

Efficiency of Alodat Sample Selection Procedure over Sen - Midzuno and Yates - Grundy Draw by Draw under Unequal Probability Sampling without Replacement Sample Size 2

O.O. DAWODU & A.A. Adewara

Department of Statistics, University of Ilorin, Ilorin, Nigeria

O. M. Olayiwola

Department of Physical Sciences, Ajayi Crowther University

Oyo, Oyo State, Nigeria

E-mail: oluwafunmilolajoshua@yahoo.com

Received: November 8, 2010 Accepted: December 7, 2010 doi:10.5539/jmr.v3n2p113

Abstract

This paper compare the efficiency of Alodat sample selection procedure over Sen - Midzuno and Yates - Grundy draw by draw using Yates - Grundy estimator under unequal probability sampling without replacement sample size 2, carried out using the data from the 2008 Demographic and health survey in Nigeria. We studied the distribution of pregnant women age 15 - 49, and children under age five in Nigeria, who use mosquito nets as a means of preventing malaria. These data sets are: (1) the number of pregnant women age 15 - 49 who slept under mosquito nets the night before the survey, and (2) the number of children under age five who slept under mosquito nets the night before the survey. (1) and (2) above are the variables of interest. The data were collected based on the six geo-political zones in Nigeria [i.e. South South, South West, South East, North West, North East, North Central]. The auxiliary variable is the number of Local Government in each geo-political zone in Nigeria. The Yates - Grundy estimate obtained using Alodat sample selection is more efficient than using Sen - Midzuno and Yates - Grundy selection procedures.

Keywords: Efficiency, Sample selection, Procedure, Probability sampling, Sample size

1. Introduction

For the purpose of this research work, interest is on the unequal probability sampling (UPS), which is also called probability proportional to size (PPS) sampling. It can be done with and without replacement of the unit selected. The abundance of works about sampling designs without replacement with fixed sample size is quite disconcerting. No less than fifty sampling procedures are presented in the famous paper of HANIF and BREWER [1980]. Several other recent papers deal with this matter as well. Among them, one can cite the CHAO updating procedure [1982] and the DEVILLE method [1992]. The study of UPS is a big task. Several people have studies the subject. They include; Horvitz and Thompson (1952), Yates and Grundy (1953), Sen (1953), Durbin (1953), Rao, Hartley and Cochran (1962), Hansen and Hurwitz (1943), etc. Horvitz and Thompson (1952) were the first to provide a complete theoretical frame work for UPS without replacement. There have been a lot of improvement on UPS over the years. The search for a sampling design with unequal probabilities is a relatively open problem. A 'good' solution should however respect the following properties.

1. The procedure should be exact in the sense that the units should be selected exactly with probabilities equaling

$$\pi_i, i \in U. \quad (1)$$

2. The procedure should be general i.e. it should be possible to apply it to any set of first - order inclusion probabilities fixed a priori, which satisfy the relation (1).
3. The $p(s)$ obtained by means of the sampling method should not depend on the order of the units on the data file.
4. The joint inclusion probabilities should be easy to compute without examining all the probabilities $p(s)$.
5. The joint inclusion probabilities should be strictly positive.
6. The joint inclusion probabilities should verify the Yates-Grundy condition:

$$\pi_{ij} \leq \pi_i \pi_j \text{ for all } i \neq j \text{ (Tille, 1996)} \quad (2)$$

Here, in dealing with UPS, we studied the distribution of pregnant women age 15 - 49 and children under age five in all house holds in the six geo-political zones in Nigeria, who slept under a mosquito net (treated or untreated), under an ever

- treated mosquito net, and under an insecticide - treated net.

This is carried out using the 2008 Demographic and Health Survey in Nigeria, which was implemented by Nigerian National Population Commission (NPC).

Malaria is endemic throughout Nigeria. The National Malaria Control Strategic Plan (NMCSP) addresses National Health Development priorities including the Roll Back Malaria Goals, and the Millennium Development Goals.

The NMCSP includes the following priorities: to reduce malaria related mortality, to reduce malaria parasite prevalence in children under age five, to increase ownership and use of insecticide - treated net and long - lasting insecticidal nets, to introduce and scale up indoor residual spraying, to increase the use of diagnostic tests for fever patients, to improve appropriate and timely treatment of malaria, and to increase coverage of intermittent preventive treatment of malaria during pregnancy.

