International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
Maths 2021; 6(6): 16-23

www.mathsjournal.com
Received: 26-07-2021
Accepted: 08-10-2021

RK Sahoo
School of Statistics, Gangadhar Meher University, Sambalpur, Odisha, India

Ajit Kumar Sabat
School of Statistics, Gangadhar Meher University, Sambalpur, Odisha, India

RK Nayak
Khallikote Higher Secondary School, Bramhapur, Ganjam, Odisha, India

LN Sahoo
Institute of Mathematics & Applications, Andharua, Bhubaneswar, Odisha, India

Corresponding Author:
RK Sahoo
School of Statistics, Gangadhar Meher University, Sambalpur, Odisha, India

On a method of estimating variance of the product estimator

RK Sahoo, Ajit Kumar Sabat, RK Nayak and LN Sahoo

Abstract
In this paper, we explore a new estimation technique for estimating variance of the conventional product estimator under simple random sampling assuming that the population mean and population variance of the auxiliary variable are known prior to sampling. Accordingly, we formulate some new variance estimators under classical as well as predictive approach. We test capabilities of our estimators empirically by means of a simulation study in respect of bias, efficiency, coverage rate based on 95% confidence interval and stability.

Keywords: Auxiliary variable, bias, confidence interval, efficiency, prediction approach, product estimator, stability

Introduction

Consider a finite population $U = \{1, 2, \dots, i, \dots, N\}$. Let y and x denote the study variable and an auxiliary variable taking values y_i and x_i respectively on the ith unit $(i = 1, 2, \dots, N)$. Let $\overline{Y} = \sum_{i=1}^{N} y_i/N$, $\overline{X} = \sum_{i=1}^{N} x_i/N$ be the population means and $S_y^2 = \sum_{i=1}^{N}\left(y_i - \overline{Y}\right)^2/(N-1)$, $S_x^2 = \sum_{i=1}^{N}\left(x_i - \overline{X}\right)^2/(N-1)$ be the population variances of y and x, and $S_{yx} = \sum_{i=1}^{N}\left(y_i - \overline{Y}\right)\left(x_i - \overline{X}\right)/(N-1)$ be the population covariance between y and x. Consider a sample s of n units drawn from U according to simple random sampling without replacement (SRSWOR) for the purpose of estimating the unknown mean $\overline{Y}$. Define $\bar{y} = \sum_{i=1}^{n} y_i/n$ and $\bar{x} = \sum_{i=1}^{n} x_i/n$ as the sample means, $s_y^2 = \sum_{i=1}^{y}(y_i - \bar{y})^2/(n-1)$ and $s_x^2 = \sum_{i=1}^{y}(x_i - \bar{x})^2/(n-1)$ as the sample variances, and $s_{yx} = \sum_{i=1}^{n}(y_i - \bar{y})(x_i - \bar{x})/(n-1)$ as the sample covariance.

In sample surveys the product method of estimation has received considerable attention when there is a negatively high correlation between y and x. In this context, the traditional or conventional product estimator of $\overline{Y}$ that utilizes known value of $\overline{X}$ is defined by

$$t_P = \frac{\bar{y}\bar{x}}{\overline{X}},$$

which yields a considerable efficiency gain over the direct estimator $\bar{y}$ when $\beta_{yx} \leq -\frac{1}{2}R$, where $\beta_{yx} = S_{yx}/S_x^2$ is the regression coefficient of y on x and $R = \overline{Y}/\overline{X}$. An approximate expression of the variance/mean square error of the said estimator is provided by

$$V(t_P) = \frac{N-n}{Nn}\left(S_y^2 + 2RS_{yx} + R^2S_x^2\right). \qquad (1)$$ [*cf.*, Cochran (1973), p.186].

As $V(t_P)$ is a function of some unknown population parameters, remains as an unknown quantity for all practical purposes. Therefore, in large scale sample surveys, estimation of $V(t_P)$ needs attention not only to access the efficiency of t_P from the survey data but also to construct confidence interval for the estimable mean $\overline{Y}$. Like ratio estimator, a consistent estimator of $V(t_P)$ can be obtained when the unknown parameters are replaced by their respective sample counterparts. But, as we are concerned with product method of estimation, we have to consider the case $\beta_{yx} \leq -\frac{1}{2}R$ for which $\hat{R} = \bar{y}/\overline{X}$ is more efficient than $r = \bar{y}/\bar{x}$

in estimating the ratio R. Hence, we feel that a plausible consistent estimator of $V(t_P)$ can be easily framed after replacing S_y^2, S_x^2, S_{yx} and R by s_y^2, s_x^2, s_{yx} and $\hat{R}$ respectively. Such an estimator is of the form

$$\hat{\mathcal{V}}_0 = \frac{N-n}{Nn}\left(s_y^2 + 2\hat{R}s_{yx} + \hat{R}^2 s_x^2\right), \tag{2}$$

which may be termed as the traditional or conventional estimator of $V(t_P)$.
With an objective of achieving improvements over $\hat{\mathcal{V}}_0$, we refer to Scott and Wu (1981) [19] and Wu (1982) [20] and accordingly consider two simple alternatives defined by

$$\hat{\mathcal{V}}_1 = \left(\frac{\bar{x}}{\bar{X}}\right)\hat{\mathcal{V}}_0 \text{ and } \hat{\mathcal{V}}_2 = \left(\frac{\bar{x}}{\bar{X}}\right)^2 \hat{\mathcal{V}}_0, \text{ respectiveely.}$$

Here, in this work, we make an attempt to develop a new approach for estimating $V(t_P)$ by giving stress on the estimation of unknown variance S_y^2 of y and assuming that the population mean $\bar{X}$ and population variance S_x^2 of x are known.

