APPENDIX - A

Estimation in Mail Surveys Under PSNR Sampling Scheme

D. Shukla and Jayant Dubey
Dr. H.S. Gour University, Sagar (M.P.)
(Received: July, 1999)

SUMMARY

In the set up of stratified scheme, when frames of each stratum are not available survey practitioners are advised to use post-stratification scheme for the estimation of population parameters. Suppose post-stratified units of the sample have a few respondents and a large number of non-respondents then estimation of population parameter is difficult. This is a most likely situation in mail surveys where questionnaires are mailed to respondents and expected to return after completion, earlier to a deadline. This paper presents a new “Post-stratified non-response (PSNR)” sampling scheme in order to handle a high rate of non-response likely to occur in mail surveys. The scheme has several steps to complete before the estimation of population mean. An unbiased estimator has been found and its precision is examined. The cost analysis is incorporated to estimate the optimal sample size required under prefix cost. Due to certain limitations, a new alternative approach is proposed and observed solution oriented for cost optimal sample size. All the derived results are numerically supported and the PSNR scheme is found useful and effective for the mean estimation.

Key words: Post-stratification, SRSWOR, Optimal, PSNR, Respondents (RS), Non-respondents (NRS).

1. Introduction

A mail survey is cheaper and faster in the present days of advanced electronic communication but, it fails to achieve the object of cent percent response. Suppose a population is stratified naturally, with known size of strata but, without the list of units of each stratum like a voter list of a city which cannot provide frames if strata are made as lower, middle and higher income groups. One could get information from other records that there are 30% lower, 60% middle and 10% higher-income group voters in the city and this needs to be utilized. It motivates to choose post-stratified sampling scheme where a random sample is drawn from the population by SRSWOR and stratified according to stratum formed by information of other records. If these post-stratified units of individuals (voters) are mailed questionnaires and are expected to return back
till the deadline there may be a few respondents and a huge amount of non-response. To cope with this, a motivation is derived from Hansen and Hurwitz [2], Agrawal and Panda [1] and a post-stratified non-response (PSNR) scheme is developed.

2. Notations

Let $N$ be the size of a population divided into $k$ strata each of size $N_i$, $(i = 1, 2, 3, ..., k)$ such that $\sum_{i=1}^{k} W_i = \frac{N_i}{N} = 1$. Further, $i^{th}$ strata contains $N'_i$ respondents (say RS) and $N^*_i$ non-respondents (say NRS) and these are suppose to be known like a stratum of literate people has 70% response while that of illiterate has only 30%. The variable $Y$ is under study with population mean $\bar{Y}_i$ of $i^{th}$ strata and $\bar{Y} = \sum_{i=1}^{k} W_i \bar{Y}_i$ alongwith population mean squares in usual notations, $S^2$ and $S^2_i$. Within the $i^{th}$ strata, response group (R) has $S^2_{ii}$ and non-response group (NR) has $S^2_{ii}$ as their population mean squares. To note that $N_i = N'_i + N^*_i$.

3. A Post Stratified Non-response (PSNR) Sampling Scheme

Step I: Select a sample of size $n$ by SRSWOR from the population $N$ and post-stratified into $k$ strata, such that $n_i$ units represent to

$$N_i \left( \sum_{i=1}^{k} n_i = n \right).$$

An auxiliary source of information or prior guess may be used for this purpose.

Step II: Mail questionnaires to all the random $n_i$ units for response over the variable $Y$ under study and wait until a deadline. If there is a possible complete response, the $\bar{Y}_i$ is sample mean from $i^{th}$ strata and

$$\bar{Y} = (n)^{-1} \sum_{i=1}^{k} n_i \bar{Y}_i.$$

Step III: Assume that non-response observed when the deadline of returning questionnaire is over, and there are $n'_i$ respondents, $n^*_i$ non-respondents in the $i^{th}$ strata $\left(n'_i + n^*_i = n_i \right)$. The $\bar{Y}'_i$ is the mean of responding $n'_i$ units.
Step IV: From non-responding \( n_i^* \), select sub-samples of size \( n_i^{**} \) by SRSWOR, maintaining a prefixed fraction \( f_i = \left( \frac{n_i^*}{n_i} \right) \) over all the \( k \) strata.

Step V: Conduct a personal interview for \( n_i^{**} \) units and assume all these responded well during that period. The \( \bar{y}_i^{**} \) is the mean based on \( n_i^{**} \).

4. Estimation Under PSNR

Deriving a motivation from Hansen and Hurwitz [2], to estimate \( \bar{Y} \), the proposed estimation strategy is

\[
\bar{y}_{PSNR} = \sum_{i=1}^{k} W_i \left[ \left( \frac{n_i^*}{n_i} \right) \bar{y}_i^* + \left( \frac{n_i}{n_i} \right) \bar{y}_i \right] \tag{4.1}
\]

**Theorem 4.1:** \( \bar{y}_{PSNR} \) is unbiased for \( \bar{Y} \).

**Proof:** The conditional expectation of \( \bar{y}_{PSNR} \) given \( n_i \) and \( (n_i^*, n_i^*) \) is denoted by \( E[()] / n_i, (n_i^*, n_i^*) \). Also, we know \( E[(\bar{y}_i^{**})/n_i, (n_i^*, n_i^*)] = \bar{y}_i^* \) which is the mean of \( n_i^* \) units.

\[
E(\bar{y}_{PSNR}) = E[E(\bar{y}_{PSNR})/n_i, (n_i^*, n_i^*)]
\]

\[
E \left[ \sum_{i=1}^{k} W_i \left( \frac{n_i^*}{n_i} \right) \bar{y}_i^* + \left( \frac{n_i^*}{n_i} \right) E(\bar{y}_i^*) \right] / n_i, (n_i^*, n_i^*)
\]

\[
E \left[ \sum_{i=1}^{k} W_i \left( \frac{n_i}{n_i} \right) \bar{y}_i + \left( \frac{n_i}{n_i} \right) \bar{y}_i \right] / n_i
\]

\[
E \left[ \sum_{i=1}^{k} W_i \{ E(\bar{y}_i) / n_i \} \right] = \sum_{i=1}^{k} W_i \bar{y}_i = \bar{Y}
\]

