International Journal of Statistics and Applications 2013, 3(3): 71-76
DOI: 10.5923/j.statistics.20130303.06

On the Variance Estimation of Regression Estimator

N. R. Das[1], R. K. Nayak[2], L. N. Sahoo[1,*]

[1]Department of Statistics, Utkal University, Bhubaneswar, 751004, India
[2]Department of Statistics, Khallikote College, Berhampur, 760001, India

Abstract In this paper, an attempt has been made to estimate variance of the classical regression estimator. Adopting some available techniques used for estimation of population variance under classical as well as predictive approach, we develop eight new variance estimators of the classical regression estimator. It is assumed that the population mean and population variance of the auxiliary variable are known prior to sampling. A simulation study has been undertaken for evaluating relative performance of the suggested estimators in respect of standard error and coverage rate based on 95% confidence interval.

Keywords Auxiliary Variable, Confidence Interval, Prediction Approach, Regression Estimator, Stability

1. Introduction

Consider a finite population $U = \{1, 2, \ldots, i, \ldots, N\}$. Let y and x denote the study variable and an auxiliary variable taking values y_i and x_i respectively on the ith unit $(i = 1, 2, \ldots, N)$. Let $\bar{Y} = \sum_{i=1}^{N} y_i/N$, $\bar{X} = \sum_{i=1}^{N} x_i/N$ be the population means and $S_y^2 = \sum_{i=1}^{N}(y_i - \bar{Y})^2/(N-1)$, $S_x^2 = \sum_{i=1}^{N}(x_i - \bar{X})^2/(N-1)$ be the population variances of y and x, and $S_{yx} = \sum_{i=1}^{N}(y_i - \bar{Y})(x_i - \bar{X})/(N-1)$ 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) to estimate the unknown mean $\bar{Y}$. Let $\bar{y} = \sum_{i\epsilon s} y_i/n$ and $\bar{x} = \sum_{i\epsilon s} x_i/n$ be the sample means, $s_y^2 = \sum_{i\epsilon s}(y_i - \bar{y})^2/(n-1)$ and $s_x^2 = \sum_{i\epsilon s}(x_i - \bar{x})^2/(n-1)$ the sample variances, and $s_{yx} = \sum_{i\epsilon s}(y_i - \bar{y})(x_i - \bar{x})/(n-1)$ be the sample covariance.

The classical regression estimator of $\bar{Y}$ that utilizes known value of $\bar{X}$ is defined by $t_{RG} = \bar{y} - b(\bar{x} - \bar{X})$, where $b = s_{yx}/s_x^2$ is the familiar least squares estimator of $\beta = S_{yx}/S_x^2$, the population regression coefficient of y on x. The main advantages associated with t_{RG} are that it can cover both the situations of positive and negative correlations between y and x, and its precision is usually higher than that of the simple expansion (direct) estimator $\bar{y}$ as well as its ratio and product counterparts. In early days regression estimator was not frequently used in practice because of its computational difficulty. But later on, due to advancement of computational facilities, the regression or regression-type estimators are of much interest to the survey statisticians.

The approximate variance of t_{RG} is given by

$$V(t_{RG}) = \frac{N-n}{Nn} S_y^2(1 - \rho^2), \tag{1.1}$$

where $\rho = S_{yx}/S_y S_x$ is the correlation coefficient between y and x. The precision of t_{RG} is usually discussed in terms of $V(t_{RG})$. But the exact value of this variance is unknown as it depends on the unknown population quantities S_y^2 and ρ. Hence, the estimation of the variance of t_{RG} seems to be an important aspect of study in the survey sampling literature. An estimate of the variance is needed to provide a measure of the error in estimation of the mean $\bar{Y}$ by t_{RG}. A variance estimate is also used to construct a confidence interval for the unknown mean. A commonly used estimator of the approximate variance $V(t_{RG})$ is its sample analogue

$$v_0(t_{RG}) = \frac{N-n}{Nn} s_y^2(1 - r^2), \tag{1.2}$$

where $r = s_{yx}/s_y s_x$ is the sample correlation coefficient, a consistent estimate of ρ.