The Suggested Estimation MethoD

Let us now rewrite the expression (1) in the following alternative form:

$$V(t_P) = \frac{N-n}{Nn}\left[S_y^2 + 2RS_x^2\left(\beta_{yx} + \frac{1}{2}R\right)\right]. \tag{3}$$

A comparison of this variance expression for t_P with the variance of $\bar{y}$ given by $V(\bar{y}) = \frac{N-n}{Nn}S_y^2$ clearly shows that the former estimator is preferable to the later one when $\beta_{yx} \le -\frac{1}{2}R$. By making use of expression (3), our new estimation mission considers the question of estimating unknown parameters S_y^2, R and β_{yx} for constructing a reasonable estimator for $V(t_P)$.
Survey sampling literature provides a number of efficient estimation techniques towards estimation of the unknown variance S_y^2 utilizing information on the concomitant variable x. In this context, we may refer to Das and Tripathi (1978) [5], Isaki (1983) [8], Kadilar and Cingi (2006), Grover (2007) [6], Yadav (2011) [21] among others who used classical approach, and Biradar and Singh (1998) [3], Nayak and Sahoo (2012) [14] among others who used predictive approach. Here, we select a suitable estimator $\hat{S}_y^2$ for S_y^2 that makes use of supplementary information on x, and simply choose $b_{yx} = s_{yx}/s_x^2$ and $\hat{R}$ as estimators of β_{yx} and R respectively. This operation obviously gives rise to a family of estimators for $V(t_P)$, defined by

$$\hat{V}(t_P) = \frac{N-n}{Nn}\left[\hat{S}_y^2 + 2\hat{R}S_x^2\left(b_{yx} + \frac{1}{2}\hat{R}\right)\right]$$

$$= \frac{N-n}{Nn}\left[\hat{S}_y^2 + \hat{\eta}S_x^2\right], \tag{4}$$

where $\hat{\eta} = 2\hat{R}\left(b_{yx} + \frac{1}{2}\hat{R}\right)$.

Construction of the new estimators

For a given sample, it is easily understood that various selections of $\hat{S}_y^2$ generates various estimators for $V(t_P)$ from the generalized estimator $\hat{V}(t_P)$ defined in (4). But in actual practice, implementation of this mechanism is not simple. Because, the variety of finite population variance estimation methods and variety of estimator selection criteria available in the survey sampling literature confusing us which estimator could be selected efficiently. However, we emphasize on the property of non-negativity of an estimator. It means that we do not consider such estimators of S_y^2 those achieve negative values very frequently under repeated sampling from a given population. For example, based on this consideration we do not recommend $\hat{S}_y^2 = s_y^2$ and $\hat{S}_y^2 = s_y^2\bar{X}/\bar{x}$ [*cf.*, Das and Tripathi (1978)] [5]. We now place below five new estimators of S_y^2 which are taken into consideration for the present work:

$v_3 = s_y^2 S_x^2 / s_x^2$ [Isaki (1983)] [8]

$v_4 = \frac{s_y^2}{s_x^2}S_x^2 + \frac{nN(\bar{x}-\bar{X})^2}{(N-n)(N-1)}\left(r^2 - \frac{s_y^2}{s_x^2}\right)$ [Biradar and Singh (1998)] [3]

$v_5 = \frac{s_y^2}{s_x^2}S_x^2 + \frac{nN(\bar{x}-\bar{X})^2}{(N-n)(N-1)}\left(b_{yx}^2 - \frac{s_y^2}{s_x^2}\right)$ [Biradar and Singh (1998)] [3]

$v_6 = \left(\frac{n-1}{N-1}\right)\left[s_y^2 + r^2\left(\frac{N-1}{n-1}S_x^2 - s_x^2\right)\right]$ [Nayak and Sahoo (2012)] [14]

$v_7 = \left(\frac{n-1}{N-1}\right)\left[s_y^2 + b_{yx}^2\left(\frac{N-1}{n-1}S_x^2 - s_x^2\right)\right].$ [Nayak and Sahoo (2012)] [14]

We use the said five estimators of S_y^2 for $\hat{S}_y^2$ in (4) and accordingly propose five new estimators for $V(t_P)$ as reported in table 1.

Table 1: Proposed Estimators of $V(t_R)$

Selection of $\hat{S}_y^2$	v_3	v_4	v_5	v_6	v_7
$\hat{V}(t_R) = \frac{N-n}{Nn}\left[\hat{S}_y^2 + \hat{\eta}S_x^2\right]$	$\hat{\mathcal{V}}_3$	$\hat{\mathcal{V}}_4$	$\hat{\mathcal{V}}_5$	$\hat{\mathcal{V}}_6$	$\hat{\mathcal{V}}_7$

Performance of the Proposed Estimators

After considering or proposing eight variance estimators in the preceding section, a desirable goal is to evaluate their performances in respect of different performance criteria. But, we see that even if our sampling scheme is simple *i.e.*, SRSWOR, the proposed variance estimators are structurally complicated. Due to this, derivation of exact results on the bias, variance/mean square error and other performance measures of the estimators under a finite population set up is not straightforward and finally, a theoretical comparison of their performances is not practicable. On the other hand, asymptotic expressions found out using Taylor's Series Expansion Method, are not so simple to rely for drawing valid conclusions regarding relative merits of different estimators. Moving from these considerations, an effort has been made here to compare the performances of the proposed estimators empirically by carrying out a simulation study.

Four useful performance criteria *viz.*, bias, standard error, coverage rate of confidence intervals and coefficient of variation, for comparing performances of the eight variance estimators $\hat{\mathcal{V}}_i\ (i = 0, 1, 2, \ldots, 7)$ of $V(t_P)$ are taken into consideration. These measures are explained as follows:

(i) ***Bias*:** Noting that the bias of an estimator is either negative or positive, for comparison purposes we consider its absolute value. For an estimator $\hat{\mathcal{V}}_i$, this measure is defined by $|Bias\ (\hat{\mathcal{V}}_i)| = |E(\hat{\mathcal{V}}_i) - V(t_P)|$.

(ii) ***Standard Error (SE)*:** Standard error of the estimator $\hat{\mathcal{V}}_i$ is defined by $SE(\hat{\mathcal{V}}_i) = +\sqrt{E[\hat{\mathcal{V}}_i]^2 - [E(\hat{\mathcal{V}}_i)]^2}$, which is of course a convenient and widely used indicator of precision or efficiency attained by the estimator.

