

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_y^2 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_y^2 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 ϕ which is highly correlated with study variable y . For example (i) y may be the use of drugs and ϕ may be the gender, (ii) y may be production of a crop and ϕ may be the particular variety, (iii) y may be the amount of milk produced and ϕ may be a particular breed of cow, (iv) y may be the yield of wheat crop and ϕ 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 ϕ , 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 ϕ . [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 $U = (U_1, U_2, \dots, U_N)$ of size N . Let y_i and ϕ_i , ($i = 1, 2, \dots, N$) be the observations on the study variable y and the auxiliary attribute ϕ respectively.

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

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 ϕ . 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 ϕ , 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 ρ_{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_ϕ^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 ϕ such as $S_\phi^2, C_p, \beta_1(\phi), \beta_2(\phi)$ and ρ_{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

$$(2.1) \quad t = [w_1 s_y^2 + w_2 (S_\phi^2 - s_\phi^2)] \left[\frac{aS_\phi^2 + b}{\alpha(as_\phi^2 + b) + (1-\alpha)(aS_\phi^2 + b)} \right]^g \exp \left[\frac{hc(S_\phi^2 - s_\phi^2)}{c(S_\phi^2 + s_\phi^2) + 2d} \right],$$

where α 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 ϕ such as $C_P, \beta_1(\phi), \beta_2(\phi), \Delta, \rho_{pb}, k_{pb} = \rho_{pb}(C_P/C_y)$, where $C_y = (S_y/\bar{Y})$ 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 $(w_1, w_2, g, h, \alpha, a, b, c, d)$. Some unknown members of the suggested class of estimators of the population variance S_y^2 are tabled in Table 2.

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

Estimator (Est.)	w_1	w_2	g	h	α	a	b	c	d
$t_1 = s_y^2 (\frac{S_\phi^2}{s_\phi^2})$ [16]Est.	1	0	1	0	1	1	0	-	-
$t_2 = s_y^2 + b_\phi(S_\phi^2 - s_\phi^2)$ [16]Est.	1	b_ϕ	0	0	-	-	-	-	-
$t_3 = s_y^2 \exp(\frac{S_\phi^2 - s_\phi^2}{S_\phi^2 + s_\phi^2})$ [16]Est.	1	0	0	1	-	-	-	1	0
$t_4 = s_y^2 (\frac{S_\phi^2 + C_P}{s_\phi^2 + C_P})$ [19]Est.	1	0	1	0	1	1	C_P	-	-
$t_5 = s_y^2 (\frac{S_\phi^2 + \beta_2(\phi)}{s_\phi^2 + \beta_2(\phi)})$ [19]Est.	1	0	1	0	1	1	$\beta_2(\phi)$	-	-
$t_6 = s_y^2 (\frac{S_\phi^2 \beta_2(\phi) + C_P}{s_\phi^2 \beta_2(\phi) + C_P})$ [19]Est.	1	0	1	0	1	$\beta_2(\phi)$	C_P	-	-
$t_7 = s_y^2 (\frac{S_\phi^2 C_P + \beta_2(\phi)}{s_\phi^2 C_P + \beta_2(\phi)})$ [19]Est.	1	0	1	0	1	C_P	$\beta_2(\phi)$	-	-
$t_8 = [s_y^2 + b_\phi(S_\phi^2 - s_\phi^2)](\frac{S_\phi^2}{s_\phi^2})$ [19]Est.	1	b_ϕ	1	0	1	1	0	-	-
$t_9 = [s_y^2 + b_\phi(S_\phi^2 - s_\phi^2)](\frac{S_\phi^2 + \beta_2(\phi)}{s_\phi^2 + \beta_2(\phi)})$ [19]Est.	1	b_ϕ	1	0	1	1	$\beta_2(\phi)$	-	-
$t_{10} = [s_y^2 + b_\phi(S_\phi^2 - s_\phi^2)](\frac{S_\phi^2 + C_P}{s_\phi^2 + C_P})$ [19]Est.	1	b_ϕ	1	0	1	1	C_P	-	-
$t_{11} = [s_y^2 + b_\phi(S_\phi^2 - s_\phi^2)](\frac{S_\phi^2 + \rho_{pb}}{s_\phi^2 + \rho_{pb}})$ [19]Est.	1	b_ϕ	1	0	1	1	ρ_{pb}	-	-
$t_{12} = [s_y^2 + b_\phi(S_\phi^2 - s_\phi^2)](\frac{S_\phi^2 \beta_2(\phi) + C_P}{s_\phi^2 \beta_2(\phi) + C_P})$ [19]Est.	1	b_ϕ	1	0	1	$\beta_2(\phi)$	C_P	-	-
$t_{13} = [s_y^2 + b_\phi(S_\phi^2 - s_\phi^2)](\frac{S_\phi^2 C_P + \beta_2(\phi)}{s_\phi^2 C_P + \beta_2(\phi)})$ [19]Est.	1	b_ϕ	1	0	1	C_P	$\beta_2(\phi)$	-	-
$t_{14} = [s_y^2 + b_\phi(S_\phi^2 - s_\phi^2)](\frac{S_\phi^2 \rho_{pb} + C_P}{s_\phi^2 \rho_{pb} + C_P})$ [19]Est.	1	b_ϕ	1	0	1	ρ_{pb}	C_P	-	-
$t_{15} = [s_y^2 + b_\phi(S_\phi^2 - s_\phi^2)](\frac{S_\phi^2 \beta_2(\phi) + \rho_{pb}}{s_\phi^2 \beta_2(\phi) + \rho_{pb}})$ [19]Est.	1	b_ϕ	1	0	1	$\beta_2(\phi)$	ρ_{pb}	-	-
$t_{16} = [s_y^2 + b_\phi(S_\phi^2 - s_\phi^2)](\frac{S_\phi^2 \rho_{pb} + \beta_2(\phi)}{s_\phi^2 \rho_{pb} + \beta_2(\phi)})$ [19]Est.	1	b_ϕ	1	0	1	ρ_{pb}	$\beta_2(\phi)$	-	-
$t_{17} = s_y^2 (\frac{S_\phi^2 - C_P}{s_\phi^2 - C_P})$ [19]Est.	1	0	1	0	1	1	$-C_P$	-	-
$t_{18} = s_y^2 (\frac{S_\phi^2 - \beta_2(\phi)}{s_\phi^2 - \beta_2(\phi)})$ [19]Est.	1	0	1	0	1	1	$-\beta_2(\phi)$	-	-
$t_{19} = s_y^2 (\frac{S_\phi^2 \beta_2(\phi) - C_P}{s_\phi^2 \beta_2(\phi) - C_P})$ [18]Est.	1	0	1	0	1	$\beta_2(\phi)$	$-C_P$	-	-
$t_{20} = s_y^2 (\frac{S_\phi^2 C_P - \beta_2(\phi)}{s_\phi^2 C_P - \beta_2(\phi)})$ [19]Est.	1	0	1	0	1	C_P	$-\beta_2(\phi)$	-	-
$t_{21} = [s_y^2 + b_\phi(S_\phi^2 - s_\phi^2)](\frac{aS_\phi^2 + b}{as_\phi^2 + b})$ [19]Est.	1	b_ϕ	1	0	1	a	b	-	-
$t_{22} = s_y^2 (\frac{\eta S_\phi^2 - \vartheta}{\alpha(\eta s_\phi^2 - \vartheta) + (1-\alpha)(\eta S_\phi^2 - \vartheta)})$ [19]or[23]Est.	1	0	1	0	α	η	$-\vartheta$	-	-
$t_{23} = w_1 s_y^2 (\frac{\eta S_\phi^2 - \vartheta}{\alpha(\eta s_\phi^2 - \vartheta) + (1-\alpha)(\eta S_\phi^2 - \vartheta)})$ [23]Est.	w_1	0	1	0	α	η	$-\vartheta$	-	-
$t_{24} = t_0 = s_y^2$ Usual unbiased Est.	1	0	0	0	-	-	-	-	-