Survey sampling literature provides a number of procedures for estimating unknown variance S_y^2 using auxiliary information based either on the classical approach [*cf.*, Das and Tripathi[1], Isaki[2], Kadilar and Cingi[3], Grover[4], Yadav[5]] or on the predictive approach [*cf.*, Bolfarine and Zacks[6], Biradar and Singh[7], Nayak and Sahoo[8]]. On the other hand, estimation of correlation coefficient although receives considerable attention in many research papers [*cf.*, Gupta and Singh[9], Shevlyakov[10], Roy[11], Shevlyakov and Smirnov[12]], the sample correlation coefficient r still remains as the most discussed estimator of ρ. Hence, our mechanism of estimating $V(t_{RG})$ in this work consists of selecting alternative estimators of S_y^2 in place of s_y^2 and selecting r as the estimator of ρ in the usual way. We assume that both $\bar{X}$ and S_x^2 are known accurately.

* Corresponding author:
lnsahoostatuu@rediffmail.com (L. N. Sahoo)
Published online at http://journal.sapub.org/statistics

2. Formulation of the Estimators

As pointed out earlier, our principal interest here is the estimation of $V(t_{RG})$ with special concentration on the use of an estimator $\hat{S}_y^2$ for S_y^2 that incorporates the available auxiliary information on x. This means that we focus attention on the creation of variance estimators having the following generalized form:

$$\hat{V}(t_{RG}) = \frac{N-n}{Nn}\hat{S}_y^2(1-r^2). \qquad (2.1)$$

This generalized variance estimator $\hat{V}(t_{RG})$ can generate a family of estimators of $V(t_{RG})$ for various selections of $\hat{S}_y^2$. This goal can be achieved by many alternatives ways. But, the variety of finite population variance estimation methods and variety of estimator selection criteria leave us wondering which estimator could be used successfully. The literature to date also offers little guidance in this choice. However, as we are concerned with the variance estimation, we give more stress on the property of non-negativity of an estimator. It means that we do not consider some estimators of S_y^2 that achieve negative values very frequently under repeated sampling from a given population. Let us now present a brief review of the estimators those are taken into account in the present work.

Two notable but simple estimators under classical approach due to Das and Tripathi[1] and Isaki[2] are very much popular in survey practice. These estimators are defined by

$v_1 = s_y^2\bar{X}/\bar{x}$ and $v_2 = s_y^2 S_x^2/s_x^2$,

respectively. It may be mentioned here that when $\hat{S}_y^2 = v_1$, $\hat{V}(t_{RG})$ is reduced to the estimator of Deng and Wu[13].

During the years that followed, emphasis has also been given on the prediction of the population variance using auxiliary information. Under this approach, the survey data are at hand *i.e.*, the sample observations are treated as fixed and unassailable. Uncertainty is then attached only to the unobserved values which need to be predicted. To take the advantage of this criterion, the population U is decomposed into two mutually exclusive domains s and r of n and $N-n$ units respectively, where $r = U - s$ denotes the collection of units in U which are not included in s. Biradar and Singh[7], Nayak and Sahoo[8] considered the following predictive equation to predict S_y^2 developed earlier by Bolfarine and Zacks[6]:

$$(N-1)\hat{S}_y^2 = (n-1)s_y^2 + (N-n-1)V_r + (1-f)n(\bar{y} - M_r)^2, \qquad (2.2)$$

where $f = n/N$, $\bar{X}_r = (N\bar{X} - n\bar{x})/(N-n)$; M_r and V_r are the implied predictors of $\bar{Y}_r = \sum_{i\epsilon r} y_i/(N-n)$ and $S_{y(r)}^2 = \sum_{i\epsilon r}(y_i - \bar{Y}_r)^2/(N-n-1)$ respectively.

For different selections of M_r and V_r, (2.2) can lead to a number of estimators of S_y^2. But, here we report below six estimators *viz*., v_3, v_4, v_5 suggested by Biradar and Singh [7], and v_6, v_7, v_8 suggested by Nayak and Sahoo[8]:

$$v_3 = \left(\frac{N-2}{N-1}\right)s_y^2$$

$$v_4 = \frac{s_y^2}{s_x^2}S_x^2 + \frac{nN(\bar{x}-\bar{X})^2}{(N-n)(N-1)}\left(r^2 - \frac{s_y^2}{s_x^2}\right)$$

$$v_5 = \frac{s_y^2}{s_x^2}S_x^2 + \frac{nN(\bar{x}-\bar{X})^2}{(N-n)(N-1)}\left(b^2 - \frac{s_y^2}{s_x^2}\right)$$

$$v_6 = \left(\frac{n-1}{n}\right)\left(\frac{N}{N-1}\right)s_y^2$$

$$v_7 = \left(\frac{n-1}{N-1}\right)\left[s_y^2 + r^2\left(\frac{N-1}{n-1}S_x^2 - s_x^2\right)\right]$$

$$v_8 = \left(\frac{n-1}{N-1}\right)\left[s_y^2 + b^2\left(\frac{N-1}{n-1}S_x^2 - s_x^2\right)\right].$$

