International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
Maths 2019; 4(6): 61-64

www.mathsjournal.com
Received: 24-09-2019
Accepted: 28-10-2019

Vyas Dubey
School of Studies in Statistics, Pt. Ravishankar Shukla University, Raipur, Chhattisgarh, India

Jaya B Rathore
MATS University, Raipur, Chhattisgarh, India

Ujjwal Dubey
Dept. of Computer Science and Engineering, Shri Shankaracharya Technical Campus, Bhilai, Chhattisgarh, India

Corresponding Author:
Vyas Dubey
School of Studies in Statistics, Pt. Ravishankar Shukla University, Raipur, Chhattisgarh, India

An almost unbiased ratio estimator of population mean in stratified sampling

Vyas Dubey, Jaya B Rathore and Ujjwal Dubey

Abstract
In this paper a generalised estimator of variance of ratio estimator has been proposed. It is seen that the suggested estimator is more efficient than existing estimator under certain conditions. Numerical illustrations have been cited to support theoretical results.

Keywords: SRSWOR, finite population correction, order of approximation, bias, MSE, relative efficiency

1. Introduction

In a sample survey the necessity of stratification is often dictated by administrative requirement or convenience. For a state wise survey, for instance, it is often convenient to draw samples independently from each district and carry out survey operations for each district separately. A statistical office may be set up for each such district to take care of the survey operations under its jurisdiction. Thus for administrative convenience, each district may be treated as a stratum. Since a stratified sample consists of units selected separately from each stratum, such a sample is expected to be better representation of the population than a simple random sample selected from the whole universe. In practice, the population often consists of heterogeneous units (with respect to the character under study). Thus, in a survey of manufacturing industries, some factories may be very large, having a huge fixed capital investment, a large number of workers, deploying sophisticated appliance, while others may be of medium size, even tiny in nature. For a socioeconomic survey, for instance, people may live in rural areas, urban localities, ordinary domestic houses, hostels, hospitals, jails, etc. It is evident that the nature of the value of y and the sampling problem will be different for these different sectors of the population and each such sector should be treated as a separate stratum. Again, administrators may require estimates for different strata separately along with the estimate for the population as a whole. This can be achieved through stratified sampling. It will be shown subsequently that by stratifying such that the units which are approximately homogeneous with respect to y, are placed in the same stratum, the variance of the estimator can be reduced from the one based on a random sample drawn from the whole population.

Let the population be divided into L homogenous groups of sizes N_1, N_2………N_L such that $N=N_1+N_2+\ldots\ldots+N_L$. Let y_{hj} and x_{hj} be the values of y and x on j-th unit in the h-th stratum, h=1,2….L;

J=1,2…..N_h. Let $\bar{Y}_h=\sum_{j=1}^{N_h}\frac{y_{hj}}{N_h}$, $\bar{X}_h=\sum_{j=1}^{N_h}\frac{x_{hj}}{N_h}$ be means of h-th strata of variables y and x. Further let

$W_h=\frac{N_h}{N}$, then $\bar{Y}=\sum_{h=1}^{L}W_hY_h$, $\bar{X}=\sum_{h=1}^{L}W_hX_h$ are populations means of y and x.

Let a sample of size n_h be taken from h-th strata by SRSWOR procedure such that $n=n_1+n_2+\ldots\ldots+n_L$. Then $\bar{y}_h=\sum_{j=1}^{n_h}\frac{y_{hj}}{n_h}$, $\bar{x}_h=\sum_{j=1}^{n_h}\frac{x_{hj}}{n_h}$ and $\bar{y}_{st}=\sum_{h=1}^{L}W_h\bar{y}_h$, $\bar{x}_{st}=\sum_{h=1}^{L}W_h\bar{x}_h$ are sample means of y and x based on h-th strata, h=1, 2,…., L.

Hansen, Hurwitz and Gurney (1946) [2] suggested a combined ratio estimate,

$$\bar{y}_{r_c} = \frac{\bar{y}_{st}}{\bar{x}_{st}} \bar{X} \quad (3.1)$$

Then E ($\bar{y}_{r_c}$) = $\sum_{h=1}^{L} W_h \bar{Y}_h \left\{ 1 + \frac{N_h - n_h}{N_h n_h} \left(C_{hx}^2 - \rho_h C_{hx} C_{hy} \right) \right\}$ (3.2)