where b_ϕ and (η, ϑ) being constants.

Table 2. Some unknown members of the proposed class of estimators t for $(g, h, \alpha)=(1,1,1)$.

Estimator	a	b	c	d
$t_1^* = [w_1s_y^2 + w_2(S_\phi^2 - s_\phi^2)](\frac{S_\phi^2}{s_\phi^2}) \exp[\frac{S_\phi^2 - s_\phi^2}{S_\phi^2 + s_\phi^2}]$	1	0	1	0
$t_2^* = [w_1s_y^2 + w_2(S_\phi^2 - s_\phi^2)](\frac{S_\phi^2 + C_p}{s_\phi^2 + C_p}) \exp[\frac{S_\phi^2 - s_\phi^2}{S_\phi^2 + s_\phi^2 + 2C_p}]$	1	C_p	1	C_p
$t_3^* = [w_1s_y^2 + w_2(S_\phi^2 - s_\phi^2)](\frac{S_\phi^2 + \beta_2(\phi)}{s_\phi^2 + \beta_2(\phi)}) \exp[\frac{S_\phi^2 - s_\phi^2}{S_\phi^2 + s_\phi^2 + 2\beta_2(\phi)}]$	1	$\beta_2(\phi)$	1	$\beta_2(\phi)$
$t_4^* = [w_1s_y^2 + w_2(S_\phi^2 - s_\phi^2)](\frac{S_\phi^2 + \rho_{pb}}{s_\phi^2 + \rho_{pb}}) \exp[\frac{S_\phi^2 - s_\phi^2}{S_\phi^2 + s_\phi^2 + 2\rho_{pb}}]$	1	ρ_{pb}	1	ρ_{pb}
$t_5^* = [w_1s_y^2 + w_2(S_\phi^2 - s_\phi^2)](\frac{S_\phi^2 C_p + \beta_2(\phi)}{s_\phi^2 C_p + \beta_2(\phi)}) \exp[\frac{C_p(S_\phi^2 - s_\phi^2)}{C_p(S_\phi^2 + s_\phi^2) + 2\beta_2(\phi)}]$	C_p	$\beta_2(\phi)$	C_p	$\beta_2(\phi)$
$t_6^* = [w_1s_y^2 + w_2(S_\phi^2 - s_\phi^2)](\frac{S_\phi^2 \beta_2(\phi) + C_p}{s_\phi^2 \beta_2(\phi) + C_p}) \exp[\frac{\beta_2(\phi)(S_\phi^2 - s_\phi^2)}{\beta_2(\phi)(S_\phi^2 + s_\phi^2) + 2C_p}]$	$\beta_2(\phi)$	C_p	$\beta_2(\phi)$	C_p
$t_7^* = [w_1s_y^2 + w_2(S_\phi^2 - s_\phi^2)](\frac{S_\phi^2 C_p + \rho_{pb}}{s_\phi^2 C_p + \rho_{pb}}) \exp[\frac{C_p(S_\phi^2 - s_\phi^2)}{C_p(S_\phi^2 + s_\phi^2) + 2\rho_{pb}}]$	C_p	ρ_{pb}	C_p	ρ_{pb}
$t_8^* = [w_1s_y^2 + w_2(S_\phi^2 - s_\phi^2)](\frac{S_\phi^2 \rho_{pb} + C_p}{s_\phi^2 \rho_{pb} + C_p}) \exp[\frac{\rho_{pb}(S_\phi^2 - s_\phi^2)}{\rho_{pb}(S_\phi^2 + s_\phi^2) + 2C_p}]$	ρ_{pb}	C_p	ρ_{pb}	C_p
$t_9^* = [w_1s_y^2 + w_2(S_\phi^2 - s_\phi^2)](\frac{S_\phi^2 \rho_{pb} + \beta_2(\phi)}{s_\phi^2 \rho_{pb} + \beta_2(\phi)}) \exp[\frac{\rho_{pb}(S_\phi^2 - s_\phi^2)}{\rho_{pb}(S_\phi^2 + s_\phi^2) + 2\beta_2(\phi)}]$	ρ_{pb}	$\beta_2(\phi)$	ρ_{pb}	$\beta_2(\phi)$
