Journal of Biometrics & Biostatistics
ISSN: 2155-6180

Anieting, J Biomet Biostat 2015, 6:1
DOI: 10.472/2155-6180.1000222

Short Communication Open Access

Comparison of Horvitz and Thompson Estimator with that of Rao, Hartley and Cochran Estimator in PPS without Replacement Scheme

Anieting AE*

Department of Mathematics and Statistics, University of Uyo. P.M.B 1017, Uyo. Nigeria

Abstract

This paper was focused on the comparison of Horvitz and Thompson estimator of population total with that of Rao, Hartley and Cochran estimator in PPS without replacement scheme when a sample of size six is taken from the same finite population. The data used were from Nigerian Bureau of Statistics bulletin. The result showed that the variances of both estimators gave positive values but the variance of the estimator by Rao, Hartley and Cochran was smaller making it a better estimator.

Keywords: PPS without replacement; Rao, Hartley and Cochran estimator; Horvitz and Thompson estimator

Introduction

Unequal probability sampling scheme or proportional to size (PPS) sampling scheme ensures that the possibility of selecting each unit into the sample is in proportion to the size or measure of size of the unit provided that the sizes of the individual units in the population are known and are correlated with the variable of interest, Horvitz and Thompson (1952) were the first to give theoretical frame work of PPS without replacement [1]. The sample variance of the population total estimator given by Horvitz and Thompson (1952) has the possibility of giving negative values for some samples. This problem has made statistician to work on this scheme. They include Yates and Grundy(1953), Sen (1953), Durbin(1953), Rao, Hartley and Cochran (1962) etc. Rao, Hartley and Cochran (1962) gave a modified estimator based on PPS without replacement scheme. The variance of the estimator for population total under Rao, Hartley and Cochran is always positive. But how do the estimators given by Horvitz and Thompson , and that given by Rao,Hartley and Cochran fare when the sample size is more than two, say six, assuming that the variance of the population total estimator by Horvitz and Thompson gives a positive value. The data for empirical verification of this study is from Nigerian Bureau of Statistics bulletin [1].

The two procedures

Horvitz and thompson estimation (procedure): First, the cumulative totals and ranges of the measure of size are formed [2]. Then to select a sample of size n without replacement, a random start between 1 and k inclusive (K=X/n) is selected using a table of random numbers. If the number selected is r then the units in the sample are those in whose ranges the umbers r,r+k,r+2k,….,r+(n-1)k fall. The probability of selecting the unit U_i in the sample of size n is

$$PrU_i = \pi_i \frac{X_i}{k} = P_i$$

If any unit in the population has its size greater than or equal to k such unit is removed before sampling and taken into the sample with probability unity. The probability that any pair of units (U_i , U_j) is together in the sample is

$$\pi_{ij} = q_{ij} / X = q_{ij} / k \qquad i \neq j$$

Where q_{ij} is the number of the random numbers between 1 and k inclusive, which will select U_i and U_j simultaneously in the sample?

An unbiased estimator of the population total for PPS without replacement as given by Horvitz and Thompson is

$$\hat{Y}_{HT} = \sum_i^n y_i / \pi_i$$

While the sample estimator of $V(\hat{Y}_{HT})$ is given as

$$V(\hat{Y}_{HT}) = \sum_i^n \frac{(1-\pi_i) y_i^2}{\pi_i^2} + 2\sum_i^n \sum_{j>i}^n \frac{\pi_{ij-\pi_i\pi_j}}{\pi_{ij}} \cdot \frac{y_i y_j}{\pi_i \pi_j} \quad \text{Provided } \pi_{ij} > 0 \ldots\ldots 2.1.1$$

Rao, hartley and cochran estimation (procedure): Divide a population of N units into n groups at random with group g containing N_g units (g=1,2,….,n) such that $N_1 + \ldots + N_n = N$. Thereafter select one unit independently from each group [3]. This will give a total of n units selected in the sample with PPS without replacement. The probability of selecting U_i in the sample in g^{th} group is $P_{i=}^* = \frac{X_i}{X_g} = \frac{P_i}{P_g}$ where $X_g = \sum_i^{N_g} X_i$

X_i's are the auxiliary variables and P_g is the sum of the initial probabilities in g^{th} group.