**Theorem 4.2:** The variance of \( \bar{y}_{PSNR} \) is

\[
V(\bar{y}_{PSNR}) = \sum_{i=1}^{k} W_i^2 \left[ \frac{1}{nW_i} + \frac{(N-n)(1-W_i)}{n^2(N-1)W_i} \right] \left\{ A_i + S_i^2 \right\} - \frac{1}{N} \sum_{i=1}^{k} W_i S_i^2 \tag{4.2.1}
\]

where \( A_i = \left[ \frac{N_i^*}{N_i} (r_i - 1) S_i^2 \right] \) \tag{4.2.2}
**Proof:** We use a standard result

\[
E \left( \frac{1}{n_i} \right) = \left[ \frac{1}{nW_i} + \frac{(N-n)(1-W_i)}{n^2 (N-1) W_i^2} \right] \tag{4.2.3}
\]

and \( V(\bar{y}_i^*/n_i, (n_i', n_i^*)) = \left[ \left( \frac{1}{n_i'} - \frac{1}{n_i} \right) S_{2i}^2 \right] \) using a double sampling set-up \( (4.2.4) \)

\[
V[\bar{y}_{PSNR}] = E[E[V(\bar{y}_{PSNR})/n_i, (n_i', n_i^*)]] + E[V[E[\bar{y}_{PSNR})/n_i, (n_i', n_i^*)]]
\]

\[+ V[E[V(\bar{y}_{PSNR})/n_i, (n_i', n_i^*)]] \tag{4.2.5}
\]

The conditional variance is a sum of three components. Consider the first one

\[
E[E[V(\bar{y}_{PSNR})/n_i, (n_i', n_i^*)]]
\]

\[
= E \left[ \sum_{i=1}^{k} \left( \frac{W_i}{n_i} \right) (n_i' \bar{y}_i') + \sum_{i=1}^{k} \left( \frac{W_i}{n_i} \right) (n_i' \bar{y}_i') \right] / n_i, (n_i', n_i^*)
\]

\[
= E \left[ 0 + \sum_{i=1}^{k} \left( \frac{W_i n_i^*}{n_i} \right)^2 V(\bar{y}_i^*) \right] / n_i, (n_i', n_i^*)
\]

\[
= E \left[ \sum \left( \frac{W_i n_i^*}{n_i} \right)^2 \left( \frac{1}{n_i^*} \right) - \left( \frac{1}{n_i^*} \right)^2 \right] E(s_{2i}^2) / n_i, (n_i', n_i^*)
\]

\[
= E \left[ \sum \left( \frac{W_i^2}{n_i} \right) \left( n_i^* \right) (f_i - 1) S_{2i}^2 \right] / n_i, (n_i', n_i^*)
\]

\[
= E \left[ \sum \left( \frac{W_i^2}{n_i} \right) (f_i - 1) S_{2i}^2 \right] / n_i
\]

\[
= \sum W_i^2 \left( E \left( \frac{1}{n_i} \right) \left( \frac{N_i^*}{N_i} \right) (f_i - 1) S_{2i}^2 \right)
\]

\[
= \sum W_i^2 \left( E \left( \frac{1}{n_i} \right) A_i \right) \tag{4.2.6}
\]

The term \( s_{2i}^2 \) is sample mean of the non-response group among \( n \) units and a result \( E \left[ \left( \frac{n_i^*}{n_i} \right) / n_i \right] = \left( \frac{N_i^*}{N_i} \right) \) is used above. The second component provides
\begin{align*}
E[V(E(\bar{y}_{PSNR})/n_i, (n_i', n_i''))] &= E \left[ V \left( \sum_{i=1}^{k} W_i \left( \left( \frac{n_i'}{n_i} \right) \bar{y}_i' + \left( \frac{n_i'^*}{n_i} \right) E(\bar{y}_{i}^*) \right) / n_i, (n_i', n_i'') \right) \right] \\
&= E \left[ V \left( \sum_{i=1}^{k} W_i \bar{y}_i \right) / n_i \right] = E \left[ \frac{\sum W_i^2}{n_i} \left( \frac{1}{n_i} - \frac{1}{N_i} \right) S_i^2 \right] / n_i \\
&= \sum W_i^2 \left( E \left( \frac{1}{n_i} \right) \right) S_i^2 - \sum \left( \frac{W_i^2 S_i^2}{N_i} \right) \quad (4.2.7)
\end{align*}

The third component vanishes and the addition of (4.2.6) and (4.2.7) with the substitution of \( E \left( \frac{1}{n_i} \right) \) provides the proof.

**Remark 4.2.1:** The variance expression be expressed in the form
\begin{equation}
V(\bar{y}_{PSNR}) = \sum W_i^2 \left( E \left( \frac{1}{n_i} \right) \right) \left( \sigma_i^2 + \frac{1}{N_i} \right) \sum W_i S_i^2 \quad (4.2.8)
\end{equation}

### 5. Cost Analysis

Assume the following components of cost, involved in a mail survey

(i) \( C_0 \): Cost of including sample units \( n_i \) of the \( i^{th} \) stratum in sample \( n \)

(ii) \( C_1 \): Cost of collecting, editing and processing per \( n_i \) unit in the \( i^{th} \) strata of the response class

(iii) \( C_2 \): Cost of personal interview and processing information per \( n_i'' \) unit in the non-response class

The cost function for \( i^{th} \) strata is
\[ C_i = C_0 n_i + C_1 n_i' + C_2 n_i'' \quad (5.1) \]

With total cost over \( k \) strata
\[ T_c = \sum_{i=1}^{k} C_i = \sum_{i=1}^{k} \left[ C_0 n_i + C_1 n_i' + C_2 n_i'' \right] \quad (5.2) \]

and the expected cost \( E(T_c) = \left( \frac{n}{N} \right) \sum_{i=1}^{k} \left[ C_0 N_i + C_1 N_i' - \frac{C_2 N_i'^*}{f_i} \right] \quad (5.3) \]
To get optimum $f_i$ and $n$, define a function $\delta$.

$$\delta = E(T_c) + \lambda \left[ V(\overline{y}_{\text{PSNR}}) - V_0 \right]$$

with $V_0$ a prefixed level of variance of $\overline{y}_{\text{PSNR}}$ and $\lambda$ a lagrange multiplier. Differentiating with respect to $f_i$, $n$ and $\lambda$ and equating to zero, the three equations are

$$\lambda \left[ Q_{li} S_{2i}^2 \right] = \left( \frac{n}{N} \right) \left( \frac{C_2 N_i^*}{f_i^2} \right)$$  \hspace{1cm} (5.4)

$$\lambda \left( n^2 (N - 1) \right) \sum_{i=1}^{k} \left[ Q_{2i} \left( \frac{N_i^*}{N_i} \right) \left( f_i - 1 \right) S_{2i}^2 + S_i^2 \right] = \left( \frac{1}{N} \right) \sum_{i=1}^{k} \left[ C_0 N_i + C_1 N_i^* + \frac{C_2 N_i^*}{f_i} \right]$$  \hspace{1cm} (5.5)

and

$$\sum_{i=1}^{k} \left[ Q_{li} \left( \frac{N_i^*}{N_i} \right) \left( f_i - 1 \right) S_{2i}^2 + Q_{li} S_i^2 - \left( \frac{W_i^2 S_i^2}{N_i} \right) \right] = V_0$$  \hspace{1cm} (5.6)

where

$$Q_{li} = \left[ \left( \frac{W_i}{N} \right) + \left( \frac{(N - n)(1 - W_i)}{n^2 (N - 1)} \right) \right], \quad Q_{2i} = \left[ (2N - n) + NW_i (n - 2) \right]$$

_CASE I:_ The (5.4) generates $k$ equations for $k$ strata, therefore in total $(k + 2)$ equations for $(k + 2)$ unknowns $f_1, f_2, f_3, \ldots, f_k, n$ and $\lambda$. But, the powers of $n$ and $f_i$ are not necessary positive unity, so the assurance of finding solution is hard.

