DOI: 10.15415/mjis.2015.41002

Generalised Synthetic Estimator Using Double Sampling Scheme and Auxiliary Information

SHASHI BAHL[1] AND SANGEETA[2]

[1]Former Head, Department of Statistics, M.D University, Rohtak

[2]Research scholar, Department of statistics, M.D. University, Rohtak

Email: sangeetpuja@gmail.com

Received: December 24, 2014| Revised: June 28, 2015| Accepted: July 19, 2015

Published online: September 30, 2015

Abstract In Small Area Estimation there are two type of estimation Direct and Indirect Estimation. The word ‘Synthetic Estimation’ was coined by National Centre for Health Statistics (1968) for the method of estimation. The term ‘Synthetic’ was used because these estimates were not derived directly from survey results. So ‘Synthetic Estimation’ is indirect method of estimation. If we want to use auxiliary information for estimation and the information is not available, this leads to the synthetic estimation under double sampling scheme. In this paper we propose a quadratic synthetic estimator and generalised ratio and product synthetic estimator under this scheme and compare their efficiency.

Keywords: ratio; product; quadratic and ratio - cum - product synthetic estimator under double sampling scheme; sampling bias; mean square error and small area domain.

1. INTRODUCTION

For any sample design let a finite population U; (1.......i.......N) divided into ‘A’ non overlapping domains. U_a of size N_a;(a=1.............A) and a sample is divided into two non overlapping domains S_a and S_a of n_a and n_a; (a=1............A)

$$\sum_{a=1}^{N} N_a = N \qquad \sum_{a=1}^{A} n_{a'} = n' \qquad \sum_{a=1}^{A} n_a = n$$

Further let $\bar{Y}$ and $\bar{X}$ are population mean for the main and the auxiliary variables. $\bar{Y}_a$ and $\bar{X}_a$ are the population mean for the main and the auxiliary

Mathematical Journal of Interdisciplinary Sciences
Vol. 4, No. 1,
September 2015
pp. 15–21

variables for the small area 'a'. $\bar{y}$, $\bar{x}'$ and $\bar{x}$ re sample mean for the main ,auxiliary variable for large sample and sub sample.

$\overline{y_a}$, $\overline{x_{a'}}$ and $\overline{x_a}$ are sample mean for the main and auxiliary variable for the large sample and sub sample of small area 'a'.

When information related to auxiliary variable is not available, then double sampling scheme is used for constructing synthetic estimator.

Under double sampling scheme 'n' is the sub sample of 'n'' the large sample. Here, we took the sample mean of size $\bar{y}$;(1............N) , $\overline{x_{a'}}$;(1.............n') and $\bar{x}$;(1............n) . The various mean squares and coefficients of variations of the population and small area 'a' for main and auxiliary variable are y, x' and x.

We introduce the usual estimator ratio synthetic estimator and product synthetic estimator under double sampling scheme.

1.1 Ratio Synthetic Estimator using double sampling:

$$\bar{y}_{syn,rd,a} = \bar{y}\left(\frac{\overline{x_{a'}}}{\bar{x}}\right)$$

If we use the synthetic assumption

$$\bar{Y}_a = \bar{Y} \quad \text{and} \quad \bar{Y}_a\,\bar{X}_a = \bar{Y}\,\bar{X} \tag{1.1}$$

$$B(\bar{y}_{syn,rd,a}) = \bar{Y}_a\left(\frac{1}{n}-\frac{1}{n'}\right)\left(\frac{S_x^2}{\bar{X}^2}-\frac{S_{xy}}{\bar{X}\bar{Y}}\right)$$

And

$$M.S.E.(\bar{y}_{syn,rd,a}) = \left(\frac{1}{n'}-\frac{1}{N}\right)S_y^2 + \left(\frac{1}{n}-\frac{1}{n'}\right)\left(S_y^2+R^2S_x^2-2R\rho S_xS_y\right) \tag{1.2}$$

1.2 Product Synthetic Estimator using double sampling:

$$\bar{y}_{syn,pd,a} = \bar{y}\left(\frac{\bar{x}}{\overline{x_{a'}}}\right)$$

After using synthetic assumption (1.1), then

$$B(\bar{y}_{syn,pd,a}) = \bar{Y}_a\left(\frac{1}{n}-\frac{1}{N}\right)\left(\frac{S_{xy}}{\bar{X}\bar{Y}}\right)$$

And

$$M.S.E.(\bar{y}_{syn,pd,a}) = \left(\frac{1}{n'} - \frac{1}{N}\right) S_y^2 + \left(\frac{1}{n} - \frac{1}{n'}\right) \left(S_y^2 + R^2 S_x^2 + 2R\rho S_x S_y\right) \quad (1.3)$$

Here, we propose the modify ratio and product synthetic estimator.