$t_{10}^* = [w_1s_y^2 + w_2(S_\phi^2 - s_\phi^2)](\frac{S_\phi^2 \beta_2(\phi) + \rho_{pb}}{s_\phi^2 \beta_2(\phi) + \rho_{pb}}) \exp[\frac{\beta_2(\phi)(S_\phi^2 - s_\phi^2)}{\beta_2(\phi)(S_\phi^2 + s_\phi^2) + 2\rho_{pb}}]$	$\beta_2(\phi)$	ρ_{pb}	$\beta_2(\phi)$	ρ_{pb}
$t_{11}^* = [w_1s_y^2 + w_2(S_\phi^2 - s_\phi^2)](\frac{S_\phi^2 + C_p}{s_\phi^2 + C_p}) \exp[\frac{\beta_2(\phi)(S_\phi^2 - s_\phi^2)}{\beta_2(\phi)(S_\phi^2 + s_\phi^2) + 2\rho_{pb}}]$	1	C_p	$\beta_2(\phi)$	ρ_{pb}
$t_{12}^* = [w_1s_y^2 + w_2(S_\phi^2 - s_\phi^2)](\frac{S_\phi^2 \beta_2(\phi) + \rho_{pb}}{s_\phi^2 \beta_2(\phi) + \rho_{pb}}) \exp[\frac{S_\phi^2 - s_\phi^2}{S_\phi^2 + s_\phi^2 + 2C_p}]$	$\beta_2(\phi)$	ρ_{pb}	1	C_p
$t_{13}^* = [w_1s_y^2 + w_2(S_\phi^2 - s_\phi^2)](\frac{S_\phi^2 + \beta_2(\phi)}{s_\phi^2 + \beta_2(\phi)}) \exp[\frac{C_p(S_\phi^2 - s_\phi^2)}{C_p(S_\phi^2 + s_\phi^2) + 2\rho_{pb}}]$	1	$\beta_2(\phi)$	C_p	ρ_{pb}
$t_{14}^* = [w_1s_y^2 + w_2(S_\phi^2 - s_\phi^2)](\frac{S_\phi^2 C_p + \rho_{pb}}{s_\phi^2 C_p + \rho_{pb}}) \exp[\frac{S_\phi^2 - s_\phi^2}{S_\phi^2 + s_\phi^2 + 2\beta_2(\phi)}]$	C_p	ρ_{pb}	1	$\beta_2(\phi)$
$t_{15}^* = [w_1s_y^2 + w_2(S_\phi^2 - s_\phi^2)](\frac{S_\phi^2 \beta_2(\phi) + C_p}{s_\phi^2 \beta_2(\phi) + C_p}) \exp[\frac{S_\phi^2 - s_\phi^2}{S_\phi^2 + s_\phi^2 + 2\rho_{pb}}]$	$\beta_2(\phi)$	C_p	1	ρ_{pb}
$t_{16}^* = [w_1s_y^2 + w_2(S_\phi^2 - s_\phi^2)](\frac{S_\phi^2 + \rho_{pb}}{s_\phi^2 + \rho_{pb}}) \exp[\frac{S_\phi^2 - s_\phi^2}{S_\phi^2 + s_\phi^2 + 2C_p}]$	1	ρ_{pb}	1	C_p
$t_{17}^* = [w_1s_y^2 + w_2(S_\phi^2 - s_\phi^2)](\frac{S_\phi^2 \rho_{pb} + C_p}{s_\phi^2 \rho_{pb} + C_p}) \exp[\frac{S_\phi^2 - s_\phi^2}{S_\phi^2 + s_\phi^2 + 2\beta_2(\phi)}]$	ρ_{pb}	C_p	1	$\beta_2(\phi)$
$t_{18}^* = [w_1s_y^2 + w_2(S_\phi^2 - s_\phi^2)](\frac{S_\phi^2 + \beta_2(\phi)}{s_\phi^2 + \beta_2(\phi)}) \exp[\frac{\rho_{pb}(S_\phi^2 - s_\phi^2)}{\rho_{pb}(S_\phi^2 + s_\phi^2) + 2C_p}]$	1	$\beta_2(\phi)$	ρ_{pb}	C_p
$t_{19}^* = [w_1s_y^2 + w_2(S_\phi^2 - s_\phi^2)](\frac{S_\phi^2 \rho_{pb} + \beta_2(\phi)}{s_\phi^2 \rho_{pb} + \beta_2(\phi)}) \exp[\frac{C_p(S_\phi^2 - s_\phi^2)}{C_p(S_\phi^2 + s_\phi^2) + 2\beta_2(\phi)}]$	ρ_{pb}	$\beta_2(\phi)$	C_p	$\beta_2(\phi)$