_CASE II:_ The limitation of case I improvise to put extra condition to determine $\lambda$. Let us consider

$$n = \sum_{i=1}^{k} n_i = \sum_{i=1}^{k} n_i^* + \sum_{i=1}^{k} n_i^*$$  or  $$\left( n - \sum_{i=1}^{k} n_i^* \right) = \sum_{i=1}^{k} n_i^* f_i$$  \hspace{1cm} (5.7)

Using (5.4),

$$f_i = \sqrt{\left( \frac{C_2 N_i^*}{Q_{li} S_{2i}} \right) \left( \frac{1}{\lambda} \right)}$$  \hspace{1cm} (5.8)

The (5.7) helps to eliminate $\lambda$ and we get

$$f_i = \left( n - \sum_{i=1}^{k} n_i^* \right) R_i \left[ \sum_{i=1}^{k} \left[ n_i^* R_i \right] \right]^{-1}$$  \hspace{1cm} (5.9)
with \[ R_i = N_i \left[ \left\{ n W_i (N - 1) N_i + (N - n)(1 - W_i) \right\} S_i^2 \right]^{1/2} \] (5.10)

The equation (5.9) provides the value of \( f_i \). We can rewrite (5.6)

\[
V_0 + \sum_{i=1}^{k} \frac{W_i^2 S_i^2}{N_i} = \sum_{i=1}^{k} Q_{ii} \left\{ \left( \frac{N_i^*}{N_i} \right) (f_i - 1) S_i^2 + S_i^2 \right\}
\] (5.11)

On substituting \( f_i \) the equation (5.11) provides solution for optimal \( n \) which is included in the term \( Q_{ii} \).

**Case III:** Another approach for solution is as under

Let \( \sum_{i=1}^{k} f_i = M \), where \( M \) is a prefixed constant. This provokes to keep the total of sub-sample fractions equal to \( M \) with no restriction on \( f_i \) and, it is used to eliminate \( \lambda \). The solution is

\[
f_i = \frac{MR_i}{\sum_{i=1}^{k} R_i}
\] (5.12)

Substituting \( f_i \) in (5.11), \( n \) could be obtained by solving the following

\[
n^2 (N - 1) \sum_{i=1}^{k} \left[ \sqrt{R_i} \left\{ V_0 + \sum_{i=1}^{k} \left( \frac{W_i^2}{N_i} \right) S_i^2 \right\} - \sum_{i=1}^{k} X_i \left\{ \sum_{i=1}^{k} \sqrt{R_i} (S_i^2) \right\} \right]
\]

\[
+ \sum_{i=1}^{k} \left[ \sqrt{R_i} (X_i S_i^2) \right] - \sum_{i=1}^{k} X_i \left\{ \sum_{i=1}^{k} f_i S_{2i}^2 \sqrt{R_i} \right\} = 0
\] (5.13)

where \( X_i = [(N - n) + NW_i (n - 1)] \)

6. An Alternative Cost Strategy

In case I, the solution of \( f_i \) and \( n \) are not certain whereas case II and III do not ensure a concrete optimal value of \( n \), therefore an alternative solution may think of for optimal choices. Consider \( i^{th} \) stratum-based approach by assuming cost \( C_0 \), \( C_{ii} \) and \( C_2 \) varying over all \( k \) strata then the reset cost strategy is

(i) \( C_0 \): cost of \( n_i \) units of the \( i^{th} \) stratum to be included into sample \( n \)

(ii) \( C_{ii} \): cost of collecting, editing and processing per \( n_i \) units in \( i^{th} \) strata of response class
(iii) $C_{2i}$ : cost of personal interview and processing per $n_i''$ units in the $i^{th}$ strata for non-response class

The $i^{th}$ stratum has total cost

$$C_i = [C_{0i}n_i + C_{1i}n_i' + C_{2i}n_i'']$$

and variance of $\bar{y}_{PSNR}$

$$V(\bar{y}_{PSNR}) = E\left(\frac{1}{n_i}\left(\frac{N_i'}{N_i}(f_i - 1)S_{2i}^2 + S_i^2\right) - \left(\frac{1}{N_i}\right)S_i^2\right)$$

The function $\delta_i'$, with lagrange multiplier $\lambda_i$ and pre-fixed level of variance $V_{0i}$

$$\delta_i' = [\text{expected cost of } i^{th} \text{ strata}] + \lambda_i \left[V(\bar{y}_{PSNR}) - V_{0i}\right]$$

$$\delta_i' = \left(\frac{n}{N}\right)\left[C_{0i}n_i + C_{1i}n_i' + C_{2i}N_i''\right]$$

$$+ \lambda_i \left[E\left(\frac{1}{n_i}\left(\frac{N_i'}{N_i}(f_i - 1)S_{2i}^2 + S_i^2\right) - \left(\frac{1}{N_i}\right)S_i^2\right) - V_{0i}\right]$$

On differentiating with respect to $f_i$, $\lambda$ and $n$, the three equations are

$$\left[\frac{n}{N}\right]\frac{C_{2i}N_i''}{f_i - \lambda_i}\left[E\left(\frac{1}{n_i}\left(\frac{N_i'}{N_i}S_{2i}^2\right)\right)\right] = 0$$

$$\left[E\left(\frac{1}{n_i}\left(\frac{N_i'}{N_i}(f_i - 1)S_{2i}^2 + S_i^2\right)\right) - V_{0i}\left(\frac{1}{N_i}\right)S_i^2\right] = 0$$

$$\left[\frac{1}{N}\right]\left[C_{0i}n_i + C_{1i}n_i' + C_{2i}N_i''\right]$$

$$- \left(\frac{\lambda_i}{n}\right)\left[E\left(\frac{1}{n_i}\right) + \frac{N_i(1 - W_i)}{n^2(N - 1)W_i^2}\right]\left(\frac{N_i'}{N_i}(f_i - 1)S_{2i}^2 + S_i^2\right) = 0$$

The equation (6.5) gives $\lambda_i = \{nC_{2i}N_i\left[f_i^2NE\left(\frac{1}{n_i}\right)S_{2i}^2\right]^{-1}\}

which on substituting (6.7), provides

$$f_i = \left[\frac{1}{2\delta_{2i}}\left|\delta_{2i} \pm \sqrt{\delta_{2i}^2 - 4\delta_i \delta_{3i}}\right|\right]$$

