Mean Estimation in Deeply Stratified Population 
Under Post-Stratification

D. Shukla and Manish Trivedi

Dr. Hari Singh Gour University, Sagar, (M.P.), 470003
(Received : August, 1999)

SUMMARY

Suppose a population is stratified according to two attributes, each having 
three levels, in particular, then it constitutes a $3 \times 3$ deep stratification and 
interesting for survey practitioners being close to reality. This paper, presents 
the problem of mean estimation under above population, when frames of each 
$3 \times 3$ stratum are assumed to be unknown. An estimation strategy has been 
proposed using the post-stratified sampling scheme. The optimum properties 
are examined and relative efficiencies are compared. Mathematical finding is 
numerically supported.

Key words : Post-stratification, SRSWOR, Optimal, Deep stratification.

1. Introduction

We assume existence of a $3 \times 3$ deeply stratified population of size $N$ in 
particular. Let $Y_{ij}$ be the $k$th value of $(i, j)$th stratum having size $N_{ij}$ of a variable $Y$ 
under study $(i = 1, 2, 3; j = 1, 2, 3$ and $k = 1, 2, \ldots, N_{ij})$. A random sample of size $n$ 
is drawn by SRSWOR and post-stratified into $n_{ij}$ units such that

$$n_{ij}\left(\sum_{i=1}^{3}\sum_{j=1}^{3} n_{ij} = n\right) \text{ comes from } N_{ij}\left(\sum_{i=1}^{3}\sum_{j=1}^{3} N_{ij} = N\right)$$

Let $\bar{Y}_{ij}$ be the mean and $S_{ij}^2$ be the population mean square of $(i, j)^{th}$ strata. 
Also, $\bar{Y}$ and $S^2$ represent entire population mean and population mean square 
along with $\bar{y}_{ij}$ and $\bar{y}$ as sample means based on $n_{ij}$ and $n$ units respectively.
Moreover, \[ N_i = \sum_{j=1}^{3} N_{ij}, \quad N_{i.} = \sum_{i=1}^{3} N_{ij}, \quad n_i = \sum_{j=1}^{3} n_{ij} \]
and
\[ n_{.j} = \sum_{i=1}^{3} n_{ij} \] are row and column totals and \( \bar{Y}_i, \bar{Y}_{i.}, \bar{Y}_{.j} \) are population and sample means based on them respectively.

1.1 An Example

In an educational survey, students are classified as per their Academic-merit and Economic-background. Let an educational institution has \( P_1 \) proportion of meritorious students, \( P_2 \) average and \( P_3 \) below average students \((P_1 + P_2 + P_3 = 1)\). Whereas same has \( P_4 \) proportion of economically poor students, \( P_5 \) from middle-class income and \( P_6 \) from above middle-income level \((P_4 + P_5 + P_6 = 1)\). This constitutes \( 3 \times 3 \) classification where \( P_m \) \((m=1,2,....6)\) are known alongwith total strength of students in the institution but, each cell frequency and cell-frames are unknown. The survey practitioner wants to estimate the average monthly expenditure of students by an effective utilization of prior information on proportions \( P_m \).

2. Derivation of Some Useful Theorems

With usual notations,
\[ W_{ij} = \left( \frac{N_{ij}}{N} \right), \quad W_{i.} = \left( \frac{N_{i.}}{N} \right), \quad W_{.j} = \left( \frac{N_{.j}}{N} \right), \quad p_{ij} = \left( \frac{n_{ij}}{n} \right), \quad p_{i.} = \left( \frac{n_{i.}}{n} \right) \quad \text{and} \quad p_{.j} = \left( \frac{n_{.j}}{n} \right) \]
assume sample \( n \) is large enough to support following
\[ p_{ij} = W_{ij} (1+\epsilon_{ij}), \quad p_{i.} = W_{i.} (1+\epsilon_{i.}), \quad p_{.j} = W_{.j} (1+\epsilon_{.j}) \]  \( (2.1) \)
where, \( E[\epsilon_{ij}] = E[\epsilon_{i.}] = E[\epsilon_{.j}] = 0; \quad i \neq i' = 1,2,3; \quad j \neq j' = 1,2,3 \)

\[ E[\epsilon_{ij}^2] = \left( \frac{1}{W_{ij}^2} \right) \left[ \frac{(N-n)W_{ij}(1-W_{ij})}{(N-1)n} \right], \quad E[\epsilon_{i.}^2] = \left( \frac{1}{W_{i.}^2} \right) \left[ \frac{(N-n)W_{i.}(1-W_{i.})}{(N-1)n} \right] \]
\[ E[\epsilon_{.j}^2] = \left( \frac{1}{W_{.j}^2} \right) \left[ \frac{(N-n)W_{.j}(1-W_{.j})}{(N-1)n} \right] \]
\[ E[\varepsilon_{ij} \varepsilon_{ij'}] = \left\{ \begin{array}{ll} \frac{-1}{W_{ij} W_{ij'}} \left[ \frac{(N-n)W_{ij}}{N-n} \right] \\
\end{array} \right. \\
\]

\[ E[\varepsilon_{ij} \varepsilon_{ij}] = \left\{ \begin{array}{ll} \frac{-1}{W_{ij} W_{ij'}} \left[ \frac{(N-n)W_{ij}}{N-n} \right] \\
\end{array} \right. \\
\]

\[ E[\varepsilon_{ij} \varepsilon_{ij'}] = \left\{ \begin{array}{ll} \frac{-1}{W_{ij} W_{ij'}} \left[ \frac{(N-n)W_{ij}}{N-n} \right] \\
\end{array} \right. \\
\]

2.1 Justification

For sample mean \( \bar{y} \) based on \( n \) units and \( E(\bar{y}) = \bar{Y} \), Sukhatme et al. [2] have used one of approximations as \( \bar{y} = \bar{Y}(1 - \varepsilon) \), \( E(\varepsilon) = 0 \), assuming sample size large and derived expressions of m.s.e. for ratio, product and regression estimators up to first and second order of approximations. In particular, for an attribute \( A \) in the same population, suppose

\[ Y_i = 1 \text{ if } i^{th} \text{ population unit possess } A \]
\[ = 0 \text{ otherwise} \]

then \( \bar{y} = w \), \( E(w) = \bar{W} \) holds where \( w \) and \( \bar{W} \) are sample and population proportions respectively with respect to \( A \). Without loss of generality, one can write \( w = \bar{W}(1 + \varepsilon') \), \( E(\varepsilon') = 0 \) for a large \( n \).