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_y^2 of the study variable y , we define

$$s_y^2 = S_y^2(1 + e_0) \text{ and } s_\phi^2 = S_\phi^2(1 + e_1).$$

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

$$E(e_0^2) = (1/n)(\lambda_{40} - 1), E(e_1^2) = (1/n)(\lambda_{04} - 1)$$

and

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

where $\lambda_{rs} = \frac{\mu_{rs}}{\binom{r/2}{\mu_{20}} \binom{s/2}{\mu_{02}}}$, $\mu_{rs} = \frac{1}{N-1} \sum_{i=1}^N (y_i - \bar{Y})^r (x_i - P)^s$;

(r, s) being non- negative integers.

Expressing Eq.(2.1) in terms of e 's we have

$$(2.2) \quad t = S_y^2[w_1(1 + e_0) - w_2re_1](1 + \alpha\tau e_1)^{-g} \exp\left\{\frac{-h\tau_0e_1}{2} (1 + \frac{\tau_0e_1}{2})^{-1}\right\},$$

where $\tau = \frac{aS_\phi^2}{(aS_\phi^2+b)}$, $\tau_0 = \frac{cS_\phi^2}{(cS_\phi^2+d)}$ and $r = \frac{S_\phi^2}{S_y^2}$.

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

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

$$\begin{aligned} t &= S_y^2[w_1(1 + e_0) - w_2re_1](1 - \alpha g\tau e_1 + \frac{g(g+1)}{2}\alpha^2\tau^2e_1^2 - \dots)\{1 - \frac{h\tau_0e_1}{2} (1 + \frac{\tau_0e_1}{2})^{-1} \\ &+ \frac{h^2\tau_0^2e_1^2}{8} (1 + \frac{\tau_0e_1}{2})^{-2} - \dots\} \\ &= S_y^2[w_1(1 + e_0) - w_2re_1](1 - \alpha g\tau e_1 + \frac{g(g+1)}{2}\alpha^2\tau^2e_1^2 - \dots)\{1 - \frac{h\tau_0e_1}{2} (1 - \frac{\tau_0e_1}{2} + \dots) + \\ &\frac{h^2\tau_0^2e_1^2}{8} - \dots\} \\ &= S_y^2[w_1(1 + e_0) - w_2re_1](1 - \alpha g\tau e_1 + \frac{g(g+1)}{2}\alpha^2\tau^2e_1^2 - \dots)\{1 - \frac{h\tau_0e_1}{2} + \frac{h\tau_0^2e_1^2}{4} + \frac{h^2\tau_0^2e_1^2}{8} - \dots\} \\ &= S_y^2[w_1(1 + e_0) - w_2re_1](1 - \alpha g\tau e_1 + \frac{g(g+1)}{2}\alpha^2\tau^2e_1^2 - \dots)\{1 - \frac{h\tau_0e_1}{2} + \frac{h(h+2)\tau_0^2e_1^2}{8} - \dots\} \\ &= S_y^2[w_1(1 + e_0) - w_2re_1]\{1 - \alpha g\tau e_1 + \frac{g(g+1)}{2}\alpha^2\tau^2e_1^2 - \frac{h\tau_0e_1}{2} + \frac{\alpha gh\tau\tau_0e_1^2}{2} + \frac{h(h+2)\tau_0^2e_1^2}{8} + \dots\} \\ &= S_y^2[w_1(1 + e_0) - w_2re_1]\{1 - (\alpha g\tau + \frac{h\tau_0}{2})e_1 + \{2(\alpha g\tau + \frac{h\tau_0}{2})^2 + (2g\alpha^2\tau^2 + h\tau_0^2)\}\frac{e_1^2}{4} - \dots\}. \end{aligned}$$

