ISSN 2249 –5460

INTERNATIONAL JOURNAL OF MATHEMATICAL SCIENCES, TECHNOLOGY AND HUMANITIES

www.internationalejournals.com

Available online at www.internationaleJournals.com

International eJournals

International Journal of Mathematical Sciences, Technology and Humanities 108 (2013) 1158–1163

Comparison of Ratio Estimators Of Population Mean

Jakkula Srinivas[1], Naveen Kumar Boiroju[2] & M. Krishna Reddy[3]

1. Department of Statistics, Kakatiya University, Warangal.

E-mail: srinivasjakkula07@yahoo.com

2. Department of Statistics, Osmania University, Hyderabad.

ABSTRACT

The ratio estimators are often employed to estimate the population mean of a study variable with the help of an auxiliary variable. In this paper, some ratio estimators of population mean are compared using simulation. The results show that Jackknife and Sahoo estimators are most efficient and almost unbiased estimators of the population mean.

Keywords: Ratio estimator, Simulation, Simple random sampling.

AMS Classification: 62D05, 68U20, 65C10

1. Introduction

The ratio estimators are frequently employed in sample surveys when estimating the population mean ($\bar{Y}$) of a variate Y with the help of the known population mean ($\bar{X}$) of correlated auxiliary variable X (Hartley and Ross, 1954). Let $\bar{x}$ and $\bar{y}$ denote the sample means of the variables X and Y respectively. Reddy et al. (2008) compared the mean per unit, ratio and regression estimators of population mean using simulating random variables from different correlated populations. It is observed from their study, the mean per unit estimator of population mean has the lowest relative bias as compared with that of the ratio and regression estimators. The linear regression estimator is more efficient than that of the mean per unit and ratio estimators. The ratio estimator of population mean is simple to calculate and is better than the mean per unit estimator. In some cases, the ratio estimator is performing on par with the regression estimator when the correlation between the study variable (Y) and auxiliary variable (X) is very high. Rao (2010) and Reddy et al. (2010) claimed that the estimators of population ratio proposed by Tin (1965), Sahoo (1987) and Beale (1962) are equally efficient and the Jackknife estimator is the best among them. The ratio estimators compared under different models using simulation by Rao (2010), Reddy et al. (2010), Tin (1965), Rao and Beegle (1967), Hutchison (1971), Sahoo and Dalabehera (1996). These studies explored that the simulation method is the best method when there are no analytical comparisons possible. Based on this review, it is noted that the ratio estimators are simple to apply, but the analytical comparisons of these estimators is very complex in nature. In this paper, some ratio estimators of population mean reviewed and compared using simulation of random variables from normal population.

2. Ratio Estimators of Population Mean

Some of the ratio estimators formulated in the literature are presented in this section. The classical ratio estimator (Cochran, 1977) is given by

$$E_1 = \hat{\bar{Y}}_1 = \frac{\bar{y}}{\bar{x}}\bar{X} \tag{1}$$

Tin's (1965) estimator is given by

$$E_2 = \hat{\bar{Y}}_2 = \frac{\bar{y}}{\bar{x}}\left(1+\theta\left(c_{yx} - c_x^2\right)\right)\bar{X} \tag{2}$$

and Sahoo (1987) estimator is given below

$$E_3 = \hat{\bar{Y}}_3 = \frac{\bar{y}}{\bar{x}}\left(1+\theta c_{yx}\right)\left(1 - \theta c_x^2\right)\bar{X} \tag{3}$$

Another estimator proposed by Sahoo (1987) is

$$E_4 = \hat{\bar{Y}}_4 = \frac{\bar{y}}{\bar{x}}\frac{\left(1 - \theta c_x^2\right)}{\left(1-\theta c_{yx}\right)}\bar{X} \tag{4}$$

where $\theta = n^{-1} - N^{-1}$, $c_{yx} = s_{xy} / \bar{x}.\bar{y}$, $c_x^2 = s_x^2 / \bar{x}^2$, $s_{xy} = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)$ and $s_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2$.

