International Journal Of Mathematics And Statistics Invention (IJMSI)
E-ISSN: 2321 – 4767 P-ISSN: 2321 - 4759
Www.Ijmsi.Org || Volume 2 Issue 09 || September. 2014 || PP-32-37

Some Unbiased Classes of Estimators of Finite Population Mean

Praveen Kumar Misra [1] and Shashi Bhushan [2]
1. Department of Statistics, Amity University, Lucknow, India.
2. Department of Applied Statistics,Babasaheb Bhimrao Ambedkar University, Lucknow, India

ABSTRACT: *In this paper, a class of biased multivariate estimators proposed by Abu – Dayyeh et al is reconsidered under the jack – knife technique of Gray and Schucany and Sukhatme et al. The proposed classes of estimators are shown to be unbiased while retaining the optimum mean square error. In order to enhance the practical utility Abu Dayyeh estimators, the class of estimators utilizing the estimated optimum values is also proposed. The comparative study shows that the proposed class of jack – knifed estimators and the proposed class of estimators based on the estimated optimum values are better than the Abu Dayyeh classes of estimators in the sense of unbiasedness and practical utility respectively.*

KEYWORDS: *Unbiasedness, generalized class of estimators and mean square error.*

I. INTRODUCTION

W.G.Cochran and R.J.Jessen[3,4] discussed the use of auxiliary information to increase the precision of estimators. The ratio estimator among the most commonly used estimator of the population mean or total of some variable of interest of a finite population with the help of a auxiliary variable when the correlation coefficient between the two variables is positive. In case of negative correlation, product estimator is used. These estimators are more efficient i.e. has smaller variance than the usual estimator of the population mean based on the sample mean of a simple random sample.

In this paper, mutivariate class of estimators given by Abu – Dayyeh et al are reconsidered under the jack – knife technique of Gray and Schucany and Sukhatme et al in section 2. In a finite population, let y denote the variable whose population mean $\bar{Y}$ is to be estimated by using information of two auxiliary variables x_1 and x_2 .assuming that the population means $\bar{X}_1$ and $\bar{X}_2$ of the auxiliary variable are known the following multivariate class of estimators of the population mean were given by Abu – Dayyeh et al(2003)

$$\bar{y}_h = \bar{y}\prod_{i=1}^{p}\left(\frac{\bar{x}_i}{\bar{X}_i}\right)^{a_i} \qquad (1)$$

where a_i; $i = 1,2,..., p$ are suitably chosen real numbers.

1.1 Definitions and results

Assume that the population means $\bar{X}_i$ of the auxiliary variables are known. Let $\bar{y}, \bar{x}_i$ and respectively denote the sample means of the variables y, x_i and based on a simple random sample without replacement (SRSWOR) of size n drawn from the population

Define:

$$e_0 = \frac{\bar{y} - \bar{Y}}{\bar{Y}}, e_i = \frac{\bar{x}_i - \bar{X}_i}{\bar{X}_i} \text{ ;i=1,2,…………p}$$

Then

$$E(e_0) = E(e_i) = 0$$

$$E(e_0^2) = \frac{f}{n} C_y^2, E(e_i^2) = \frac{f}{n} C_{x_i}^2$$

$$E(e_0 e_i) = \frac{f}{n} \rho_{yx_i} C_y C_{x_i}, E(e_i e_j) = \frac{f}{n} \rho_{x_i x_j} C_{x_i} C_{x_j}$$

Where

$$f = \frac{N-n}{N}, C_y^2 = \frac{S_y^2}{\bar{Y}^2}, C_x^2 = \frac{S_x^2}{\bar{X}^2}, S_y^2 = \left(N-1\right)^{-1} \sum_{i=1}^{N} \left(Y_i - \bar{Y}\right)^2,$$

$$\rho_{yx} = \frac{S_{yx}}{\sqrt{S_y^2 \sqrt{S_x^2}}} S_{yx} = \left(N-1\right)^{-1} \sum_{i=1}^{N} \left(Y_i - \bar{Y}\right)\left(X_i - \bar{X}\right)$$

Also, ρ_{yx_i} and $\rho_{x_i x_j}$ denote the correlation coefficient between Y and X_i , X_i and X_j respectively and $C_y^2, C_{x_i}^2$ denote the coefficient of variation of Y , X_i respectively.