Theorem 2.1: Using (2.1) and avoiding terms of higher order, an approximate result for \( j \neq j' \), is

\[ A_{ij(j')} = E \left[ \frac{p_{ij'}}{p_{ij}} \right] = \frac{W_{ij'} / W_{ij}}{1 + \frac{\text{Var}(p_{ij})}{W_{ij}^2} - \frac{\text{Cov}(p_{ij}, p_{ij'})}{W_{ij} W_{ij'}}} \]

Proof:

\[ E \left[ \frac{p_{ij'}}{p_{ij}} \right] = E \left[ \frac{W_{ij'} (1 + \varepsilon_{ij'})}{W_{ij} (1 + \varepsilon_{ij})} \right] = \frac{W_{ij'} / W_{ij}}{1 + \varepsilon_{ij'} - \varepsilon_{ij} + \varepsilon_{ij'}^2 - \varepsilon_{ij}^2 + \varepsilon_{ij'}^3 - \varepsilon_{ij}^3 + \varepsilon_{ij'}^4 - \varepsilon_{ij}^4 + \ldots} \]

Avoiding all higher order terms \( \left[ (\varepsilon_{ij}, \varepsilon_{ij'})^r (\varepsilon_{ij}, \varepsilon_{ij'})^s \right] \) for \( r, s > 2 \), theorem holds.
Theorem 2.2: Using (2.1), an approximate result, for \( j \neq j' \) is

\[
B_{ij(j')} = E \left[ \frac{P_{ij}^2}{\hat{p}_{ij}} \right] = \frac{W_{ij}^2}{W_{ij}'} \left[ 1 + \frac{\text{Var}(p_{ij})}{W_{ij}^2} + \frac{\text{Var}(p_{ij'})}{W_{ij'}^2} - 2\frac{\text{Cov}(p_{ij}, p_{ij'})}{W_{ij} W_{ij'}} \right]
\]

Proof:

\[
E \left[ \frac{\hat{p}_{ij}^2}{p_{ij}} \right] = E \left[ \frac{W_{ij}^2 (1 + \epsilon_{ij})^2}{W_{ij} (1 + \epsilon_{ij})} \right] = \frac{W_{ij}^2}{W_{ij}} E \left[ 1 + \epsilon_{ij}^2 + 2 \epsilon_{ij} - \epsilon_{ij} - 2 \epsilon_{ij} \epsilon_{ij} + \epsilon_{ij}^2 + ... \right]
\]

Avoiding \( (\epsilon_{ij})^r (\epsilon_{ij'})^s \) for \( r + s > 2 \), theorem holds.

Theorem 2.3: Using (2.1), an approximate result, for \( i \neq i', j \neq j' \), is

\[
C_{ij(i'j')} = E \left[ \frac{P_{ij'} P_{i'j}}{\hat{p}_{ij'}} \right] = \frac{W_{ij} W_{i'j}}{W_{ij'}} \left[ 1 + \frac{\text{Var}(p_{ij})}{W_{ij}^2} + \frac{\text{Cov}(p_{ij}, p_{ij'})}{W_{ij} W_{ij'}} \right]
\]

\[
- \frac{\text{Cov}(p_{ij}, p_{ij'})}{W_{ij} W_{i'j'}} - \frac{\text{Cov}(p_{ij'}, p_{ij'})}{W_{ij'} W_{i'j'}}
\]

Proof:

\[
E \left[ \frac{P_{ij'} P_{i'j}}{\hat{p}_{ij}} \right] = E \left[ \frac{W_{ij} (1 + \epsilon_{ij}) W_{i'j} (1 + \epsilon_{i'j})}{W_{ij} (1 + \epsilon_{ij})} \right]
\]

\[
= \frac{W_{ij} W_{i'j}}{W_{ij}} E \left[ 1 + \epsilon_{ij} + \epsilon_{ij'} + \epsilon_{ij} \epsilon_{ij'} - \epsilon_{ij} - \epsilon_{ij'} - \epsilon_{ij} \epsilon_{ij'} + \epsilon_{ij}^2 + ... \right]
\]

On avoiding terms \( (\epsilon_{ij})^r (\epsilon_{ij'})^s (\epsilon_{i'j})^t \) for \( r + s + t > 2 \), we get result.

2.2 Some Symbols

\[
D_{ij} = E \left[ \frac{1}{n_{ij}} \right] = \frac{1}{n W_{ij}} + \frac{(N - n)(1 - W_{ij})}{(N - 1) n^2 W_{ij}^2}
\]
\[
 F_{i} = E \left[ \frac{p_{i}^2}{N_{ij}} \right] = \left( \frac{1}{N_{ij}} \right) \left\{ \frac{(N-n)}{(N-1)} \frac{W_{i} \cdot (1-W_{i})}{n} + W_{i}^2 \right\}
\]

\[
 F_{j} = E \left[ \frac{p_{j}^2}{N_{ij}} \right] = \left( \frac{1}{N_{ij}} \right) \left\{ \frac{(N-n)}{(N-1)} \frac{W_{j} \cdot (1-W_{j})}{n} + W_{j}^2 \right\}; M_{i} = \sum_{j=1}^{3} \bar{Y}_{ij}
\]

\[
 F_{ij} = E \left[ \frac{p_{i} \cdot p_{j}}{N_{ij}} \right] = \left( \frac{1}{N_{ij}} \right) \left\{ \text{Cov}(p_{i}, p_{j}) + E(p_{i})E(p_{j}) \right\}; M_{j} = \sum_{i=1}^{3} \bar{Y}_{ij}
\]

3. Proposed Estimation Strategy

To recall assumptions are (a) a setup of \( 3 \times 3 \) deeply stratified population \( N \) (b) frame of \( N \) units available for non-stratifying variable (c) sample size \( n \) is large (d) stratum sizes \( N_{ij} \) are unknown but information about \( N_{i} \) and \( N_{j} \) are known by some other sources.

To estimate \( \bar{Y} \) a Deeply stratified Post-stratified estimator is

\[
 \bar{Y}_{dps} = \sum_{i=1}^{3} \sum_{j=1}^{3} W_{\alpha ij} \bar{Y}_{ij}
\]

(3.1)

where

\[
 W_{\alpha ij} = \left[ \left( \frac{\alpha}{2} \right) \left( \frac{n_{i}}{n} \right) + \left( \frac{N_{i} - n_{i}}{N} \right) \right] + \left[ \left( \frac{1-\alpha}{2} \right) \left( \frac{n_{j}}{n} \right) + \left( \frac{N_{j} - n_{j}}{N} \right) \right]
\]

The constant \( \alpha \) be suitably chosen such that \( 0 \leq \alpha \leq 1 \).