After identifying eight estimators of S_y^2, we then utilize equation (2.1) to produce estimators of $V(t_{RG})$ by substituting an estimator $v_i, i = 1, 2, \ldots, 8$ for $\hat{S}_y^2$. This operation generates eight estimators corresponding to different but specific choices of $\hat{S}_y^2$ as shown in Table 1. To save space, the detail expressions of proposed estimators are not given.

3. Performance of the Proposed Estimators

Generally we judge estimators by their design-based qualities, such as design expectation and design variance, under repeated sampling with a given design from the fixed finite population. The regression estimator is no exception. We are thus interested in evaluating the statistical properties of the eight variance estimators *i.e.*, $v_i(t_{RG}), i = 1, 2, \ldots, 8$, compared to the traditional estimator $v_0(t_{RG})$. But, this cannot be done exactly, because of the complex nature of the considered estimators. However, here we rely on a Monte Carlo simulation in which 5000 independent samples for $n = 6, 8$ and 10 are drawn from 20 populations available in various text books and research papers on survey sampling. The following performance measures of an estimator $v_i(t_{RG})$ $(i = 0, 1, 2, \ldots, 8)$ are taken into consideration:

(i) *Standard Error (SE)* : This performance measure of $v_i(t_{RG})$ is defined by

$$SE(v_i(t_{RG})) = +\sqrt{E[v_i(t_{RG})]^2 - [E(v_i(t_{RG}))]^2},$$

which is a convenient and widely used indicator of the precision attained by the variance estimator.

Table 1. Proposed Estimators of $V(t_{RG})$

Selection of $\hat{S}_y^2$	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8
Estimator of $V(t_{RG})$	$v_1(t_{RG})$	$v_2(t_{RG})$	$v_3(t_{RG})$	$v_4(t_{RG})$	$v_5(t_{RG})$	$v_6(t_{RG})$	$v_7(t_{RG})$	$v_8(t_{RG})$

(ii) *Coverage Rate (CR) Based on 95% Confidence Interval for Estimating* $\bar{Y}$: We consider an approximate 95% confidence interval for $\bar{Y}$ based on t_{RG} and its variance estimator $v_i(t_{RG})$ defined by $t_{RG} \pm 1.96\sqrt{v_i(t_{RG})}$ under the assumption that sampling distribution of t_{RG} is approximately a normal distribution. This performance measure gives us an idea about which percentage of the so constructed confidence intervals based on the variance estimator covers the true value of $\bar{Y}$ under repeated draws of independent samples from a population.

4. Description of the Monte Carlo Simulation

Our Monte Carlo study involves repeated draws of simple random without replacement samples from 20 natural populations. Table 2 presents the source, size (N), definitions of the variables y and x in respect of these populations. 5,000 independent samples, for $n = 6, 8$ and 10, were selected from each population and for each sample numerical values of the comparable estimators were calculated. Then, considering 5,000 such combinations, simulated values of the performance measures *viz*., SE and CR were computed and summarized in Tables 3 and 4. Results for $n = 8$ are not shown, as they confirm more or less the tendencies found in the cases of $n = 6$ and 10. Major findings of the study are discussed in subsections 4.1 and 4.2.

4.1. Results Based on the Standard Error

Simulation results on the SE of the different estimators are provided in Table 3. From these results, as is usually expected, we note that the SE of an estimator diminishes with enlargement of sample size. In respect of this criterion, the performance of $v_6(t_{RG})$ seems to be very poor, and the estimators $v_3(t_{RG}), v_4(t_{RG})$ and $v_5(t_{RG})$ behave very much erratically and there is no clear indication that which one of them would have a decidedly better overall performance than other two estimators. The estimator $v_7(t_{RG})$ emerged out as the best performer as it is decidedly more efficient than the rest of the estimators in 12 and 9 populations for $n = 6$ and 10 respectively, and ranked as second in 6 and 9 populations for $n = 6$ and 10 respectively. However, on this consideration we may choose $v_8(t_{RG})$ and $v_2(t_{RG})$ as the second best and third best performers respectively.