(iii) ***Coverage Rate (CR) Based on 95% Confidence Interval for Estimating $\overline{Y}$*:** When t_P is considered as a point estimator for the unknown mean $\overline{Y}$, a 95% confidence interval for $\overline{Y}$ based on the variance estimator $\hat{\mathcal{V}}_i$ is given by $t_R \pm 1.96\sqrt{\hat{\mathcal{V}}_i}$.
This interval will contain the unknown mean $\overline{Y}$ for an approximate proportion of 95% of repeated independent samples drawn from a given population if it is assumed that t_P is asymptotically normally distributed with mean $\overline{Y}$ and variance $V(t_P)$.

(iv) ***Coefficient of Variation (CV)*:** This is taken as one of the best measures of the stability of a survey estimator *i.e.*, as a standardized statistical measure of variability of the probability distribution of the estimator [*cf.*, Rao (1969) [15], Rao and Bayless (1969) [16], Bayless and Rao (1970)] [1]. The coefficient of variation of $\hat{\mathcal{V}}_i$ is given by $CV(\hat{\mathcal{V}}_i) = 100 \times \sqrt{\frac{E[\hat{\mathcal{V}}_i - V(t_R)]^2}{[V(t_R)]^2}}$, is also known as its relative standard error.

Description of the simulation study

The simulation study reported here uses data of 25 bivariate natural populations in which the variables y and x are negatively correlated. Source, size (N) and definitions of y and x for these populations are presented in table 2. We consider 3,000 independent samples of sizes $n = 2, 4, 6$ and 8 drawn without replacement from each population when $\binom{N}{n} \geq 3000$, otherwise all $\binom{N}{n}$ possible samples are considered. Based on the observed values of (y, x) of each realized sample, the estimates t_P and $\hat{\mathcal{V}}_i\ (i = 0, 1, 2, \ldots, 7)$ are calculated. Then considering such values for the considered samples, the measures $|Bias\ (\hat{\mathcal{V}}_i)|, SE(\hat{\mathcal{V}}_i)$ and $CV(\hat{\mathcal{V}}_i)$ are computed in the usual manner. For each sample, the confidence interval is calculated and then counting the number of intervals that contain the true value of $\overline{Y}$, the coverage rates are finally calculated and expressed in percentage.

The numerical results in favor of |bias|, SE, CR and CV for all competing estimators and all populations are summarized in tables 3, 4, 5 and 6 respectively. The entries for the best performer cases are boldly printed and those for the second best performer cases are underlined. Results for $n = 2$ and 8 are not given because the pattern is similar to those of $n = 4$ and 6 respectively. Discussions on the empirical findings are briefly described in subsections 5.1, 5.2, 5.3 and 5.4.

Findings on the Bias

Computed values on the |bias| of the competing estimators displayed in table 3 revel that, for $n = 4$, the conventional estimator $\hat{\mathcal{V}}_0$ is the least biased in 12 populations whereas $\hat{\mathcal{V}}_6$ and $\hat{\mathcal{V}}_7$ are respectively the same in 4 and 8 populations. A similar trend was also noticed for $n = 2$. But, for $n = 6$, $\hat{\mathcal{V}}_7$ emerges out as the best performer as it is decidedly the least biased, second least biased and third least biased in 12, 7 and 2 populations respectively. On the other hand, $\hat{\mathcal{V}}_6$ turns out as the second best performer as it is the least biased, second least biased and third least biased in 6, 4 and 6 cases respectively. On the same consideration, for $n = 6$, we accept $\hat{\mathcal{V}}_0$ as the third best performer because it is ranked as first or second or third in 12 populations. The said findings indicate that for increased sample size the estimators $\hat{\mathcal{V}}_6$ and $\hat{\mathcal{V}}_7$ improve gradually compared to $\hat{\mathcal{V}}_0$ and rest of the estimators. The estimators $\hat{\mathcal{V}}_1$ and $\hat{\mathcal{V}}_2$ also appear to be less biased than others in some cases whereas $\hat{\mathcal{V}}_3$, $\hat{\mathcal{V}}_4$ and $\hat{\mathcal{V}}_5$ behave very much erratically and turn out as the worst performer.

Findings on the Standard Error

From the summarized simulated results given in table 4, we see that the SE of all estimators (with few exceptions) usually decrease with enlargement of sample size. On the average, $\hat{\mathcal{V}}_6$ is emerged as the best performer because on the magnitude of SE it occupies first, second and third positions in 11, 10 and 3 cases respectively for $n = 4$, and in 14, 9 and 2 cases respectively for $n = 6$. On the same ground, we identify $\hat{\mathcal{V}}_7$ and $\hat{\mathcal{V}}_0$ as the second and third best performers which are the most efficient in 7 and 5 populations for $n = 4$, and 8 and 3 populations for $n = 6$ respectively. In about 6 cases, for both $n = 4$ and $n = 6$, $\hat{\mathcal{V}}_1$ happens to be more efficient than others whereas $\hat{\mathcal{V}}_2$, $\hat{\mathcal{V}}_3$, $\hat{\mathcal{V}}_4$ and $\hat{\mathcal{V}}_5$ are consistently less efficient than others.

