European Journal of Statistics and Probability
Vol.4, No.4, pp.12-17, August 2016
Published by European Centre for Research Training and Development UK (www.eajournals.org)

AN IMPROVED RATIO ESTIMATOR FOR POPULATION MEAN IN STRATIFIED RANDOM SAMPLING

Etebong P. Clement
Department of Mathematics and Statistics, University of Uyo, Uyo, Nigeria

ABSTRACT: *This paper suggests an improved ratio estimator in stratified random sampling. Analytically, the bias and mean square error (MSE) expressions for the proposed estimator are obtained. Efficiency comparisons are made to evaluate the relative performance of the proposed estimator. In addition, an empirical study is provided to support the analytical study. Analytical and numerical results showed that the proposed estimator is more efficient than the estimator under study.*
KEYWORDS: analytical study, efficiency comparisons, empirical study, ratio estimator, relative performance, stratified sampling.
Mathematics Subject Classification: 62D05, 62G05, 62H12

INTRODUCTION

Hansen, Hurwitz and Gurney (1946) proposed the combined ratio estimator for $(\bar{y})$ in SRSWOR as:

$$\bar{y}_{RC} = \frac{\bar{y}}{\bar{x}} \bar{X} \qquad (1)$$

With bias expression given by:

$$B(\bar{y}_{RC}) = \left(\frac{1-f}{n}\right) \bar{Y}\, C_x^2\, (1-K) \qquad (2)$$

and MSE expression given by:

$$MSE(t(\alpha,\delta)) = \left(\frac{1-f}{n}\right) \bar{Y}^2 \left[C_y^2 + C_x^2 (1-2K)\right] \qquad (3)$$

Also in SRSWOR Solanki, Singh and Rathour (2012) proposed a ratio-type estimator for $(\bar{y})$ as:

$$t(\alpha,\delta) = \bar{y} \left\{2 - \left(\frac{\bar{x}}{\bar{X}}\right)^{\alpha} exp\left[\frac{\delta(\bar{x}-\bar{X})}{(\bar{x}+\bar{X})}\right]\right\} \qquad (4)$$

With bias expression given by:

$$E(t(\alpha,\delta) - \bar{Y}) = -\frac{3}{8}\left(\frac{1-f}{n}\right) \bar{Y}\, C_x^2\, (1+4K) \qquad (5)$$

and MSE expression given by:

$$MSE(t(\alpha,\delta)) = \left(\frac{1-f}{n}\right) \bar{Y}^2 \left[C_y^2 + \frac{3C_x^2}{4}(3-4K)\right] \qquad (6)$$

Let

$$\left.\begin{aligned} K &= (\rho_{xy} C_y / C_x) \\ C_y^2 &= (S_y^2 / \bar{Y}^2) \\ C_x^2 &= (S_{hx}^2 / \bar{X}^2) \\ S_{xy} &= \rho_{xy} S_x S_y \end{aligned}\right\} \qquad (7)$$

Using equation (7), the bias and MSE in equations (5) and (6) respectively becomes

ISSN 2055-0154(Print), ISSN 2055-0162(Online)

$$B\big(t(\alpha,\delta)\big) = -\frac{3}{8}\left(\frac{1-f}{n}\right)\frac{1}{\bar{X}}\left(RS_x^2 - 4S_{xy}\right) \qquad (8)$$

$$MSE\big(t(\alpha,\delta)\big) = \left(\frac{1-f}{n}\right)\left(S_y^2 + \frac{9}{4}R^2S_x^2 - 3RS_{xy}\right) \qquad (9)$$

In sample surveys the scientific technique for selecting a sample is that of selecting a probability sample that is usually based upon a stratification of the population. It is well known that stratification is one of the design tools that give increased precision. In the progression for improving the performance of the ratio estimators, authors have proposed various improved ratio-type estimators in stratified sampling. Notably among them include Chaudhary et al (2009), Koyuncu and Kadilar (2009), Vishwakarma, Singh and Singh (2010), Tailor, Sharma and Kim (2011), Vishwakarma and Singh (2011), Malik and Singh (2012) and Clement and Enang (2015).

