td>3253.05</td>
<td>X</td>
<td>$t_{12}^*$</td>
<td>2865.64</td>
<td>XV</td>
</tr>
<tr>
<td>$t_{3}^*$</td>
<td>2856.35</td>
<td>XVI</td>
<td>$t_{13}^*$</td>
<td>5490.66</td>
<td>V</td>
</tr>
<tr>
<td>$t_{4}^*$</td>
<td>5289.32</td>
<td>VII</td>
<td>$t_{14}^*$</td>
<td>5368.64</td>
<td>VI</td>
</tr>
<tr>
<td>$t_{5}^*$</td>
<td>3375.65</td>
<td>IX</td>
<td>$t_{15}^*$</td>
<td>5879.30</td>
<td>IV</td>
</tr>
<tr>
<td>$t_{6}^*$</td>
<td>6257.46</td>
<td>II</td>
<td>$t_{16}^*$</td>
<td>6184.06</td>
<td>III</td>
</tr>
<tr>
<td>$t_{7}^*$</td>
<td>4798.68</td>
<td>VIIX</td>
<td>$t_{17}^*$</td>
<td>3003.95</td>
<td>XII</td>
</tr>
<tr>
<td>$t_{8}^*$</td>
<td>3083.84</td>
<td>XI</td>
<td>$t_{18}^*$</td>
<td>231.79</td>
<td>XIV</td>
</tr>
<tr>
<td>$t_{9}^*$</td>
<td>2789.78</td>
<td>XVI</td>
<td>$t_{19}^*$</td>
<td>2976.46</td>
<td>XIII</td>
</tr>
<tr>
<td>$t_{10}^*$</td>
<td>1517.02</td>
<td>XVIII</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

is better than the rest of the estimators discussed in [19] which may not be true in general. Therefore, we have tried to find out the members of the proposed class of estimators $t$ which are better than $t_{M1}$ and such estimators are listed in Table 4. The estimator is a member of the [19] class of estimators which can be obtained just by putting $(h, c, d) = (1, 1, 0)$ in Eq. (3.9).

It is observed from Tables 3 and 4 that the performance of the estimators $t_{j}^*$ ($j = -1, 2, ..., 19$), which are members of the proposed class of estimators $t$, are better than the estimators $t_i$ ($i = -1, 2, ..., 24$) and $t_{M1}$ due to [16], [19] and [23] with larger gain in efficiency. It is interesting to note that the member $t_{11}^*$ of the proposed class of estimator $t$ is the best with largest $PRE = 13011.54$ among the entire estimators discussed in Tables 3 and 4. Thus the estimator $t_{j}^*$ ($j = -1, 2, ..., 19$) tabled in Table 2 belonging to the proposed class of estimators $t$ are to be preferred over existing estimators listed in Table 1. However this conclusion should not be extrapolated in general, due to limited empirical study.
5. Conclusion

We have developed a new class of estimators for estimating the population variance $S^2_y$ of the study variable $y$ in the presence of auxiliary attribute $\phi$. It is shown that in addition to [16], [19] and [23] estimators, many other estimators as indicated in Table 2 are members of the proposed class of estimators. We have obtained the bias and MSE of the proposed class of estimators $t$. It has been shown theoretically as well as empirically that the suggested class of estimators $t$ is better than usual unbiased estimator $t_0 = s^2_y$, usual ratio estimator $t_1$, ratio-type exponential estimator $t_3$, difference-type estimator $t_d$ and various other estimators listed in Table 1.

In the Table 4, we have computed the percent relative efficiencies of the members $t^*_j$ ($j = 1, 2, \ldots, 19)$ of the proposed class of estimators $t$ with respect to the usual unbiased estimator $s^2_y$ for different values of parameters. It is observed that the estimator $t^*_1$ has the largest PRE = 13011.54 which is very large as compared to the estimator $t_{M1}$ (whose PRE = 284.57) reported by [19]. Thus there is enough scope of obtaining estimators from the proposed class of estimators $t$ better than the existing estimators for suitable values of the scalars involved in the suggested class of estimators $t$. The major advantage of the proposed class of estimators $t$ is that it unifies several results at one place. Thus the proposal of such a class of estimators is justified.

Acknowledgements Authors are thankful to the referees for their valuable suggestions regarding improvement of the paper.

References


