Pak J. Statist.
2004 Vol. 20(2) pp 297 – 302

MODIFICATION OF MURTHY'S ESTIMATOR USING GENERAL SELECTION PROCEDURE FOR UNEQUAL PROBABILITY SAMPLING WITHOUT REPLACEMENT

Nadeem Shafique Butt
Department of Social and Preventive Pediatrics,
K.E.M.C. Mayo Hospital, Lahore (Pakistan)
Email: nadeemshafique@hotmail.com

and

Muhammad Qaiser Shahbaz
Department of Statistics, Govt. College University
Lahore (Pakistan)
Email: drshahbaz@stats.gcu.edu.pk

ABSTRACT

A new estimator of population total has been developed following the method of Murthy (1957) by using the Shahbaz and Hanif (2003) General selection procedure. Two special cases have been obtained of the general estimator. Empirical study has been carried out to obtain the most suitable value of the constant involved.

KEY WORDS

Unequal probability sampling, Murthy's (1957) Estimator, Shahbaz and Hanif (2003) selection procedure.

1. INTRODUCTION

An estimator of population total in case of unequal probability sampling without replacement, proposed by Murthy (1957), was:

$$t_{symm} = \frac{\sum_{i=1}^{n} P(S|i) y_i}{P(s)} \tag{1.1}$$

where P(s|i) = conditional probability of getting the set of units that was drawn given that the ith unit was drawn first. Also P(s) = unconditional probability of getting the set of units that was drawn.

For sample of size 2 the estimator in (1.1) under Yates-Grundy (1953) selection procedure takes the form:

$$t_{symm} = \frac{1}{2-p_i-p_j}\left[(1-p_j)\frac{y_i}{p_i}+(1-p_i)\frac{y_i}{p_j}\right] \tag{1.2}$$

The sampling variance of (1.2) is given by:

$$V(t_{symm}) = \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{P_iP_j(1-P_i-P_j)}{2-P_i-P_j}\left(\frac{Y_i}{P_i}-\frac{Y_j}{P_j}\right)^2 \tag{1.3}$$

Shahbaz and Hanif (2003a) has obtained following estimator using Brewer (1963) selection procedure in (1.1)

$$t_{symm} = \frac{b(1-2p_i)(1-2p_j)\left[\frac{y_i}{p_i}(1-p_j)+\frac{y_i}{p_i}(1-p_i)\right]}{4(1-p_i-p_j)(1-p_i)(1-p_j)} \tag{1.4}$$

where

$$b = 1+\sum_{i=1}^{N}\frac{p_i}{1-2p_i}$$

Shahbaz and Hanif (2003a) has obtained the following the design based variance of estimator in (1.4):

$$Var(t_{MM(1)}) = \frac{1}{2}\sum_{i\neq j=1}^{N}\sum\frac{P_iP_j}{4(1-P_i-P_j)}\left[\frac{Y_i^2}{P_i^2}Aij+\frac{Y_j^2}{P_j^2}B_{ij}-2\frac{Y_iY_j}{P_iP_j}C_{ij}\right] \tag{1.5}$$

$$A_{ij} = \frac{b(1-2P_i)(1-2P_j)}{(1-P_i)^2}-\frac{4P_i(1-P_i-P_j)}{(1-P_i)} \tag{1.6}$$

$$B_{ij} = \frac{b(1-2P_i)(1-2P_j)}{(1-P_j)^2}-\frac{4P_j(1-P_i-P_j)}{(1-P_j)} \tag{1.7}$$

$$C_{ij} = 4(1-P_i-P_j)-\frac{b(1-2P_i)(1-2P_j)}{(1-P_i)(1-P_j)} \tag{1.8}$$

2. MODIFIED MURTHY ESTIMATIOR

In this section we have developed the modified Murthy estimator by using the Shahbaz and Hanif (2003b) selection procedure.

Murthy(1957) used the Yates-Grundy(1953) draw-by-draw method and develop the estimator for population total given in (1.2). Using Shahbaz and Hanif (2003b) selection procedure we have

$$p(s|i) = \frac{p_j}{1-p_i} \quad \text{and} \quad p(s|j) = \frac{p_i}{1-p_j}$$

also p(s) for this selection procedure is given as:

$$p(s) = \frac{1}{\frac{k}{2}}\left[\frac{a\,p_i\,p_j}{1-2a\,p_i} + \frac{a\,p_i\,p_j}{1-2a\,p_j}\right] \qquad \text{for } a > 0.5 \text{ and } p_i < \frac{1}{2a}$$

where k is constant given as

$$k = 1 + (2a-1)\sum_{i=1}^{N}\frac{p_i}{1-2a\,p_i}$$

$$t_{symm\,(mod)} = \frac{k(1-2ap_i)(1-2ap_j)}{4a(1-ap_i-ap_j)}\left[\frac{y_i}{p_i(1-p_i)} + \frac{y_j}{p_j(1-p_j)}\right]$$

3. DESIGN BASED VARIANCE OF THE MODIFIED MURTHY ESTIMATOR

The unbiasedness of Modified Murthy estimator may be proved in a straightforward way. The variance of Modified Murthy Estimator is derived as:

$$\text{var}\left(t_{symm\,(mod)}\right) = E\left(t_{symm\,(mod)}\right)^2 - \left[E\left(t_{symm\,(mod)}\right)\right]^2$$