Keeping this in view, this paper suggests an improved ratio estimator for population mean in stratified random sampling based on Solanki et al (2012) ratio estimator.

The suggested estimator

Let rewrite equation (4) as:

$$t(\alpha,\delta) = \lambda\bar{y} \qquad (10)$$

where the coefficient $\lambda = \left\{2 - \left(\frac{\bar{x}}{\bar{X}}\right)^{\alpha} exp\left[\frac{\delta(\bar{x}-\bar{X})}{(\bar{x}+\bar{X})}\right]\right\}$

In stratified random sampling, this estimator is proposed as:

$$t_{st}(\alpha_h,\delta_h) = \lambda^*\bar{y}_h \qquad (11)$$

Bias expression for the proposed estimator

Let

$$t_{st}(\alpha_h,\delta_h) - \bar{Y} = \lambda^*\bar{y}_h - \bar{Y} \qquad (12)$$

Taking expectation of equation (12) gives

$$E(t_{st}(\alpha_h,\delta_h) - \bar{Y}) = E(\lambda^*\bar{y}_h - \bar{Y})$$

$$E(t_{st}(\alpha_h,\delta_h) - \bar{Y}) = \bar{X}E(\lambda^*\hat{R} - R)$$

$$E(t_{st}(\alpha_h,\delta_h) - \bar{Y}) = \bar{X}E\left(\frac{(\lambda^*\bar{y}_h - R\bar{x}_h)}{\bar{x}_h}\right)$$

But

$$\frac{1}{\bar{x}_h} = \frac{1}{\bar{X}}\left(1 + \frac{\bar{x}_h - \bar{X}}{\bar{X}}\right)^{-1}$$

And by Taylor's series expansion of the first order approximation

$$\frac{1}{\bar{x}_h} = \frac{1}{\bar{X}}\left(1 - \frac{\bar{x}_h - \bar{X}}{\bar{X}}\right)$$

So that

$$E(t_{st}(\alpha_h,\delta_h) - \bar{Y}) = E\left((\lambda^*\bar{y}_h - R\bar{x}_h)\left(1 - \frac{\bar{x}_h - \bar{X}}{\bar{X}}\right)\right)$$

$$E(t_{st}(\alpha_h,\delta_h) - \bar{Y}) = E\left((\lambda^*\bar{y}_h - R\bar{x}_h) - \lambda^*\bar{y}_h\left(\frac{\bar{x}_h - \bar{X}}{\bar{X}}\right) + R\bar{x}_h\left(\frac{\bar{x}_h - \bar{X}}{\bar{X}}\right)\right)$$

$$E(t_{st}(\alpha_h,\delta_h) - \bar{Y}) = \lambda^*E(\bar{y}_h) - RE(\bar{x}_h) - \frac{\lambda^*}{\bar{X}}E(\bar{y}_h(\bar{x}_h - \bar{X})) + \frac{R}{\bar{X}}E(\bar{x}_h(\bar{x}_h - \bar{X}))$$

ISSN 2055-0154(Print), ISSN 2055-0162(Online)

$$E(t_{st}(\alpha_h,\delta_h) - \bar{Y}) = \lambda^*\bar{Y} - R\bar{X} - \frac{\lambda^*}{\bar{X}}E((\bar{x}_h - \bar{X})(\bar{y}_h - \bar{Y})) + \frac{R}{\bar{X}}E(\bar{x}_h - \bar{X})^2$$

$$E(t_{st}(\alpha_h,\delta_h) - \bar{Y}) = \lambda^*\bar{Y} - R\bar{X} - \frac{\lambda^*}{\bar{X}}E((\bar{x}_h - \bar{X})(\bar{y}_h - \bar{Y})) + \frac{R}{\bar{X}}E(\bar{x}_h - \bar{X})^2$$

$$E(t_{st}(\alpha_h,\delta_h) - \bar{Y}) = (\lambda^* - 1)\bar{Y} + \sum_{h=1}^{H} w_h\gamma_h\bar{Y}_hC_{hx}^2\left[-\frac{3}{2}\left(\frac{1+\lambda^*4K_h}{4}\right)\right]$$

$$E(t_{st}(\alpha_h,\delta_h) - \bar{Y}) = (\lambda^* - 1)\bar{Y} - \frac{3}{8}\sum_{h=1}^{H} w_h\gamma_h\bar{Y}_hC_{hx}^2(1+\lambda^*4K_h) \qquad (13)$$

