International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Simple Estimation Method, Ratio Method Is Much More Efficient for Other Sampling Scheme with Ratio Estimate Iin Stratified Random Sampling

Dr. Syed Qaim Akbar Rizvi

Department Of Statistics, Shiap.G.College Lucknow, Lucknow University, Lucknow

ABSTRACT

In ratio method of estimation sampling with equal probability of selection for ratio estimate in stratified random Sampling, the variance of estimate is slightly smaller than that of estimate this means that using separate ratios for the two groups gives a little more accurate result than using one common ratio, so that the improvement is very small. Both methods also give the same best distribution of the population between the two groups, compared with the simple estimation method, the ratio method is much more efficient, this means that the ratio method gives a more accurate estimate of the total population with the same sample size, separate ratios provide only a small improvement over a common ratio, but the ratio method performs much better than simple estimation obtain the most accurate estimate sampling variance, smaller variance means greater accuracy, compare the three methods to determine which gives the most accurate estimate for the same sample size for example which is true for values are given and graphical location.

KEYWORDS : Ratio estimator, stratified, bias, stratum, precision, optimum allocation, optimum allocation

INTRODUCTION AND RESEARCH METHODOLOGY

Let R_{n_h} denote the estimate of the population ratio for the h-th stratum and Y_{R_h} denote the ratio estimate of the population total of y in the h-th stratum then clearly the estimate of the population total of y over all the strata is given by

$$Y_R = \sum_{h=1}^{k} Y_{R_h}$$

$$= \sum_{h=1}^{k} R_{\mathrm{h}} \times N_h \bar{x}_{N_h} \qquad \ldots\ldots\ldots\ldots\ldots\ldots \quad (01)$$

Now

$$E(Y_R) = \sum_{h=1}^{k} N_h \bar{x}_{N_h} E\left(R_{n_h}\right) \qquad (02)$$

Expected value of the ratio estimate based on the simple mean, in ratio estimate numerator and denominator are both random variables.

The derivation of the expected value of R_{n_h} by using ratio method of estimation we have

$$E(Y_R) = \sum_{h=1}^{k} N_h \bar{y}_{N_h} \left\{1 + \frac{N_h - n_h}{N_h n_h}\left(\frac{\mathbf{S}_{hz}{}^2}{\bar{x}_{N_h}{}^2} - \rho_h \frac{\mathbf{S}_h \mathbf{S}_{hx}}{\bar{y}_{N_h}\bar{x}_{N_h}}\right)\right\} \quad \ldots\ldots\ldots\ldots\ldots \quad (03)$$

$\frac{k}{n}\left(C_z^2 - \rho C_y C_z\right)$.....(04)

This means that the estimate is biased, but it becomes more accurate as the sample size increases to understand how the bias decreases when the sample size increases, we make a few assumptions, we assume that the finite multiplier is approximately equal to 1, we also assume that two quantities have approximately the same order of magnitude across all strata, we represent them by suitable symbols under these assumptions, we can calculate the relative bias of the estimate. It follows that in order that Y_R should provide a satisfactory estimate of the population total, the sample size within each stratum should be sufficiently large the variance of Y_R since to a first approximation

$$E\left(R_{n_h}\right) = R_{N_h}$$

and sampling is done independently in the different strata; the product term is zero and we obtain

$$V_1(Y_R) = \sum_{h=1}^{k} N_h{}^2 \bar{x}_{N_h}{}^2 V_1\left(R_{n_h}\right) \quad (05)$$

where $p_h = N_h/N$, this can also be written as

$$V_1(Y_R) = N \sum_{h=1}^{k} \frac{p_h(N_h - n_h)}{n_h} \cdot \frac{1}{N_h - 1}\left\{\sum_{i=1}^{N_h} \left(y_{hi} - R_{N_h} x_{hi}\right)^2\right\} \quad (06)$$

