Int. J. Open Problems Comp. Math., Vol. 2, No. 1, March 2009

On Unequal Probability Sampling Without Replacement Sample Size 2

Naser .A. Alodat

Department of Mathematics, Irbid National University, Jordan.
e-mail: n_odat@yahoo.com

Communicated by Esam Eldin A. Rakha

Abstract

A new selection procedure has been developed for use with the Hurvits-Thompson estimator. Some results have been verified for first and second order inclusion probabilities.

Keywords: *Unequal probability sampling, Hurvits-Thompson estimator, variance estimator.*

1 Introduction

The concept of sampling with unequal probability without replacement was first introduced by Meadow (2). Horvitz and Thompson (1) were the first to give theoretical frame work of unequal probability sampling without replacement. The estimation proposed by H.T (1) was:

$$\hat{y} = \sum_{i \in s} \frac{y_i}{\pi_i} \quad (1.1)$$

Where π_i is probability of inclusion of ith unit in the sample.

The variance of Horvitz-Thompson estimator was:

$$V\left(\hat{y}_{HT}\right) = \sum_{i=1}^{N} \left(\frac{1-\pi_i}{\pi_i}\right) y_i^{\,2} + \sum_{i,j=1, i \neq j} \sum \left(\frac{\pi_{ij} - \pi_i \pi_j}{\pi_i \pi_j}\right) y_j y_j \quad (1.2)$$

with an unbiased variance estimator

$$v\left(\hat{y}_{HT}\right)=\sum_{i=1}^{n}\frac{1-\pi_i}{\pi_i^2}y_i^2+\sum_{i=1}^{n}\sum_{i\neq j=1}\left(\frac{\pi_{ij}-\pi_i\pi_j}{\pi_{ij}}\right)\left(\frac{y_i y_j}{\pi_i\pi_j}\right). \tag{1.3}$$

The expression for variance of Hurvits-Tompson given by Yates Grundy (4) is:

$$V_{YG}\left(\hat{y}_{HT}\right)=\sum_{i=1}^{N}\sum_{i<j}^{N}\left(\pi_i\pi_j-\pi_{ij}\right)\left(\frac{y_i}{\pi_i}-\frac{y_j}{\pi_j}\right)^2 \tag{1.4}$$

With an unbiased variance estimator:

$$v_{YG}\left(\hat{y}_{HT}\right)=\sum_{i=1}^{n}\sum_{i<j}^{n}\left(\frac{\pi_i\pi_j-\pi_{ij}}{\pi_{ij}}\right)\left(\frac{y_i}{\pi_i}-\frac{y_j}{\pi_j}\right)^2 \tag{1.5}$$

Shahbaz and Hanif (3) suggested a new procedure where selecting a sample of size two where first unit selected by probability proportional to $\frac{p_i}{(1-2p_i)}$ and second with probability proportional to p_i .

2 New Selection Procedure

In this section, we have given a new selection procedure for use with the Hurvits-Thompson (1) estimator with sample size 2.

a) Select first unit with probability proportional to $\frac{p_i}{2-p_i}$ and without replacement.

b) Select second unit with probability $\frac{p_i}{1-p_j}$

The first inclusion probability is

$$\pi_i=\frac{p_i}{B\left(2-p_i\right)}+\sum_{j\neq i=1}\frac{p_j}{B\left(2-p_j\right)}\times\frac{p_i}{1-p_j}$$

$$=\frac{p_i}{B}\left[\frac{1}{2-p_i}+\sum_{j=1}^{N}\frac{p_j}{\left(2-p_j\right)\left(1-p_j\right)}-\frac{p_i}{\left(2-p_i\right)\left(1-p_i\right)}\right]$$

where,

$$B=\sum_{i=1}^{N}\frac{p_i}{\left(2-p_i\right)}$$

$$\pi_i=\frac{p_i}{B}\left[\frac{1-2p_i}{\left(2-p_i\right)\left(1-p_i\right)}+\sum_{j=1}^{N}\frac{p_j}{\left(2-p_j\right)\left(1-p_j\right)}\right] \tag{2.1}$$

$$\pi_{ij} = p_i p\left(j/i\right) + p_j p\left(i/j\right)$$

$$\pi_{ij} = \frac{p_i p_j}{B}\left[\frac{1}{\left(2-p_i\right)\left(1-p_i\right)} + \frac{1}{\left(2-p_j\right)\left(1-p_j\right)}\right] \qquad (2.2)$$

Some results for new selection procedure.

Results (1): $\sum_{i=1}^{N} \pi_i = n$

Proof: summing both side of (2.1)

$$\sum_{i=1}^{N} \pi_i = \sum_{i=1}^{N} \frac{p_i}{B}\left[\frac{1-2p_i}{2-p_i} + \sum_{j=1}^{N} \frac{p_j}{\left(2-p_j\right)\left(1-p_j\right)}\right]$$

$$= \frac{1}{B}\left[\sum_{i=1}^{N} \frac{\left(1-2p_i\right)p_i}{\left(2-p_i\right)\left(1-p_i\right)} + \sum_{i=1}^{N} p_i \sum_{j=1}^{N} \frac{p_j}{\left(2-p_j\right)\left(1-p_j\right)}\right]$$

$$= \frac{1}{B}\left[\sum_{i=1}^{N} \frac{\left(1-2p_i\right)p_i + p_i}{\left(2-p_i\right)\left(1-p_i\right)}\right]$$

$$= \frac{1}{B}\sum_{i=1}^{N} \frac{p_i\left(1-2p_i+1\right)}{\left(2-p_i\right)\left(1-p_i\right)}$$

$$= \frac{2}{B}\sum_{i=1}^{N} \frac{p_i\left(1-p_i\right)}{\left(2-p_i\right)\left(1-p_i\right)} = \frac{2B}{B} = 2. \qquad (2.3)$$

Since n=2, therefore equation (2.3) can be written as $\sum_{i=1}^{N} \pi_i = n$.