Table 2: Populations under Study

Pop. No.	Source	N	y	x
1	Gujarati and Porter (2009, p.406) [7]	35 passenger cars (1-35)	average miles per gallon	engine HP
2	Gujarati and Porter (2009, p.406) [7]	35 passenger cars (36-70)	average miles per gallon	engine HP
3	Gujarati and Porter (2009, p.168) [7]	32 countries (1-32)	child mortality	female literacy rate
4	Gujarati and Porter (2009, p.168) [7]	32 countries (33-64)	child mortality	female literacy rate
5	Gujarati and Porter (2009, p.51) [7]	27 years	average hourly earnings	civilian labor force participation rate
6	Maddala (1992, p.194) [11]	34 rural lands (1-34)	sale price of land	distance from airport
7	Maddala (1992, p.194) [11]	33 rural lands (35-67)	sale price of land	distance from airport
8	Morrison (1990, p.470) [13]	26 lighter and heavier under wt. young males	pigment creatinine	phosphate (mg/mL)
9	Morrison (1990, p.470) [13]	19 lighter and heavier obese young males	pigment creatinine	phosphate (mg/mL)
10	Bhuyan (2005, p.4) [2]	28 married couples of middle class families	fertility level (no. of ever born children)	education level of father (in completed years)
11	Bhuyan (2005, p.224) [2]	20 married couples (1-20)	no. of ever born children	education of mother (in completed years of schooling)
12	Bhuyan (2005, p.76) [2]	28 two times milking cows	daily milk production	wt. of cow after lactation period
13	Bhuyan (2005, p.76) [2]	28 three times milking cows	daily milk production	wt. of cow after lactation period
14	Johnson and Wichern (2007, p.215) [9]	20 healthy females	sweat rate	potassium content
15	Rawlings, Pantula and Dickey (1998, p.396) [17]	20 plots (depth 1)	sand percentage	silt percentage
16	Rawlings, Pantula and Dickey (1998, p.396) [17]	20 plots (depth 2)	clay percentage	sand percentage
17	Montgomery, Peck and Vining (2012, p.556) [12]	32 automobiles	miles/gallon	horsepower
18	Montgomery, Peck and Vining (2012, p.15) [12]	20 obs.	shear strength	age of propellant
19	Montgomery, Peck and Vining (2012, p.291)	16 obs.	conversion of $n-$ heptane to acetylene (%)	contact time (sec)
20	Montgomery, Peck and Vining (2012, p.483) [12]	20 time periods	selling price of toothpaste per pound	market share of toothpaste
21	Montgomery, Peck and Vining (2012, p.572) [12]	32 young red wines	quality rating (20 maximum)	total SO_2 (ppm)
22	Montgomery, Peck and Vining (2012, p.558) [12]	27 Belle Ayr liquefaction runs	CO_2	space time (min)
23	Montgomery, Peck and Vining (2012, p.558) [12]	27 Belle Ayr liquefaction runs	oil yield	coal total
24	Rencher (2002, p.269) [18]	23 obs. (1-23)	evaporation	minimum daily relative humidity
25	Rencher (2002, p.269) [18]	23 obs.(24-46)	evaporation	minimum daily relative humidity

Table 3: |Bias| of the Variance Estimators

n	Pop. No.	$\hat{V}_0$	$\hat{V}_1$	$\hat{V}_2$	$\hat{V}_3$	$\hat{V}_4$	$\hat{V}_5$	$\hat{V}_6$	$\hat{V}_7$
4	1	21.0327	17.4496	15.3541	20.2842	21.9744	21.6323	51.9432	45.5674
	2	38.3323	49.2247	47.7124	65.9809	86.6754	102.987	45.9870	40.2546
	3	675.016	647.057	713.345	1018.33	945.815	944.119	71.7328	221.875
	4	114.092	138.145	307.905	2401.58	2085.25	2141.76	108.321	46.7659
	5	2.1627	2.4931	2.5081	2.8765	3.5647	4.9871	2.7809	2.3421
	6	23497.6	58862.8	13296.0	43875.2	78654.2	54432.0	3568.22	3998.54
	7	33225.4	31010.5	31898.1	30297.9	31957.4	32010.7	19119.1	6811.17
	8	1717.12	30795.9	16093.2	89765.8	4580.09	24509.3	1925.73	4690.01
	9	1.5004	2.3588	4.3132	4.2051	4.4558	3.5922	9.4456	2.1762
	10	14447.8	10756.0	5538.06	976.543	19870.3	888.125	780.765	565.982
	11	0.2982	0.3971	0.4066	0.9887	1.5564	9.0988	0.7677	0.6578
	12	18114.0	9969.52	3652.33	41211.8	40240.9	41041.2	23753.5	351.792
	13	18200.2	10058.9	3736.84	13714.9	12806.2	13591.7	23458.9	2376.64
	14	76.8207	783.866	3758.98	11658.4	11518.6	11556.6	1421.31	4184.06
	15	401.545	2564.72	11222.4	56463.8	55837.5	56060.1	2388.10	24216.4
	16	15.7458	5474.30	23907.6	4188.67	4328.67	4137.67	12931.3	738.814
	17	19000.2	57974.8	21054.8	91954.3	72405.6	44987.5	15235.8	11000.2
	18	3397.59	23676.4	72430.5	17837.8	18880.3	17695.5	50372.4	17030.7
	19	173.327	4206.90	52408.8	16300.4	16021.5	16009.8	2345.57	3741.51
	20	2.5443	39.1617	154.203	129.115	126.763	127.678	10.2913	50.4369