It is restricted to the use of mosquito nets as a way of preventing malaria among pregnant women age 15 - 49 years and children under age five years in Nigeria.

The main objective of carrying out this research is to determine the best selection procedure out of the three procedures under consideration for estimating the population of pregnant women age 15 - 49 and children under age five in Nigeria, who use mosquito nets to prevent malaria.

2. Data Used

The data used for this research is secondary data on the number of Local Government in each geo - political zone in Nigeria collected from the National Population Commission, Ibadan. Nigeria on the use of mosquito nets extracted from the 2008 Demographic and Health Survey in Nigeria as shown in tables 1 - 3(b).

3. Sample Selection Procedures Considered

(i) Sen - Midzuno: This selection procedure is reported by Horvitz and Thompson (1952) and is applicable for a sample of any size. This selection procedure is stated as - Select first unit with P_i probability - Select a sample of size $n - 1$ from remaining units with equal probability and without replacement. The quantities π_i and π_{ij} for this selection procedure are given as:

$$\pi_i = P_i + \frac{(n-1)}{(N-1)}(1 - P_i) \quad (3)$$

$$\pi_{ij} = \frac{(n-1)}{(N-1)} \frac{(N-n)}{(N-2)} \left[(P_i + P_j + \frac{(n-2)}{(N-2)}) \right] \quad (4)$$

therefore, for a sample of size $n = 2$ and $N = 6$, which is the basis of this research,

$$\pi_{ij} = P_i + 0.2(1 - P_i) \quad (5)$$

$$\pi_i = 0.2(P_i + P_j) \quad (6)$$

(ii) Yates - Grundy Draw - by - Draw: This selection procedure was developed by Yates and Grundy (1953) and also reported by Durbin (1953) and Hajek (1964). The selection procedure is stated as:

- Select first unit with probability proportional to size

- Select second unit with probability proportional to size of remaining units. The quantities π_i and π_{ij} for this selection procedure are given as:

$$\pi_i = P_i \left[1 + S - \frac{P_i}{1 - P_i} \right]. \quad (7)$$

where

$$S = \sum_{j=1}^N \frac{P_j}{1 - P_j} \quad (8)$$

$$\pi_{ij} = P_i P_j \left[\frac{1}{1 - P_i} + \frac{1}{1 - P_j} \right] \quad (9)$$

(iii) Alodat The procedure is given for the use with Horvitz and Thompson estimator with sample size 2.

- Select first unit with probability proportional to.

$$\frac{P_i}{2 - P_i} \quad (10)$$

and without replacement.

- Select second unit with probability

$$\frac{P_i}{1 - P_j} \quad (11)$$

The first inclusion probability is.

$$\Pi_i = \frac{P_i}{B(2 - P_i)} + \sum_{j \neq i=1} \frac{P_j}{(2 - P_i)} x \frac{P_i}{1 - P_j} \quad (12)$$

$$\Pi_i = \frac{P_i}{B} \left[\frac{1}{2 - P_i} + \sum_{j=1} \frac{P_j}{(2 - P_j)(1 - P_j)} - \frac{P_i}{(2 - P_i)(1 - P_i)} \right] \quad (13)$$

where

$$B = \sum_{i=1} \frac{P_i}{2 - P_i} \quad (14)$$

$$\Pi_i = \frac{P_i}{B} \left[\frac{1 - 2P_i}{(2 - P_i)(1 - P_i)} + \sum \frac{P_j}{(2 - P_j)(1 - P_j)} \right] \quad (15)$$

$$\Pi_{ij} = P_i P(j/1) + P_j P(i/j) \quad (16)$$

$$\Pi_{ij} = \frac{P_i P_j}{B} \left[\frac{1}{(2 - P_i)(1 - P_j)} + \frac{1}{(2 - P_j)(1 - P_i)} \right] \quad (17)$$

Our intension is to apply this sample selection procedure to Yate and Grundy estimator instead of the Horvitz and Thompson estimator with sample size 2 which was initially proposed for.

4. Estimator Considered

Yate and Grundy Estimator.