After slight algebraic manipulation we have:

$$\text{var}(t_{symm\,(mod)}) = \frac{k}{8}\sum_{j\neq i}^{N}\sum D_{ij}\left[A_{ij}\frac{Y_i^2}{P_i^2} + B_{ij}\frac{Y_j^2}{P_j^2} - C_{ij}\frac{2Y_iY_j}{P_iP_j}\right] \quad (3.1)$$

where

$$A_{ij} = \left\{ \frac{1}{(1-P_i)^2} - \frac{4aP_i(1-aP_i-aP_j)}{k(1-P_i)(1-2aP_i)(1-2aP_j)} \right\} \tag{3.2}$$

$$B_{ij} = \left\{ \frac{1}{(1-P_j)^2} - \frac{4aP_j(1-aP_i-ap_j)}{k(1-P_j)(1-2aP_i)(1-2aP_j)} \right\} \tag{3.3}$$

$$C_{ij} = \left\{ -\frac{1}{(1-P_i)(1-P_j)} + \frac{4a(1-aP_i-aP_j)}{k(1-2aP_i)(1-2aP_j)} \right\} \tag{3.4}$$

$$D_{ij} = \frac{P_iP_j(1-2aP_i)(1-2aP_j)}{a(1-aP_i-aP_j)} \tag{3.5}$$

It can be easily seen that the estimator given in (2.1) and the variance given in (3.1) transformed to the result of Murthy (1957) for a=0.5 and to Shahbaz and Hanif (2003a) for a=1.0.

4. EMPIRICAL STUDY AND DISCUSSION

In this section the empirical study has been given for various values of "a". To carry out study fifty natural populations have been used, which are given in standard texts on sampling techniques. The sampling variance of the estimator given in section (3.1) has been obtained for various values of "a" for all populations. After evaluating the sampling variance, ranking has been done for each estimator according to sampling variance. An estimator with smallest variance has been given a rank of 1, the estimator with second smallest given a rank of 2 and so on. After obtaining the rank of 50 populations, the average rank of each value of "a" has been calculated for various ranges of coefficient of variation of measure of size and correlation coefficient between actual variable and measure of size. The result of this empirical study has been given in tables 1 through 3. From table 1 we can readily see that a = – 0.5 clearly outperform all other competing values. From table 2 we can see that for small coefficient of variation the values of a = 1.0 and a = – 1.0 are equal in performance. For moderate and large coefficient of variation a = – 0.5 clearly outperform other values. Table 3 also shows similar sort of picture. In general we can see that the value of a = – 0.5 is the best for this estimator.

ACKNOWLEDGEMENT

We are thankful for the referee of suggesting useful points, which improve the text of the paper. We are also thankful to Prof. Dr. Muhammad Hanif who supported us a lot throughout this project.

REFERENCES

1. Brewer, K. R. W. (1963) A model of systematic sampling with unequal probabilities. *Australian Jour. Stat.,* 5, 5 – 13.
2. Murthy, M. N. (1957) Ordered and unordered estimators in sampling without replacement, *Sankhya*, 18, 379 – 390.
3. Shahbaz, M. Q. and Hanif, M. (2003a) A modification of Murthy's estimator under various selection procedures for unequal probability sampling without replacement. *Pak. J. Statist.* Vol. 19(1) 151 – 160.
4. Shahbaz, M. Q. and Hanif, M. (2003b) A general procedure for unequal probability without replacement and sample size 2. *J. Applied Statistical Sciences.* Vol. 13 to appear (USA).
5. Yates, F. and Grundy, P. M. (1953) Selection without replacement from within strata with probability proportional to size. *J. Roy. Stat. Soc.*, B, 15, 153 – 161.

Table 1:
Frequency table of ranks of various values of "a"

Rank	**A**							
	-2	**-1.5**	**-1**	**-0.5**	**0.5**	**1**	**1.5**	**2**
1	16	1	4	6	6	10	3	4
2	1	19	6	7	4	6	5	2
3	1	4	20	7	10	3	4	1
4	4	5	7	22	1	9	1	1
5	6	7	4	8	1	22	1	1
6	6	3	9	0	28	0	2	2
7	5	11	0	0	0	0	33	1
8	11	0	0	0	0	0	1	38
Average	**4.52**	**4.02**	**3.56**	**3.38**	**4.42**	**3.54**	**5.7**	**6.86**

Table 2:
Frequency table of ranks of various values of "a" for various ranges of C.V.(X)

CV	A							
	-2.0	**-1.5**	**-1.0**	**-0.5**	**0.5**	**1.0**	**1.5**	**2.0**
1 – 10	4.50	3.90	3.60	3.70	4.20	3.60	5.70	6.80
11 – 20	3.70	3.50	3.40	3.20	5.10	3.40	6.40	7.30
21 – 30	5.20	4.70	4.20	4.00	3.50	3.60	4.80	6.00
31 – 40	4.50	3.90	3.40	2.90	5.10	3.60	5.90	6.70
41 – 50	4.70	4.10	3.20	3.10	4.20	3.50	5.70	7.50
Average	**4.52**	**4.02**	**3.56**	**3.38**	**4.42**	**3.54**	**5.70**	**6.86**

Table 3:
Frequency table of ranks of various values of "a" for various ranges of ρ_{xy}

Rxy	A							
	-2	**-1.5**	**-1**	**-0.5**	**0.5**	**1**	**1.5**	**2**
1 – 10	4.8	3.9	3.5	3.5	4.3	3.5	5.6	6.9
11 – 20	4.8	4.2	3.6	3.7	4.3	3.7	5.3	6.4
21 – 30	3.2	3.2	3.4	3.4	5.1	4	6.4	7.3
31 – 40	4.3	4.2	3.8	3.9	4.5	4	5	6.3
41 – 50	5.5	4.6	3.5	2.4	3.9	2.5	6.2	7.4
Average	**4.52**	**4.02**	**3.56**	**3.38**	**4.42**	**3.54**	**5.7**	**6.86**