	21	1049.82	671.188	52947.5	657.610	520.921	600.637	2415.86	205.823
	22	28.3902	47.4825	82.1418	151.183	144.699	146.197	85.4957	63.6644
	23	3582.34	2746.06	1305.13	1209.16	783.466	875.091	860.532	168.325
	24	26.6403	38.2102	55.5485	221.616	216.799	217.824	1.6208	7.38476
	25	51.8282	88.0455	143.816	2014.16	1990.82	1995.82	49.3434	567.967
	1	151.647	126.043	104.771	51.7976	42.2323	43.8569	40.1466	19.9311
	2	133.062	112.337	96.8650	623.177	562.263	566.212	6.4158	28.0157
	3	365.746	296.534	105.674	516.341	470.950	418.930	259.705	90.0237
	4	708.017	743.815	747.287	761.228	744.851	731.839	183.934	238.651
	5	68.1716	69.6040	71.1186	3484.99	3382.37	3418.10	32.3621	1135.49
	6	343.041	73.1895	556.826	74952.6	73333.7	73484.1	5360.91	248.721
	7	1137.95	1101.36	1122.88	2003.54	1976.41	1973.59	809.576	434.327
	8	26115.9	24814.7	22206.8	15729.7	15764.9	15830.8	15780.0	15628.7
	9	1.8240	2.8816	3.5967	6.0263	6.2981	5.3671	8.1913	1.2153
	10	10863.7	9320.19	7287.21	8535.27	7294.75	6841.64	11855.1	6580.28
	11	0.3450	0.0325	0.3253	0.1434	0.1353	0.1292	0.0140	0.0635
	12	9207.60	8153.85	7324.89	28415.0	27535.8	27771.1	13074.2	8322.12
6	13	7613.87	8064.27	7332.44	91152.2	90012.8	90529.9	14120.7	4792.34
	14	962.873	909.512	470.968	7464.91	7314.99	7319.60	161.885	254.053
	15	145.777	474.885	1981.12	17460.9	17149.4	17144.9	317.652	310.382
	16	447.033	837.899	3932.17	1599.55	1683.13	1557.42	498.605	406.265
	17	3019.75	6582.80	4049.00	8294.47	5384.50	8332.74	11368.3	9378.57
	18	1345.32	2810.84	10930.7	12980.8	13096.8	12708.3	1242.59	552.673
	19	382.235	637.023	3298.20	597.669	9327.81	560.938	17646.8	1141.82
	20	1.9327	4.8421	20.609	208.70	202.70	203.82	4.1664	3.5353
	21	3698.22	5201.44	6271.31	4511.08	3587.19	4239.05	13886.3	491.034
	22	9.2127	3.1344	4.6936	45.7729	45.8426	44.5801	4.8722	12.6746
	23	10619.1	3141.81	4696.35	12651.2	12445.0	12510.1	7739.37	2100.07
	24	21.616	20.156	19.920	34.945	32.746	31.506	4.0664	10.478
	25	146.71	85.124	39.384	770.84	762.60	762.71	14.953	14.662

Table 4: Standard Error of the Variance Estimators

n	Pop. No.	$\hat{\mathcal{V}}_0$	$\hat{\mathcal{V}}_1$	$\hat{\mathcal{V}}_2$	$\hat{\mathcal{V}}_3$	$\hat{\mathcal{V}}_4$	$\hat{\mathcal{V}}_5$	$\hat{\mathcal{V}}_6$	$\hat{\mathcal{V}}_7$
	1	149.307	153.640	151.713	209.907	178.033	196.867	61.877	128.392
	2	142.500	152.062	172.841	163.876	155.987	176.043	103.764	115.902
	3	1755.02	1801.44	2032.99	6754.08	6035.18	6212.26	1724.08	1789.64
	4	2375.57	2271.01	3136.84	20180.4	16490.7	18072.7	1573.94	1488.42
	5	40.5210	71.4817	72.6410	98.2354	75.4532	102.333	34.0987	45.0084
	6	13950.3	27425.5	54357.6	36791.0	47861.7	30987.6	14987.8	21098.4
	7	607.638	611.362	696.181	12655.8	8936.48	9580.39	386.681	538.483
	8	6592.27	17836.8	48354.8	3389.17	8954.34	12093.3	1456.93	1276.98
	9	33.0883	26.1249	27.1620	36.3235	32.4134	33.4569	23.2602	25.1957
	10	31163.7	41743.9	17943.5	30987.5	42098.6	39876.9	18154.4	27654.3
	11	0.3847	0.4174	0.4511	0.5512	0.6351	0.6340	0.3901	0.2453
	12	93117.6	73232.6	51948.3	55248.2	55125.6	55196.9	11639.2	33827.5
4	13	9297.24	7308.88	15776.1	14043.9	14001.4	14021.1	6464.90	7207.64
	14	451.312	1899.07	7986.12	34840.6	34352.7	34540.7	1269.32	1769.34
	15	924.168	1768.05	2623.72	47453.1	46897.5	47350.9	1027.48	39390.1
	16	4801.30	16084.7	53817.6	9017.07	8941.35	8925.78	5742.27	3224.36
	17	17926.7	66166.8	59402.7	31245.6	37852.7	35051.2	23454.7	9834.66
	18	3143.02	7476.73	17758.4	3564.28	3760.24	3747.85	5686.16	2134.42
	19	761.726	240.897	940.129	27186.5	26559.7	26672.4	596.344	5546.80
	20	39.2013	122.223	379.802	479.641	469.572	473.548	33.5076	24.7859
	21	78516.2	76445.5	88272.3	80880.0	79473.5	68153.5	20883.2	61455.7
	22	117.360	867.059	653.880	749.769	655.578	688.362	249.558	445.117
	23	4414.75	7434.98	13233.0	32651.5	18417.9	21103.2	5073.92	14847.4
	24	382.456	414.144	659.666	525.340	497.440	509.714	129.325	305.979
	25	492.776	586.812	625.160	7664.30	547.234	7598.73	385.427	4088.08
	1	59.052	66.9972	52.6688	139.140	105.905	117.405	51.4699	52.2680
	2	51.8673	66.4237	79.3297	3460.70	3001.78	3050.85	41.3831	44.0003
	3	551.515	576.703	559.439	1132.61	983.264	970.816	386.469	471.623
	4	1417.22	1712.54	1787.01	2000.44	1968.66	1977.92	1370.93	1644.00
	5	22.1831	44.9148	416.385	3503.93	3366.19	3442.19	20.8241	24.4205
6	6	362.707	551.473	848.027	14644.1	14193.1	14239.9	284.461	280.846
	7	19807.1	19086.8	20070.4	39799.5	39680.2	39723.3	10422.8	8601.22
	8	3313.72	2625.86	3326.67	29649.0	29306.8	29648.4	2309.91	2968.66
	9	3.8308	4.0559	4.5451	10.1363	10.0115	9.0721	2.2200	1.6526
	10	15369.6	19736.7	25709.0	28117.3	25843.4	27293.2	12743.3	22607.1
	11	0.1331	0.1565	0.1883	0.1794	0.1392	0.1375	0.0856	0.0424