4.2. Results Based on the Coverage Rate

Using several variance estimators, the coverage rates of nominal 95% confidence intervals for $\bar{Y}$ are shown in Table 4. The results on the CR give clear indication of improvement in the performance of an estimator as the sample size increases. The CR of the estimators (except some few cases) usually bears no resemblance to the nominal rates aimed at. The three estimators $v_3(t_{RG}), v_4(t_{RG})$ and $v_5(t_{RG})$ perform equally well. However, on the ground of the achieved CR we may consider $v_7(t_{RG})$, $v_2(t_{RG})$ and $v_8(t_{RG})$ as the best, second best and third best performers respectively.

Table 2. Populations Under Study

Pop. No.	Source	N	y	x
1	Cochran[14] p.152	49	no of inhabitants in 1930	no. of inhabitants in 1920
2	Sukhatme and Sukhatme[15] p.185	34	area under wheat in1937	area under wheat in1936
3	Sukhatme and Sukhatme[15] p.185	34	area under wheat in1937	area under wheat in1931
4	Samford[16] p.61	35	acreage under oats in 1957	acreage of crops and grass in 1947
5	Wetherill[17] p.104	32	percent yield of petroleum	petroleum fraction end point
6	Murthy[18] p.398	43	no of absentees	no of workers
7	Murthy[18] p.399	34	area under wheat in 1964	cultivated area in 1961
8	Murthy[18] p.399	34	area under wheat in 1964	area under wheat in 1963
9	Steel and Torrie[19] p.282	30	leaf burn in secs.	percentage of potassium
10	Shukla[20]	50	fiber yield	height of plant
11	Shukla[20]	50	fiber yield	base diameter
12	Dobson[21] p.83	30	cholesterol	age in years
13	Dobson[21] p.83	30	cholesterol	body mass
14	Yates[22] p.159	25	measured volume of timber	eye estimated volume of timber
15	Yates[22] p.159	43	no. of absentees	total no. of persons
16	Panse and Sukhatme[23] p.118	25	progeny mean	parental plant value
17	Panse and Sukhatme[23] p.118	25	progeny mean	parental plot mean
18	Dobson[21] p.69	20	total calories from carbohydrate	calories as protein
19	Horvitz and Thompson[24]	20	actual no. of households	eye estimated number of households
20	Dobson[21] p.69	20	carbohydrate	body weight