There is no attempt is made to study the design properties of these estimators (Dalabehra and Sahoo, 1994). Suppose a simple random sample of size n is selected in the form of m independent subsamples of k units each. Let $R^{(i)} = \frac{\bar{y}^{(i)}}{\bar{x}^{(i)}}$ where $\bar{x}^{(i)}$ and $\bar{y}^{(i)}$ are unbiased estimators of $\bar{X}$ and $\bar{Y}$ obtained after omitting the i^{th} subsample. Quenouille (1956) suggested the following estimator based on resampling methods.

$$E_5 = \hat{\bar{Y}}_5 = \left(m\frac{\bar{y}}{\bar{x}} - \frac{m-1}{m}\sum_{i=1}^{m} R^{(i)}\right)\bar{X} \tag{5}$$

The above estimator is popularly known as Jackknife ratio estimator (Reddy et al. 2010). A simulation approach adopted for the comparison of the above ratio estimators as the analytical comparison of these estimators is not possible.

3. Simulation Study

This section presents the computational procedure for the comparison of ratio estimators. It is required to generate a bivariate population with a specified correlation between the study and auxiliary variables for the comparison of the ratio estimators. Reddy et al. (2010) explained a simple procedure to generate such correlated populations with known marginal distributions. The same methodology adopted herewith and implemented on two different normal populations. The following algorithm explains the simulation procedure to compare the ratio estimators of the population mean.

Step-1: Generate two independent random variables X from $N(\mu, \sigma^2)$ and X_1 from $N(\mu_1, \sigma_1^2)$ using Box-Muller method (Jhonson, 1987).

Step-2: Set $Y = \rho X + \sqrt{1-\rho^2}\, X_1$ where $0 < \rho = 0.4,\ 0.6\ and\ 0.8 < 1$.

Step-3: Return the pair (Y, X).

Step-4: Consider the population-I with the parameters μ=5, σ=3, μ_1=5 and σ_1=3 in step-1 and repeat the steps 1 to 3 for 2000 times. This population will contain the same variance for the variables Y and X.

Step-5: Similarly, generate the population-II with the parameters μ=3, σ=2, μ_1=5 and σ_1=3 in step-1 and repeat the steps 1 to 3 for 2000 times. This population will have different variances for the variables Y and X.

Step-6: From the population of size N= 2000, draw 1000 simple random samples (y_i, x_i) (i=1, 2, ...,n) without replacement of size n=10, 30 and 50. For each of the sample, the five ratio estimators are computed.

Step-7: Compute the estimator of population mean as

$$E\left(\hat{\bar{Y}}_k\right) = \frac{\sum_{j=1}^{1000} \hat{\bar{Y}}_k}{1000}, \qquad \text{for k=1, 2,..., 5.} \tag{6}$$

The relative bias (RB) of $\hat{\bar{Y}}_k$ is defined by

$$RB\left(\hat{\bar{Y}}_k\right) = \frac{\left|\hat{\bar{Y}}_k - \bar{Y}\right|}{\bar{Y}}, \qquad \text{for k= 1, 2, …, 5.} \tag{7}$$

The mean squared error (MSE) of the estimator is defined by

$$MSE\left(\hat{\bar{Y}}_k\right) = \frac{1}{1000}\sum_{j=1}^{1000}\left(\hat{\bar{Y}}_{kj} - \bar{Y}\right)^2, \qquad \text{for k= 1, 2, …, 5.} \tag{8}$$

The coefficients of skewness (β_1) and kurtosis (β_2) computed for each of the estimators to know the behaviour of the distributions of ratio estimators of population mean.

4. Results and Discussion

From this simulation study, it is observed that the ratio estimators proposed by Tin (1965) and Sahoo (1987) are equally almost unbiased estimators of population of mean when the sample size is large. As the sample size and the correlation between the two variables increases, the mean squared error and relative bias decreasing. The Jackknife estimator is an efficient and almost unbiased estimator of population mean for the large samples having high correlation between the study variable and auxiliary variable. If the study variable and auxiliary variable have the same variance, then the Jackknife estimator is the efficient and almost unbiased estimator of the population mean.