Applying equation (7) in equation (13) gives the bias as:

$$B(t_{st}(\alpha_h,\delta_h)) = (\lambda^* - 1)\bar{Y} - \frac{3}{8}\left[\frac{1}{\bar{X}}\sum_{h=1}^{H} w_h\gamma_h\left(RS_{hx}^2 - 4\lambda^*S_{hxy}\right)\right]$$

Mean square error (MSE) expression for the proposed estimator

Squaring both sides of (12) and taking expectation gives

$$MSE(t_{st}(\alpha_h,\delta_h)) = E(\lambda^*\bar{y}_h - \bar{Y})^2 \qquad (14)$$
$$= E(\lambda^{*2}\bar{y}_h^2 - 2\lambda^*\bar{y}_h\bar{Y} + \bar{Y}^2)$$
$$= \lambda^{*2}E(\bar{y}_h^2) - 2\lambda^*\bar{Y}^2 + \bar{Y}^2 + \lambda^{*2}\bar{Y}^2 - \lambda^{*2}[E(\bar{y}_h)]^2$$
$$= \lambda^{*2}[E(\bar{y}_h^2) - [E(\bar{y}_h)]^2] + \bar{Y}^2(\lambda^* - 1)^2$$
$$= \lambda^{*2}\,Var(\bar{y}_h)^2 + \bar{Y}^2(\lambda^* - 1)^2 \qquad (15)$$

$$MSE(t_{st}(\alpha_h,\delta_h)) = \lambda^{*2}\sum_{h=1}^{L} W_h^2\gamma_h\,\bar{Y}^2\left[C_{hy}^2 + \frac{3C_{hx}^2}{4}(3 - 4K_h)\right] + \bar{Y}^2(\lambda^* - 1)^2$$

Applying equation (7) gives

$$MSE\left(t_{st}(\alpha_h,\delta_h)\right) = \lambda^{*2}\sum_{h=1}^{L} W_h^2\gamma_h\left(S_{hy}^2 + \frac{9}{4}R^2S_{hx}^2 - 3RS_{hxy}\right) + \bar{Y}^2(\lambda^* - 1)^2 \quad (16)$$

Minimizing equation (16) with respect to λ^* and equating to zero gives

$$\lambda^* = \frac{\bar{Y}^2}{\bar{Y}^2 + \varphi} \qquad (17)$$

Where

$\varphi = \sum_{h=1}^{L} W_h^2\gamma_h\left(S_y^2 + \frac{9}{4}R^2S_x^2 - 3RS_{xy}\right)$ and $0 < \lambda^* < 1$.

$MSE\left(t_{st}(\alpha_h,\delta_h)\right)$ in equation (16) is minimized when $\lambda^* = \frac{\bar{Y}^2}{\bar{Y}^2+\varphi}$

Therefore, substituting equation (17) into equation (16), gives

$$MSE(t_{st}(\alpha_h,\delta_h))_{opt} = \frac{\bar{Y}^2\varphi}{\bar{Y}^2 + \varphi} \qquad (18)$$

Efficiency comparison

If the MSE of the Solanki et al (2012) ratio estimator is compared with the MSE of proposed estimator, the following conditions are observed.