209
\[ \delta_{li} = \left( \frac{1}{N} \right) \left[ C_{0i} N_i + C_{li} N_i^* \right], \delta_{2i} = C_{2i} N_i^* \left[ \frac{1 - W_i}{n^2 (N - 1) W_i^2} - \left( \frac{(l - n)}{N} \right) \right] \]

\[ \delta_{3i} = \left[ \left( \frac{N_i S_i^2}{N S_{2i}^2} \right) - \left( \frac{N_i^*}{N_i} \right) + \left( \frac{(1 - W_i)}{n^2 (N - 1) W_i^2} \right) \right] \left\{ N_i^* - \left( \frac{\frac{S_i^2}{S_{2i}^2}}{E(l/n) \bar{S}_{2i}^2} \right) \right\} \]

The positive value of \( f_i \) will occur if \( \delta_{2i}^2 > (\delta_{li}, \delta_{3i}) \)

**Remark 6.1:** Suppose \( f_i \) is prefixed (say \( f_i = 1.5 \)) for all \( i \) then (6.9) gives optimum selection of \( C_{0i} \), \( C_{li} \) and \( C_{2i} \) provided the ratio \( \frac{S_i^2}{S_{2i}^2} \) is guessed well for all \( k \) strata.

**Remark 6.2:** Using (6.6), the optimum \( n \) is

\[
\left( n_{opt} \right)_i = \left[ (NW_i - 1) \pm \sqrt{(NW_i - 1)^2 + 4 \gamma_i (N - 1) W_i^2 N (1 - W_i)} \right] \left[ 2 \gamma_i (N - 1) W_i^2 \right]^{-1}
\]

where \( \gamma_i = \left[ \left( V_{0i} + \frac{S_i^2}{N_i} \right) / \left( \frac{N_i^*}{N_i} \right) \left( f_i - 1 \right) S_{2i}^2 + S_i^2 \right] \)

If \( N \) is sufficiently large then

(i) \( (N - 1) = N \), and

(ii) \( (NW_i - 1) = NW_i \)

and then (6.10) reduces into \( \left( n_{opt} \right)_i = N \left[ 1 \pm \sqrt{1 + 4 \gamma_i (1 - W_i)} \right] \left[ 2 N_i \gamma_i \right]^{-1} \)

This equation would provide the optimum \( (n_{opt})_i \), when \( f_i \) from (6.9) is substituted along with a suitable choice of \( V_{0i} \).

**Remark 6.3:** The (6.11) varies over \( i \), therefore, provide \( k \) values of optimum sample size \( (n_{opt})_i \), over different strata. One can choose any one of these for practical purpose. Another way may be to choose an average integer value of these \( (n_{opt})_i \), or minimum/maximum of \( (n_{opt})_i \), over all \( k \) strata.

**Remark 6.3.1:** The justification of remark 6.3 is based on unique selection of \( a \) value over availability of \( k \) different optimal \( n \). A criteria is to be designed by the survey practitioner for single-valued choice like
(i) selection of any one value on the basis of experience
(ii) selection of an average integer value
(iii) selection of a minimum value, due to cost constraint
(iv) selection of a maximum value, due to objective of high precision

All these depends on the situation and circumstances to be faced.

Remark 6.4: In (6.11), the suitable \((n_{op})\), exists only if choice of \(\gamma_i\) lies in a certain range. This helps to choose a prefixed level of variance \(V_{0i}\) for \(i^{th}\) strata using (6.10.1).

7. Numerical Illustrations

Suppose a population of 400 units, has four strata I, II, III, IV and each one is divided into two groups as response (R) and non-response (NR). The detail of these is in appendix A and population parameters are given below

\[ N = 400; \quad \bar{Y} = 81.65; \quad S^2 = 1857.025 \]

<table>
<thead>
<tr>
<th>Type</th>
<th>Division</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size of strata</td>
<td>Response Class (R)</td>
<td>(N'_1 = 30)</td>
<td>(N'_2 = 40)</td>
<td>(N'_3 = 60)</td>
<td>(N'_4 = 70)</td>
</tr>
<tr>
<td></td>
<td>Non-Res. Class (NR)</td>
<td>(N''_1 = 30)</td>
<td>(N''_2 = 40)</td>
<td>(N''_3 = 60)</td>
<td>(N''_4 = 70)</td>
</tr>
<tr>
<td>Total (N_i)</td>
<td></td>
<td>(N_1 = 60)</td>
<td>(N_2 = 80)</td>
<td>(N_3 = 120)</td>
<td>(N_4 = 140)</td>
</tr>
<tr>
<td>Mean (\bar{Y}_i)</td>
<td></td>
<td>(\bar{Y}_1 = 13.48)</td>
<td>(\bar{Y}_2 = 48.47)</td>
<td>(\bar{Y}_3 = 85.50)</td>
<td>(\bar{Y}_4 = 126.88)</td>
</tr>
<tr>
<td>Response Class (R)</td>
<td>(S^2_{11} = 67.06)</td>
<td>(S^2_{12} = 156.67)</td>
<td>(S^2_{13} = 228.01)</td>
<td>(S^2_{14} = 240.20)</td>
<td></td>
</tr>
<tr>
<td>Mean Squares</td>
<td>Non-Res. Class (NR)</td>
<td>(S^2_{21} = 62.97)</td>
<td>(S^2_{22} = 204.87)</td>
<td>(S^2_{23} = 248.79)</td>
<td>(S^2_{24} = 254.74)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(S^2_1 = 64.38)</td>
<td>(S^2_2 = 180.46)</td>
<td>(S^2_3 = 225.63)</td>
<td>(S^2_4 = 245.74)</td>
</tr>
<tr>
<td>Weights (W_i)</td>
<td></td>
<td>(W_1 = 0.15)</td>
<td>(W_2 = 0.20)</td>
<td>(W_3 = 0.30)</td>
<td>(W_4 = 0.35)</td>
</tr>
</tbody>
</table>
A sample of size \( n = 80 \) is drawn by SRWOR and post-stratified into \( n_i \) units. When a deadline of receiving mailed questionnaire is over, \( n_i' \) and \( n_i'' \) units are as below along with prefixed \( n_i''' \) and \( f_i \)

\[
\begin{array}{|c|c|c|c|c|c|}
\hline
\text{Type} & \text{Strata I} & \text{Strata II} & \text{Strata III} & \text{Strata IV} & \text{Total} \\
\hline
\text{Sample size} & n_i & 16 & 18 & 22 & 24 & 80 \\
\hline
n_i' & 06 & 08 & 09 & 10 & 33 \\
\hline
n_i'' & 10 & 10 & 13 & 14 & 47 \\
\hline
n_i''' & 03 & 04 & 04 & 05 & 16 \\
\hline
f_i = n_i''' / n_i'' & 3.33 & 2.50 & 2.50 & 2.00 & 10.33 \\
\hline
\end{array}
\]

(i) We have \( V(\bar{y}_{PSNR}) = 2.228 \) which is quite small showing a high precision of \( \bar{y}_{PSNR} \).