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

$$t \cong S_y^2 [w_1\{1 - (\alpha g\tau + \frac{h\tau_0}{2})e_1 + e_0 - (\alpha g\tau + \frac{h\tau_0}{2})e_0e_1 + (2(\alpha g\tau + \frac{h\tau_0}{2})^2 + (2g\alpha^2\tau^2 + h\tau_0^2))\frac{e_1^2}{4}\} - w_2r\{e_1 - (\alpha g\tau + \frac{h\tau_0}{2})e_1^2\}]$$

or

$$(2.3) \quad (t - S_y^2) \cong S_y^2 [w_1\{1 + e_0 - \theta e_1 - \theta e_0e_1 + (2\theta^2 + 2g\alpha^2\tau^2 + h\tau_0^2)(\frac{e_1^2}{4})\} + w_2r\{(\theta e_1^2 - e_1) - 1\}],$$

where $\theta = (\alpha g\tau + \frac{h\tau_0}{2})$.

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

$$(2.4) \quad B(t) = S_y^2[w_1\{1 + \frac{1}{n}[\frac{1}{4}(2\theta^2 + 2g\alpha^2\tau^2 + h\tau_0^2)(\lambda_{04} - 1) - \theta(\lambda_{22} - 1)]\} + w_2(\frac{r\theta}{n})(\lambda_{04} - 1) - 1].$$

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

$$(t - S_y^2)^2 = S_y^4[1 + w_1^2\{1 + 2e_0 - 2\theta e_1 + e_0^2 - 4\theta e_0e_1 + \{2\theta^2 + g\alpha^2\tau^2 + (\frac{h\tau_0^2}{2})\}e_1^2\} + w_2^2r^2e_1^2 + 2w_1w_2r$$

$$(2.5) \quad (2\theta e_1^2 - e_1 - e_0e_1) - 2w_2r(\theta e_1^2 - e_1) - 2w_1\{1 + e_0 - \theta e_1 - \theta e_0e_1 + (2\theta^2 + 2g\alpha^2\tau^2 + h\tau_0^2)(\frac{e_1^2}{4})\}]$$

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

$$(2.6) \quad MSE(t) = S_y^4[1 + w_1^2A_1 + w_2^2A_2 + 2w_1w_2A_3 - 2w_1A_4 - 2w_2A_5],$$

where

$$A_1 = [1 + \frac{1}{n}\{(\lambda_{40} - 1) - 4\theta(\lambda_{22} - 1) + (2\theta^2 + g\alpha^2\tau^2 + \frac{h\tau_0^2}{2})(\lambda_{04} - 1)\}],$$

$$A_2 = \frac{r^2(\lambda_{04} - 1)}{n},$$

$$A_3 = \frac{r}{n}\{2\theta(\lambda_{04} - 1) - (\lambda_{22} - 1)\},$$

$$A_4 = [1 + \frac{1}{n}\{\frac{1}{4}(2\theta^2 + 2g\alpha^2\tau^2 + h\tau_0^2)(\lambda_{04} - 1) - \theta(\lambda_{22} - 1)\}],$$

$$A_5 = \frac{\theta r(\lambda_{04} - 1)}{n}.$$

Differentiating Eq.(2.6) partially with respect to w_1 and w_2 and equating to zero, we get

the equation:

$$(2.7) \quad \begin{bmatrix} A_1 & A_3 \\ A_3 & A_2 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} A_4 \\ A_5 \end{bmatrix}.$$

Solving (2.7) we get the optimum values of w_1 and w_2 respectively as

$$(2.8) \quad \begin{cases} w_1 = \frac{(A_2 A_4 - A_3 A_5)}{(A_1 A_2 - A_3^2)} = w_{10}(\text{say}), \\ w_2 = \frac{(A_1 A_5 - A_3 A_4)}{(A_1 A_2 - A_3^2)} = w_{20}(\text{say}). \end{cases}$$

Putting Eq. (2.8) in (2.6) we get the minimum MSE of 't' as

$$(2.9) \quad \min.MSE(t) = S_y^4(1 - A),$$

$$\text{where } A = \frac{(A_2 A_4^2 - 2A_3 A_4 A_5 + A_1 A_5^2)}{(A_1 A_2 - A_3^2)}.$$

Thus, we established the following theorem.

