Comparison of Some Selected Values of the Constants of Basit and Shahbaz Selection Procedure under Unequal Probability Sampling Without Replacement

O.O. Dawodu
Department of Physical and Computer Sciences, McPherson University,
Seriki – Sotayo Ogun State, NIGERIA.
E-mail: omotola_dawodu@yahoo.com

Abstract

Probability sampling involves random selection of units from the population. It can be categorized as either equal probability sampling or unequal probability sampling (UPS). UPS however, can be with replacement or without replacement. My interest is on UPS without replacement. It occurs when there is a fixed sample size to be selected from finite population with the intention of not returning the sample already selected into the population. This is the idea used in this research work. In order to select the desired sample, the sample selection procedure, or simply put, sampling scheme proposed by Basit and Shahbaz is applied here which deals with sample size two. Their selection procedure consists of formulae for calculating the probabilities of inclusion of a unit in the sample, and the joint probabilities of inclusion of two different units in the sample. These formulae consist of two constants with varying values which can be positive or negative. I also used the Modified Murthy estimator, which Basit and Shahbaz proposed to estimate the variance. They carried out empirical studies based on their proposed method and came up with results. In this research, a sample size of two is selected from the population of size six. One of the constants is fixed while the other varies. I am particularly interested in determining the best value(s) of the constant that is allowed to vary. This will help in reducing the stress of varying the values of the constant. The results obtained in this research work are approximately in support of the empirical studies carried out by Basit and Shahbaz.

Keywords: modified murthy estimator, sample size two, sampling scheme.

1. Introduction

Horvitz and Thompson (1952) were the first to give theoretical framework of UPS without replacement, in accordance with the unified theory. Therefore, this research deals with UPS without replacement, for sample size 2, using the sample selection procedure proposed by Basit and Shahbaz (2007), and the modified Murthy estimator of variance, which they proposed.

Here, I deal with the distribution of the number of youngest children age 6 - 35 months living the mother, who consumed foods rich in vitamin A, in the past 24 hours to the interview.

This research is carried out using the 2008 Nigeria Demographic and Health Survey, which was implemented by National Population Commission, Nigeria.

Vitamin A is essential micronutrients for the immune system and plays an important role in maintaining the epithelial cause of eye damage. Vitamin A Deficiency can also increase severity of infections such as measles and diarrheal diseases in children and slow recovery from illness. Vitamin A is found in breast milk, other milks, liver, eggs, fish, butter, red palm oil, mangoes, papayas, carrots, pumpkins, and dark green leafy vegetables.

2. Sample selection procedure considered

Basit and Shahbaz (2007) gave a sample selection procedure for unequal probability sampling without replacement, sample size 2. The procedure is stated as follows:

- Select first unit with probability proportional to $P_i^{\alpha}(1-P_i^{\beta})$ / (1-2P_i^{β}), and without replacement,
- Select a random sample of size $(n-1)$ from the remaining $(N-1)$units.

The probability of inclusion of the i^{th} unit in the sample is given as

$$\pi_i = \frac{1}{d}\left[\frac{P_i^{\alpha}(1-P_i^{\beta})}{(1-2P_i^{\beta})}\left\{\frac{N-n}{N-1}\right\} + \frac{(n-1)}{N-1}d\right],$$

where $\boldsymbol{d = \sum_{i=1}^{N} \frac{P_i^{\alpha}(1-P_i^{\beta})}{(1-2P_i^{\beta})}}$

and the joint probability of inclusion for both i^{th} and j^{th} units in the sample for this selection procedure is

$$\boldsymbol{\pi_{ij} = \frac{(n-1)(N-n)}{d(N-1)(N-2)}\left[\frac{P_i^{\alpha}(1-P_i^{\beta})}{1-2P_i^{\beta}} + \frac{P_i^{\alpha}(1-P_i^{\beta})}{1-2P_i^{\beta}}\right] + \frac{(n-1)(n-2)}{(N-1)(N-2)}}$$

where α and β are constants and $\boldsymbol{n}$ is the sample size.