(ii) For calculation purpose alternative strategy is considered with cost \( C_{0i}, C_{1i} \) and \( C_{2i} \). Moreover, the cost optimal \( f_i \) is calculated using (6.9) and shown below

\[
\begin{array}{|c|c|c|c|c|}
\hline
\text{Strata I} & \text{Strata II} & \text{Strata III} & \text{Strata IV} \\
\hline
C_{01} = 01.0 & C_{02} = 01.0 & C_{03} = 01.0 & C_{04} = 02.0 \\
\hline
C_{11} = 03.0 & C_{12} = 03.0 & C_{13} = 03.0 & C_{14} = 03.0 \\
\hline
C_{21} = 0.50 & C_{22} = 01.0 & C_{23} = 01.0 & C_{24} = 01.0 \\
\hline
f_1 = 8.00 & f_2 = 15.80 & f_3 = 15.78 & f_4 = 12.27 \\
\hline
\end{array}
\]

Table 3 reveals that \( C_{0i} \) and \( C_{2i} \) are nearly same but \( C_{1i} \) is required much higher to these two.

(iii) The optimal sample size (\( n_{opt} \)), and values of \( V_{0i} \) are tabulated below

\[
\begin{array}{|c|c|c|c|}
\hline
\text{Strata} & \text{Strata I} & \text{Strata II} & \text{Strata III} & \text{Strata IV} \\
\hline
V_{0i} & 100 & 150 & 275 & 125 \\
(n_{opt}) & 23 & 59 & 27 & 40 \\
\hline
\end{array}
\]

As per remark (6.3), an average integer value of cost optimal \( n \) is 37 taken over all the four strata.

(iv) The calculation of population mean \( \bar{Y} \) using estimator \( \bar{y}_{PSNR} \) is performed over four samples and displayed below.
<table>
<thead>
<tr>
<th>Strata</th>
<th>Sample 1</th>
<th>Sample 2</th>
<th>Sample 3</th>
<th>Sample 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>n^1 6(9, 6, 11, 15, 19, 7)</td>
<td>n^1 6(4, 13, 14, 30, 7, 10)</td>
<td>n^1 6(9, 12, 14, 8, 1, 29)</td>
<td>n^1 6(2, 17, 22, 19, 7, 20)</td>
</tr>
<tr>
<td></td>
<td>10^1 3(22, 4, 15)</td>
<td>10^1 3(26, 30, 15)</td>
<td>10^1 3(22, 2, 20)</td>
<td>10^1 3(20, 8, 15)</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>11.167</td>
<td>13.00</td>
<td>12.167</td>
<td>14.5</td>
</tr>
<tr>
<td>II</td>
<td>n^2 8(59, 35, 47, 26, 34, 42, 60, 58)</td>
<td>n^2 8(35, 50, 62, 56, 43, 64, 47, 60)</td>
<td>n^2 8(49, 36, 42, 31, 64, 26, 58, 34)</td>
<td>n^2 8(38, 62, 42, 55, 43, 48, 39, 63)</td>
</tr>
<tr>
<td></td>
<td>n^2 10</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>n^2 4(26, 70, 50, 69)</td>
<td>4(36, 66, 69, 58)</td>
<td>4(66, 28, 53, 44)</td>
<td>4(36, 70, 62, 53)</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>18</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>45.125</td>
<td>52.125</td>
<td>42.500</td>
<td>48.75</td>
</tr>
<tr>
<td></td>
<td>53.750</td>
<td>57.250</td>
<td>47.75</td>
<td>55.25</td>
</tr>
<tr>
<td>III</td>
<td>n^3 9(93, 68, 62, 80, 84, 59, 66, 100, 70)</td>
<td>n^3 9(88, 71, 91, 86, 100, 106, 74, 68, 49)</td>
<td>n^3 9(84, 66, 92, 84, 59, 102, 70, 81, 47)</td>
<td>n^3 9(62, 91, 78, 66, 84, 96, 111, 105, 47)</td>
</tr>
<tr>
<td></td>
<td>n^3 13</td>
<td>13</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>n^3 4(109, 67, 89, 95)</td>
<td>4(107, 99, 65, 72)</td>
<td>4(109, 67, 89, 72)</td>
<td>4(109, 67, 83, 96)</td>
</tr>
<tr>
<td></td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>75.780</td>
<td>81.440</td>
<td>76.110</td>
<td>82.220</td>
</tr>
<tr>
<td></td>
<td>90.000</td>
<td>85.750</td>
<td>84.250</td>
<td>88.750</td>
</tr>
<tr>
<td></td>
<td>n^4 14</td>
<td>14</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>n^4 5(139, 146, 106, 111, 121)</td>
<td>5(127, 145, 113, 106, 132)</td>
<td>5(149, 127, 109, 132, 110)</td>
<td>5(109, 147, 113, 132, 127)</td>
</tr>
<tr>
<td></td>
<td>24</td>
<td>24</td>
<td>24</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>121.600</td>
<td>129.300</td>
<td>131.1</td>
<td>129.1</td>
</tr>
<tr>
<td></td>
<td>125.200</td>
<td>124.600</td>
<td>125.4</td>
<td>125.6</td>
</tr>
<tr>
<td></td>
<td>80.3608</td>
<td>83.3530</td>
<td>80.0590</td>
<td>79.3960</td>
</tr>
</tbody>
</table>

| $\bar{y}_{PSN}$ | 299 | 213 |
(v) Table 5 reveals that sample estimates $\bar{y}_{PSNR}$ are very close to the true value of the population mean $\bar{Y}$. Clearly, the variability $\text{V}(\bar{y}_{PSNR})$ would be very small.

ACKNOWLEDGEMENT

Authors are thankful to the referee for useful comments and suggestion.

REFERENCES


## APPENDIX-A

Population (N = 400)

### Strata – I

<table>
<thead>
<tr>
<th>Response (R)</th>
<th>05</th>
<th>04</th>
<th>02</th>
<th>09</th>
<th>18</th>
<th>03</th>
<th>06</th>
<th>12</th>
<th>11</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non Response (NR)</td>
<td>04</td>
<td>22</td>
<td>14</td>
<td>26</td>
<td>08</td>
<td>30</td>
<td>20</td>
<td>15</td>
<td>06</td>
<td>08</td>
</tr>
<tr>
<td></td>
<td>02</td>
<td>04</td>
<td>16</td>
<td>08</td>
<td>10</td>
<td>22</td>
<td>14</td>
<td>06</td>
<td>18</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>01</td>
<td>13</td>
<td>25</td>
<td>07</td>
<td>19</td>
<td>21</td>
<td>15</td>
<td>15</td>
<td>27</td>
<td>09</td>
</tr>
</tbody>
</table>