Table 1. Comparison of ratio estimators of population mean for Population-I

r_{yx}	n	Estimator	Estimate	MSE	RB	Skewness	Kurtosis
0.4	10	E_1	6.985	2.445	0.037	2.414	6.999
		E_2	6.751	1.747	0.002	1.175	4.956
		E_3	6.747	1.738	0.001	1.172	4.962
		E_4	6.749	1.740	0.001	1.172	4.961
		E_5	6.745	1.718	0.001	1.129	4.890
	30	E_1	6.823	0.515	0.012	0.585	4.470
		E_2	6.764	0.480	0.004	0.536	4.370
		E_3	6.764	0.479	0.004	0.536	4.369
		E_4	6.764	0.479	0.004	0.536	4.370
		E_5	6.760	0.478	0.003	0.532	4.363
	50	E_1	6.774	0.297	0.005	0.380	3.495
		E_2	6.742	0.287	0.001	0.360	3.474
		E_3	6.742	0.287	0.001	0.360	3.474
		E_4	6.742	0.287	0.001	0.360	3.474
		E_5	6.739	0.286	0.000	0.358	3.472
0.6	10	E_1	7.100	1.201	0.015	0.657	4.654
		E_2	6.914	0.887	0.011	0.198	3.696
		E_3	6.906	0.876	0.013	0.187	3.678
		E_4	6.909	0.879	0.012	0.189	3.677
		E_5	6.913	0.881	0.012	0.194	3.708
	30	E_1	7.062	0.431	0.010	0.591	4.789
		E_2	7.011	0.393	0.002	0.484	4.596
		E_3	7.010	0.392	0.002	0.482	4.590
		E_4	7.011	0.393	0.002	0.482	4.592
		E_5	7.008	0.391	0.002	0.477	4.579
	50	E_1	7.031	0.229	0.005	0.431	4.871
		E_2	7.003	0.218	0.001	0.382	4.798
		E_3	7.003	0.218	0.001	0.381	4.797
		E_4	7.003	0.218	0.001	0.381	4.797
		E_5	7.000	0.217	0.001	0.376	4.789
0.8	10	E_1	6.886	0.692	0.010	0.971	5.204
		E_2	6.761	0.567	0.008	0.535	3.919
		E_3	6.753	0.561	0.009	0.525	3.898
		E_4	6.757	0.563	0.008	0.525	3.895
		E_5	6.756	0.558	0.009	0.499	3.823
	30	E_1	6.813	0.183	0.000	0.119	2.633
		E_2	6.780	0.178	0.005	0.116	2.616
		E_3	6.779	0.178	0.005	0.116	2.616
		E_4	6.780	0.178	0.005	0.116	2.616
		E_5	6.778	0.178	0.005	0.116	2.616
	50	E_1	6.821	0.107	0.001	0.067	2.470
		E_2	6.803	0.105	0.002	0.062	2.450
		E_3	6.802	0.105	0.002	0.062	2.450
		E_4	6.802	0.105	0.002	0.062	2.450
		E_5	6.801	0.105	0.002	0.062	2.448

Table 2. Comparison of ratio estimators of population mean for Population-II

r_{yx}	n	Estimator	Estimate	MSE	RB	Skewness	Kurtosis
0.4	10	E_1	6.050	2.482	0.037	1.702	5.962
		E_2	5.769	1.756	0.011	0.892	4.723
		E_3	5.763	1.741	0.012	0.875	4.697
		E_4	5.764	1.744	0.012	0.876	4.695
		E_5	5.756	1.725	0.013	0.823	4.514
	30	E_1	5.893	0.614	0.010	0.500	4.392
		E_2	5.826	0.581	0.001	0.468	4.287
		E_3	5.826	0.580	0.001	0.467	4.283
		E_4	5.826	0.580	0.001	0.467	4.284
		E_5	5.822	0.579	0.002	0.464	4.274
	50	E_1	5.840	0.350	0.001	0.261	2.956
		E_2	5.803	0.341	0.005	0.254	2.933
		E_3	5.803	0.341	0.005	0.254	2.933
		E_4	5.803	0.341	0.005	0.254	2.933
		E_5	5.799	0.341	0.006	0.253	2.930
0.6	10	E_1	6.096	1.752	0.025	1.788	7.058
		E_2	5.827	1.197	0.020	0.567	4.309
		E_3	5.818	1.183	0.021	0.542	4.253
		E_4	5.821	1.186	0.021	0.545	4.254
		E_5	5.818	1.184	0.021	0.545	4.283
	30	E_1	6.021	0.531	0.013	0.623	3.776
		E_2	5.951	0.476	0.001	0.503	3.643
		E_3	5.950	0.476	0.001	0.502	3.640
		E_4	5.951	0.476	0.001	0.502	3.641
		E_5	5.947	0.473	0.000	0.496	3.636
	50	E_1	5.994	0.292	0.008	0.802	3.764
		E_2	5.956	0.273	0.002	0.742	3.683
		E_3	5.956	0.273	0.002	0.741	3.682
		E_4	5.956	0.273	0.002	0.741	3.682
		E_5	5.952	0.271	0.001	0.735	3.673
0.8	10	E_1	5.337	0.912	0.023	2.671	4.641
		E_2	5.169	0.687	0.009	1.191	5.656
		E_3	5.160	0.676	0.011	1.104	5.383
		E_4	5.164	0.679	0.010	1.116	5.427
		E_5	5.159	0.674	0.011	1.086	5.329
	30	E_1	5.236	0.242	0.003	0.012	2.925
		E_2	5.192	0.231	0.005	0.011	2.903
		E_3	5.192	0.231	0.005	0.011	2.903
		E_4	5.192	0.231	0.005	0.011	2.903
		E_5	5.189	0.231	0.006	0.011	2.902
	50	E_1	5.242	0.152	0.004	-0.011	3.296
		E_2	5.217	0.148	0.000	-0.012	3.269
		E_3	5.217	0.148	0.000	-0.012	3.269
		E_4	5.217	0.148	0.000	-0.012	3.269
		E_5	5.215	0.147	0.001	-0.012	3.266