Result (2): $\sum_{i \neq j}^{N} \pi_{ij} = \left(n-1\right)\pi_i$.

Proof:

Summing both sides of (2.2), we get:

$$\sum_{j \neq i}^{N} \pi_{ij} = \sum \frac{p_i p_j}{B}\left[\frac{1}{\left(2-p_i\right)\left(1-p_j\right)} - \frac{1}{\left(2-p_j\right)\left(1-p_i\right)}\right]$$

$$= \frac{1}{B}\left[\frac{p_i}{\left(2-p_i\right)\left(1-p_i\right)}\sum_{j \neq i} p_j + p_i \sum \frac{p_j}{\left(2-p_j\right)\left(1-p_j\right)}\right]$$

$$= \frac{1}{B}\left[\frac{p_i\left(1-p_i\right)}{\left(2-p_i\right)\left(1-p_i\right)} + p_i\sum\frac{p_j}{\left(2-p_j\right)\left(1-p_j\right)} - \frac{p_i}{\left(2-p_i\right)\left(1-p_i\right)}\right]$$

$$= \frac{p_i}{B}\left[\frac{1-2p_i}{\left(2-p_i\right)\left(1-p_i\right)} + \sum_{j=1}^{N}\frac{p_j}{\left(2-p_j\right)\left(1-p_j\right)}\right] \qquad (2.4)$$

Comparing (2.4) with (2.1), it can be seen that

$$\sum_{j\neq 1}^{N}\pi_{ij} = \pi_i$$

Result (3): $\sum_{i=1, j\neq i}\sum_{j=1}^{N}\pi_{ij} = n\left(n-1\right)$

Applying double summation on both side of (2.2)

$$\sum_{i=1, j\neq 1}\sum_{j=1}^{N}\pi_{ij} = \sum\sum\frac{p_i p_j}{B}\left[\frac{1}{\left(2-p_i\right)\left(1-p_i\right)} + \frac{1}{\left(2-p_j\right)\left(1-p_j\right)}\right]$$

$$= \sum_{i=1}^{N}\left[\sum_{j=1, j\neq i}^{N}\frac{p_i p_j}{B}\left(\frac{1}{\left(2-p_i\right)\left(1-p_i\right)} + \frac{1}{\left(2-p_j\right)\left(1-p_i\right)}\right)\right] \qquad (2.5)$$

also

$$\sum_{j=1, j\neq i}\pi_{ij} = \pi_i \qquad (2.6)$$

Substituting (2.5) in (2.6) we get:

$$\sum_{i=1, j\neq i}\sum_{j=1}\pi_{ij} = \sum_{i=1}^{N}\pi_i = 2 \qquad (2.7)$$

Since n=two therefore equation (2.7) can be written as: $\sum_{i=1}\sum_{j=1, i\neq j}\pi_{ij} = n\left(n-1\right)$.

Result (4): $p_i = p_j = \frac{1}{N}$.

The value of π_i and π_{ij} reduces to the standard results of simple random sampling.

Putting $p_i = p_j = \frac{1}{N}$ in (2.1), we get:

$$\pi_i = \frac{\frac{1}{N}}{B}\left[\frac{1-\frac{2}{N}}{\left(2-\frac{1}{N}\right)\left(1-\frac{1}{N}\right)} + \sum_{j}^{N}\frac{\frac{1}{N}}{\left(2-\frac{1}{N}\right)\left(1-\frac{1}{N}\right)}\right] \qquad (2.8)$$

$$B = \sum_{i=1}^{N}\frac{p_i}{2-p_i} = \sum_{i=1}^{N}\frac{\frac{1}{N}}{2-\frac{1}{N}} = \frac{N}{2N-1} \qquad (2.9)$$

From (2.8) and (2.9) we get $\pi_i = \frac{2}{N}$.

This is for simple random sampling without replacement for a sample of size equal two.

For π_{ij} substituting $p_i = p_j = \frac{1}{N}$ and the value of B, in (2.2) we get:

$$\pi_{ij} = \frac{\frac{1}{N^2}}{B}\left[\frac{1}{\left(2-\frac{1}{N}\right)\left(1-\frac{1}{N}\right)}+\frac{1}{\left(2-\frac{1}{N}\right)\left(1-\frac{1}{N}\right)}\right]=\frac{2}{N\,(N-1)}$$

Which is the joint inclusion probability for a sample of size two in case of simple random sampling without replacement.

3 Open Problems

In this paper, we suggest the use of new procedure for selection a sample of size two by unequal probability without replacement. The problem considered in this paper can be extend for a sample of size greater than two. Moreover, there is needed to compare between proposed estimator and the extended on.

References

[1] Hurvitz, D.G.and Thompson, D.J. (1952).A generalization of sampling without replacement from a finite universe, J.Amer.Stat.Assoc.47, 663-685.

[2] Madow, W.G. (1949).On the theory of systematic sampling II.Ann.Math.Stat.20, 333-354.

[3] Shahbas, M.Q.and Hanif, M. A sample procedure for unequal probability sampling without replacement sample size 2.Pak.J.Statist.2003 Vol.19 (1), 151-160.

[4] 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.