ISSN 1684-8403
J. Stat.
Vol. 10, No.1 (2003) PP 9-15

A MODIFIED MURTHY ESTIMATOR OF POPULATION TOTAL IN UNEQUAL PROBABILITY SAMPLING

Hina Muzaffar and Muhammad Qaiser Shshbaz

ABSTRACT A Modified Murthy Estimator of population total has been developed for use with unequal probability sampling. Design based expectation and variance of the Modified Murthy estimator has been obtained. The empirical study has also been carried out to see the relative performance of this estimator. It has been found that the t_{MM} (Modified Murthy) estimator performs better for populations having negative skewness and negative kurtosis.

KEY WORDS: Murthy Estimator, Unequal Probability Sampling.

1. INTRODUCTION:

A lot of work on unequal probability sampling estimator has been done since early forties. Hansen and Hurwitz (1943) were the first to introduce the use of unequal probability sampling estimators. Midzuno (1952), Sen (1952a, 1952b) generalized the Hansen and Hurwitz (1943) scheme to sampling a combination of *n* elements from a stratum with probability proportional to size (pps) of the combination. Sen (1952c) further generalized the scheme for obtaining an unbiased estimate of the population total when the first *r* units are selected with pps and the remaining *n-r* units are selected with equal probability and without replacement. He also derived expression for an estimate of the variance of the estimate. Horvitz and Thomson (1952) presented another technique for dealing with the problem of selecting *n* p.s.u.'s without replacement and with varying probabilities from a finite population. Sen (1954) derived expression for the unbiased estimate of the variance referred by Sen (1953) and showed that this estimate was always positive. He also gave a biased estimator for the sampling variance, which was more efficient than Horvitz and Thompson's unbiased estimate of the variance. Hartley and Rao (1962) derived expression for the variance of the estimates of the population total together with variance estimates for the selection procedure. Rao (1972) showed that the generalized π ps sampling strategy consisting of the design with π_i, the probability of inclusion of the ith unit in the sample, proportional to the modified size together with the corresponding Horvitz-Thompson estimator, is superior to the symmeterized. Hanif and Brewer (1979) presented a general theory of sampling with unequal probability, which allowed population units to appear more than once in sample. They presented two estimators for use in both single stage and multi-stage sample design.

Department of Statistics Government College University, Lahore.

Shahbaz and Hanif (2003a) developed a new estimator of population total and obtained design based expectation and variance of the estimator. Shahbaz and Hanif (2003b) developed approximate formulae for the variance of Horvitz-Thompson estimator by using the first order inclusion probabilities. They also checked the validity of the formulae through empirical study.

Shahbaz and Hanif (2004) has given a selection procedure for use with Horvitz-Thompson estimator by merging Brewer (1963) and Durbin (1967) selection procedures.

2. MURTHY ESTIMATOR

A Modified Murthy Estimator of population total have been developed by using the estimator of Murthy (1957), which is defined as:

$$t_{symm} = \frac{1}{P(S))}\sum_{i=1}^{n} P(S| \ i)y_i \qquad (2.1)$$

where $P(s|i)$ is the probability of obtaining a sample "s" given that ith unit has been already selected and $P(s)$ is the probability of obtaining a sample "s".

Murthy (1957) used the Yates–Grundy (1953) draw–by–draw procedure to obtain following unbiased estimator of population total for a sample of size 2

$$t_{symm} = \frac{\left[\frac{y_i}{p_i}(1-p_j)+\frac{y_j}{p_j}(1-p_i)\right]}{(2-p_i-p_j)} \qquad (2.2)$$

2.1 Modified Murthy Estimator

The Modified Murthy (t_{MM}) estimator has been developed by using Shahbaz (2002) selection procedure. The Murthy estimator for sample size 2 is:

$$t_{symm} = \frac{P(s|i)y_i + P(S|j)y_j}{P(S)} \qquad (2.3)$$

Modified Murthy estimator has been obtained by using Shahbaz (2002) selection procedure given as follows:

Select first unit with probability proportional to $\frac{2p_i(1-p_i)}{1-4p_i}=q_i$

Select second unit with probability proportional of remaining units.