2. QUADRATIC SYNTHETIC ESTIMATOR USING DOUBLE SAMPLING:

First, we propose the quadratic synthetic estimator as

$$\bar{y}_{syn,qd,a} = \bar{y} \left(\frac{\bar{x}}{\bar{\bar{x}}_{a'}}\right)^2 \quad (2.1)$$

Here , we took the various properties like as mean square and coefficients of variations of the population and small area 'a' for main and auxiliary variable for the sample and sub- sample y, x' and x for proposed estimators.
Let

$$\bar{y} = \bar{Y}(1+e) \ , \ \bar{x} = \bar{X}(1+e_1) \text{ and } \bar{x}_{a'} = \bar{X}_a(1+e_2) \quad (2.2)$$

$$E(e) = E(e_1) = E(e_2) = 0 \quad (2.3)$$

If we use the equation (2.2), then the sampling bias and mean square error of this estimator is as

$$\bar{y}_{syn,qd,a} = \bar{Y}(1+e) \left(\frac{\bar{X}(1+e_1)}{\bar{X}_a(1+e_2)}\right)^2$$

$$B(\bar{y}_{syn,qd,a}) = \bar{Y}_a \left(\frac{1}{n} - \frac{1}{n'}\right)\left(\frac{S_x^2}{\bar{X}^2} + 2\frac{S_{xy}}{\bar{X}\bar{Y}}\right) \quad (2.4)$$

($\therefore$ In equation (2.4), neglect the terms of order greater than two and use the assumption (1.1)) and

$$M.S.E.(\bar{y}_{syn,qd,a}) = E\left[\bar{y}_{syn,qd,a} - \bar{Y}_a\right]^2$$

($\therefore$ We use the assumption (1.1) and equation (2.2) then neglect the terms order greater than one)

$$M.S.E.\ (\bar{y}_{syn,qd,a}) = \bar{Y}_a^2\ E\left(e + 2e_1 - 2e_2\right)^2$$

($\therefore$ by equation (2.3))

$$M.S.E.\ (\bar{y}_{syn,qd,a}) = \left(\frac{1}{n'} - \frac{1}{N}\right) S_y^2 + \left(\frac{1}{n} - \frac{1}{n'}\right)\left(S_y^2 + 4R^2 S_x^2 + 4R\rho S_x S_y\right) \quad (2.5)$$

3. GENERALISED RATIO - CUM - PRODUCT SYNTHETIC ESTIMATOR USING DOUBLE SAMPLING:

We propose the second generalised ratio – cum – product synthetic estimator as

$$\bar{y}_{syn,rpd,a} = \bar{y}\left[\left(\frac{\bar{x}}{\bar{x}_{a'}}\right)^{1+\beta} + \left(\frac{\bar{x}}{\bar{x}_{a'}}\right)^{\beta-1}\right] \Big/ 2 \quad , \quad -1<\beta<1 \qquad (3.1)$$

Here the cases arise according to value of ‘β’ are as

(i)
$$\bar{y}_{syn,rpd,a} = \bar{y}\left[\left(\frac{\bar{x}}{\bar{x}_{a'}}\right) + \left(\frac{\bar{x}_{a'}}{\bar{x}}\right)\right] \Big/ 2 \quad , \text{ for } \beta = 0$$
$$= \left[\bar{y}_{syn,pd,a} + \bar{y}_{syn,rd,a}\right] \Big/ 2$$

(ii)
$$\bar{y}_{syn,rpd,a} = \bar{y}\left[1 + \left(\frac{\bar{x}_{a'}}{\bar{x}}\right)^2\right] \Big/ 2 \quad , \text{ for } \beta = -1$$

(iii)
$$\bar{y}_{syn,rpd,a} = \bar{y}\left[\left(\frac{\bar{x}}{\bar{x}_{a'}}\right)^2 + 1\right] \Big/ 2 \quad , \text{ for } \beta = 1$$

Here, we find the sampling bias and mean square error of estimator (3.1) using equation (2.2)

$$B(\bar{y}_{syn,rpd,a}) = E\left(\bar{y}_{syn,rpd,a}\right) - \bar{Y}_a$$

(Neglect terms of order greater than two)
Here, we using the equation (1.1) & (2.3), then

$$B\left(\bar{y}_{syn,rpd,a}\right) = \frac{\bar{Y}_a}{2}\left(\frac{1}{n}-\frac{1}{n'}\right)\left(2\beta\frac{\rho S_x S_y}{\bar{X}\bar{Y}}+\left(\beta^2-\beta+1\right)\frac{S_x^2}{\bar{X}^2}\right) \qquad (3.2)$$