### Strata – II

<table>
<thead>
<tr>
<th>Response (R)</th>
<th>25</th>
<th>67</th>
<th>59</th>
<th>41</th>
<th>53</th>
<th>35</th>
<th>47</th>
<th>49</th>
<th>50</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-Response (NR)</td>
<td>41</td>
<td>26</td>
<td>36</td>
<td>64</td>
<td>66</td>
<td>66</td>
<td>70</td>
<td>48</td>
<td>49</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>62</td>
<td>25</td>
<td>46</td>
<td>28</td>
<td>40</td>
<td>62</td>
<td>34</td>
<td>63</td>
<td>62</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>28</td>
<td>47</td>
<td>24</td>
<td>55</td>
<td>40</td>
<td>69</td>
<td>47</td>
<td>50</td>
<td>53</td>
</tr>
<tr>
<td></td>
<td>55</td>
<td>58</td>
<td>61</td>
<td>66</td>
<td>53</td>
<td>49</td>
<td>67</td>
<td>44</td>
<td>42</td>
<td>70</td>
</tr>
</tbody>
</table>

### Strata – III

<table>
<thead>
<tr>
<th>Response (R)</th>
<th>60</th>
<th>88</th>
<th>67</th>
<th>93</th>
<th>84</th>
<th>85</th>
<th>60</th>
<th>71</th>
<th>68</th>
<th>79</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>62</td>
<td>66</td>
<td>65</td>
<td>77</td>
<td>92</td>
<td>80</td>
<td>91</td>
<td>62</td>
<td>72</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>89</td>
<td>98</td>
<td>84</td>
<td>86</td>
<td>78</td>
<td>90</td>
<td>59</td>
<td>74</td>
<td>96</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>66</td>
<td>100</td>
<td>84</td>
<td>102</td>
<td>78</td>
<td>76</td>
<td>70</td>
<td>96</td>
<td>106</td>
<td>88</td>
</tr>
<tr>
<td></td>
<td>80</td>
<td>84</td>
<td>74</td>
<td>86</td>
<td>68</td>
<td>81</td>
<td>111</td>
<td>113</td>
<td>105</td>
<td>68</td>
</tr>
<tr>
<td></td>
<td>96</td>
<td>71</td>
<td>62</td>
<td>83</td>
<td>94</td>
<td>90</td>
<td>96</td>
<td>47</td>
<td>98</td>
<td>49</td>
</tr>
<tr>
<td>Non-Response (NR)</td>
<td>61</td>
<td>103</td>
<td>105</td>
<td>107</td>
<td>109</td>
<td>70</td>
<td>63</td>
<td>84</td>
<td>98</td>
<td>95</td>
</tr>
<tr>
<td>------------------</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td></td>
<td>90</td>
<td>99</td>
<td>98</td>
<td>67</td>
<td>96</td>
<td>85</td>
<td>74</td>
<td>83</td>
<td>72</td>
<td>101</td>
</tr>
<tr>
<td></td>
<td>84</td>
<td>65</td>
<td>67</td>
<td>89</td>
<td>70</td>
<td>106</td>
<td>98</td>
<td>79</td>
<td>105</td>
<td>87</td>
</tr>
<tr>
<td></td>
<td>89</td>
<td>91</td>
<td>87</td>
<td>96</td>
<td>72</td>
<td>75</td>
<td>97</td>
<td>80</td>
<td>81</td>
<td>102</td>
</tr>
<tr>
<td></td>
<td>108</td>
<td>92</td>
<td>63</td>
<td>90</td>
<td>104</td>
<td>107</td>
<td>88</td>
<td>95</td>
<td>109</td>
<td>120</td>
</tr>
<tr>
<td></td>
<td>120</td>
<td>83</td>
<td>62</td>
<td>105</td>
<td>74</td>
<td>107</td>
<td>66</td>
<td>98</td>
<td>109</td>
<td>112</td>
</tr>
</tbody>
</table>

Strata- IV

<table>
<thead>
<tr>
<th>Response (R)</th>
<th>100</th>
<th>101</th>
<th>112</th>
<th>123</th>
<th>104</th>
<th>108</th>
<th>146</th>
<th>127</th>
<th>148</th>
<th>108</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>149</td>
<td>122</td>
<td>112</td>
<td>133</td>
<td>110</td>
<td>145</td>
<td>126</td>
<td>117</td>
<td>108</td>
<td>129</td>
</tr>
<tr>
<td></td>
<td>120</td>
<td>124</td>
<td>124</td>
<td>136</td>
<td>148</td>
<td>130</td>
<td>102</td>
<td>134</td>
<td>136</td>
<td>148</td>
</tr>
<tr>
<td></td>
<td>121</td>
<td>133</td>
<td>105</td>
<td>127</td>
<td>149</td>
<td>131</td>
<td>133</td>
<td>135</td>
<td>137</td>
<td>149</td>
</tr>
<tr>
<td></td>
<td>149</td>
<td>121</td>
<td>122</td>
<td>143</td>
<td>144</td>
<td>135</td>
<td>146</td>
<td>107</td>
<td>108</td>
<td>149</td>
</tr>
<tr>
<td></td>
<td>121</td>
<td>112</td>
<td>134</td>
<td>106</td>
<td>128</td>
<td>108</td>
<td>106</td>
<td>109</td>
<td>102</td>
<td>140</td>
</tr>
<tr>
<td></td>
<td>131</td>
<td>138</td>
<td>143</td>
<td>111</td>
<td>115</td>
<td>150</td>
<td>137</td>
<td>138</td>
<td>129</td>
<td>146</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Non Response (NR)</th>
<th>139</th>
<th>148</th>
<th>137</th>
<th>130</th>
<th>149</th>
<th>138</th>
<th>127</th>
<th>116</th>
<th>145</th>
<th>124</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>109</td>
<td>150</td>
<td>147</td>
<td>106</td>
<td>145</td>
<td>144</td>
<td>113</td>
<td>112</td>
<td>141</td>
<td>110</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>104</td>
<td>106</td>
<td>108</td>
<td>140</td>
<td>132</td>
<td>124</td>
<td>146</td>
<td>108</td>
<td>128</td>
</tr>
<tr>
<td></td>
<td>121</td>
<td>103</td>
<td>105</td>
<td>127</td>
<td>149</td>
<td>111</td>
<td>118</td>
<td>125</td>
<td>147</td>
<td>109</td>
</tr>
<tr>
<td></td>
<td>130</td>
<td>149</td>
<td>108</td>
<td>147</td>
<td>106</td>
<td>125</td>
<td>124</td>
<td>125</td>
<td>102</td>
<td>121</td>
</tr>
<tr>
<td></td>
<td>119</td>
<td>107</td>
<td>145</td>
<td>104</td>
<td>143</td>
<td>121</td>
<td>132</td>
<td>108</td>
<td>146</td>
<td>144</td>
</tr>
<tr>
<td></td>
<td>132</td>
<td>134</td>
<td>126</td>
<td>148</td>
<td>110</td>
<td>112</td>
<td>134</td>
<td>146</td>
<td>128</td>
<td>150</td>
</tr>
</tbody>
</table>
APPENDIX - B

Applied Science Periodical

A new sampling scheme for coping with complete non-response in sample surveys

by D. Shukla & Jayant Dubey, Department of Mathematics and Statistics,
Dr. H.S. Gour University, Sagar-470003

(Received February 04, 2000)

Abstract

This paper presents a sampling scheme named “Poststratified Completely Non-Response (PSC NR)” scheme which takes into account the case of complete non-response in addition to the consideration of non-availability of frames in surveys. The population mean is under object of estimation for which two estimators are constructed and compared in precision.