	12	13173.7	25753.3	21878.7	14909.1	14983.0	14698.2	10578.2	12064.8
	13	29730.4	24813.8	20638.6	22565.4	22483.4	22398.2	2953.28	34613.8
	14	1129.52	3820.32	4066.52	7964.69	7784.60	7837.54	2656.75	3711.01
	15	866.799	4096.51	5051.87	22448.7	22017.0	22126.6	2469.23	3499.89
	16	1778.96	4269.06	10220.5	2182.61	2161.08	2156.93	1976.02	1329.02
	17	25487.0	4924.60	22838.3	9308.72	9008.10	9183.33	3503.89	4164.67
	18	9723.60	18588.1	34985.5	13122.5	12993.8	13033.9	10067.2	5767.45
	19	879.333	229.552	599.751	65107.7	62648.8	62768.9	213.489	9651.16
	20	11.2000	25.3532	55.8530	169.260	163.332	164.677	17.1275	46.5067
	21	83895.5	69917.4	67860.5	52828.9	49450.8	52041.9	8881.86	41415.1
	22	36.4568	46.4374	62.3225	78.6556	78.3908	76.9139	33.1600	26.8915
	23	4384.78	3919.78	5309.69	3396.61	3324.14	3320.15	2135.37	3034.70
	24	39.6100	47.3533	56.4868	44.6487	41.2821	40.2092	35.8571	21.1180
	25	187.504	1115.17	1151.09	1444.48	1426.87	1431.47	100.015	555.095

Table 5: Coverage Rate of the Variance Estimators

n	Pop. No.	$\hat{v}_0$	$\hat{v}_1$	$\hat{v}_2$	$\hat{v}_3$	$\hat{v}_4$	$\hat{v}_5$	$\hat{v}_6$	$\hat{v}_7$
4	1	87.00	67.00	45.00	64.00	58.00	69.00	86.00	72.00
	2	80.00	68.00	56.00	49.00	73.00	69.00	93.00	89.00
	3	88.00	68.00	67.00	67.00	56.00	66.00	72.00	84.00
	4	63.00	52.00	71.00	80.00	54.00	68.00	83.00	89.00
	5	92.00	61.00	62.00	77.00	86.00	87.00	91.00	97.00
	6	78.00	86.00	74.00	82.00	79.00	82.00	90.00	97.00
	7	96.00	90.00	89.00	86.00	87.00	86.00	95.00	92.00
	8	51.00	65.00	45.00	41.00	41.00	46.00	55.00	61.00
	9	97.00	85.00	84.00	96.00	92.00	91.00	100.00	87.00
	10	33.00	24.00	39.00	29.00	24.00	24.00	23.00	44.00
	11	95.00	81.00	84.00	86.00	85.00	85.00	90.00	88.00
	12	29.00	31.00	26.00	17.00	12.00	17.00	18.00	37.00
	13	26.00	19.00	16.00	17.00	12.00	16.00	24.00	22.00
	14	20.00	20.00	20.00	85.00	86.00	85.00	95.00	90.00
	15	87.00	82.00	60.00	78.00	79.00	78.00	89.00	80.00
	16	20.00	20.00	70.00	67.00	69.00	57.00	89.00	75.00
	17	42.00	22.00	32.00	30.00	31.00	31.00	37.00	38.00
	18	88.00	66.00	57.00	80.00	81.00	79.00	91.00	86.00
	19	25.00	25.00	25.00	90.00	60.00	89.00	100.00	98.00
	20	66.00	36.00	45.00	30.00	35.00	39.00	90.00	28.00
	21	39.00	33.00	44.00	41.00	57.00	37.00	49.00	62.00
	22	95.00	92.00	79.00	85.00	86.00	85.00	100.00	78.00
	23	51.00	37.00	40.00	61.00	61.00	60.00	73.00	69.00
	24	91.00	80.00	78.00	70.00	80.00	76.00	87.00	79.00
	25	92.00	89.00	72.00	67.00	80.00	86.00	90.00	88.00
6	1	91.00	96.00	87.00	90.00	89.00	89.00	98.00	74.00
	2	89.00	89.00	90.00	83.00	79.00	87.00	96.00	92.00
	3	94.00	85.00	91.00	89.00	89.00	88.00	90.00	65.00
	4	81.00	94.00	80.00	92.00	93.00	91.00	95.00	98.00
	5	97.00	98.00	93.00	93.00	95.00	95.00	94.00	96.00
	6	60.00	52.00	46.00	80.00	76.00	79.00	95.00	88.00
	7	98.00	77.00	76.00	88.00	85.00	88.00	94.00	89.00
	8	8.00	5.00	12.00	14.00	14.00	19.00	18.00	23.00
	9	78.00	71.00	65.00	96.00	87.00	86.00	100.00	89.00
	10	30.00	31.00	38.00	11.00	15.00	19.00	22.00	41.00
	11	99.00	79.00	69.00	77.00	77.00	76.00	86.00	83.00
	12	97.00	90.00	94.00	88.00	88.00	86.00	99.00	95.00
	13	95.00	90.00	91.00	83.00	85.00	88.00	98.00	100.00
	14	79.00	81.00	91.00	97.00	93.00	98.00	100.00	94.00
	15	21.00	21.00	21.00	76.00	88.00	80.00	83.00	91.00
	16	65.00	97.00	31.00	98.00	96.00	96.00	100.00	76.00
	17	16.00	25.00	39.00	25.00	17.00	19.00	35.00	31.00
	18	71.00	55.00	46.00	94.00	83.00	86.00	99.00	91.00
	19	34.00	34.00	34.00	81.00	91.00	90.00	100.00	93.00
	20	99.00	90.00	24.00	91.00	70.00	70.00	98.00	100.00
	21	19.00	39.00	29.00	27.00	32.00	34.00	37.00	43.00
	22	87.00	84.00	81.00	89.00	95.00	91.00	100.00	98.00
	23	94.00	94.00	95.00	91.00	91.00	93.00	99.00	97.00
	24	87.00	90.00	77.00	76.00	76.00	76.00	94.00	78.00
	25	92.00	61.00	70.00	83.00	84.00	83.00	88.00	85.00