Then bias diminishes with the size of the sample, we assume that n_h is proportional to N_h and $S_{hx} / \bar{X}$, $S_{yx} / \bar{Y}$ and ρ_h are same over all strata say C_x, C_y and ρ respectively, the relative bias of $\bar{y}_{r_c}$ given by is

$$B(\bar{y}_{r_c}) = \frac{N - n}{Nn} (C_x^2 - \rho C_y C_x) \quad (3.3)$$

And MSE of $\bar{y}_{r_c}$ is

$$M(\bar{y}_{r_c}) = \sum_{h=1}^{L} W_h^2 \frac{N_h - n_h}{n_h} \bar{Y}_h^2 (C_{hy}^2 + C_{hx}^2 - 2\rho C_{hy} C_{hx}) \quad (3.4)$$

Sukhatme and Sukhatme (2004) defined separate ratio estimator

$$\bar{y}_{r_s} = \sum_{h=1}^{L} W_h \frac{\bar{y}_h}{\bar{x}_h} \bar{X}_h \quad (3.5)$$

Let $S_{hy}^2 = \frac{1}{N_h - 1} \sum_{j=1}^{N_h} (y_{jh} - \bar{Y}_h)^2$, $S_{hx}^2 = \frac{1}{N_h - 1} \sum_{j=1}^{N_h} (x_{hj} - \bar{X}_h)^2$ be variance of y and x for the h-th stratum; ρ_h be correlation coefficient between y and x for the h-th stratum, $C_{hy}^2 = \frac{S_{hy}^2}{\bar{Y}_h^2}$ and $C_{hx}^2 = \frac{S_{hx}^2}{\bar{Y}_h^2}$ are coefficient of variation of y and x for h-th stratum then

E ($\bar{y}_{r_s}$) = $\sum_{h=1}^{L} W_h \bar{Y}_h \left\{ 1 + \frac{N_h - n_h}{N_h n_h} \left(C_{hx}^2 - \rho_h C_{hx} C_{hy} \right) \right\}$ (3.6)

It follows that in order to $\bar{y}_{r_s}$ should provide a satisfactory estimate of the population mean, the sample size within each stratum should be sufficiently large the MSE of $\bar{y}_{r_s}$ to O (n^{-1}) of approximation is given as

$$M_1(\bar{y}_{r_s} = \sum_{h=1}^{L} W_h^2 \frac{N_h - n_h}{N_h n_h} \bar{Y}_h^2 (C_{hy}^2 + C_{hx}^2 - 2\rho C_{hy} C_{hx}) \quad (3.8)$$

Respectively.

Further Dubey and Rathore (2012) proposed bias adjusted separate ratio type estimator of $\bar{Y}$ as

$$\bar{y}_{ur_s} = \bar{y}_{r_s} + \sum_{h=1}^{L} W_h \frac{N_h - n_h}{N_h n_h} \bar{y}_h \left(\frac{s_{hyx}}{\bar{y}_h \bar{x}_h} - \frac{s_{hx}^2}{\bar{x}_h^2} \right) \quad (3.10)$$

Where $s_{hx}^2 = \frac{\sum_{j=1}^{n_h} \left(x_{hj} - \bar{X}_h \right)^2}{n_h - 1}$ and $s_{hyx} = \frac{\sum_{j=1}^{n_h} (y_{hj} - \bar{Y}_h)(x_{hj} - \bar{X}_h)}{n_h - 1}$

Bias and MSE of $\bar{y}_{ur_s}$ are given by

$$B(\bar{y}_{u_{r_s}}) = -3\left[\sum_{h=1}^{L} \bar{y}_h W_h \left(\frac{N_h - n_h}{N_h n_h}\right)^2 C_{hx}^2 \left(C_{hx}^2 - C_{hyx}\right)\right]$$

and

$$M(\bar{y}_{u_{r_s}}) = M_1(\bar{y}_{r_s}) + \sum_{h=1}^{L} W_h^2 \bar{y}_h^2 \left(\frac{N_h - n_h}{N_h n_h}\right)^2 (2C_{hx}^4 - 4C_{hx}^2 C_{hyx} + C_{hyx}^2 + C_{hx}^2 C_{hy}^2)$$

Where

$$M_1(\bar{y}_{u_{r_s}}) = \sum_{h=1}^{L} W_h^2 \frac{N_h - n_h}{N_h n_h} \bar{y}_h^2 (C_{hy}^2 + C_{hx}^2 - 2C_{hyx})$$

Is MSE of separate ratio estimator $\bar{y}_{r_s}$ upto first of approximation.

Sukhatme *et al*. (1992) [4] have discussed that the combined ratio estimate $\bar{y}_{r_c}$ will have a lower precision than that based on separate ratio estimate $\bar{y}_{r_s}$ for h-th strata. On the other hand, the bias in the former estimate will be smaller than latter. Unless, therefore, the population ratios in the different strata vary considerably, the use of a combined ratio would provide an estimate which has a negligible bias and whose precision is almost as high as that of the estimate based on separate ratios, so in section 3.2 we propose bias adjusted separate ratio estimator and study its properties.