Theorem 1: To the first degree of approximation,
 $MSE(t) \geq S_y^4(1 - A)$

with equality holding if

$$\begin{cases} w_1 = w_{10}, \\ w_2 = w_{20}. \end{cases}$$

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

$$(2.10) \quad t_{(1)} = [s_y^2 + d(S_\phi^2 - s_\phi^2)] \left[\frac{aS_\phi^2 + b}{\{\alpha(as_\phi^2 + b) + (1 - \alpha)(aS_\phi^2 + b)\}} \right]^g \exp\left[\frac{hc(S_\phi^2 - s_\phi^2)}{c(S_\phi^2 + s_\phi^2) + 2d} \right]$$

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

$$(2.11) \quad B(t_{(1)}) = \left(\frac{S_y^2}{n} \right) [\{2\theta^2 + 2g\alpha^2\tau^2 + h\tau_0^2\} \frac{(\lambda_{04} - 1)}{4} - \theta(\lambda_{22} - 1) + dr\theta(\lambda_{04} - 1)]$$

and

$$(2.12) \quad MSE(t_{(1)}) = S_y^4 [1 + A_1 - 2A_4 + d^2 A^2 - 2d(A_5 - A_3)].$$

The $MSE(t_{(1)})$ at Eq.(12) is minimum when $d = \frac{(A_5 - A_3)}{A_2}$

$$(2.13) \quad = \left[\left(\frac{\rho^*}{\theta} \right) \sqrt{\left(\frac{\lambda_{40} - 1}{\lambda_{04} - 1} \right) - 1} \right] = d_0(\text{say})$$

$$\text{where } \rho^* = \frac{(\lambda_{22} - 1)}{(\sqrt{\lambda_{40} - 1})(\lambda_{04} - 1)}.$$

Thus, the resulting minimum MSE of $t_{(1)}$ is given by

$$(2.14) \quad MSE(t_{(1)}) = S_y^4 \left[1 + A_1 - 2A_4 - \frac{(A_5 - A_3)^2}{A_2} \right],$$

$$(2.15) \quad = \left(\frac{S_y^4}{n} \right) [(\lambda_{40} - 1)(1 - \rho^{*2})].$$

The minimum MSE of $t_{(1)}$ Eq. (2.14) (or (2.15)) is equal to the minimum MSE of the difference-type estimator $t_2 = [s_y^2 + b_\phi(S_\phi^2 - s_\phi^2)]$

$$(2.16) \quad = t_d(\text{say})$$

which is due to [19].

Now, we state the following corollary. **Corollary 1:** To the first degree of approximation,

$$MSE(t_{(1)}) \geq \frac{S_y^4[(\lambda_{40}-1)(1-\rho^{*2})]}{n}$$

with equality holding if

$$d = [(\frac{\rho^*}{\theta})\sqrt{\frac{(\lambda_{40}-1)}{(\lambda_{04}-1)}} - 1].$$

3. Efficiency comparisons

To compare the proposed class of estimators t with usual unbiased estimator s_y^2 , ratio-type estimator $t_1 = s_y^2(S_\phi^2/s_\phi^2)$, exponential ratio-type estimator t_3 and difference type estimator $t_2 = s_y^2 + b_\phi(S_\phi^2 - s_\phi^2) = t_d$ (say), we write the MSEs of s_y^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

$$(3.1) \quad MSE(s_y^2) = (\frac{S_y^4}{n})(\lambda_{40} - 1),$$

$$(3.2) \quad MSE(t_1) = (\frac{S_y^4}{n})[(\lambda_{40} - 1) + (\lambda_{04} - 1) - 2(\lambda_{22} - 1)],$$

$$(3.3) \quad MSE(t_3) = (\frac{S_y^4}{n})[(\lambda_{40} - 1) + (1/4)(\lambda_{04} - 1) - (\lambda_{22} - 1)],$$

$$(3.4) \quad \min.MSE(t_2 \text{ or } t_d) = (\frac{S_y^4}{n})[(\lambda_{40} - 1) - \frac{(\lambda_{22} - 1)^2}{(\lambda_{04} - 1)}].$$

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

$$(3.5) \quad MSE(s_y^2) - \min.MSE(t_d) = (\frac{S_y^4}{n})\frac{(\lambda_{22} - 1)^2}{(\lambda_{04} - 1)},$$

$$(3.6) \quad MSE(t_1) - \min.MSE(t_d) = (\frac{S_y^4}{n})\frac{(\lambda_{04} - \lambda_{22})^2}{(\lambda_{04} - 1)}.$$

$$(3.7) \quad MSE(t_3) - \min.MSE(t_d) = (\frac{S_y^4}{n})\frac{(\lambda_{04} - 2\lambda_{22} + 1)^2}{(\lambda_{04} - 1)}.$$