Table 6: Coefficient of Variation of the Variance Estimators

n	Pop. No.	$\hat{\mathcal{V}}_0$	$\hat{\mathcal{V}}_1$	$\hat{\mathcal{V}}_2$	$\hat{\mathcal{V}}_3$	$\hat{\mathcal{V}}_4$	$\hat{\mathcal{V}}_5$	$\hat{\mathcal{V}}_6$	$\hat{\mathcal{V}}_7$
	1	180.831	183.131	180.405	266.907	227.037	250.659	102.248	172.426
	2	368.929	287.127	326.773	354.897	391.876	395.909	270.985	412.098
	3	122.802	165.056	140.744	446.722	398.948	410.368	112.691	131.524
	4	679.294	485.737	504.957	676.743	553.506	606.027	422.519	382.562
	5	215.321	218.243	221.780	176.776	168.990	196.006	111.983	125.908
	6	321.556	637.575	1271.96	453.221	1006.89	601.987	678.320	534.321
	7	948.308	443.757	495.715	8196.32	488.598	6205.18	250.618	1966.86
	8	250.592	400.136	1126.58	1099.09	987.098	998.546	361.445	157.097
	9	371.772	488.783	188.663	199.212	278.249	283.321	136.771	137.973
	10	127.317	159.779	215.749	140.009	197.123	213.449	109.332	146.095
	11	158.369	262.181	276.377	322.981	480.443	984.022	132.666	123.665
	12	388.697	302.835	337.912	2270.06	2264.75	2267.91	100.791	321.077
4	13	399.904	511.432	444.322	595.641	593.479	594.636	100.951	306.297
	14	1316.52	1420.47	6102.70	25401.7	25051.2	25182.8	1327.53	12570.6
	15	217.701	786.260	2770.88	46401.7	45858.8	46298.6	1126.53	38319.8
	16	500.867	710.824	2463.67	475.952	415.599	491.589	303.299	220.740
	17	428.655	561.775	523.755	567.908	612.564	543.987	550.147	233.762
	18	233.262	578.680	1415.12	307.359	310.463	305.814	260.516	165.982
	19	535.764	547.409	588.861	1400.59	1456.71	1418.66	403.690	441.52
	20	254.655	831.985	657.245	3219.94	3152.95	3179.38	264.652	439.66
	21	444.945	431.666	474.384	475.309	495.837	424.232	171.306	329.857
	22	139.364	700.456	607.984	882.800	774.879	812.228	304.474	518.982
	23	307.358	349.666	251.095	616.990	348.105	398.840	190.945	180.367
	24	160.373	222.775	312.875	1055.24	1004.27	1025.88	139.368	582.549
	25	538.411	558.319	636.038	9464.37	9322.07	9383.07	464.077	492.344
	1	317.786	272.988	237.972	301.296	231.377	254.337	143.332	213.520
	2	417.328	381.363	465.869	10275.4	8924.33	9067.40	267.696	404.238
	3	69.8137	71.1789	99.7343	131.316	115.014	111.545	77.4286	67.9349
	4	84.2123	95.4075	98.3790	115.134	113.223	113.445	74.4056	89.3608
	5	241.913	349.128	357.115	24773.0	23921.0	24317.1	207.545	312.591
	6	133.088	301.471	310.452	6009.51	5835.95	5853.67	241.918	1408.15
	7	238.239	249.824	239.853	464.709	462.329	462.600	137.641	100.492
	8	96.0217	93.6830	94.3535	122.250	121.381	122.251	95.5910	122.543
	9	100.006	86.9133	104.652	111.193	111.526	199.391	56.2921	63.7669
	10	114.157	152.384	162.074	175.146	162.871	170.661	105.566	146.565
	11	288.465	286.699	293.244	144.012	112.044	109.639	67.6463	59.5962
	12	211.523	381.122	254.698	215.154	240.188	251.675	88.4507	105.275
6	13	211.988	176.444	146.352	840.310	834.052	836.327	99.6487	204.398
	14	154.713	459.663	1384.88	12938.4	12661.1	12710.7	1381.45	4606.30
	15	146.310	357.818	903.264	4734.01	665.452	659.397	771.835	1663.57
	16	131.551	312.013	785.378	194.070	196.451	190.803	147.873	99.6699
	17	225.819	212.858	204.078	109.708	108.278	109.105	100.123	98.7487
	18	124.141	237.795	463.630	233.478	233.363	230.263	212.265	73.2870
	19	894.421	1234.59	873.537	6472.28	6237.48	6239.81	727.63	857.047
	20	126.302	786.836	661.589	2986.17	2892.90	2911.97	215.878	649.195
	21	409.535	478.386	364.037	459.463	429.650	452.471	142.844	358.914
	22	209.255	168.054	198.509	172.561	172.194	188.569	164.754	110.521
	23	378.038	433.057	519.907	404.884	399.611	401.531	249.066	352.409
	24	140.009	159.681	185.842	175.919	163.490	158.494	111.967	73.1456
	25	377.647	336.038	312.643	278.233	3239.36	3247.59	227.311	249.594

Findings on the Coverage Rate

The values of the CRs of the nominal 95% confidence intervals for $\overline{Y}$ based on different variance estimators are compiled in table 5. The results on the achieved CRs give clear indication of improvement in the performance of an estimator as the sample size increases. In the light of this performance measure, the estimators in most of the cases (except $\hat{\mathcal{V}}_6$ and $\hat{\mathcal{V}}_7$) appear to be unpredictable and usually bear no resemblance to the nominal rate aimed at. However, interestingly CR achievements in favor of $\hat{\mathcal{V}}_6$ are 100% for about 8 cases. On the consideration of closeness of the calculated CR to the nominal value 95%, $\hat{\mathcal{V}}_6$ is sorted out as the best estimator as its achieved CR is the highest in 10 and 12 populations, second highest in 8 and 6 populations and third highest in 5 populations for $n = 4$ and $n = 6$ respectively. On this ground, $\hat{\mathcal{V}}_7$ and $\hat{\mathcal{V}}_0$ may be identified as the second and third best estimators respectively.