1.2 Properties of the estimator $\bar{y}_h$

Bias of the Abu Dayyeh classes of estimators

$$Bias(\bar{y}_h) = \bar{Y} \frac{\left(1-f\right)}{n} \left\{ \sum_{i=1}^{p} a_i \rho_{yx_i} C_y C_{x_i} + \sum_{i<j=1}^{p} a_i a_j \rho_{x_i x_j} C_{x_i} C_{x_j} + \sum_{i=1}^{p} \frac{a_i \left(a_i - 1\right)}{2} C_{x_i}^2 \right\} \tag{2}$$

MSE of the Abu Dayyeh class of estimators

The MSE of $\bar{y}_h$ *is* given by

$$MSE(\bar{y}_h) = \bar{Y}^2 \frac{\left(1-f\right)}{n} \left\{ C_y^2 + \sum_{i=1}^{p} a_i^2 C_{x_i}^2 + 2\sum_{i=1}^{p} a_i \rho_{yx_i} C_y C_{x_i} + 2\sum_{i<j=1}^{p} a_i a_j \rho_{x_i x_j} C_{x_i} C_{x_j} \right\}$$

$$= \bar{Y}^2 \frac{\left(1-f\right)}{n} \left\{ C_y^2 + \sum_{i=1}^{p}\sum_{j=1}^{p} a_i a_j \rho_{x_i x_j} C_{x_i} C_{x_j} + 2\sum_{i=1}^{p} a_i \rho_{yx_i} C_y C_{x_i} \right\}$$

$$= \frac{\left(1-f\right)}{n} S_y^2 \left\{1 + b' A b + 2b' e\right\} \tag{3}$$

where $A = (a_{ij}) = \left(\rho_{ij} \dfrac{C_i C_j}{C_y^2} \right),\ b'_{1Xp} = (a_1, \ldots, a_p)$

$e'_{1Xp} = (e_1, e_2, ..., e_p),\ e_i = \rho_{yi} C_i / C_y$

The optimum value minimizing the MSE is $b = -e' A^{-1}$.

The minimum MSE of the class of estimators is given by

$$MSE(\bar{y}_h)_{\min} = \bar{Y}^2 \frac{(1-f)}{n} C_y^2 \left(1 - \rho_{y.12\ldots p}^2\right)$$

where $\rho_{y.12\ldots p}^2$ is the multiple correlation coefficient.

II. THE PROPOSED JACK-KNIFE ESTIMATOR $\bar{y}_{hj}$

Let a simple random sample of size n=2m is drawn without replacement from the population of size N. this sample of size n=2m is then split up at random in to two sub samples each of size m

$$\bar{y}_h = \bar{y} \prod_{i=1}^{p} \left(\frac{\bar{x}_i}{\bar{X}_i} \right)^{a_i} \tag{4}$$

Let us define $\bar{y} = \bar{Y}(1+e_0),\ \bar{x}_i = \bar{X}_i(1+e_i)$

Substituting these values and simplifying we get

$$Bias(\bar{y}_h) = \bar{Y} \frac{(1-f)}{n} \left\{ \sum_{i=1}^{p} a_i \rho_{yx_i} C_y C_{x_i} + \sum_{i<j=1}^{p} a_i a_j \rho_{x_i x_j} C_{x_i} C_{x_j} + \sum_{i=1}^{p} \frac{a_i (a_i - 1)}{2} C_{x_i}^2 \right\} = B_1(\text{say}) \tag{5}$$

Using the jack - knife technique we propose to estimate the population mean by

$$\bar{y}_{hJ} = \frac{\bar{y}_h^3 - R\bar{y}_h'}{1-R} \tag{6}$$

$$\text{where } R = \frac{B_1}{B_2};\ B_1 = Bias\left(\bar{y}_h^3\right) \text{ and } B_2 = Bias\left(\bar{y}_h'\right) \tag{7}$$

$\bar{y}_h^3 =$ estimator based on sample of size $n = 2m$

$$\bar{y}_h' = \frac{\bar{y}_h^1 + \bar{y}_h^2}{2} = \text{pooled estimator based on sub samples of size } m \text{ each} \tag{8}$$