Table 3. Standard Error of the Estimators

n	Pop. No.	$v_0(t_{RG})$	$v_1(t_{RG})$	$v_2(t_{RG})$	$v_3(t_{RG})$	$v_4(t_{RG})$	$v_5(t_{RG})$	$v_6(t_{RG})$	$v_7(t_{RG})$	$v_8(t_{RG})$
6	1	43.34	48.05	36.87	82.44	245.72	242.88	244.33	42.44	108
	2	248.71	170.52	213.54	241.17	127.82	124.04	127.43	123.43	102.57
	3	266.28	224.43	228.63	258.21	404.88	405.12	406.12	213.93	227.66
	4	25.31	41.06	21.71	64.56	109.36	107.2	121.56	24.87	42.32
	5	5.3	5.27	4.56	5.13	102.66	102.67	102.89	1.01	14.87
	6	1.47	1.46	1.25	1.43	2.06	2.06	2.08	0.86	0.87
	7	197.43	197.4	169.51	201.75	447.52	447.52	448.19	191.45	280.43
	8	62.8	49.25	53.92	60.9	63.24	63.55	63.69	52.9	58.9
	9	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0	0
	10	0.07	0.08	0.06	0.07	0.52	0.53	0.56	0.01	0.21
	11	0.05	0.05	0.06	0.05	0.23	0.23	0.24	0.01	0.04
	12	0.05	0.06	0.1	0.05	0.04	0.14	0.14	0.05	0.05
	13	0.08	0.09	0.07	0.08	0.17	0.17	0.18	0.02	0.05
	14	255.31	266	221.62	244.67	179.27	174.07	182.28	148.29	70.51
	15	5.6	5.54	4.78	5.47	7.8	7.79	7.88	8.15	3.3
	16	0.05	0.05	0.08	0.04	0.07	0.06	0.06	0.04	0.02
	17	0.06	0.06	0.15	0.05	0.15	0.05	0.19	0.03	0.03
	18	1.96	1.91	1.72	1.86	3.55	3.59	4.07	1	1.21
	19	1.74	1.56	1.53	1.65	2.51	2.55	2.62	1.26	1.68
	20	2.4	2.5	4.11	2.28	15.33	15.84	18.55	1.26	1.47
10	1	14.85	15.94	13.64	14.54	14.05	14.03	14.03	13.36	17.25
	2	120.77	88.94	111.99	117.11	77.92	74.71	77.05	61.54	92.31
	3	126.26	96.15	117.08	122.44	118.39	120.74	124.2	83.05	93.12
	4	9.51	17.75	8.81	9.23	53.37	52.48	66.81	21.61	25.15
	5	2.21	2.19	0.67	2.13	6.89	6.89	6.91	1.16	3.05
	6	0.8433	0.8016	0.7771	0.8233	0.5673	0.5681	0.5711	0.3298	0.3082
	7	71.03	61.9	85.87	68.88	188.27	188.54	190.42	64.5	85.76
	8	29.31	21.97	27.18	28.42	20.12	20.31	20.4	19.27	17.52
	9	0.0041	0.0044	0.0038	0.0039	0.005	0.005	0.0055	0.0016	0.0015
	10	0.0362	0.0366	0.0332	0.0354	0.0382	0.0382	0.0383	0.0188	0.0084
	11	0.0184	0.0192	0.0169	0.018	0.068	0.0688	0.0813	0.024	0.008
	12	7	0.0214	0.0202	0.0209	0.0298	0.03	0.0303	0.0164	0.0279
	13	0.0331	0.0331	0.0308	0.0319	0.0062	0.063	0.0634	0.0181	0.012
	14	79.35	91.03	74.39	76.04	71.06	63.77	72.63	33.95	65.6
	15	3.32	3.17	3.06	3.25	2.22	2.22	2.23	1.29	1.15
	16	0.0193	0.019	0.0181	0.0226	0.0235	0.0209	0.0226	0.0086	0.0282
	17	0.0178	0.0175	0.0167	0.0171	0.0522	0.0504	0.0759	0.0101	0.012
	18	0.4936	0.478	0.4676	0.4676	0.6296	0.6293	0.6561	0.2822	0.3088
	19	0.5542	0.548	0.525	0.525	0.5715	0.5783	0.5837	0.4858	0.4081
	20	0.697	0.708	0.6603	0.6603	0.6155	0.6188	0.6244	0.3318	0.4229

Table 4. Coverage Rate of the Estimators

n	Pop. No.	$v_0(t_{RG})$	$v_1(t_{RG})$	$v_2(t_{RG})$	$v_3(t_{RG})$	$v_4(t_{RG})$	$v_5(t_{RG})$	$v_6(t_{RG})$	$v_7(t_{RG})$	$v_8(t_{RG})$
6	1	26	26	39	25	54	40	24	39	37
	2	72	74	77	72	76	77	70	77	77
	3	65	65	69	65	61	66	63	69	68
	4	40	45	50	40	52	44	39	51	50
	5	58	58	57	69	26	40	54	69	68
	6	58	59	61	58	58	50	57	61	61
	7	74	74	78	73	69	76	71	78	78
	8	59	61	64	58	63	64	56	64	64
	9	65	65	66	65	36	40	62	67	66
	10	38	37	54	38	24	31	35	54	54
	11	60	61	76	60	52	55	56	77	76
	12	76	76	83	75	95	65	72	84	85
	13	79	79	82	78	64	65	82	83	75
	14	63	67	59	62	64	50	60	60	60
	15	60	61	63	59	59	53	59	63	63
	16	77	83	77	82	93	68	80	83	80
	17	73	73	72	72	61	56	70	73	72
	18	72	72	75	71	66	61	70	75	75
	19	58	59	64	67	65	63	96	74	65
	20	82	82	84	81	78	66	84	85	80
10	1	99	99	99	99	98	98	99	99	100
	2	99	99	98	100	99	99	100	100	98
	3	67	63	67	70	70	71	69	71	71
	4	39	37	43	30	45	45	29	45	30
	5	4	2	17	17	35	35	15	36	17
	6	16	20	28	27	24	24	26	24	28
	7	87	77	80	82	90	90	81	91	83
	8	99	99	99	99	99	99	99	99	99
	9	35	33	60	58	57	57	57	58	59
	10	19	13	33	32	40	40	30	40	33
	11	59	49	67	65	79	79	64	81	66
	12	83	99	94	93	96	95	92	95	93
	13	79	86	96	96	98	98	95	96	96
	14	53	65	67	61	63	60	61	65	62
	15	16	18	28	27	23	23	26	23	27
	16	68	87	79	78	77	75	78	79	76
	17	64	72	80	79	85	85	79	83	80
	18	77	80	86	84	86	85	84	86	85
	19	61	61	58	56	68	68	56	62	57
	20	77	84	89	88	85	85	88	85	89