Basit and Shahbaz (2007) have already shown in their empirical studies that α could be fixed to be 1 while the values of β can be allowed to vary.

3. Estimator considered

Murthy (1957) proposed the estimator of the population total

$$t_m = \frac{1}{P(S)}\sum_{i=1}^{n} P(S/i)\ y_i$$

where $P(S/i)$ is probability of sample given that i^{th} unit is selected at the first draw, and $P(S)$ is the probabilities of the samples.

The modified Murthy estimator of the variance which Basit and Shahbaz (2007) proposed is:

$$V(Y_{BS}) = \frac{1}{2(N-1)^2} \sum_{i=1}^{N} \sum_{\substack{j=1 \\ i \neq j}}^{N} \left[\frac{(y_i - y_j)^2}{\pi_{ij}} - (N-1)\{(y_i + y_j)^2 - 2N y_i y_j\} \right].$$

This is the variance estimator used in this research work.

Table 1: Distribution of the number of local government in each geo-political zone in Nigeria.

Geo – political zones	Total number of local government in each geo-political zone (X_i)
North Central	121
North East	112
North West	186
South East	95
South South	123
South West	137
Total	774

Source: National Population Commission, Nigeria.

Table 2: Distribution of the number of youngest children age 6 - 35 months living with the mother, who consumed foods rich in Vitamin A in past 24 hours to the interview (in percentage)

Geo – political zones	Percentage of children under consideration	Number of children
North Central	77.3	1442
North East	60.1	1679
North West	53.5	3390
South East	81.2	955
South South	88.1	1364
South West	82.5	1812

Source: 2008 Demographic and Health Survey, Nigeria

Table 3: Distribution of number of youngest children age 6 - 35 months living with the mother, who consumed foods rich in vitamin A, in past 24 hours to the survey.

Geo – political zones	Number of children under consideration (Y_i)
North Central	1115
North East	1009
North West	1814
South East	776
South South	1202
South West	1495

Source: 2008 Demographic and Health Survey in Nigeria

The correlation coefficient ρ = 0.969, so $X_i's$ and $Y_i's$ are highly correlated, and the data is suitable for use.

Table 4: Probabilities of selection

Geo – political zones	Probabilities of selection (P_i)
North Central	0.1563
North East	0.1447
North West	0.2403
South East	0.1228
South South	0.1589
South West	0.1770
Total	1

Source: Researcher's analysis

Table 5: Summary of the probabilities of inclusion of the i^{th} unit in the sample varying the values of β

β	-4	-3	-2	-1	1	2	3	4
π_1	0.3251	0.3252	0.3258	0.2569	0.3197	0.3310	0.3247	0.3250
π_2	0.3158	0.3159	0.3164	0.3304	0.3087	0.3208	0.3154	0.3157
π_3	0.3920	0.3915	0.3896	0.4027	0.4194	0.3636	0.3937	0.3926
π_4	0.2983	0.2985	0.2991	0.3121	0.2891	0.3019	0.2978	0.2981
π_5	0.3272	0.3273	0.3275	0.3418	0.3223	0.3333	0.3269	0.3270
π_6	0.3416	0.3416	0.3416	0.3561	0.3408	0.3494	0.3415	0.3416
Total	2	2	2	2	2	2	2	2

Source: Researcher's analysis

Table 6: Summary of the joint probabilities of inclusion of the i^{th} and j^{th} units in the sample when α = 1 and varying the values of β