The Jackknife estimator and the Sahoo (1987) estimators are most efficient estimators when the sample size is small. The most common relation among the estimators was observed as $MSE\left(\hat{\hat{Y}}_5\right) < MSE\left(\hat{\hat{Y}}_3\right) < MSE\left(\hat{\hat{Y}}_4\right) < MSE\left(\hat{\hat{Y}}_2\right) < MSE\left(\hat{\hat{Y}}_1\right)$

and $RB\left(\hat{\hat{Y}}_2\right) \leq RB\left(\hat{\hat{Y}}_3\right) \leq RB\left(\hat{\hat{Y}}_4\right) \leq RB\left(\hat{\hat{Y}}_5\right) < RB\left(\hat{\hat{Y}}_1\right)$. Since the skewness value of the estimators is near to zero and the kurtosis is near to three, which indicates that the estimators are normally distributed for the large samples whereas the same is not true for the small samples under the population-II which has different variances for the study and auxiliary variables. Tin (1965) estimator is an efficient and less biased estimator as compared with the classical ratio estimator.

REFERENCES

1. Beale, E.M.L. (1962), Some Uses of Computer in Operational Research, *Industielle Organisation*, 31, 51-52.
2. Cochran, W.G. (1977), Sampling Techniques, Third Edition, Wiley Ltd.
3. Dalabehra, M. And Sahoo, L.N. (1994), Comparison of Six Almost Unbiased Ratio Estimators, *Questiio*, Vol. 18(3), 369-375.
4. Hartley, H.O. and Ross, A. (1954), UnbaisedRatio Estimators, *Nature*, Vol. 174, 270-271.
5. Hutchison, M.C. (1971), A Monte Carlo Comparison Of Some Ratio Estimators, *Biometrika*, 58, 313 – 321.
6. Johnson, M.E. (1987), Multivariate Statistical Simulation, John Wiley & Sons, New York.
7. Quenouille, M.N. (1956), Note on Bias Estimation, *Biometrika*, 43, 353-360.
8. Rao, J.N.K. and Beegle, L. D. (1967), A Monte Carlo Study Of Some Ratio Estimators, *Sankhya,*B29, 47- 56.
9. Rao, K.R. (2010), Comparison of Ratio Estimators Using Monte Carlo Simulation, Unpublished Ph.D. Thesis, Department of Statistics, Osmania University, Hyderabad.
10. Reddy, M.K., Rao, K. R. and Boiroju, N.K. (2010), Comparison of Ratio Estimators Using Monte Carlo Simulation, *International Journal of Agriculture and Statistical Sciences*.
11. Reddy, M.K., Srinivas, J. and Kumar, B.N. (2008), Comparison Of Mean Per Unit, Ratio And Regression Estimators In Simple Random Sampling Using Simulation, *Proceedings of the VI International symposium on optimization and statistics*, Editors: Abdul Hamid Khan and HaseebAthar, Department of Statistics and operations research, Alighar Muslim University.
12. Sahoo, L.N. (1987), On a Class of Almost Unbiased Estimator for Population Ratio, *Statistics*, 18, 119-121.
13. Sahoo, L.N. and Dalabehera, M. (1996), A Monte Carlo Comparison of Six Almost Unbiased Ratio Estimators, *Monte Carlo Methods and Appl.,* 2, 237-249.
14. Tin, M. (1965), Comparison of Some Ratio Estimators, *Jour. Amer. Stat. Assoc.,* 60, 294-307.