For this selection procedure:

$P(S|i)=\frac{p_i}{1-p_i}$, $P(S|j)=\frac{p_i}{1-p_i}$,

$$P(S)=\frac{2p_ip_j}{b}\left[\frac{1}{1-4p_i}+\frac{1}{1-4p_j}\right]$$

where $b=\sum_{i=1}^{N}\frac{2p_i(1-p_i)}{1-4p_i}$

Substituting these values in Murthy estimator given in (2.3) a modified estimator is defined as:

$$t_{mm}=\frac{b(1-4p_i)(1-4p_j)}{4(1-2p_i-2p_j)(1-p_i)(1-p_j)}\left[\frac{y_i(1-p_j)}{p_i}+\frac{y_j(1-p_i)}{p_j}\right]$$

4)

This estimator is a slight modification of the Murthy (2.4) estimator and it requires verification:

In case of $p_i=p_j=\frac{1}{N}$ for sample size n=2, the estimator obtained in (2.1.1), reduce to the estimator for the total in simple random sampling, i.e.

t_{MM} $$t_{mm}=\frac{N}{2}\left(y_i+y_j\right)$$

2.2 Variance Of The Modified Murthy Estimator

The estimator is an unbiased estimator of the population total i.e $E(t_{MM})$ =Y.The variance of the estimator is obtained as:

$Var\ (t_{MM})= E(t_{MM}{}^2) - [E(t_{MM})]^2$

$$=\frac{1}{2}\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N}\left[\frac{\frac{P_i}{1-P_i}Y_i+\frac{P_i}{1-P_j}Y_j}{\frac{2P_iP_j}{b}\left[\frac{1}{1-4P_i}+\frac{1}{1-4P_j}\right]}\right]^2 * \frac{2P_iP_j}{b}\left[\frac{1}{1-4P_i}+\frac{1}{1-4P_j}\right]$$

$$-Y^2$$

After simplification we get

$$V(t_{mm})=\frac{b}{8}\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N}D_{ij}\left[A_{ij}\frac{Y_i^2}{P_i^2}+B_{ij}\frac{Y_j^2}{P_j^2}-2C_{ij}\frac{Y_iY_j}{P_iP_j}\right] \tag{2.5}$$

where

$$A_{ij}=\frac{b(1-4P_i)(1-4P_j)-4P_i(1-2P_i-2P_j)(1-P_i)P_i}{(1-P_i)^2(1-4P_i)(1-4P_j)b} \tag{2.6}$$

$$B_{ij}=\frac{b(1-4P_i)(1-4P_j)-4(1-2P_i-2P_j)(1-P_j)P_j}{b*(1-P_j)^2(1-4P_i)(1-4P_j)} \tag{2.7}$$

$$C_{ij}=\frac{-b(1-4P_i)(1-4P_j)+4(1-2P_i-2P_j)(1-P_i)(1-P_j)}{b*(1-P_i)(1-P_j)(1-4P_i)(1-4P_j)} \tag{2.8}$$

$$D_{ij}=\frac{P_iP_j(1-4P_i)(1-4P_j)}{1-2P_i-2P_j} \tag{2.9}$$

2.3 Verification Of Variance

In this section it has been verified that $V(t_{MM})$ to variance of SRS for $P_i=P_j=\frac{1}{N}$ which is an essential condition for correctness of the expression developed.

For $P_i=P_j=\frac{1}{N}$, A_{ij}, B_{ij} and C_{ij} are given as:

$$A_{ij}=\frac{N(N-2)}{(N-1)^2}$$

Since $P_i = P_j = \frac{1}{N} \Rightarrow B_{ij} = \frac{N(N-2)}{(N-1)^2}$

$$C_{ij} = \frac{N(N-2)}{(N-1)^2}$$

$$D_{ij} = \frac{1}{N^2}\frac{N-4}{N}$$

Substituting the values of A_{ij}, B_{ij}, C_{ij} and D_{ij} in (2.2.1)

$$V(t_{MM}) \quad \frac{2}{8}\left(\frac{N-1}{N-4}\right)\left(\frac{1}{N^2}\right)\left(\frac{N-4}{N}\right)$$

$$\left[\frac{N^2N(N-2)}{(N-1)^2}\sum_{i=1}^{N}Y_i^2 + \frac{N^2N(N-2)}{(N-1)^2}\sum_{j=1}^{N}Y_j^2 - \frac{2N^2N(N-2)}{(N-1)^2}\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N}Y_iY_j\right]$$

$$= \frac{N^2}{4}\frac{1}{N^2}\left(\frac{N-2}{N-1}\right)\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N}(Y_i - Y_j)^2$$

In case of simple random sampling, for n=2 the variance of the total is

$$= \frac{N^2(N-2)}{2N}S^2$$

Where

$$S^2 = \frac{1}{2N}\left(\frac{1}{N-1}\right)\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N}(Y_i - Y_j)^2$$

It can be seen that the variance formula reduces to the expression of Simple Random Sampling for $P_i = P_j = \frac{1}{N}$. Hence the formula is true.

3. EMPIRICAL STUDY

The Modified Murthy estimator is developed. This estimator has been compared with the following estimators.