β	-4	-3	-2	-1	1	2	3	4
π_{12}	0.0602	0.0603	0.0606	0.0468	0.0571	0.0629	0.0600	0.0602
π_{13}	0.0793	0.0792	0.0788	0.0649	0.0848	0.0736	0.0796	0.0794
π_{14}	0.0558	0.0559	0.0562	0.0423	0.0522	0.0582	0.0556	0.0558
π_{15}	0.0631	0.0631	0.0633	0.0497	0.0605	0.0661	0.0629	0.0630
π_{16}	0.0667	0.0667	0.0669	0.0533	0.0651	0.0701	0.0666	0.0666
π_{23}	0.0770	0.0769	0.0764	0.0833	0.0820	0.0711	0.0773	0.0771
π_{24}	0.0535	0.0536	0.0539	0.0606	0.0495	0.0557	0.0533	0.0535
π_{25}	0.0607	0.0608	0.0609	0.0680	0.0577	0.0635	0.0606	0.0607
π_{26}	0.0644	0.0644	0.0645	0.0716	0.0624	0.0676	0.0642	0.0643
π_{34}	0.0726	0.0725	0.0722	0.0787	0.0771	0.0664	0.0729	0.0727
π_{35}	0.0798	0.0797	0.0793	0.0861	0.0854	0.0742	0.0835	0.0799
π_{36}	0.0834	0.0833	0.0828	0.0897	0.0900	0.0783	0.0838	0.0836
π_{45}	0.0564	0.0564	0.0567	0.0634	0.0529	0.0588	0.0562	0.0563
π_{46}	0.0599	0.0600	0.0602	0.0671	0.0575	0.0628	0.0598	0.0599
π_{56}	0.0672	0.0672	0.0673	0.0745	0.0658	0.0707	0.0671	0.0671
Total	1	1	1	1	1	1	1	1

Source: Researcher's analysis

Table 7: Summary of the variances and the corresponding standard errors.

β	$V(Y_{BS})$	$S(Y_{BS})$
-4	18793,243.52	4335.12
-3	18793748.80	4335.18
-2	18795115.30	4335.41
-1	18743298.15	4329.35
1	18770564.57	4332.52
2	18832212.26	4339.61
3	19079106.45	4367.96
4	18792621.14	4335.05

Source: Researcher's analysis

4. Summary and Discussion of Results

Consider the positive value of β, that is, β = 1, 2, 3, 4. β = 1 produced the least standard error of estimate, followed by β = 2, then β = 4, and lastly β = 3 is the highest.

Consider the negative values of β, that is, β = -1, -2, -3, -4. β = -1 produced the least standard error of estimate, followed by β = -4, then β = -3, and lastly β = -2. The value of the standard error of estimate for β = -2 and β = -3 and β = -4, are approximately the same.

Based on the analysis and estimates obtained, the least standard error was obtained when α = 1 and β = -1 for the negative value of β, and also, α = 1 and β = 1 for the positive value of β.

To a reasonable extent, the results obtained is in support of the empirical study carried out by Basit and Shahbaz (2007) for positive values of β, and fixing the value of α to be 1.

For the negative values of β, the results are approximately in support of the empirical study carried out by Basit and Shahbaz (2007).

5. Conclusion

Basit and Shahbaz sample selection procedure could be used for estimating the number of children age 6-35 months in the six geo-political zones in Nigeria, who are living with the mother and consume food rich in vitamin A, 24 hours to the interview.

While fixing the values of α to be 1, the value of β could also be fixed to 1 or -1, and thereby reducing the stress of varying the values of β.

References

Basit, A. and Shahbaz, Q. (2007) "A general class of selection procedures and modified Murthy estimator," *Journal of Statistics*, 14, 13-19.

Dawodu, O.O., Adewara, A.A., and Olayiwola, O.M. (2011) "Efficiency of Alodat sample selection procedure over Sen-Midzuno and Yates-Grundy Draw by Draw under unequal probability sampling without replacement, sample size 2 ," *Journal of Mathematical Research*, 3, 113 – 118.

Tikkiwal, B.D., (1965) "Some aspects of sampling with varying probabilities from a finite population," *University of Rajasthan Studies (Statistics)*, 1 – 21.