And

$$M.S.E.\left(\bar{y}_{syn,rpd,a}\right) = \mathrm{E}\left[\bar{y}_{syn,rpd,a}-\bar{Y}_a\right]^2$$

$$M.S.E.\left(\bar{y}_{syn,rpd,a}\right) = \bar{Y}_a^2\ \mathrm{E}\left[e+\beta e_1-\beta e_2\right]^2$$

(After using assumption (1.1) and equation (2.2) then neglect the terms of order greater than one and use equation (2.3))

$$M.\mathrm{S.E.}\left(\bar{y}_{syn,rpd,a}\right) = \left(\frac{1}{n'}-\frac{1}{N}\right)\mathrm{S}_y^2+\left(\frac{1}{n}-\frac{1}{n'}\right)\left(S_y^2+\beta^2R^2S_x^2+2\beta R\rho S_xS_y\right) \qquad (3.3)$$

3.1. Optimum Generalised Estimator

For finding the optimum generalised estimator, minimising the equation (3.3) w . r . t . β , we have

$$\beta = -\frac{\rho S_y}{RS_x}$$

For finding the minimum variance of generalised ratio - cum – product synthetic estimator , put the value of ‘β’ in equation (3.3) then minimum variance is as

$$Mini.\ \mathrm{V}\left(\bar{y}_{syn,rpd,a}\right) = \left(\frac{1}{n'}-\frac{1}{N}\right)S_y^2+\left(\frac{1}{n}-\frac{1}{n'}\right)\left(1-\rho^2\right)S_y^2 \qquad (3.4)$$

4. COMPARISON OF PROPOSED ESTIMATORS WITH USUAL ESTIMATORS:

(i) Quadratic synthetic estimator is better than ratio synthetic estimator under double sampling, when

$$V(\bar{y}_{syn,qd,a}) < V(\bar{y}_{syn,rd,a})$$

$$\rho\frac{S_y}{S_x} < -\frac{R}{2} \qquad (4.1)$$

(ii) Quadratic synthetic estimator is better than product synthetic estimator under double sampling, when

$$V(\bar{y}_{syn,qd,a}) < V(\bar{y}_{syn,pd,a})$$
$$\rho \frac{S_y}{S_x} < -\frac{3}{2} R \tag{4.2}$$

(iii) Generalised ratio - cum - product synthetic estimator is better than ratio synthetic estimator under double sampling, when

$$V(\bar{y}_{syn,rpd,a}) < V(\bar{y}_{syn,rd,a})$$
$$\rho \frac{S_y}{S_x} < -\frac{\beta - 1}{2} R \tag{4.3}$$

(iv) Generalised ratio - cum - product synthetic estimator is better than product synthetic estimator under double sampling, when

$$V(\bar{y}_{syn,rpd,a}) < V(\bar{y}_{syn,pd,a})$$
$$\rho \frac{S_y}{S_x} < -\frac{\beta + 1}{2} R \tag{4.4}$$

(v) Generalised ratio - cum - product synthetic estimator is better than Quadratic synthetic estimator under double sampling, when

$$V(\bar{y}_{syn,rpd,a}) < V(\bar{y}_{syn,qd,a})$$
$$\rho \frac{S_y}{S_x} < -\frac{\beta + 2}{2} R \tag{4.5}$$

5. CONCLUSION

- After comparing proposed estimator under Double Sampling with usual estimators such as ratio synthetic estimator and product synthetic estimator, we find that the Generalised ratio-cum-product synthetic estimator and quadratic synthetic estimator are more efficient than usual estimators. But according to the value of β, results are change.
- When β=0, then the generalised ratio-cum-product synthetic estimator under double sampling becomes simple synthetic estimator.
- When β=-1, then the generalised ratio-cum-product synthetic estimator and the ratio synthetic estimator both are equal efficient under sampling.

- When β=1, then the generalised ratio-cum-product synthetic estimator and product synthetic estimator both are equal efficient under double sampling.
- When β=2, then the generalised ratio-cum-product synthetic estimator and quadratic synthetic estimator are equal efficient. Otherwise the generalised ratio-cum-product synthetic estimator is more efficient than quadratic synthetic estimator under double sampling.

REFERENCES:

[1] Chochran, W.G. (1963): Sampling Techniques, John Wiley and Sons inc., N.Y.

[2] Ghangurde, P.D. and Singh, M.P. (1977). Synthetic Estimators in Periodic Household Surveys, Survey Methodology, **3**, (151-181)

[3] Gonzalez, M.E and Waksberg, j.(1973), Estimation of the error of Synthetic Estimates, Paper Presented at the first meeting of the International association of Survey Statisticians, Vienna, Austria.

[4] Rao, J.N.K. (2003) Small Area Estimation. Wiley, New York. http://dx.doi.org/10.1002/0471722189