Now,

$$E(\bar{y}_{hj}) = \frac{E(\bar{y}_h) - RE(\bar{y}_h')}{1-R} \tag{9}$$

$= \bar{Y}$ (i.e $\bar{y}_{hj}$ is an unbiased estimator of population mean $\bar{Y}$) (10)

2.1 Mean square error of $\bar{y}_{aj}$

$$MSE(\bar{y}_{hj}) = E(\bar{y}_{hj} - \bar{Y}) = E\left\{ \frac{\bar{y}_h - R\bar{y}_h'}{1-R} - \bar{Y} \right\}^2$$

$$= E\left\{\frac{(\bar{y}_h - \bar{Y}) - R(\bar{y}'_h - \bar{Y})}{1-R}\right\}^2$$

$$= \frac{E(\bar{y}_h - \bar{Y})^2 + R^2 E(\bar{y}'_h - \bar{Y}) - 2R(\bar{y}_h - \bar{Y})(\bar{y}'_h - \bar{Y})}{(1-R)^2} \qquad (11)$$

$$MSE(\bar{y}_h) = E(\bar{y}_h - \bar{Y})^2 = \bar{Y}\,\frac{(1-f)}{n}\left\{\sum_{i=1}^{p} a_i \rho_{yx_i} C_y C_{x_i} + \sum_{i<j=1}^{p} a_i a_j \rho_{x_i x_j} C_{x_i} C_{x_j} + \sum_{i=1}^{p} \frac{a_i(a_i-1)}{2} C_{x_i}^2\right\}$$

$$= \bar{Y}^2 \frac{f}{n} A \qquad (12)$$

where A= $\bar{Y}\,\frac{(1-f)}{n}\left\{\sum_{i=1}^{p} a_i \rho_{yx_i} C_y C_{x_i} + \sum_{i<j=1}^{p} a_i a_j \rho_{x_i x_j} C_{x_i} C_{x_j} + \sum_{i=1}^{p} \frac{a_i(a_i-1)}{2} C_{x_i}^2\right\}$

The MSE of the proposed class of jack-knifed estimators is given by

$$MSE(\bar{y}_h) = \frac{(1-f)}{n} S_y^2 \{1 + b'Ab + 2b'e\}$$

which can again be minimized so that the minimum MSE is

$$MSE(\bar{y}_{hJ*})_{\min} = \bar{Y}^2 \frac{(1-f)}{n} C_y^2 \left(1 - \rho_{y.12\ldots p}^2\right) \qquad (13)$$

$$= MSE(\bar{y}_h)_{\min}$$

where $A = (a_{ij}) = \left(\rho_{ij} \frac{C_i C_j}{C_y^2}\right)$, $b'_{1\mathrm{X}p} = (a_1, \ldots, a_p)$, $e'_{1\mathrm{X}p} = (e_1, e_2, \ldots, e_p)$, $e_i = \rho_{yi} C_i / C_y$

The optimum value minimizing the MSE is $b = -e'A^{-1}$.

The minimum MSE of the class of estimators is given by

$$MSE(\bar{y}_h)_{\min} = \bar{Y}^2 \frac{(1-f)}{n} C_y^2 \left(1 - \rho_{y.12\ldots p}^2\right) \qquad (14)$$

where $\rho_{y.12\ldots p}^2$ is the multiple correlation coefficient.

III. COMPARISON OF THE ESTIMATORS

In this section, we compare the proposed estimators with some known estimator. The comparison will be in terms of the bias and the mean square error up to the order of n^{-1}. We consider the following known estimators.

1. The SRS mean $\bar{y}$. It is unbiased estimator and its variance is given by

Var($\bar{y}$)= $\frac{f}{n} S_y^2$

2. The ratio estimator

$$\bar{Y}_R = \bar{y}\frac{\bar{X}}{\bar{x}}$$

Where x may be chosen as x_1 or x_2.