2. Proposed estimator

Following Tin (1965) [5] and extending the idea to stratified random sampling, we propose bias adjusted separate type ratio estimator of $\bar{Y}$ as

$$t^* = \bar{y}_{r_s}\left[1 + \sum_{h=1}^{L} W_h \theta_h \left(\frac{m'_{h11}(y,x)}{\bar{X}_h \bar{y}_h} - \frac{m'_{h2}(x)}{\bar{x}_h^{\,2}}\right)\right] \tag{3.9}$$

Where

$$m'_{h2}(x) = \frac{1}{n_h}\sum_{i=1}^{n_h} x_{hi}^2 \quad \text{and} \quad m'_{h11}(y,x) = \frac{1}{n_h}\sum_{i=1}^{n_h} x_{hi} y_{hi}$$

be unbiased estimator of $\mu'_{h2}(x) = \frac{1}{N_h}\sum_{i=1}^{N_h} x_{hi}^2$ and $\mu'_{h11}(y,x) = \frac{1}{N_h}\sum_{i=1}^{N_h} y_{hi} x_{hi}$ respectively

3. Properties of proposed estimator

For simplicity we assume that for each strata y and x follow Bivariate normal population where odd order moments reduces to zero. Then

$$B(t^*) = 2\sum_{h=1}^{L} \bar{Y}_h W_h \left(\frac{N_h - n_h}{N_h n_h}\right)^2 \left(2C_{hxy} - 3C_{hx}^4\right) \tag{3.10}$$

and

$$M(t^*) = M(\bar{y}_{r_s}) + \sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \left(\frac{N_h - n_h}{N_h n_h}\right)^2 \Big[2(C_{hxy} - C_{hx}^2) - 2C_{hx}^2 (C_{hx}^2 - C_{hxy}) + (C_{hxy} - C_{hx}^2)(7C_{hx}^2 - 3C_{hxy})\Big] \tag{3.11}$$

Where

$$M(\bar{y}_{r_s}) = M_1(\bar{y}_{r_s}) + 3\sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \left(\frac{N_h - n_h}{N_h n_h}\right)^2 \left[3C_{hx}^4 - 6C_{hx}^2 C_{hxy} + 2C_{hxy}^2 + C_{hx}^2 C_{hy}^2 \right] \quad (3.12)$$

Therefore,

$$M(\bar{y}_{r_s}) - M(t^*) = \sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \left(\frac{N_h - n_h}{N_h n_h}\right)^2 \left[2(C_{hxy} - C_{hx}^2) - 2C_{hx}^2 (C_{hx}^2 - C_{hxy})\right.$$

$$\left. + (C_{hxy} - C_{hx}^2)(7C_{hx}^2 - 3C_{hxy}) \right] > 0$$

If $\dfrac{C_{hyx}}{C_{hx}^2} < 1$ (3.13)

4. Numerical Illustration: Let us consider the data taken from Cochran (1977) where

Table: Type of data, the relative efficiency (%) for $\bar{y}_{r_s}$ and t*

Strata	S_{hx}^2	S_{hx}^2	S_{hx}^2	S_{hx}^2	S_{hx}^2	$\bar{X}$	$\bar{Y}$	N_h	n_h
1	5186	6462	8699	72.01	93.27	53.8	69.48	47	18
2	2367	3100	4614	48.65	67.93	31.07	43.64	118	46
3	4877	4817	7311	69.83	85.51	56.97	66.39	91	36

For the data, the relative efficiency (%) for $\bar{y}_{r_s}$ and t* is computed.

Table 1: Type of Estimator and Efficiency

Estimator	Efficiency
$\bar{y}_{r_c}$	478.32
$\bar{y}_{r_s}$	527.96
$\bar{y}_{u_{r_s}}$	679.78
t^*	797.34

From the above table, we find that proposed estimator is better than separate ratio type estimator.

5. References

1. Cochran WG. Sampling Techniques, John Wile and Sons., New York, III edition, 1977.
2. Hansen MH, Hurwitz WN, Gurney M. Problems and methods of the sample survey of business. J Amer. Stat. Assoc. 1946; 41:173-189.
3. Dubey V, Rathore J. A note on bias adjustment in ratio estimator, Applications of Reliability Theory and Survival Analysis, Editors-Chandra, N. and Gopal, G. Bonfring Publication, Coimbtore, India, 2012, 210-223.
4. Sukhatme PV, Sukhatme BV, Sukhatme S, Asok C. Sampling theory of surveys with applications, lowa State, 1992.
5. Tin M. Comparison of some ratio estimators. J Amer. Stat. Assoc. 1965; 60:294-307.