1. Hansen- Hurwitz Estimator
2. Horvitz-Thompson Estimator [YG-bdb]
3. Horvitz -Thompson Estimator [BDS Proc]
4. Rao -Hartly -Cochran Estimator.
5. Raj Estimator
6. Murthy Estimator

3.1 Populations Selected For The Study:

For the empirical study of the t_{MM} estimator twenty populations are selected from standard textbooks with regard to different measure; all the populations are natural.

The efficiencies of above developed populations has been compared with various estimators by using some standard populations. The complete result of these analyses has been given in the following table.

Table 1: Population Variance Of Different Estimators

Pop	Hansen-Hurwitz	Horvitz-Thompson YG	Horvitz-Thompson BDS	Rao-Hartly Cochran	Raj	Murthy	Modified Murthy	Best
1	34424.67	32568.22	31917.29	32129.72	32067.75	31891.73	31051.92	Mod.Murthy
2	105.996	94.615	94.101	90.854	93.543	91.543	91.852	RHC
3	816.35	766.37	754.03	753.56	758.22	753.77	660.52	Mod.Murthy
4	862.73	826.82	829.02	825.22	824.53	822.60	849.23	Murthy
5	199.49	177.165	177.41	177.32	179.54	177.32	194.55	RHC
6	1317.32	1248.41	1215.89	1248.04	1234.53	1228.40	1044.86	Mod.Murthy
7	391832.13	376569.28	370586.0	371209.38	371006.19	369816.72	331138.84	Mod.Murthy
8	44782804	43200928	42875176	42425816	42707620	42605616	41551304	Mod.Murthy
9	42668.53	35870.34	132359.26	133157.29	133291.91	132622.94	111366.20	Mod.Murthy
10	1618.26	1387.12	1443.10	1493.78	1465.73	1449.36	2122.43	HT YG

From above results we can see that the Modified Murthy Estimators performs better than the other estimators in 6o% of populations. The Rao-Hartly-Cochran estimators performs better than the other estimators in 20% of populations whereas as the Horvitz-Thompson estimator under the Yates-Grundy draw by draw and Murthy estimator performs better in 10% of populations.

4. CONCLUSIONS:

We conclude that the Modified Murthy estimator give smaller variance for the populations having negative skewness and negative kurtosis when compared to the other estimators. While Modified Murthy estimator provide larger variance for the populations having positive skewness and negative kurtosis on the comparison.

ACKNOWLEDGEMENT

The authors are thankful to Prof. Dr. Afzal Baig his kind cooperation and useful comments during the course of my research.

REFERENCES

1. Hanif, M and Brewer, K. R. W. (1979). Generalization of the Horvitz and Thompson estimator. Punjab university. J. Math. Vol.12-13, 123-134.
2. Hansen, M. H. and Hurwitz, W. N (1943). On the theory of sampling from finite population.Ann.Math.Stat.,14.333-362
3. Hartly, H .O and Rao, J.N, K (1962). Sampling with un equal probability cal. Ann. Maths. stat. 33
4. Horvitz, D. G and Thompson, D. J. Ageneraliztion of sampling without replacement from a finite Universe. J.A.SA.,47,663-685
5. Midzuno, H. (1952). On the sampling system with probability proportional to sums of sizes. Ibid, 3, 99-107
6. Murthy, M. N. (1957). Ordered and unordered estimators in sampling without replacement, Sankhya, 18, 379 – 390.
7. Rao, T. J (1972). Horvitz-Thompson and Des Raj estimators Revisited. The Australian Journal of statistics, 14. 227-230.
8. Samiuddin, M., Kattan, A. K. A., Hanif, M. and Asad, H. (1992). Some remarks on models, sampling schemes and estimators in unequal probability sampling, *Pak. J. Stat.*, 8(1), A, 1 – 18.
9. Sen A.R (1952c). On the selection of n primary sampling units from stratum structure (n=2).
10. Sen A.R (1953). Recent advances in sampling with varying probability cal.stat.Assoc.bull.5,1-15.
11. Sen, A. R. (1952a). Further developments of the theory and application of the selection of primary sampling unit's special reference to N. C. Agricultural population. Thesis. Library, N.C. State College.
12. Shahbaz, M. Q. Hanif, M. and Ahmad, M. (2001). Some estimators for unequal probability sampling 'Proc. ICCS-VII.Vol.12, 297-302.
13. Shahbaz and Hanif (2003a). Modifications of Murthy estimator under some selection procedures,Pak. J. Stat, Vol.19(2), 201-208
14. Shahbaz, M.Q, Hanif M.and Samiuddin, M. (2003). On some estimators in unequal probability sampling, J. Stat. Theo. & App. To appear.