3.1 Motivation

I. The usual post-stratified estimator for a \( 3 \times 3 \) set-up is

\[
 \bar{Y}_{ps} = \sum_{i=1}^{3} \sum_{j=1}^{3} W_{ij} \bar{Y}_{ij}
\]

(3.2)

with \( W_{ij} = \left( \frac{N_{ij}}{N} \right) \) which essentially requires a knowledge of \( N_{ij} \)

II. When only information of \( N_{i} \) and \( N_{j} \) available but not \( N_{ij} \), the usual estimator (3.2) fails to perform estimation.

III. The information \( N_{i} \) and \( N_{j} \) are more common to be priorly known.
IV. An effective utilization of known $N_i$ and $N_j$ for estimation of $\bar{Y}$, is required.

V. A contribution by Agrawal and Panda [1] supports for choosing $W_{aij}$ in the present form.

3.2 Properties of Strategy

(I) At $\alpha = 1$, estimator $(\bar{y}_{dps})_1$ with $W_{1ij} = \left( \frac{1}{2} \right) \left[ \frac{n_i}{N} + \frac{N_i}{n} \right]$

(II) At $\alpha = 0$, estimator $(\bar{y}_{dps})_0$ with $W_{0ij} = \left( \frac{1}{2} \right) \left[ \frac{n_j}{N} + \frac{N_j}{n} \right]$

(III) At $\alpha = \frac{1}{2}$, estimator $(\bar{y}_{dps})_{1/2}$ with

$$W_{1/2ij} = \left( \frac{1}{4} \right) \left[ \left( \frac{n_i}{N} + \frac{N_i}{n} \right) + \left( \frac{n_j}{N} + \frac{N_j}{n} \right) \right]$$

We have $(\bar{y}_{dps})_1$ purely based on row totals, $(\bar{y}_{dps})_0$ on column totals and $(\bar{y}_{dps})_{1/2}$ on an average of these two.

Theorem 3.1: The estimator $\bar{y}_{dps}$ is biased for $\bar{Y}$.

Proof: Denote $E[(\cdot)/n_{ij}]$ as a conditional expectation given $n_{ij}$

$$E(\bar{y}_{dps}) = E \left[ E \left( \frac{(\bar{y}_{dps})_1}{n_{ij}} \right) \right] = E \left[ \left( \sum_{i=1}^{3} \sum_{j=1}^{3} W_{aij} E(\bar{y}_{ij}) \right) \frac{1}{n_{ij}} \right]$$

$$= \sum_{i=1}^{3} \sum_{j=1}^{3} E(W_{aij}) \bar{y}_{ij} = \bar{Y} + \alpha V_1 + (1-\alpha)V_2$$

where, $V_1 = \sum_{i=1}^{3} \sum_{j=1}^{3} \sum_{j'=1}^{3} W_{ij} \bar{y}_{ij'}$; $V_2 = \sum_{i=1}^{3} \sum_{i'=1}^{3} \sum_{j=1}^{3} W_{ij} \bar{y}_{i'j}$
Theorem 3.2: The Mean square error of $y_{dps}$ is

$$M(y_{dps}) = \left(\frac{1}{4}\right)\left[\left(U_1 + U_2\right) + \alpha^2 \left(R_1 + R_2 + 4V_1^2\right) + (1 - \alpha)^2 \left(S_1 + S_2 + 4V_2^2\right) + 2\alpha (1 - \alpha) \left(T_1 + T_2 + 4V_1 V_2\right)\right]$$

where,

$U_1 = \left[\left(\frac{11}{n}\right) - \left(\frac{15}{N}\right)\right]\sum_{i=1}^{3} \sum_{j=1}^{3} W_{ij} S_{ij}^2$, $U_2 = 0$ (considered for symmetry)

$$R_1 = \sum_{i=1}^{3} \sum_{j \neq j'}^{3} \left(\frac{1}{n}\right) B_{ij(j')} S_{ij}^2 + \sum_{i=1}^{3} \sum_{j=1}^{3} W_{i}^2 D_{ij} S_{ij}^2 + \sum_{i \neq i'}^{3} \sum_{j \neq j'}^{3} \left(\frac{2}{n}\right) C_{ij(i')} S_{ij}^2$$

$$+ \left(\frac{2}{n}\right) \sum_{i=1}^{3} \sum_{j \neq j'}^{3} W_{ij} A_{ij(j')} S_{ij}^2 - \sum_{i=1}^{3} \sum_{j=1}^{3} F_{ij} S_{ij}^2$$

$$- \left(\frac{3}{N}\right) \sum_{i=1}^{3} \sum_{j \neq j'}^{3} \left\{ \left(\frac{W_{ij} S_{ij}^2}{W_{ij}}\right) \right\} - \left(\frac{6}{N}\right) \sum_{i \neq i'}^{3} \sum_{j \neq j'}^{3} \left(\frac{W_{ij} W_{ij'}}{W_{ij}}\right) S_{ij}^2$$

$$R_2 = \left\{ \left(\frac{N - n}{(N - 1)n}\right) \left[ \sum_{i=1}^{3} W_{i} (1 - W_{i}) M_{i}^2 - \sum_{i=1}^{3} \sum_{i' \neq i} M_{i} M_{i'} W_{i} W_{i'} \right] \right\}$$

$$S_1 = \sum_{i=1}^{3} \sum_{j \neq j'}^{3} \left(\frac{1}{n}\right) B_{ij(j')} S_{ij}^2 + \sum_{i=1}^{3} \sum_{j=1}^{3} W_{i}^2 D_{ij} S_{ij}^2 + \sum_{i \neq i'}^{3} \sum_{j \neq j'}^{3} \left(\frac{2}{n}\right) C_{ij(i')} S_{ij}^2$$

$$+ \left(\frac{2}{n}\right) \sum_{i=1}^{3} \sum_{j \neq j'}^{3} W_{ij} A_{ij(j')} S_{ij}^2 - \sum_{i=1}^{3} \sum_{j=1}^{3} F_{ij} S_{ij}^2 - \left(\frac{3}{N}\right) \sum_{i=1}^{3} \sum_{j \neq j'}^{3} \left\{ \left(\frac{W_{ij} S_{ij}^2}{W_{ij}}\right) \right\}$$