The above formula is based on the assumption that n_h is large. This, however, is not always true in practice to get over this difficulty, Hansen, Hurwitz and Gurney suggest a single combined ratio, say that

$$\frac{\sum_{h=1}^{k} p_h \bar{y}_{n_h}}{\sum_{h=1}^{k} p_h \bar{x}_{n_h}} \quad (07)$$

now to estimates based on separate strata. To obtain the expected value of (o7), we write

$$\sum_{h=1}^{k} p_h \bar{y}_{nh} = \bar{y}_N + \bar{\epsilon}_n \text{ and } \sum_{h=1}^{k} p_h \bar{x}_{nh} = \bar{x}_N + \bar{\epsilon}_n' \quad (08)$$

Where

$$E(\dot{\epsilon}_n) = 0, E(\dot{\epsilon}_n{}') = 0 \quad (09)$$

$$E(\dot{\epsilon}_n{}^2) = \sum_{h=1}^{k} \frac{N_h - n_h}{N_h n_h} p_h{}^2 \, \mathrm{S}_{hz}{}^2 \quad (10)$$

Then to a first approximation

$$E\left(R_{n_c}\right) = \frac{\bar{y}_N}{\bar{x}_N}\left\{1 + \frac{E(\bar{\epsilon}_n{}'^2)}{\bar{x}_N{}^2} - \frac{E(\bar{\epsilon}_n \bar{\epsilon}_n{}')}{\bar{y}_N \bar{x}_N}\right\} \quad (11)$$

Hence, the relative bias in R_{n_c} is given by bias (R_{n_c}) = $E\left(R_{n_c}\right)$-mean

$$\sum_{h=1}^{k} \frac{1-W_h}{n_h} \cdot p_h{}^2 \left(\frac{\mathbf{S}_{hx}{}^2}{\bar{x}_N{}^2} - \frac{\rho_h \mathbf{S}_{hy} \mathbf{S}_{hx}}{\bar{y}_N \bar{x}_N} \right) \tag{12}$$

where $\frac{W_h}{} = \frac{n_h}{N_h}$, To have an idea as to how rapidly the bias diminishes with the size of the sample, we will suppose that n_h is proportional to N_h and $\mathbf{S}_{hx}, \mathbf{S}_{hy}$ and ρ_h are constant. The relative bias will then be seen to be given by

$$\left(\frac{N-n}{N} \right) \cdot \frac{1}{n} \cdot \left(C_x{}^2 - \rho C_y C_x \right) \tag{13}$$

(14)

$$V_1(\gamma_{R_c}) = N \sum_{h=1}^{k} \frac{N_h - n_h}{n_h} \cdot p_h \{ \mathbf{S}_{hy}{}^2 + R_N{}^2 \mathbf{S}_{hy}{}^2 - 2R_N \rho_h \mathbf{S}_{hy} \mathbf{S}_{hz} \}$$

It is interesting to note that the sampling variance of the combined ratio has the same form as that of the ratio estimate based on separate strata except that there is now a single ratio R_N in place of R_{N_h}.The difference between the sampling variance of Y_{Rc} in (14) and of Y_R in (06) is given by

$$= \sum_{h=1}^{k} \frac{N p_h (N_h - n_h)}{n_h} \{ \mathbf{S}_{hx}{}^2 (R_N - R_{Nh})^2 + 2(R_N - R_{N_h}) \times (R_{Nh} \mathbf{S}_{hx}{}^2 - \rho_h \mathbf{S}_{hy} \mathbf{S}_{hx}) \} \tag{14}$$

$$(R_{Nh} \mathbf{S}_{hx}{}^2 - \rho_h \mathbf{S}_{hy} \mathbf{S}_{hx})$$

It will be seen that (14) depends upon the magnitude of the variation between the strata ratios and the value of the sample among different strata is determined when the ratio estimation method is used and the survey cost is proportional to the sample size. The aim is to distribute the sample among the strata in such a way as to obtain It follows that the combined estimate will have a lower precision than that based on separate strata. It means that if the ratios in different strata are nearly the same, using one common ratio is almost as good as using separate ratios for each stratum, one ratio is used for all strata, different ratio is used for each stratum. The difference between the expected estimate and the true population value which is near about or consistent the estimates are the optimum allocation of the most precise estimate at the minimum cost.