Yates and Grundy (1953) suggested an alternative estimator of variance which is believed to be "less often negative than Horvitz and Thompson estimator of variance" and its unbiased estimator of variance is.

$$V(Y_{YG}) = \sum_{i=1}^n \sum_{j>1}^n \left[\frac{\Pi_i \Pi_j - \Pi_{ij}}{\Pi_{ij}} \right] \left[\frac{y_i}{\Pi_i} - \frac{y_j}{\Pi_j} \right]^2 \quad (18)$$

In other to find the estimates of the three selection procedures under consideration and Yates and Grundy estimator, we then have these summary of inclusion probabilities shown in table 4 from the data set to be used in other to make our computations easier. The following conditions are met

< Table 4 >

$$\sum_{i=1}^n P_i = 1 \quad (19)$$

$$\sum_{i=1}^n \prod i = 2 \quad (20)$$

$$\sum_{j=1}^n \prod_{i \neq j} i = \prod i \quad (21)$$

$$\sum \sum \prod_{i \neq j} i = 1 \quad (22)$$

5. Summary of Results

Using the inclusion probabilities shown in table 4, we then compute the variance of Yates and Grundy estimator based on the three selection procedures for the two data sets used as shown in table 5.

6. Discussion of Results

Considering the analysis and estimates obtained, Alodat selection procedure gave the least variance under both data sets. This is followed by Yates and Grundy Draw - by Draw procedure. Sen - Midzuno procedure gave the highest variance using both data sets. Alodat (2009) selection procedure is the best selection procedure in this paper.

7. Conclusion

Finally, for estimating the number of pregnant women age 15 - 49 and children under five years in the six geo - political zones in Nigeria, who use mosquito nets as a means of preventing malaria, using unequal probability sampling without replacement on Yates and Grundy estimator, among all the three selection procedures considered in this paper, Alodat (2009) selection procedure is the most efficient of them all.

References

- Brewer, K.W. R. (1963). A model of systematic sampling with unequal probabilities. *Australian Jour. Stat.*, 5, pp. 5-13.
- Chao, M.T. (1982). A General Purpose Unequal Probability Sampling Plan. *Biometrika*, 69, pp. 653-656.
- Deville, J.C. (1992). Une nouvelle (encore une!) methode de triage a probabilités inegales. Manuscript. Paris. INSAE.
- Durbin, J. (1953). Some results in sampling theory when the units are selected with unequal probabilities. *Journal Royal Stat. Soc.*, B15, pp. 262 - 269.
- Hajek, J. (1981). *Sampling from Finite Population*. New York. Marcel Dekker.
- Hanif, M., Brewer, K.R.W. (1980). Sampling with Unequal Probabilities without Replacement: a review. *International Statistical Review*, 48, pp. 217 - 335.
- Hanson, M.H. and Hurwitz, W.N. (1943). On the theory of sampling from finite populations. *Ann. Math. Stat.*, 14, pp. 333 - 362.
- Horvitz, D. G. and Thompson, D. J. (1952). A generalization of sampling without replacement from a finite population. *Jour. American Statistics Ass.*, Vol. 47, 663 -685.
- Naser. A. Alodat. (2009). On unequal probability sampling without Replacement (Sample Size 2). *Int. J. Open Problems Comp. Maths.*, Vol. 2, No. 1, pp. 109 - 112.
- Rao, J.N.K., Hartley, H.O. and Cochran, W.G. (1962). A simple procedure of unequal probability sampling without replacement. *Journal Roy. Stat. Soc.*, B24, pp. 482 - 491.
- Sen, A. R. (1953). On the estimate of the variance in sampling with varying probability. *Jour. Ind. Soc. Agric. Statistic*, Vol. 5, pp. 119 - 129.
- Yates, F. and Grundy P. M. (1953). Selection without replacement from within strata with probability proportional to size. *Jour. Roy Statist. Soc. B*, Vol. 15, pp. 253 - 261.
- Yves Tille. (1996). Some remarks on unequal probability sampling designs without replacement. *Annales D'conomie et de Statistique*, No. 44.

Table 1. Distribution of the number of Local Government in each Geo - Political Zones in Nigeria

Zones	Total no. of Local Govt. in each Zone (X_i)
NC	121
NE	112
NW	186
SE	95
SS	123
SW	137
Total	774

Source: 2008 Demographic and Health Survey in Nigeria.

NC = North Central

NE = North East

NW = North West

SE = South East

SS = South South

SW = South West

Table 2a. Distribution of the number of children under age five who slept under mosquito nets (in percentage).