Let

$$\left.\begin{aligned}\beta_h &= \left(S_{hy}^2 + \frac{9}{4}R^2 S_{hx}^2 - 3RS_{hxy}\right)\\ \beta &= \left(S_y^2 + \frac{9}{4}R^2 S_x^2 - 3RS_{xy}\right)\end{aligned}\right\} \quad (19)$$

$$MSE\left(t_{st}(\alpha_h, \delta_h)\right) < MSE(t(\alpha, \delta))$$

$$\lambda^{*2}\sum_{h=1}^{L} W_h^2 \gamma_h \beta_h + \bar{Y}^2(\lambda^* - 1)^2 - \left(\frac{1-f}{n}\right)\beta < 0$$

$$\lambda^{*2}\left(\sum_{h=1}^{L} W_h^2 \gamma_h \beta_h + \bar{Y}^2\right) - (2\bar{Y}^2)\lambda^* - \left(\left(\frac{1-f}{n}\right)\beta - \bar{Y}^2\right) < 0$$

Where there are two conditions as follows

(i) When $(\lambda^* - 1) < 0$

$$\lambda^* > \frac{\bar{Y}^2 - \left(\bar{Y}^4 + (\sum_{h=1}^{L} W_h^2 \gamma_h \beta_h + \bar{Y}^2)\left(\left(\frac{1-f}{n}\right)\beta - \bar{Y}^2\right)\right)^{\frac{1}{2}}}{(\sum_{h=1}^{L} W_h^2 \gamma_h \beta_h + \bar{Y}^2)}$$

(ii) When $(\lambda^* - 1) > 0$

$$\lambda^* < \frac{\bar{Y}^2 + \left(\bar{Y}^4 + (\sum_{h=1}^{L} W_h^2 \gamma_h \beta_h + \bar{Y}^2)\left(\left(\frac{1-f}{n}\right)\beta - \bar{Y}^2\right)\right)^{\frac{1}{2}}}{(\sum_{h=1}^{L} W_h^2 \gamma_h \beta_h + \bar{Y}^2)}$$

When any of the conditions is satisfied, the suggested estimator is more efficient than the Solanki et al (2012) ratio estimator.

Application

To judge the relative performances of the proposed ratio estimator over the Solanki et al (2012) ratio estimator, data set from Clement and Enang (2015 pp. 90) given in table 1 was considered.

Using Table 1, the following sample information and the MSE values of estimators in Table 2 are obtained.

$$\sum_{h=1}^{L} W_h^2 \gamma_h \beta_h = 5449.0267$$

$$\beta = 1194190.1670$$

$$\left(\frac{1-f}{n}\right) = 0.06$$

$$\lambda^* = 0.9687$$

$$MSE(t_{st}(\alpha_h, \delta_h))_{opt} = 5257.0541$$

ISSN 2055-0154(Print), ISSN 2055-0162(Online)

Table 1: Data statistics

Parameter	Stratum 1	Stratum 2	Stratum 3	Total
N_h	6	8	11	$N = 25$
n_h	3	3	4	$n = 10$
$\bar{X}_h$	6.813	10.12	7.967	$\bar{X} = 8.3792$
$\bar{Y}_h$	417.33	503.375	340.00	$\bar{Y} = 410.84$
S_{hx}^2	15.9712	132.66	38.438	$S_x^2 = 59.7368$
S_{hy}^2	74775.467	259113.70	65885.60	$S_y^2 = 1237702$
S_{hxy}^2	1007.0547	5709.1629	1404.71	$S_{xy}^2 = 2524.79$
ρ_{hxy}	0.9215	0.9738	0.8827	$\rho = 0.9285$
γ_h	0.1667	0.2083	0.1591	$R = 49.8610$
w_h^2	0.0576	0.1024	0.1936	$\rho^* = 0.9409$

Table 2
MSE values of estimators

S/No.	Estimator	MSE
1.	Proposed	5278.6174
2.	Solanki et al (2012)	71651.4100

In Table 2, the values of MSE are given. It is observed that the proposed estimator has the minimum MSE and therefore it is the best estimator for the data.