The bias and the mean square error of this estimator are respectively given by

$$B(\bar{y}_R) = \bar{Y}\frac{f}{n}C_x(C_x - \rho C_y)$$

$$MSE(\bar{y}_R) = \bar{Y}^2\frac{f}{n}(C^2{}_y + C^2{}_x - 2\rho C_y C_x)$$

3. Product estimator

$$\bar{Y}_R = \bar{y}\frac{\bar{x}}{\bar{X}}$$

The bias and the mean square error of this estimator x_1 or x_2 where x may be chosen as are respectively given by

$$B(\bar{y}_P) = \bar{Y}\frac{f}{n}\rho C_y C_x$$

4. The estimator $\bar{y}_a = \bar{y}\left(\frac{\bar{x}}{\bar{X}}\right)^a$ where x may be chosen as x_1 or x_2.

5. If $a_1= a_2=1$, then the estimator reduces to

$$\bar{y} = \bar{y}\frac{\bar{x}_1}{\bar{X}_1}\frac{\bar{x}_2}{\bar{X}_2}$$

6. If $a_1= a_2=-1$ then the estimator (1) reduces to

$$\bar{y} = \bar{y}\frac{\bar{X}_1}{\bar{x}_1}\frac{\bar{X}_2}{\bar{x}_2}$$

The proposed jack-knife estimators $\bar{y}_{hj}$ are unbiased and mean square error less or equal as compared with the above estimators.

IV. NUMERICAL ILLUSTRATIONS

The comparison among these estimators is given by using a real data set.
The data for this illustration has been taken from [1], District Handbook of Aligarh, India. The population that we like to study contains 332 villages. We consider the variables Y, X_1, X_2 where Y is the number of cultivators, X_1 is the area of villages and X_2 is the number of household in a village. Throughout this study, two types of calculations are used. The first type depends on the population data, and the second type depends on simulation study of 30,000 repeated samples without replacement of sizes 80 from the data. We compute the bias and the MSE for all estimators.

Numerical computation
The following values were obtained using the whole data set:

$\bar{Y}$ =1093.1, $\overline{X_1}$ =181.57, $\overline{X_2}$ =143.31

C_y =0.7625, C_{x_1} =0.7684, C_{x_2} =0.7616

ρ_{yx_1} =0.973, ρ_{yx_2} =0.862, $\rho_{x_2x_1}$ =0.842

Using the above results we calculated the MSE, bias, and the efficiency for all the estimators in section-4.
In the Table1 , we present the MSE, efficiency, and the bias all for the estimators given in Section 4. As we can see that our suggested estimators $\overline{y}_{hj}$ and dominate all other estimators with efficiency 24.51 . They are superior over all other estimators and are unbiased.

Table1: Bias and MSE of estimators under study based on population data

Estimator	Auxiliary variable	MSE	Efficiecy	Bias
$\overline{y}$	none	11807.6	1	0
$\overline{y}_{a^*}$	x_1, x_2	482.337	24.5098	0.265
$\overline{y}_{w^*}$	x_1, x_2	620.5	18.1818	-0.40110
$\widehat{y}_a$	x_1	3118.804	3.7878	0.78127
$\widehat{y}_a$	x_2	1629.45	7.2463	0.63602
$\overline{y}_R$	x_1	6143.05	1.9223	0.3753
$\overline{y}_R$	x_2	3254.39	3.6284	1.4736
$\overline{y}_P$	x_1	46942.5	0.2520	10.589
$\overline{y}_P$	x_2	43910.4	0.2689	9.2983
$\overline{y}_{aj}$	x_1, x_2	99050.6	0.1192	0
$\overline{y}_{wj}$	x_1, x_2	12095.1	0.9765	0
$\overline{y}_{MR}$	x_1, x_2	655.2	18.0214	9.2983

REFERENCES

[1] W.A.Abu-Dayyeh et al. ,Some estimators of a finite population mean using auxiliary information, Applied Mathematics and Computation 139 (2003) 287-298.
[2] Jack – knife technique of Gray and Schucany and Sukhatme et al.
[3] W.G.Cochran, Sampling Techniques, third ed. John Wiley &sons, New York, 1977.
[4] R.S.Jessen, Statistical Survey Techniques, John Wiley &sons, New York, 1978.
[5] D.Singh,F.S.Chaudhary, Theory and Analysis of sample Survey Design, New Age Publication, New Delhi, India 1986.
[6] S.K.Srivastava, An estimator using auxiliary information in sample surveys, Calcutta Statistical Association Bulletin 16(1967) 121-132.