$$- \left(\frac{6}{N}\right) \sum_{i \neq i'}^{3} \sum_{j \neq j'}^{3} \left(\frac{W_{ij} W_{ij'}}{W_{ij}}\right) S_{ij}^2$$

$$S_2 = \left\{ \left(\frac{N - n}{(N - 1)n}\right) \left[ \sum_{j=1}^{3} W_{j} (1 - W_{j}) M_{j}^2 - \sum_{j=1}^{3} \sum_{j' \neq j} M_{j} M_{j'} W_{j} W_{j'} \right] \right\}$$
\[ T_1 = \sum_{i=1}^{3} \sum_{j=1}^{3} W_i W_j D_{ij} S_{ij}^2 + \sum_{i \neq i'} \sum_{j \neq j'} \left( \frac{1}{n} \right) C_{ij(i', j')} S_{ij}^2 - \sum_{i=1}^{3} \sum_{j=1}^{3} F_j S_{ij}^2 \]

\[ + \left( \frac{1}{n} \right) \sum_{i=1}^{3} \sum_{j \neq j'} \left( W_i + W_j \right) A_{ij(j')} S_{ij}^2 - \left( \frac{3}{N} \right) \sum_{i \neq i'} \sum_{j \neq j'} \left( \frac{W_{ij} W_{ij'}}{W_{i'}} \right) S_{ij}^2 \]

\[ T_2 = - \left( \frac{(N-n)}{(N-1)n} \right) \sum_{i=1}^{3} \sum_{j=1}^{3} W_i W_j M_i M_j \]

Proof: \[ M(\bar{y}_{dps}) = \mathbb{E} \left[ \frac{\mathbb{V}(\bar{y}_{dps})}{n_{ij}} \right] + \mathbb{V} \left[ \frac{\mathbb{E}(\bar{y}_{dps})}{n_{ij}} \right] + [\text{Bias}(\bar{y}_{dps})]^2 \]

\[ \mathbb{E} \left[ \frac{\mathbb{V}(\bar{y}_{dps})}{n_{ij}} \right] = \mathbb{E} \left[ \sum_{i=1}^{3} \sum_{j=1}^{3} \left( \frac{1}{n_{ij}} \right) W_{ij}^2 S_{ij}^2 \right] - \mathbb{E} \left[ \sum_{i=1}^{3} \sum_{j=1}^{3} \left( \frac{1}{N_{ij}} \right) W_{ij}^2 S_{ij}^2 \right] \] (3.3)

For further derivation of (3.3) following are used

\[ a_1 : \mathbb{E} \left[ \sum_{i=1}^{3} \sum_{j=1}^{3} \left( \frac{1}{n_{ij}} \right) \left( \left( \frac{n_{ij}}{n} \right) + \left( \frac{N_{ij}}{N} \right) \right) S_{ij}^2 \right] = \left( \frac{11}{n} \right) \sum_{i=1}^{3} \sum_{j=1}^{3} W_{ij} S_{ij}^2 \]

\[ + \left( \frac{1}{n} \right) \sum_{i=1}^{3} \sum_{j \neq j'} \left( B_{ij(j')} S_{ij}^2 + \frac{2}{n} \sum_{i \neq i'} \sum_{j \neq j'} \left( C_{ij(i', j')} S_{ij}^2 \right) \right) \]

\[ + \sum_{i=1}^{3} \sum_{j=1}^{3} W_{ij}^2 D_{ij} S_{ij}^2 \]

\[ a_2 : \mathbb{E} \left[ \sum_{i=1}^{3} \sum_{j=1}^{3} \left( \frac{1}{n_{ij}} \right) \left( \left( \frac{n_{ij}}{n} \right) + \left( \frac{N_{ij}}{N} \right) \right) S_{ij}^2 \right] = \left( \frac{11}{n} \right) \sum_{i=1}^{3} \sum_{j=1}^{3} W_{ij} S_{ij}^2 \]

\[ + \left( \frac{1}{n} \right) \sum_{i=1}^{3} \sum_{j \neq j'} \left( B_{ij(j')} S_{ij}^2 + \frac{2}{n} \sum_{i \neq i'} \sum_{j \neq j'} \left( C_{ij(i', j')} S_{ij}^2 \right) \right) \]

\[ + \sum_{i=1}^{3} \sum_{j=1}^{3} W_{ij}^2 D_{ij} S_{ij}^2 \]
\[ a_3 : \operatorname{E} \left[ \sum_{i=1}^{3} \sum_{j=1}^{3} \left( \frac{1}{n_{ij}} \right) \left( \frac{n_{ij} + N_i}{n} \right) + \left( \frac{n_{ij} + N_j}{n} \right) \right] S_{ij}^2 = \left( \frac{11}{n} \right) \sum_{i=1}^{3} \sum_{j=1}^{3} W_{ij} S_{ij}^2 \]

\[ + \sum_{i=1}^{3} \sum_{j=1}^{3} W_i W_j D_{ij} S_{ij}^2 \left( \frac{1}{n} \right) \sum_{i \neq i' = 1}^{3} \sum_{j \neq j' = 1}^{3} C_{ij(i',j')} S_{ij}^2 \]

\[ + \left( \frac{1}{n} \right) \sum_{i=1}^{3} \sum_{j \neq j' = 1}^{3} \left( W_i + W_j \right) A_{ij(j')} S_{ij}^2 \]