Let

$$\phi = N \sum_{h=1}^{k} \frac{N_h - n_h}{n_h} \{ p_h \cdot \cdot \mathbf{S}_{hy}{}'^2 \} + \mu \left(\sum_{h=1}^{k} n_h \right) \tag{15}$$

where μ is a constant, and

$$S_{hy}'^2 = \frac{\sum_{i=1}^{N_h} (y_{hi} - R_{Nh} x_{hi})^2}{N_h - 1} \tag{16}$$

So that it clears the function ϕ will be directly proportional N_h ,$S_{hy}{}'$ then the function ϕ can be written as

$$\phi = \sum_{h=1}^{k} \left\{ \frac{N_h \, S_{hy}{}'}{\sqrt{n_h}} - \sqrt{\mu n_h} \right\}^2 + \text{terms independent of } n_h \qquad (17)$$

It follows that the optimum value of n_h is given by

$$n_h \propto N_h \, \mathbf{S}_{hy}{}' \qquad (18)$$

This result is thus analogous to that for the simple arithmetic mean estimate, except that instead of the variance of y within a stratum, viz., $\mathbf{S}_{hy}{}^2$, we now have the residual variance of y about the ratio line in the stratum. Researchers want to estimate the total number of livestock in a subdivision by selecting 73 villages instead of studying all the villages. The villages are divided into two groups: those with agricultural land up to 1,000 acres and those with agricultural land greater than 1,000 acres. The three methods are used which are given, ratio method with a common ratio, ratio method with separate ratios, simple estimation, optimum allocation, sampling variance, Comparison of Different Methods of Estimation, Ratio estimate with a common ratio for both strata, the number of villages selected from Stratum 1 and Stratum 2 is 54 and 19, respectively. Compare the three methods to determine which method gives the most accurate estimate? The sampling variance is 8,707,000, ratio estimate with separate ratios for the two strata, the number of villages selected from Stratum 1 and Stratum 2 is 54 and 19, respectively, the sampling variance is 8,688,000, simple estimate within each stratum, the number of villages selected from Stratum 1 and Stratum 2 is 53 and 20, respectively.

Comparison of Different Methods of Estimation for Graphical Representation

The bar graph compares the percentage standard error obtained under the three estimation methods. A lower percentage standard error indicates a more precise estimate.