The Rao, Hartley and Cochran estimator of the population total is

$$\hat{Y}_{RHC} = \frac{\sum_{g=1}^n y_{g_i}}{P_i^*}$$

Where y_{gi} is the value of the study variate for the i^{th} unit in g^{th} group and $P_i^* = \frac{P_i}{P_g}$

The estimator of the variance is given as

$$V(\hat{Y}_{RHC}) = \frac{N-n}{N(n-1)} \sum_{g=1}^n P_g \frac{y_{g_i}^2}{P_i^2} - \hat{Y}_{RHC}^2 \qquad (2.2.1)$$

***Corresponding author:** Anieting AE, Department of Mathematics and Statistics,University of Uyo. P.M.B 1017, Uyo. Nigeria, E-mail: akananieting@yahoo.com

Received April 21, 2015; **Accepted** May 22, 2015; **Published** May 29, 2015

Citation: Anieting AE (2015) Comparison of Horvitz and Thompson Estimator with that of Rao, Hartley and Cochran Estimator in PPS without Replacement Scheme. J Biomet Biostat 6: 222. doi:10.4172/2155-6180.1000222

S/N	States	X_i	Range	y_i
1	Abia	18	1-18	22
2	Adamawa	21	19-39	41
3	Akwa Ibom	31	40-70	65
4	Anambra	20	71-91	52
5	Bauchi	8	92-111	38
6	Bayelsa	23	112-119	37
7	Benue	27	120-142	27
8	Borno	18	143-169	42
9	Cross River	25	170-187	25
10	Delta	12	188-212	47
11	Ebonyi	18	213-224	29
12	Edo	16	225-242	31
13	Ekiti	16	243-258	21
14	Enugu	11	259-274	21
15	Gombe	27	275-285	14
16	Imo	27	286-312	63
17	Jigawa	27	313-339	65
18	Kaduna	23	340-362	59
19	Kano	45	363-407	123
20	Katsina	34	408-442	38
21	Kebbi	20	443-461	9
22	Kogi	21	462-482	40
23	Kwara	16	483-498	7
24	Lagos	20	499-518	43
25	Nassarawa	13	519-531	60
26	Niger	26	532-557	39
27	Ogun	20	558-577	25
28	Ondo	18	578-595	36
29	Osun	30	596-625	50
30	Oyo	33	626-658	47
31	Plateau	17	659-675	20
32	Rivers	23	676-698	34
33	Sokoto	23	699-721	3
34	Taraba	16	722-737	71
35	Yobe	17	738-754	19
36	Zamfara	14	755-768	35

Table 1: The correlation coefficient Are between the auxiliary variable x_i and the variable of interest y_i.

s/n	sample	π_i	y_i	y_i/π_i
1	Bauchi	20/128	38	243.2
2	Edo	18/128	31	220.4
3	Kaduna	23/128	59	**328.3**
4	Kwara	16/128	7	56
5	Osun	30/128	50	213.3
6	Yobe	17/128	19	143.1
				1204.3

Table 2: The population total estimate using Horvitz and Thompson estimator.

Selected states	P_i^*	P_i	P_g	y_{gi}	y_{gi}/P_i^*
Bornu	27/137	27/768	137/768	42	213.11
Benue	23/126	23/768	126/768	27	147.91
Edo	18/133	18/768	133/768	31	229.06
Taraba	16/101	16/768	101/768	71	448.19
Delta	25/113	25/768	113/768	47	212.44
Ogun	20/158	20/768	158/768	25	197.5
					1448.21

Table 3: Using eqn 2.1.1., the variance gives 690317.62 also using 2.2.1 the variance gives 102892.4.

Analysis

The correlation coefficient r between the auxiliary variable x_i (the number of local government in each state) and the variable of interest y_i (the number of police stations in Nigeria in each state) is 0.60. Other results of the analysis are summarized in Tables 1-3.

Results and discussions

Based on a sample size of six from a population of 36 units, Tables 2 and 3 show that the population total estimate using Horvitz and Thompson estimator is 1204 which is smaller than 1448, the population total estimate using Rao,Hartley and Cochran estimator. Also the variance of the estimate of the population total by Horvitz and Thompson is 690317.62 which is bigger than 102892.4, the variance of the estimate of the population total by Rao, Hartley and Cochran [4].

Conclusion

Since the variance of the estimator given by Rao, Hartley and Cochran gives a value smaller than that given by Horvitz and Thompson with respect to sample size of six,it shows that the estimator given by Rao,Hartley and Cochran is better than that given by Horvitz and Thompson.

References

1. Alodat NA (2009) on unequal probability Sampling Without Replacement Sample Size 2. Int J Open Problems Comp Maths 2: 1 108-112.
2. Nigerian Bureau of Statistics bulletin (2008 Edition).
3. Dawodu OO, Adeware AA, Oshundare IO (2013) Rao, Hartley and Cochran's Sampling Scheme: Application. Mathematical Theory and Modeling 3: 8 101-103.
4. Okafor FC (2013). Sample Survey Theory with Applications. Afro-Orbis Publication Ltd, 2002. Open Journal of Statistics.