To obtain results in \( a_1, a_2, a_3 \) theorems 2.1, 2.2 and 2.3 are used wherever required. With \( a_1, a_2, a_3, \alpha \) and other terms the resultant expression is

\[ \operatorname{E} \left[ \sum_{i=1}^{3} \sum_{j=1}^{3} \left( \frac{1}{n_{ij}} \right) W_{aij} S_{ij}^2 \right] \]

\[ = \left( \frac{11}{4} \right) \sum_{i=1}^{3} \sum_{j=1}^{3} W_{ij} S_{ij}^2 + \alpha^2 \left( \frac{1}{n} \right) \sum_{i=1}^{3} \sum_{j \neq j' = 1}^{3} B_{ij(j')} S_{ij}^2 \]

\[ + \left( \frac{2}{n} \right) \sum_{i=1}^{3} \sum_{j \neq j' = 1}^{3} C_{ij(i',j')} S_{ij}^2 + \sum_{i=1}^{3} \sum_{j=1}^{3} W_i^2 D_{ij} S_{ij}^2 \]

\[ + \left( \frac{2}{n} \right) \sum_{i=1}^{3} \sum_{j \neq j' = 1}^{3} W_i A_{ij(j')} S_{ij}^2 \left( \frac{1-\alpha^2}{4} \right) \sum_{i=1}^{3} \sum_{j \neq j' = 1}^{3} B_{ij(j')} S_{ij}^2 \]

\[ + \left( \frac{2}{n} \right) \sum_{i=1}^{3} \sum_{j \neq j' = 1}^{3} C_{ij(i',j')} S_{ij}^2 + \sum_{i=1}^{3} \sum_{j=1}^{3} W_j^2 D_{ij} S_{ij}^2 \]

\[ + \left( \frac{2}{n} \right) \sum_{i=1}^{3} \sum_{j \neq j' = 1}^{3} W_j A_{ij(j')} S_{ij}^2 \left( \frac{\alpha(1-\alpha)}{2} \right) \sum_{i \neq i' = 1}^{3} \sum_{j \neq j' = 1}^{3} C_{ij(i',j')} S_{ij}^2 \]

\[ + \sum_{i=1}^{3} \sum_{j=1}^{3} W_i W_j D_{ij} S_{ij}^2 + \left( \frac{1}{n} \right) \sum_{i=1}^{3} \sum_{j \neq j' = 1}^{3} \left( W_i + W_j \right) A_{ij(j')} S_{ij}^2 \]
We also have:

\[ a_4 : E \left[ \sum_{i=1}^{3} \sum_{j=1}^{3} \left( \frac{1}{N_{ij}} \right) \left( \left( \frac{n_i}{n} \right) + \left( \frac{N_i}{N} \right) \right) S_{ij}^2 \right] = \left( \frac{15}{N} \right) \sum_{i=1}^{3} \sum_{j=1}^{3} W_{ij} S_{ij}^2 \]

\[ + \sum_{i=1}^{3} \sum_{j=1}^{3} F_{ij} S_{ij}^2 \]

\[ a_5 : E \left[ \sum_{i=1}^{3} \sum_{j=1}^{3} \left( \frac{1}{N_{ij}} \right) \left( \left( \frac{n_j}{n} \right) + \left( \frac{N_j}{N} \right) \right) S_{ij}^2 \right] = \left( \frac{15}{N} \right) \sum_{i=1}^{3} \sum_{j=1}^{3} W_{ij} S_{ij}^2 \]

\[ + \sum_{i=1}^{3} \sum_{j=1}^{3} F_{ij} S_{ij}^2 \]

\[ a_6 : E \left[ \sum_{i=1}^{3} \sum_{j=1}^{3} \left( \frac{1}{N_{ij}} \right) \left( \left( \frac{n_i}{n} \right) + \left( \frac{N_i}{N} \right) \right) \left( \left( \frac{n_j}{n} \right) + \left( \frac{N_j}{N} \right) \right) S_{ij}^2 \right] = \left( \frac{15}{N} \right) \sum_{i=1}^{3} \sum_{j=1}^{3} W_{ij} S_{ij}^2 \]

\[ + \sum_{i=1}^{3} \sum_{j=1}^{3} F_{ij} S_{ij}^2 \]

Theorems 2.1, 2.2 and 2.3 are also used to derive \( a_4 \), \( a_5 \), and \( a_6 \) and, we get

\[ E \left[ \sum_{i=1}^{3} \sum_{j=1}^{3} \left( \frac{1}{N_{ij}} \right) W_{aij}^2 S_{ij}^2 \right] = \]

\[ \left( \frac{15}{4N} \right) \sum_{i=1}^{3} \sum_{j=1}^{3} W_{ij} S_{ij}^2 + \left( \frac{\alpha^2}{4} \right) \sum_{i=1}^{3} \sum_{j=1}^{3} F_{ij} S_{ij}^2 + \left( \frac{6}{N} \right) \sum_{i=1}^{3} \sum_{j=1}^{3} \frac{W_{ij} W_{ij}'}{W_{ij} S_{ij}^2} \]

\[ + \left( \frac{\alpha(1-\alpha)}{2} \right) \sum_{i=1}^{3} \sum_{j=1}^{3} F_{ij} S_{ij}^2 + \left( \frac{3}{N} \right) \sum_{i=1}^{3} \sum_{j=1}^{3} \frac{W_{ij} W_{ij}'}{S_{ij}^2} \]

\[ \tag{3.3.2} \]
\[
\begin{align*}
V\left(\frac{E(\bar{y}_{dps})}{n_{ij}}\right) &= V\left[\sum_{i=1}^{3} \sum_{j=1}^{3} W_{\alpha ij} \bar{Y}_{ij}\right] \\
&= \left(\frac{\alpha^2}{4}\right) \sum_{i=1}^{3} V(p_{i}) \bar{Y}_{i}^2 + \left(\frac{(1-\alpha)^2}{4}\right) \sum_{j=1}^{3} V(p_{j}) \bar{Y}_{j}^2 \\
&+ \left(\frac{\alpha(1-\alpha)}{2}\right) \sum_{i=1}^{3} \sum_{j=1}^{3} \text{Cov}(p_{i}, p_{j}) \bar{Y}_{i} \bar{Y}_{j} \\
&= \left(\frac{\alpha^2}{4}\right) R_2 + \left(\frac{(1-\alpha)^2}{4}\right) S_2 + \left\{\frac{\alpha(1-\alpha)}{2}\right\} T_2 \\
&= \left[\text{Bias}(\bar{y}_{dps})\right]^2 = \left[\alpha^2 V_1^2 + (1-\alpha)^2 V_2^2 + 2\alpha(1-\alpha)V_1 V_2\right]
\end{align*}
\]

Use of (3.3.1), (3.3.2), (3.4) and (3.5) provides the proof of theorem.