Zones	WSUAN	WSUAETN	WSUAITN	TPWSUNPZ	TNOCPZ
NC	9.7	9.4	3.8	22.9	3607
NE	12.8	12.5	3.60	28.9	4118
NW	11.6	11.2	4.10	26.9	7792
SE	14.3	13.9	10.5	38.7	2490
SS	16.3	15.8	9.40	41.5	3399
SW	8.8	8.50	5.00	22.3	4377

Source: 2008 Demographic and Health Survey in Nigeria.

WSUAN = who slept under any net

WSUAETN = who slept under an ever-treated net

WSUAITN = who slept under an insecticide treated net

TPWSUNPZ = total percentage who slept under net per zone

TNOCPZ = total number of children per zone

Table 2b. Distribution of the total number of children under age five who slept under any net, an ever-treated net or an insecticide treated net [Data 1].

Zones	total number of children under age five who slept under nets per zone (Yi)
NC	826
NE	1190
NW	2096
SE	964
SS	1411
SW	976
Total	7463

Source: 2008 Demographic and Health Survey in Nigeria.

Table 3a. Distribution of the number of pregnant women age 15 – 49 who slept under mosquito nets (in percentage).

Zones	WSUAN	WSUAETN	WSUAI	TNOPWPZ	TPWSUNPZ
NC	9.4	9.3	3.4	481	22.1
NE	17.6	17.2	5.6	527	40.4
NW	12.4	12.0	4.2	1051	28.6
SE	10.2	9.4	6.4	342	26.0
SS	11.3	11.1	7.2	444	29.6
SW	8.9	8.6	3.4	553	20.9

Source: 2008 Demographic and Health Survey in Nigeria.

WSUAN = who slept under any net

WSUAETN = who slept under an ever-treated net

WSUAI = who slept under an ITN

TNOPWPN = total number of pregnant women per zone

TPWSUNPZ = total percentage who slept under net per zone

Table 3b. Distribution of the total number of pregnant women age 15 – 49 who slept under any net, an ever-treated net or an insecticide treated net [Data 2].

Zones	total number of pregnant women age 15 – 49 who slept under nets(Y_i)
NC	106
NE	213
NW	301
SE	89
SS	131
SW	116
Total	956

Source: 2008 Demographic and Health Survey in Nigeria.

Table 4. Summary of inclusion probabilities.

Procedure	P_1	P_2	P_3	P_4	P_5	P_6	π_1	π_2	π_3	π_4	π_5	π_6
Alodat	0.1563	0.1447	0.2403	0.1228	0.1589	0.1770	0.3161	0.2942	0.4625	0.2520	0.3209	0.3543
Sen-Mem	0.1563	0.1447	0.2403	0.1228	0.1589	0.1770	0.3250	0.3158	0.3922	0.2983	0.3271	0.3416
YateandGrundy	0.1563	0.1447	0.2403	0.1228	0.1589	0.1770	0.3172	0.2960	0.4562	0.2548	0.3219	0.3540

π_{12}	π_{13}	π_{14}	π_{15}	π_{16}	π_{23}	π_{24}	π_{25}
0.0525	0.0953	0.0438	0.0584	0.0661	0.0874	0.0401	0.0535
0.0602	0.0793	0.0558	0.0630	0.0667	0.0770	0.0535	0.0607
0.0533	0.0940	0.0445	0.0590	0.0664	0.0864	0.0410	0.0542

π_{26}	π_{34}	π_{35}	π_{36}	π_{45}	π_{46}	π_{56}
0.0606	0.0730	0.0971	0.1098	0.0446	0.0505	0.0673
0.0643	0.0726	0.0798	0.0835	0.0563	0.0600	0.0672
0.0611	0.0725	0.0957	0.1077	0.0454	0.0512	0.0676

Table 5. Variance of Yates and Grundy estimator based on the three selection procedures.

Selection Procedure	Estimator	VOYAGE Using data 1	VOYAGE Using data2
Alodat (2009)	Yates and Grundy	13, 443, 938.96	626, 018.964
Sen-Midzuno(1952)	Yates and Grundy	20, 385, 287.82	788, 188.9463
Yates and Grundy(1953)	Yates and Grundy	13, 601, 913.64	629, 997.7596

Source: 2008 Demographic and Health Survey in Nigeria.

VOYAGR =Variance of Yates and Grundy Estimator