5. Conclusions

Our Monte Carlo simulation study shows that the estimator $v_7(t_{RG})$ is preferable to its competitors on the grounds of SE and CR. It means that this variance estimator is more efficient than others and can also produce shorter confidence intervals for the population mean. On the other hand, the estimators $v_8(t_{RG})$ and $v_2(t_{RG})$ may be considered as the second best choice on the consideration of SE and CR respectively.

Although the conclusions of this Monte Carlo investigation may not be applicable to all situations, they provide certain guidelines on the overall performance of the variance estimators under consideration. 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] Das, A. and Tripathi, T. P. (1978): Use of auxiliary information in estimating the finite population variance, Sankhyā, C40, 139-148.

[2] Isaki, C. T. (1983): Variance estimation using auxiliary information, Journal of the American Statistical Association, 78, 117-123.

[3] Kadilar, C. and Cingi, H. (2006): Improvement in variance estimation using auxiliary information, Hacettepe Journal of Mathematics and Statistics, 35, 111-115.

[4] Grover, L. K. (2007): A wide and efficient class of estimators of population variance under sub-sampling scheme, Model Assisted Statistics and Applications, 2, 41-51.

[5] Yadav, S. K. (2011): Efficient estimators for population variance using auxiliary information, Global Journal of Mathematical Sciences: Theory and Practical, 3, 369-376.

[6] Bolfarine, H. and Zacks, S. (1992): Prediction Theory for Finite Populations, Springer-Verlag.

[7] Biradar, R. S. and Singh, H. P. (1998): Predictive estimation of finite population variance, Calcutta Statistical Association Bulletin, 48, 229-235.

[8] Nayak, R. K. and Sahoo, L. N. (2012): Some alternative predictive estimators of population variance, Revista Colombiana de Estadistica, 13, 1-13.

[9] Gupta, P. C. and Singh, R. (1978): On the estimation of the finite population correlation coefficient-I. Sankhyā, C40, 38-59.

[10] Shevlyakov, G. L. (1997): On robust estimation of a correlation coefficient, Journal of Mathematical Sciences, 83, 434-438.

[11] Roy, Anuradha (2006): Estimating correlation coefficient between two variables with repeated observations using mixed effects model, Biometrical Journal, 48, 286-301.

[12] Shevlyakov, G. and Smirnov, P. (2011): Robust estimation of the correlation coefficient: An attempt of survey, Austrian Journal of Statistics, 40, 147-156.

[13] Deng, L. Y. and Wu, C. F. J. (1987): Estimation of variance of the regression estimator, Journal of the American Statistical Association, 82, 568-576.

[14] Cochran , W. G. (1977), Sampling Techniques (3rd Edition), Wiley Eastern Limited.

[15] Sukhatme, P. V. and Sukhatme, B. V. (1977): Sampling Theory of Surveys with Applications, Asia Publishing House, New Delhi.

[16] Sampford, M. R. (1962): An Introduction to Sampling Theory, Oliver & Boyd, Edinburg.

[17] Wetherill, G. B. (1981): Intermediate Statistical Methods, Chapman & Hall, London.

[18] Murthy, M. N. (1967): Sampling Theory and Methods, Statistical Publishing Society, Kolkata.

[19] Steel, Robert G. D. and Torrie, James H. (1960): Principles and Procedures of Statistics, Mc. Graw Hill Book Company.

[20] Shukla, G. K. (1966): An alternative multivariate ratio estimate for finite population, Calcutta Statistical Association Bulletin, 15, 127-134.

[21] Dobson, A. J. (1990): An Introduction to Generalized Linear Models, Chapman & Hal, New York.

[22] Yates, F. (1960): Sampling Methods for Censuses and Surveys, Charls & Griffin, London.

[23] Panse, V. G. and Sukhatme, P. V. (1985): Statistical Methods for Agricultural Workers, Indian Council for Agricultural Research, New Delhi.

[24] Horvitz, D. G. and Thompson, D. J. (1952): A generalization of sampling without replacement from a finite universe, Journal of the American Statistical Association, 47, 663-685.