4. Optimum Choice

\[
\alpha_{opt} = \left[\frac{(S_1 + S_2 + 4V_2^2) - (T_1 + T_2 + 4V_1 V_2)}{(R_1 + R_2 + 4V_2^2) + (S_1 + S_2 + 4V_2^2) - 2(T_1 + T_2 + 4V_1 V_2)}\right]
\]

\[
\begin{align*}
M\left[\bar{y}_{dps, opt}\right] &= \left(\frac{1}{4}\right) \left[U_1 + U_2\right] \\
&+ \left\{\frac{(R_1 + R_2 + 4V_1^2)(S_1 + S_2 + 4V_2^2) - (T_1 + T_2 + 4V_1 V_2)^2}{(R_1 + R_2 + 4V_1^2) + (S_1 + S_2 + 4V_2^2) - 2(T_1 + T_2 + 4V_1 V_2)}\right\}
\end{align*}
\]

5. Efficiency Comparison

I. \(\bar{y}_{dps}\) is efficient over \(\bar{y}_{dps, 0}\) if \(R_1 + R_2 + 4V_1^2 \leq S_1 + S_2 + 4V_2^2\)

II. \(\bar{y}_{dps}\) is efficient over \(\bar{y}_{dps, 1/2}\) if
\[
(R_1 + R_2 + 4V_1^2) \leq \frac{1}{3} \left[ (S_1 + S_2 + 4V_2^2) + 2(T_1 + T_2 + 4V_1 V_2) \right]
\]

III. \( (\bar{y}_{dps})_0 \) is efficient over \( (\bar{y}_{dps})_{1/2} \) if

\[
(S_1 + S_2 + 4V_2^2) \leq \frac{1}{3} \left[ (R_1 + R_2 + 4V_1^2) + 2(T_1 + T_2 + 4V_1 V_2) \right]
\]

6. Numerical Illustrations

Consider two populations of size \( N=650 \) and \( N=490 \) and samples of size 260 and 196 by SRSWOR respectively and post-stratified according to \( 3 \times 3 \) classification.

Parameters of population are given in Table 6.1 and Table 6.2.

(a) For data set -I

\[
M\left[ (\bar{y}_{dps})_1 \right] = 132.5416 \quad B\left[ (\bar{y}_{dps})_1 \right] = 9.0127
\]

\[
M\left[ (\bar{y}_{dps})_0 \right] = 118.6221 \quad B\left[ (\bar{y}_{dps})_0 \right] = 8.8626
\]

\[
M\left[ (\bar{y}_{dps})_{1/2} \right] = 14.8312 \quad B\left[ (\bar{y}_{dps})_{1/2} \right] = 8.6274
\]

\[
M\left[ (\bar{y}_{dps})_{opt} \right] = 7.0314 \quad B\left[ (\bar{y}_{dps})_{opt} \right] = 8.1721
\]

with \( \alpha_{opt} = 0.4613 \)

For data set -II

\[
M\left[ (\bar{y}_{dps})_1 \right] = 98.1312 \quad B\left[ (\bar{y}_{dps})_1 \right] = 6.7285
\]

\[
M\left[ (\bar{y}_{dps})_0 \right] = 87.6234 \quad B\left[ (\bar{y}_{dps})_0 \right] = 6.5578
\]

\[
M\left[ (\bar{y}_{dps})_{1/2} \right] = 13.1394 \quad B\left[ (\bar{y}_{dps})_{1/2} \right] = 6.4432
\]

\[
M\left[ (\bar{y}_{dps})_{opt} \right] = 4.3122 \quad B\left[ (\bar{y}_{dps})_{opt} \right] = 5.7055
\]

with \( \alpha_{opt} = 0.4913 \)

(b) This is to recall that it is not possible to get estimate of \( \bar{Y} \) from usual sample mean estimator since \( N_j \)'s are assumed unknown.

(c) It seems that estimator \( (\bar{y}_{dps})_{1/2} \) is more efficient than \( (\bar{y}_{dps})_0 \) and \( (\bar{y}_{dps})_1 \)