1. Some notations

Let \( N \) be size of a population divided into \( k \) strata each of size \( N_i \) such that \( \Sigma N_i = N \). Further assume that \( i^{th} \) strata has \( N_i^r \) respondents \( (R) \) and \( N_i^* \) non-respondents \( (NR) \). The variable \( Y \) is under study with \( \bar{Y}_i \) as a population mean of \( i^{th} \) strata and \( \bar{Y} = \Sigma W_i \bar{Y}_i \) as a grand mean \( (i=1,2,3,...,k) \). The population mean squares are \( S^2 \) and \( S_i^2 \) for entire population and for \( i^{th} \) strata respectively. Within the \( i^{th} \) strata, response group \( (R) \) has \( S_{1i}^2 \) and non-response group \( (NR) \) has \( S_{2i}^2 \) having their mean squares.

2. Poststratified complete non-response (PSCNR) scheme

Step I : Draw a random sample of size \( n \) from \( N \) by SRSWOR and Poststratified it into \( k \) groups of size \( n_i \) from \( N_i (\Sigma n_i = n) \) using known auxiliary variable.

Step II : Questionnaires are mailed to all these \( n \) respondents stratawise in order to get observations on variable \( Y \).

Step III : Assume that the complete non-response was observed when a deadline of returning questionnaires is over.
Step IV: To cope with this, select sub-samples of size \( n_i^{''} \) from \( n_i \) by SRSWOR, by maintaining a preassigned ratio \( f_i = \left( \frac{n_i^{''}}{n_i} \right) \).

Step V: Conduct a personal interview for the data collection over these \( n_i^{''} \) units and let \( \bar{y}_i^{''} \) is sample mean based on these.

3. Estimation strategy

Under PSCNR [see Hansen and Hurwitz (1946)] for estimating population mean \( \bar{Y} \) proposed estimators are

\[
\bar{Y}_{PSCNR} = \sum_{i=1}^{k} W_i \bar{y}_i^{''}
\]

(3.1)

and

\[
\bar{Y}_{PSNCR} = \sum_{i=1}^{k} W_i \bar{y}_i^{'}
\]

(3.2)

where \( W_i = (N_i N) / f_i = (n_i / n) \)

\( W_i = (\alpha f_i + (1 - \alpha) (W_i)) \) and \( 0 \leq \alpha \leq 1 \) is a suitably chosen constant.

Note: In particular, \( \alpha = 0 \) leads to \( W_i = W_i \) [see Agrawal and Panda (1993)]

Theorem 3.1: Estimators \( \bar{Y}_{PSCNR} \) and \( \bar{Y}_{PSNCR} \) are unbiased for \( \bar{Y} \).

Proof: \( E \left\{ \bar{Y}_{PSCNR} \right\} = E \left[ E \left\{ \left( \frac{1}{n_i^{''}} \sum_{i=1}^{k} W_i \bar{y}_i^{''} / n_i, n_i^{''} \right) \right\} \right] = \sum_{i=1}^{k} W_i E(Y_i|n_i^{''}) = \bar{Y} \),

Here \( E \left( \cdot \mid n_i, n_i^{''} \right) \) denotes conditional expectation given \( n_i, n_i^{''} \).

And \( E \left( \bar{Y}_{PSNCR} \right) = E \left\{ E \left\{ \left( \frac{1}{n_i} \sum_{i=1}^{k} W_i \bar{y}_i^{'} / n_i, n_i^{''} \right) \right\} \right\} = \sum_{i=1}^{k} (W_i n_i) \bar{y}_i = \bar{Y} \).

Theorem 3.2: The variance of \( \bar{Y}_{PSCNR} \) is

\[
\sum_{i=1}^{k} W_i^2 \left( \left( \frac{1}{n_i} \right) (f_i - 1) S_i^2 + E \left\{ \left( \frac{1}{n_i} \right) S_i^2 \right\} - \left( \frac{1}{N} \right) S^2 \right) \]

(3.3)
where \( E\left( \frac{1}{n_i} \right) = \left[ \frac{1}{n_i W_i} + \frac{(N-n)(1-W_i)}{(N-1)n^2 W_i^2} \right] \) (3.4)

**Proof:** \( V\left( \overline{y}_{\text{PSCNR}} \right) = E\left[ E \left\{ V \left( \overline{y}_{\text{PSCNR}} \mid n_t, n_i^{\prime \prime} \right) \right\} \right] + E\left[ E \left\{ V \left( \overline{y}_{\text{PSCNR}} \mid n_t, \overline{y}_i \right) \right\} \right] + E\left[ E \left\{ V \left( \overline{y}_{\text{PSCNR}} \mid n_t, n_i^{\prime \prime} \right) \right\} \right] \) (3.4.1)

Here \( V \{ (.)/ n_t, n_i^{\prime \prime} \} \) denotes conditional variance given \( n_t, n_i^{\prime \prime} \)

\[
E\left[ E \left\{ V \left( \overline{y}_{\text{PSCNR}} \mid n_t, n_i^{\prime \prime} \right) \right\} \right] = E\left[ E \left\{ V \left( \sum_{i=1}^{k} W_i \overline{y}_i \mid n_t, n_i^{\prime \prime} \right) \right\} \right] = \sum_{i=1}^{k} W_i^2 \left[ \left( \frac{1}{n_t} \right) S_i^2 \right] \] (3.5)

The third term of (3.4.1) vanishes whereas the second term is

\[
E\left[ E \left\{ V \left( \overline{y}_{\text{PSCNR}} \mid n_t, n_i^{\prime \prime} \right) \right\} \right] = E\left[ E \left\{ \sum_{i=1}^{k} W_i^2 \left( \frac{1}{n_t} \right) S_i^2 \right\} \right] = \sum_{i=1}^{k} W_i^2 E \left( \frac{1}{n_t} \right) S_i^2 - \sum_{i=1}^{k} W_i^2 \left( \frac{1}{N_i} \right) \overline{y}_i^2 \] (3.6)

On adding (3.5) and (3.6) one can get proof of the theorem.