All the result draw on graphical representation given below

V (γ_{R_c}) = 6935000 + 175300 = 8688000, S.E. = 2948

Y = (151.9) (364) = 55300, % S.E. = (2948 / 55300) × 100 = 5.3

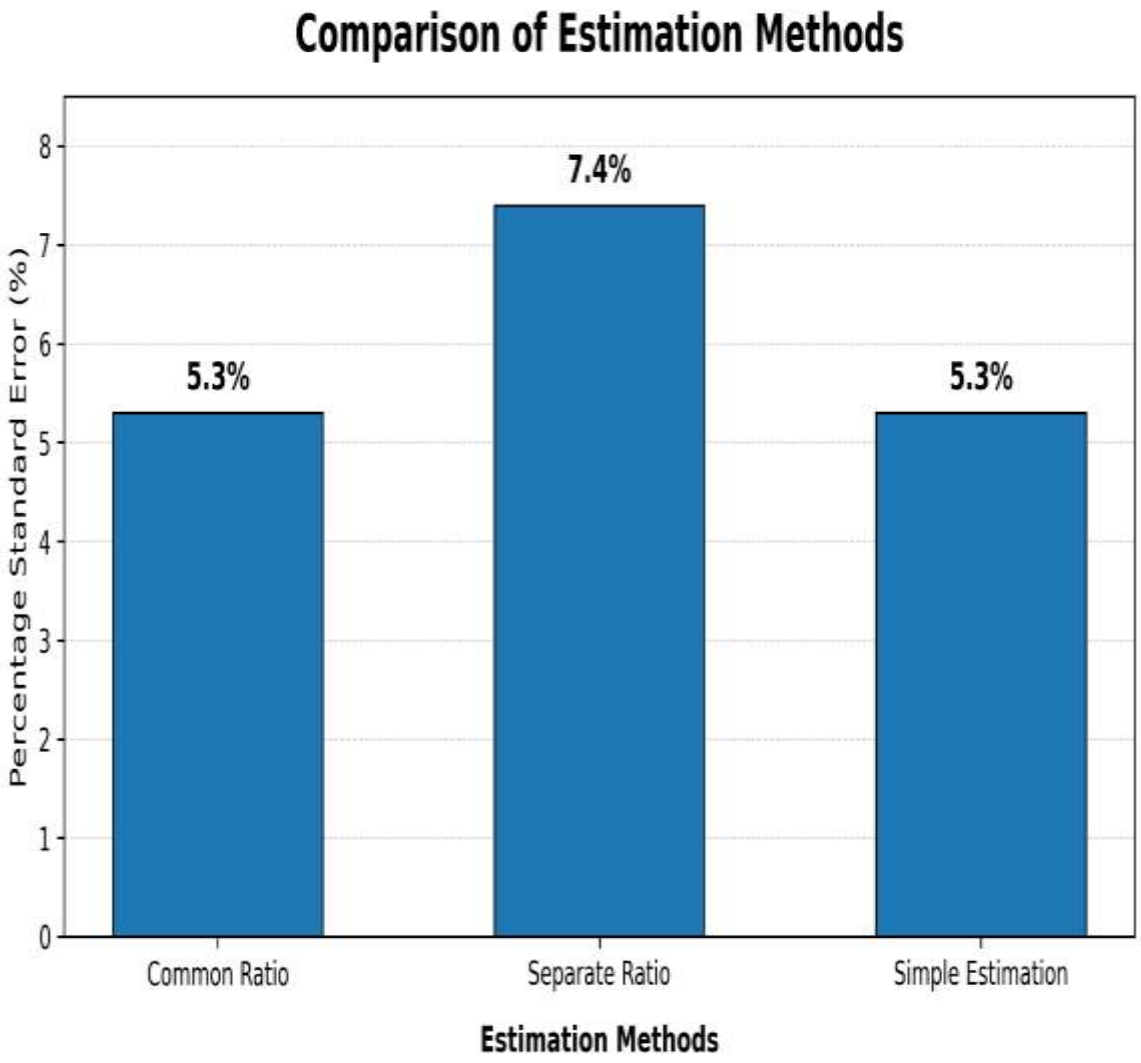
V (γ_{R_c}) = 16677000 Y = 55300 S.E. = 4084

% S.E. = (4084 / 55300) × 100 = 7.4

V (γ_{R_c}) = 6963000 + 1744000 = 8707000, S.E. = 2951

Y = (151.9) (364) = 55300, % S.E. = (2951 / 55300) × 100 = 5.3

Comparison of Different Methods of Estimation for graphical representation. The graph compares the percentage standard error obtained under the three estimation methods. Therefore, the first and third methods provide more precise estimates than the second method. Among the three methods, the first and third methods are equally efficient based on their percentage standard errors.



From the above calculations, the percentage standard error for the ratio method is found to be 5.3%. For the second method, the percentage standard error is 7.4%, while for the ratio method with a common ratio, it is 5.3%. Comparing the results, the first and third methods give the same percentage standard error, which is lower than that of the second method.

REFERENCES

1. Neyman, J. (1934) "On the Two Different Aspects of the Representative Method: The Method of Stratified Sampling and the Method of Purposive Selection," Jour. Roy. Statist. Soc., 97, 558-606.
2. Tschuprow, A. A. (1923) "On the Mathematical Expectation of the Moments of Frequency Distributions in the Case of Correlated Observations," Metron, 2, No. 4.
3. Sukhatme, P. V. (1935) 137 Contribution to the Theory of the Representative Method,"" Jour. Roy. Statist. Soc. Suppl., 2, 253-68
4. Evans, W. D. (1951) stratified sampling references. On Stratification and Optimum Allocations," Jour. Amer. Statist. Assoc, 46, 95-104
5. David, F. N. and Neyman, J. (1938) "Extension of the Markoff Theorem on Least Squares," Statist. Res. Mem., 2, 105-16.
6. Cochran, W. G. (1940) "The Estimation of the Yields of Cereal Experiments by Sampling for the Ratio

of Grain to Total Produce," Jour. Agr. Sci., 30, 262-75.
7. Hansen, M. H., Hurwitz, W. N. and Gurney, M. (1946) "Problems and Methods of the Sample Survey of Business," Jour. Amer. Statist. Assoc., 41, 173-89.