171
Table 6.1 (for data set 1)

<table>
<thead>
<tr>
<th>B</th>
<th>A</th>
<th>Attribute A</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low</td>
<td>Medium</td>
<td>High</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>N_{11} = 71, n_{11} = 28</td>
<td>N_{12} = 65, n_{12} = 26</td>
<td>N_{13} = 68, n_{13} = 27</td>
</tr>
<tr>
<td></td>
<td>\bar{Y}_{11} = 48.7464</td>
<td>\bar{Y}_{12} = 147.6923</td>
<td>\bar{Y}_{13} = 247.4264</td>
</tr>
<tr>
<td></td>
<td>W_{11} = 0.10923</td>
<td>W_{12} = 0.1</td>
<td>W_{13} = 0.10461</td>
</tr>
<tr>
<td></td>
<td>S_{11}^2 = 857.187</td>
<td>S_{12}^2 = 791.2476</td>
<td>S_{13}^2 = 876.9092</td>
</tr>
<tr>
<td>Medium</td>
<td>N_{21} = 77, n_{21} = 31</td>
<td>N_{22} = 74, n_{22} = 30</td>
<td>N_{23} = 80, n_{23} = 32</td>
</tr>
<tr>
<td></td>
<td>\bar{Y}_{21} = 346.8831</td>
<td>\bar{Y}_{22} = 431.9054</td>
<td>\bar{Y}_{23} = 549.525</td>
</tr>
<tr>
<td></td>
<td>W_{21} = 0.11846</td>
<td>W_{22} = 0.11384</td>
<td>W_{23} = 0.12307</td>
</tr>
<tr>
<td></td>
<td>S_{21}^2 = 866.6208</td>
<td>S_{22}^2 = 964.5512</td>
<td>S_{23}^2 = 829.1359</td>
</tr>
<tr>
<td>High</td>
<td>N_{31} = 73, n_{31} = 29</td>
<td>N_{32} = 70, n_{32} = 28</td>
<td>N_{33} = 72, n_{33} = 29</td>
</tr>
<tr>
<td></td>
<td>\bar{Y}_{31} = 654.315</td>
<td>\bar{Y}_{32} = 737.957</td>
<td>\bar{Y}_{33} = 846.597</td>
</tr>
<tr>
<td></td>
<td>W_{31} = 0.1123</td>
<td>W_{32} = 0.10769</td>
<td>W_{33} = 0.11076</td>
</tr>
<tr>
<td></td>
<td>S_{31}^2 = 787.885</td>
<td>S_{32}^2 = 1044.759</td>
<td>S_{33}^2 = 780.469</td>
</tr>
<tr>
<td>Total</td>
<td>N_1 = 221, n_1 = 88</td>
<td>N_2 = 209, n_2 = 84</td>
<td>N_3 = 220, n_3 = 88</td>
</tr>
<tr>
<td></td>
<td>\bar{Y}_1 = 352.6515</td>
<td>\bar{Y}_2 = 446.01912</td>
<td>\bar{Y}_3 = 553.3727</td>
</tr>
<tr>
<td></td>
<td>W_1 = 0.3399</td>
<td>W_2 = 0.32153</td>
<td>W_3 = 0.3384</td>
</tr>
<tr>
<td>A</td>
<td>Attribute A</td>
<td></td>
<td>Total</td>
</tr>
<tr>
<td>-------------------</td>
<td>-------------</td>
<td>-----------------</td>
<td>------------------</td>
</tr>
<tr>
<td>B</td>
<td>Low</td>
<td>Medium</td>
<td>High</td>
</tr>
<tr>
<td></td>
<td>N_{i1} = 56, n_{i1} = 22</td>
<td>N_{i2} = 50, n_{i2} = 20</td>
<td>N_{i3} = 52, n_{i3} = 21</td>
</tr>
<tr>
<td></td>
<td>\bar{Y}_{i1} = 37.232</td>
<td>\bar{Y}_{i2} = 113.02</td>
<td>\bar{Y}_{i3} = 188.8846</td>
</tr>
<tr>
<td></td>
<td>W_{i1} = 0.1143</td>
<td>W_{i2} = 0.102</td>
<td>W_{i3} = 0.1061</td>
</tr>
<tr>
<td></td>
<td>S^2_{i1} = 504.1068</td>
<td>S^2_{i2} = 434.947</td>
<td>S^2_{i3} = 505.163</td>
</tr>
<tr>
<td></td>
<td>N_{i21} = 48, n_{i21} = 19</td>
<td>N_{i22} = 62, n_{i22} = 25</td>
<td>N_{i23} = 58, n_{i23} = 23</td>
</tr>
<tr>
<td></td>
<td>\bar{Y}_{i21} = 267.3333</td>
<td>\bar{Y}_{i22} = 321.258</td>
<td>\bar{Y}_{i23} = 413.776</td>
</tr>
<tr>
<td></td>
<td>W_{i21} = 0.09796</td>
<td>W_{i22} = 0.1265</td>
<td>W_{i23} = 0.11836</td>
</tr>
<tr>
<td></td>
<td>S^2_{i21} = 500.926</td>
<td>S^2_{i22} = 768.06</td>
<td>S^2_{i23} = 466.716</td>
</tr>
<tr>
<td></td>
<td>N_{i31} = 54, n_{i31} = 22</td>
<td>N_{i32} = 60, n_{i32} = 24</td>
<td>N_{i33} = 50, n_{i33} = 20</td>
</tr>
<tr>
<td></td>
<td>\bar{Y}_{i31} = 483.037</td>
<td>\bar{Y}_{i32} = 553.666</td>
<td>\bar{Y}_{i33} = 625.000</td>
</tr>
<tr>
<td></td>
<td>W_{i31} = 0.1102</td>
<td>W_{i32} = 0.12245</td>
<td>W_{i33} = 0.102</td>
</tr>
<tr>
<td></td>
<td>S^2_{i31} = 564.56</td>
<td>S^2_{i32} = 529.17</td>
<td>S^2_{i33} = 562.53</td>
</tr>
<tr>
<td></td>
<td>N_{i1} = 158, n_{i1} = 63</td>
<td>N_{i2} = 172, n_{i2} = 69</td>
<td>N_{i3} = 160, n_{i3} = 64</td>
</tr>
<tr>
<td></td>
<td>\bar{Y}_{i1} = 259.4999</td>
<td>\bar{Y}_{i2} = 341.796</td>
<td>\bar{Y}_{i3} = 406.694</td>
</tr>
<tr>
<td></td>
<td>W_{i1} = 0.32246</td>
<td>W_{i2} = 0.35095</td>
<td>W_{i3} = 0.32646</td>
</tr>
</tbody>
</table>
(d) The estimator \( \hat{\bar{Y}}_{dps} \) has made possible to estimate \( \bar{Y} \) in a 3x3 set-up even without the prior knowledge of \( N_i \) and frames. It has an effective utilization of row and column totals \( N_i \) and \( N_j \).

(e) The estimator is found most efficient at optimal selection of \( \alpha = 0.4613 \) for set-I and \( \alpha = 0.4913 \) for set-II.

(f) On the basis of data considered herein, one can think of choosing \( \alpha \) to a value near to 0.5 which reveals that almost a fifty percent fraction of row sum of size-proportions \[ \left( \frac{n_i}{n} \right) + \left( \frac{N_i}{N} \right) \] and rest fifty percent same from column generates an ideal, quick and easy choice of \( \alpha \). Thus, the proposed estimator provides an easy optimum choice \( \alpha = \frac{1}{2} \) or very close to it.

REFERENCES


A NEW ESTIMATION FOR POSTSTRATIFICATION SAMPLING SCHEME

D. SHUKLA & MANISH TRIVEDI
SCHOOL OF STATISTICS, D.A.V. UNIVERSITY, KHANOWA ROAD INDORE, MP, INDIA
DEPARTMENT OF MATHEMATICS AND STATISTICS
DHAN SINGH GARG UNIVERSITY
GAGAN (MP) INDIA

ABSTRACT

The post stratification is used in sampling surveys when mainly frames as per each strata are not available. The usual post stratified estimator incorporates stratum sizes as weights for combining \( k \) different means of stratum units in sample \( n \). In this paper an attempt has been made to use the “varying proportion “ of unsampled population units in combination with strata sizes as appropriate combining weights of sample means. A new estimator has been proposed and its properties are examined. It is found that the estimator is efficient than the existing estimator under optimal conditions. Moreover it is observed that the optimal choices are not hard to achieve in reality. Reasults are numerically supported in the form of numerical study.
A NEW ESTIMATION FOR ..... 1 + 1

1.0 INTRODUCTION:

When population contains high level heterogeneity among units of study variable, the survey practitioners are being advised to use appropriate techniques of stratification for efficient estimation. But, while random sample selection from strata, the prior knowledge of strata sizes, strata frames and a possible variability within stratum are all essential requirements. In practical situations, strata sizes are manageable but lists of stratum units are hard to get. Moreover stratum frames may be incomplete or overlapping. Several population units may fall under multiple strata while their classification. Under these circumstances it is advised to use post stratification techniques consisting of selecting a random sample from the entire population and classifying units later according to their representation from different stratum. This technique is useful, effective, and more close to real life situation [see Holt and Smith (1979), Jaqars et al (1986), and Agrawal & Pandya (1993)].