**Theorem 3.3:** The variance of \( \overline{y}_{\text{PSCNR}} \) is

\[
\sum_{i=1}^{k} \left\{ \frac{a^2}{n} \frac{E(p_i)}{n} + (1-a)^2 \frac{W_i^2}{N_i} E \left( \frac{1}{n_t} \right) + 2a(1-a) \left( \frac{W_i^2}{N_i} \right) \right\} \] + \sum_{i=1}^{k} \left\{ \left( \frac{a}{n} \right) E(p_i) + (1-a)^2 \frac{W_i^2}{N_i} E \left( \frac{1}{n_t} \right) + 2a(1-a) \left( \frac{W_i^2}{N_i} \right) - \left( \frac{a^2}{N_i} \right) E(p_i^2) \right\} \] - \left( 1-a \right)^2 \left( \frac{W_i^2}{N_i} \right) - 2a(1-a) \left( \frac{W_i^2}{N_i} \right) \overline{y}_i^2 \] + \sum_{i=1}^{k} \left( \frac{W_i}{N_i} \right) \left( \frac{W_i}{N_i} \right) \] (3.7)

Where, \( p_i = (ni/n) \) with \( E(p_i) = W_i \)

\[
E(p_i^2) = \left\{ \left( 1 - \frac{n_i}{N_i} \right) \right\} \] + \left( 1 - \frac{n_i}{N_i} \right) \left( \frac{W_i}{N_i} \right) \] (3.8)

and \( \sum_{i=1}^{k} \left( \frac{W_i}{N_i} \right) \] (3.9)
Proof: \[
E \left[ E \left\{ V \left( \tilde{y}'_{\text{PSGNR}} \mid n_t, n''_i \right) \right\} \right] = E \left[ \sum_{i=1}^{k} W_i^2 \left( \frac{1}{n_i} \right) S_{2i}^2 \right] / n_t \\
= \sum_{i=1}^{k} \left[ \left\{ \left( \frac{2^2}{n} \right) E(\rho_i) + (1-\alpha)^2 \frac{W_i^2}{n} \right\} E \left( \frac{1}{n_i} \right) + 2\alpha(1-\alpha) \left( \frac{W_i}{n} \right) \right] \left( f_i - 1 \right) S_{2i}^2 \right] \tag{3.10}
\]

And, \[
E \left[ E \left\{ V \left( \tilde{y}'_{\text{PSGNR}} \mid n_t, n''_i \right) \right\} \right] = E \left[ \left\{ \sum_{i=1}^{k} W_i^2 \tilde{y}_i T(\tilde{y}_i) \right\} \right] / n_t \\
= \sum_{i=1}^{k} \left[ \left\{ \left( \frac{2^2}{n} \right) E(\rho_i) + (1-\alpha)^2 \frac{W_i^2}{n} \right\} E \left( \frac{1}{n_i} \right) + 2\alpha(1-\alpha) \left( \frac{W_i}{n} \right) \right] S_i^2 \right] \tag{3.11}
\]

\[
E \left[ E \left\{ V \left( \tilde{y}'_{\text{PSGNR}} \mid n_t, n''_i \right) \right\} \right] = E \left[ \left\{ \sum_{i=1}^{k} W_i T_i \tilde{y}_i \right\} \right] / n_t = \sum_{i=1}^{k} \frac{T_i^2}{T_i} \left( \tilde{y}'_{\text{PSGNR}} \right) \tag{3.12}
\]

On adding the proof of theorem exists.

Remark 3.1: At \(\alpha = 0\), \(\tilde{y}'_{\text{PSGNR}} = \tilde{y}'_{\text{PSGNR}}\) holds.

Theorem 3.4: The optimum choice of \(x\) is
\[
x_{\text{opt}} = \frac{Q_1 + Q_2 + Q_3}{Q_1 + Q_2 + Q_3} \tag{3.13}
\]

Where \(Q_1 = \sum_{i=1}^{k} \left\{ \left( \frac{(N-n)W_i}{n(N-1)} \right) \left( \frac{1}{n_i} \right) \right\} \left\{ \left( \frac{W_i^2}{n_i} \right) \right\} \}
\]

\[
Q_2 = \sum_{i=1}^{k} \left\{ \left( \frac{(N-n)W_i}{n(N-1)} \right) \left( \frac{1}{n_i} \right) \right\} \left\{ \left( \frac{W_i^2}{n_i} \right) \right\} \}
\]

\[
Q_3 = \sum_{i=1}^{k} \left\{ \left( \frac{(N-n)W_i}{n(N-1)} \right) \left( \frac{1}{n_i} \right) \right\} \}
\]

Remark 3.2: On substituting optimum value \(x_{\text{opt}}\), the expression of optimum variance would be
\[
\tilde{y}'_{\text{opt}} \left( \tilde{y}'_{\text{PSGNR}} \right) = \sum_{i=1}^{k} \left( \frac{W_i^2}{n_i} \right) + \left( 1 - x_{\text{opt}} \right) x Q_1
\]

220
\[ + \frac{k}{n} \sum_{i=1}^{\infty} \left\{ \frac{a_{opt}^2 (N-1)W_i (1-W_i)}{n(N-1)+\bar{W}^2} \right\} \left\{ \left( \frac{1-W_i}{nW_i N-1} \right) + \left( \frac{1}{N} \right) \right\} \]

\[ \cdot \left\{ \frac{(N-1)S_i^2}{W_i} \right\} + a_{opt}^2 \sum_{i=1}^{\infty} \left( \frac{(N-1)W_i (1-W_i)}{n(N-1)} \right) \]

4. Comparison of estimators

The expression of variance of \( \tilde{\gamma} \) could also be expressed as

\[ \Gamma \left( \tilde{\gamma}_\text{PSCNR} \right) - \Gamma \left( \bar{\gamma}_\text{PSCNR} \right) = \left[ a^2 \phi_1 - 2a \phi_2 \right] \]

Where \( \phi_1 = Q_1 + Q_2 + Q_3 \)

and \( \phi_2 = Q_1 \)

Clearly, \( \tilde{\gamma}_\text{PSCNR} \) will be efficient than \( \bar{\gamma}_\text{PSCNR} \) if \( a < \left[ \frac{2\phi_2}{\phi_1} \right] \) or \( 0 < a \leq a_{opt} \) (3.15)

Obviously, the optimum value of \( a \) would lie in the efficient range.

References

To,

Mr. Jayant Dubey,
Research Scholar,
Deptt. Of Maths & Statistics,
Dr. Hari Singh Gour Vishwavidyalaya,
Sagar (M.P.)

Dear Mr. Dubey,

It is my pleasure to inform you that your research paper entitled “On Estimation Under PSTPNR Sampling Scheme For Mail Surveys” (Jointly with D. Shukla) has been accepted for publication in the proceedings of the seminar on “National Science Day – 2000”, organized by the Faculty of Science, Dr. H.S. Gour Vishwavidyalaya, Sagar on February 28, 2000.

With best wishes,

Yours Sincerely,

(Dr. K.K.N. Sharma)
Organizing Secretary
Seminar on National Science Day – 2000
Dr. H.S. Gour Vishwavidyalaya,
Sagar (M.P.)