Findings on the Coefficient of Variation

The results given in table 6 reveal that, the behaviors of the estimators in respect of the CV are more or less similar to those of the SE. Considering higher CV value as less stability *i.e.*, less accuracy or variability of an estimator towards sampling fluctuations, $\hat{\mathcal{V}}_6$ may be chosen as the most stable estimator because according to the CV values it remains respectively as the best, second best and third best estimator in 13, 7 and 3 populations for $n = 4$, and in 13, 8 and 4 populations for $n = 6$. In respect of this criterion, we select $\hat{\mathcal{V}}_7$ and $\hat{\mathcal{V}}_0$ respectively as the second and third stable estimators. The estimator $\hat{\mathcal{V}}_1$ is considerably more stable than $\hat{\mathcal{V}}_2$, $\hat{\mathcal{V}}_3$, $\hat{\mathcal{V}}_4$ and $\hat{\mathcal{V}}_5$ as it is ranked as either first or second or third in about 6 and 7 cases for $n = 4$ and 6 respectively. We also see on the average that the stabilities of $\hat{\mathcal{V}}_2$, $\hat{\mathcal{V}}_3$, $\hat{\mathcal{V}}_4$ and $\hat{\mathcal{V}}_5$ are very poor compared to others.

Conclusions

Findings of the simulation study clearly indicate that no estimator is uniformly better than others on the ground of all four aforesaid performance criteria. However, an analysis of the simulation results shows that out of eight comparable estimators overall performances of only three estimator's *viz*., $\hat{\mathcal{V}}_0, \hat{\mathcal{V}}_6$ and $\hat{\mathcal{V}}_7$ may be taken into consideration. Although in certain cases $\hat{\mathcal{V}}_1$ exhibits good results, these are not so significant compared to those of the three competing estimators $\hat{\mathcal{V}}_0, \hat{\mathcal{V}}_6$ and $\hat{\mathcal{V}}_7$. Tentatively we may conclude that from the point of view of bias, the conventional estimator $\hat{\mathcal{V}}_0$ is superior to others followed by $\hat{\mathcal{V}}_7$ when the sample size is small whereas for larger sample size $\hat{\mathcal{V}}_7$ is the most preferred estimator followed by $\hat{\mathcal{V}}_6$. On the other hand, in terms SE, CR and CV criteria, it may also be warranted to conclude that $\hat{\mathcal{V}}_6, \hat{\mathcal{V}}_7$ and $\hat{\mathcal{V}}_0$ are respectively preferable as the best, second best and third best estimators.

In many survey sampling situations, the bias of an estimator is not a matter of great concern. If it is so, then the results of the present study leads to a major conclusion that the estimator $\hat{\mathcal{V}}_6$ may be recommended for our purpose. However, no general comments may be derived from this numerical investigation in the sense that the conclusions drawn may not fit to other situations. The study only provides certain guidance on the overall performance of the comparable estimators. Therefore, we suggest that further investigations in this direction with the help of other performance measures may be more useful for better understanding of the statistical properties of the estimators.

References

1. Bayless DL, Rao JNK. An empirical study of stabilities of estimators and variance estimators in unequal probability sampling. *Journal of the American Statistical Association* 1970;65:1645-1667.
2. Bhuyan KC. *Multivariate Analysis and its Applications*. New Central Book Agency (P) Ltd, Kolkata 2005.
3. Biradar RS, Singh HP. Predictive estimation of finite population variance. *Calcutta Statistical Association Bulletin,* 1998;48:229-235.
4. Cochran WG. *Sampling Techniques* (3rd Edition). Wiley Eastern Limited 1977.
5. Das A, Tripathi TP. Use of auxiliary information in estimating the finite population variance. *Sankhyā,* 1978;C40:139-148.
6. Grover LK. A wide and efficient class of estimators of population variance under sub-sampling scheme. *Model Assisted Statistics and Applications* 2007;2:41-51.
7. Gujarati DN, Porter DC. *Basic Econometrics* (5th Edition). McGraw-Hill, New York, NY, 10020, 2009.
8. Isaki CT. Variance estimation using auxiliary information. *Journal of the American Statistical Association* 1983;78:117-123.
9. Johnson RA, Wichern DW. *Applied Multivariate Statistical Analysis*. Pearson Education, Inc., New Jersey 2007.
10. Kadilar C, Cingi H. Ratio estimators in stratified random sampling. *Biometrical Journal* 2003;45:218-225.
11. Maddala GS. *Introduction to Econometrics* (2nd Edition). Macmillan Publishing Company, New York 1992.
12. Montgomery DC, Peck EA, Vining GG. *Introduction to Linear Regression Analysis* (5th Edition). John Wiley & Sons, Inc 2012.
13. Morrison DF. *Multivariate Statistical Methods* (3rd Edition). McGraw-Hill Publishing Company, New York 1990.
14. Nayak RK, Sahoo LN. Some alternative predictive estimators of population variance. *Revista Colombiana de Estadistica*, 2012;13:1-13.
15. Rao JNK. Ratio and regression estimators. *New Developments in Survey Sampling*. N.L. Johnson and H. Smith Jr. (eds.), Chapel Hill, North Carolina, 1969, 213-234.
16. Rao JNK, Bayless DL. An empirical study of the stabilities of estimators and variance estimators in unequal probability sampling of two units per stratum. *Journal of the American Statistical Association* 1969;64:540-549.
17. Rawlings JO, Pantula SG, Dickey DA. *Applied Regression Analysis* (2nd Edition). Springer-Verlag, New York, Inc 1998.
18. Rencher AC. *Methods of Multivariate Analysis* (2nd Edition). John Wiley & Sons, Inc 2002.
19. Scott AJ, Wu CF. On the asymptotic distribution of ratio and regression estimators. *Journal of the American Statistical Association* 1981;76:98-102.
20. Wu CF. Estimation of variance of the ratio estimators. *Journal of the American Statistical Association* 1982;69:183-89.
21. Yadav SK. Efficient estimators for population variance using auxiliary information. *Global Journal of Mathematical Sciences: Theory and Practical* 2011;3:369-376.