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_y^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

$$(3.8) \quad \min.MSE(t_{(1)} \text{ or } t_d) = \frac{S_y^4[A_2(A_1 - A_4) - A_3(A_3 - A_5)]^2}{A_2(A_1A_2 - A_3^2)}$$

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_y^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 (or t_2). [19] suggested another class of estimators for the population variance S_y^2 as

$$(3.9) \quad t_M = s_y^2[m_1 + m_2(S_\phi^2 - s_\phi^2)]\exp[\frac{hc(S_\phi^2 - s_\phi^2)}{c(S_\phi^2 + s_\phi^2) + 2d}],$$

where c , d and h are same as defined earlier and (m_1 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

$$(3.10) \quad MSE(t_M) = S_y^4[1 + m_1^2R_1 + m_2^2R_2 + 2m_1m_2R_3 - 2m_1R_4 - 2m_2R_5]$$

which is minimized for

$$(3.11) \quad \begin{cases} m_1 = \frac{(R_2 R_4 - R_3 R_5)}{(R_1 R_2 - R_3^2)}, \\ m_2 = \frac{(R_1 R_5 - R_3 R_4)}{(R_1 R_2 - R_3^2)}. \end{cases}$$

where

$$\begin{aligned} R_1 &= [1 + \frac{1}{n}\{(\lambda_{40} - 1) - 2h\tau_0(\lambda_{22} - 1) + \frac{h(h+1)}{2}\tau_0^2(\lambda_{04} - 1)\}], \\ R_2 &= \frac{S_\phi^4}{n}(\lambda_{04} - 1), \\ R_3 &= \frac{S_\phi^4}{n}[h\tau_0(\lambda_{04} - 1) - 2(\lambda_{22} - 1)], \\ R_4 &= [1 + (\frac{h\tau_0}{2n})\{\frac{\tau_0(h+2)}{4}(\lambda_{04} - 1) - (\lambda_{22} - 1)\}], \\ R_5 &= \frac{S_\phi^2}{n}[\frac{h\tau_0}{2}(\lambda_{04} - 1) - (\lambda_{22} - 1)], \\ \tau_0 &= \frac{cS_\phi^2}{cS_\phi^2 + d}. \end{aligned}$$

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

$$(3.12) \quad \min.MSE(t_M) = S_y^4(1 - R),$$

where

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

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

$$(3.13) \quad \min.MSE(t_M) - \min.MSE(t) = S_y^4(A - R),$$

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_y^2 by using the formula:

$$(4.1) \quad PRE(t, s_y^2) = \frac{MSE(s_y^2)}{\min.MSE(t)} \times 100.$$

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

y : is the number of villages in the circle,

ϕ : represents a circle consisting more than five villages;

$N=89, n=23, S_y^2 = 4.074, S_\phi^2 = 0.110, C_y = 0.601, C_\phi = 2.678, \rho_{y\phi} = 0.766,$

$\beta_2(\phi) = 6.162, \lambda_{22} = 3.996, \lambda_{40} = 3.811, \lambda_{04} = 6.162.$

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_y^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:

$$(4.2) \quad t_{M1} = s_y^2[m_1 + m_2(S_\phi^2 - s_\phi^2)] \exp\left\{\frac{(S_\phi^2 - s_\phi^2)}{(S_\phi^2 + s_\phi^2)}\right\}$$

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 .

Estimator	PRE
t_0	100.000
t_1	141.898
t_3	254.270
t_4	108.834
t_5	103.823
t_6	155.191
t_7	110.306
t_{17}	91.351
t_{18}	96.214
t_{19}	51.726
t_{20}	89.956
$t_2, t_8, t_9, t_{11}, t_{12}, t_{13}, t_{14}, t_{15}, t_{16}$ and t_{21}	262.187
t_{M1}	284.570

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 .

Estimator	PRE	RANK	Estimator	PRE	RANK
t_1^*	356.62	XIX	t_{11}^*	13011.54	I
t_2^*	3253.05	X	t_{12}^*	2865.64	XV
t_3^*	2856.55	XVI	t_{13}^*	5490.66	V
t_4^*	5289.32	VII	t_{14}^*	5368.64	VI
t_5^*	3375.65	IX	t_{15}^*	5879.30	IV
t_6^*	6257.46	II	t_{16}^*	6184.06	III
t_7^*	4798.68	VIIX	t_{17}^*	3003.95	XII
t_8^*	3083.84	XI	t_{18}^*	2931.79	XIV
t_9^*	2789.78	XVI	t_{19}^*	2976.46	XIII
t_{10}^*	1517.02	XVIII			