In the same way, when comparing the MSE of proposed estimator with MSE of Solanki et al estimator, it is observed that the conditions

$$\lambda^* = 0.9687 > \frac{\bar{Y}^2 - \left(\bar{Y}^4 + (\sum_{h=1}^{L} W_h^2 \gamma_h \beta_h + \bar{Y}^2)\left(\left(\frac{1-f}{n}\right)\beta - \bar{Y}^2\right)\right)^{\frac{1}{2}}}{(\sum_{h=1}^{L} W_h^2 \gamma_h \beta_h + \bar{Y}^2)} = 0.3515$$

and

$$\lambda^* = 0.9687 < \frac{\bar{Y}^2 + \left(\bar{Y}^4 + (\sum_{h=1}^{L} W_h^2 \gamma_h \beta_h + \bar{Y}^2)\left(\left(\frac{1-f}{n}\right)\beta - \bar{Y}^2\right)\right)^{\frac{1}{2}}}{(\sum_{h=1}^{L} W_h^2 \gamma_h \beta_h + \bar{Y}^2)} = 1.5859$$

are satisfied. Thus, the proposed estimator is more efficient than the Solanki et al (2012) ratio estimator.

CONCLUSION

This paper has derived an improved ratio estimator for population mean in stratified random sampling from the ratio estimator of Solanki, Singh and Rathour (2012) and obtained its MSE equation. By this equation, the MSE of proposed estimator has been compared with that of Solanki, Singh and Rathour (2012) in theory and by this comparison it has been found that in all conditions the proposed estimator has a smaller MSE than the Solanki, Singh and Rathour (2012) estimator. These theoretical results are also satisfied by the results of an empirical study. In this empirical study, it is concluded that the proposed estimator is more efficient than Solanki, Singh and Rathour (2012) estimator. It is observed that the proposed estimator

ISSN 2055-0154(Print), ISSN 2055-0162(Online)

is very attractive and should be preferred in practice as it provides consistent and more precise parameter estimates. Also, authors are encouraged to propose estimators not only in simple random sampling but also in stratified random sampling.

REFERENCES

Chaudhary, M. K., Singh, R.., Shukla, R. K., Kumar, M. & Smarandache, F. (2009): A family of estimators for estimating population mean in stratified sampling under non-response, *Pakistan Journal of Statistics and Operation Research.* 5(1), 47-54.

Clement, E. P. & Enang, E. I. (2015): Calibration approach alternative ratio estimator for population mean in stratified sampling. *International Journal of Statistics and Economics,* 16(1), 83-93.

Hansen, M.H., Hurwitz, W.N. & Gurney, M. (1946): The problems and methods of the Sample survey of business, *Journal of the American Statistical Association*, 41,173-189.

Koyuncu, N. & Kadilar, C. (2009): Family of estimators of population mean using two auxiliary variables in stratified random sampling. *Communications in Statistics - Theory and Methods* 38(14), 2398-2417

Malik, S. & Singh, R. (2012): Some improved multivariate-ratio-type estimators using geometric and harmonic means in stratified random sampling. *ISRN Probability and Statistics*, Article ID 509186

Solanki, R. S., Singh, H. P. & Rathour, A. (2012): An alternative estimator for estimating the finite population mean using auxiliary information in sample surveys. *ISRN Probability and Statistics,* Article ID 657682, doi:10.5402/2012/657682

Tailor, R.., Sharma, B. & Kim, J. (2011): A generalized ratio-cum-product estimator of finite population mean in stratified random sampling. *Communications of the Korean Statistical Society*, 18(1), 111-118.

Vishwakarma, G. K., Singh, H. P. & Singh, S. (2010): A family of estimators of population mean using multi-auxiliary variate and post-stratification. *Nonlinear Analysis*: *Modelling and Control*, 15 (2), 233-253.

Vishwakarma, G. K. & Singh, H. P. (2011): Separate ratio-product estimator for estimating population mean using auxiliary information. *Journal of Statistical Theory and Applications*, 10 (4), 653-664.

ISSN 2055-0154(Print), ISSN 2055-0162(Online)