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^* ($j = 1, 2, \dots, 19$), which are members of the proposed class of estimators t , are better than the estimators t_i ($i = 1, 2, \dots, 24$) 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^* ($j = 1, 2, \dots, 19$) 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_y^2 of the study variable y in the presence of auxiliary attribute ϕ . 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 $t_0 = s_y^2$, 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^* ($j = 1, 2, \dots, 19$) of the proposed class of estimators t with respect to the usual unbiased estimator s_y^2 for different values of parameters. It is observed that the estimator t_{11}^* 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

- [1] Abd-Elfattah, A.M., El-Sherpieny, E. A., Mohamed, S.M. and Abdou, O.F. (2010): *Improvement in estimating the population mean in simple random sampling using information on auxiliary attribute*. Appl. Math. Comp., 215, 4198-4202.
- [2] Adichwal, N. K., Sharma, P. and Singh, R. (2015): *Estimation of finite population variances using auxiliary attribute*. Jour. Adv. Comp., 4(2), 88-100.
- [3] Adichwal, N. K., Sharma, P., Verma, H.K. and Singh, R. (2015): *Generalized class of estimators for estimating population variances using auxiliary attribute*. Inter. Jour. Comp. Math., 1(3), 1-10.
- [4] Asghar, A., Sanaullah, A. and Hanif, M. (2014): *Generalized exponential-type estimator for population variance in survey sampling*. Rev. Colomb. Estad., 37(1), 211-222.
- [5] Jhajj, H.S., Sharma, M. K. and Grover, L.K. (2006): *Families of estimators of population mean using information on auxiliary attribute*. Pak. Jour. Statist., 22(1), 43-50.
- [6] Kendall, M. and Stuart, A. (1977): *The advanced theory of statistics*. I, Charles, Griffin and Co. London.
- [7] Koyuncu, N. (2012): *Efficient estimators of population mean using auxiliary attributes*. Appl. Math. Comp., 218, 10900-10905.
- [8] Malik, S. and Singh, R. (2013): *An improved estimator using two auxiliary attributes*. Appl. Math. Comp., 219, 10983-10986.
- [9] Naik, V.D. and Gupta, P.C. (1996): *A note on estimation of mean with known population proportion of an auxiliary character*. Jour. Ind. Soc. Agri. Statist., 48(2), 151-158.
- [10] Shabbir, J. and Gupta, S. (2007): *On improvement in variance estimation using auxiliary information*. Commun. Statist. Theo. Meth., 36(12), 2177 -2185.
- [11] Sharma, P. and Singh, R. (2013): *Efficient estimators of population mean in stratified random sampling using auxiliary attribute*. World Appl. Sci. Jour., 27(12), 1786-1791.
- [12] Sharma, P. and Singh, R. (2014): *Improved ratio type estimators using two auxiliary variables under second order approximation*. Math. Jour. Inter. Sci., 2(2), 193-214.

- [13] Sharma, P., Singh, R. and Kim, J. M. (2013b): *Study of some improved ratio-type estimators using information on auxiliary attributes under second order approximation*. Jour. Sci. Res., 57, 138-146.
- [14] Sharma, P., Verma, H., Sanaullah, A. and Singh, R. (2013a): *A study of some exponential ratio-product type estimators using information on auxiliary attributes under second order approximation*. Inter. Jour. Statist. Eco., 12(3), 58-66.
- [15] Sharma, P., Verma, V. K. and Singh, R. (2015): *An efficient class of estimators using two auxiliary attribute*. Chilean Jour. Statist., 6(2), 59-68.
- [16] Singh, R. and Kumar, M. (2011): *A family of estimators of population variance using information on auxiliary attribute*. Studies in Sampling Techniques and Time Series Analysis. Zip Publish., 63-70.
- [17] Singh, H.P. and Solanki, R.S. (2012): *Improved estimation of population mean in simple random sampling using information on auxiliary attribute*. Appl. Math. Comp., 218, 7798-7812.
- [18] Singh, H.P. and Solanki, R.S. (2013): *Improved estimation of population variance using auxiliary information*. Commun. Statist. Theo. Meth., 42, 2718-2730.
- [19] Singh, R. and Malik, S. (2014): *Improved estimation of population variance using information on auxiliary attribute in simple random sampling*. Appl. Math. Comp., 235, 43-49.
- [20] Singh, R., Chauhan, P., Swain, N. and Smarandache, F. (2008): *Ratio estimators in simple random sampling using information on attribute*. Pak. Jour. Statist. Opera. Res., 4(1), 47-53.
- [21] Sukhatme P.V. and Sukhatme B.V. (1970): *Sampling theory of surveys with applications*. Iowa State University Press, Ames.
- [22] Verma, H.K., Sharma, P. and Singh, R. (2015): *Some families of estimators using two auxiliary variables in stratified random sampling*. Rev. Invest. Opera., 36(2), 140-150.
- [23] Yadav, S. K. and Pandey, H. (2012): *Improved family of estimators for population variance using qualitative auxiliary information*. Assam Statist. Rev., 26 (1-2), 63-70.