2.0 NOTATIONS:

Consider a finite population of \( N \) units divided into \( k \) strata, the
size of $i^{th}$ stratum is $N_i$ such that $\sum_{i=1}^{k} N_i = N$. It is presumed that a random sample of size $n$ is drawn from population $N$ using SRSWOR. In follows notations are as under:

$Y$ : Variable under study

$Y_{ij}$ : $j^{th}$ unit of the study variable in $i^{th}$ stratum of population $[i = 1, 2, \ldots k ; j = 1, 2, \ldots N_i]$

$y_{ij}$ : $j^{th}$ unit in sample of size $n$ coming from $i^{th}$ stratum $[i = 1, 2, \ldots k ; j = 1, 2, \ldots N_i]$

$n_i$ : representation of $i^{th}$ stratum unit in sample $n$.

$$\bar{y}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} y_{ij}$$

$S_i^2$: Population mean square for the $i^{th}$ stratum

$$= \frac{1}{(N_i - 1)} \sum_{j=1}^{N_i} (y_{ij} - \bar{y}_i)^2$$

$S^2$: Population Mean Square

$$= \frac{1}{(N - 1)} \sum_{j=1}^{k} \sum_{i=1}^{N_i} (Y_{ij} - \bar{Y})^2$$

$$\bar{Y} = \frac{1}{N} \sum_{i=1}^{k} \sum_{j=1}^{N_i} (Y_{ij} - \bar{Y})^2$$  Population mean of study variable

$$Y_i = \frac{1}{N_i} \sum_{j=1}^{N_i} y_{ij}$$  Stratum mean for $i^{th}$ stratum of the study variable
A New Estimation for .......

\[ W_i = N_i / N \] Population Proportion

\[ P_i = n_i / n \] Sample Proportion

\[ f = n / N \] Sample Fraction

3.0 Proposed Estimator in Poststratification:

The sample takes values \( n_i \), such that \( n = \sum_{i=1}^{k} n_i \) and each \( n_i \) is random for prefixed "n". Under assumption that probability of \( n_i \) being zero is very small, the usual unbiased estimator of population mean \( \bar{Y} \) under post stratification is

\[
\bar{Y}_{ps} = \sum_{i=1}^{k} W_i \bar{y}_i \quad \quad (3.1)
\]

where 'ps' stands for post-stratification. Here \( W_i \)'s are weights used to pool up different means in sample \( n \). Agrawal and Pandya(1993) have suggested to utilize the proportion \( p_i = (n_i / n) \) in addition to \( W_i \) in order to design estimator.

Let us define \( p_{ia} = \frac{N_i - an_i}{N - an} \), where \( a \) is constant laying between \(-\infty < a < \frac{N}{n}\). Clearly, the \( p_{ia} \) is varying proportion over constant \( a \), and when \( a = 1 \), \( p_{ia} = \frac{N_i - an_i}{N - n} \) which is prefix proportion of the \( i \)th unsampled stratum units over the remaining part of the entire population. Moreover it is obvious that \( p_i \), \( p_{ia} \) are unbiased estimators of \( W_i \).

Using the "varying proportion" \( p_{ia} \), we propose the estimator

\[
\bar{Y}_{psm} = \sum_{i=1}^{k} p_{ia} \bar{y}_i
\]
where
\[ W_{ia} = \left[ \frac{n}{N} P_{ia} + \frac{N-n}{N} W_i \right] \]

The justification for designing the estimator in above way is mainly to use "varying proportion" instead of "prefix proportion" as suggested by earlier authors. The "varying proportion", for suitable choice of constant, may give an appropriate value of unsampled units of population. In other way one can use best proportion required for making the estimator efficient.

At \( a = 0 \), \( P_{ia} = W_i \) and \( \bar{y}_{psm} = \bar{y}_{ps} \)

**THEOREM 1.0**

The estimator \( \bar{y}_{psm} \) is unbiased for \( \bar{y} \) with variance in terms of \( 0(n^2) \) given by

\[
V(\bar{y}_{psm}) = \left( \frac{1}{n} - \frac{1}{N} \right)^k \sum_{i=1}^{k} W_i S_i^2 + \frac{(N-n)(1+r)^2}{(N-1)n^2} \sum_{i=1}^{k} (1-W_i) S_i^2 + \frac{(N-n)}{N.n} f^2 (S^2 - \sum_{i=1}^{k} W_i S_i^2)
\]

where \( f = \left\{ \frac{an}{N-an} \right\} \)

**PROOF:**

To evaluate the expected value of \( \bar{y}_{psm} \) we first note that

\[
E(\bar{y}_{ia}) = E\{ E(W_{ia} / n_i) \} = E\left\{ \frac{N_i - an_i}{N - an} + (1-f) \frac{N_i}{N} \right\} / n_i
\]
A new estimation for .....  

\[ E\left[ \left( 1 + \frac{anf}{N-an} \right) \frac{N_i}{n} - \left( \frac{anf}{N-an} \right) \frac{n_i}{n} \right] = \frac{1}{n} \sum_{i=1}^{k} E(W_i) \bar{Y}_i = \frac{1}{n} \sum_{i=1}^{k} W_i \bar{Y}_i = \bar{Y} \]

Hence \( \bar{Y}_{psm} \) is the unbiased estimator of \( \bar{Y} \).

Looking to variability of the estimator, one can start with

\[ V(\bar{Y}_{psm}) = E\left[ V(\bar{Y}_{psm}) \right] + V\left[ E(\bar{Y}_{psm}) \right] \]

(3.4)

To evaluate the above quantity, we assume that \( n_i > 0 \) for all \( i \), and following standard results are used.

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

(3.4.1)

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

(3.4.2)

\[ \gamma\left( \frac{n_i}{n} \right) = \frac{N-n}{(N-1)} \frac{W_i(1-W_i)}{n} \]

(3.4.3)

\[ \text{Cov}\left( \frac{n_i}{n}, \frac{n_j}{n} \right) = \frac{N-n}{N-1} \frac{\bar{Y}_i \bar{Y}_j}{n} \quad \text{when } i \neq j \]

(3.4.4)
and

\[ S^2 = \frac{1}{(N-1)} \left[ \sum_{i=1}^{k} (N_i - 1)S_i^2 + \sum_{i=1}^{k} N_i (\bar{Y}_i - \bar{Y})^2 \right] \quad (3.4.5) \]

Now expanding the term on r. h. s. of expression 3.4 we have

\[ E\left[ \mathcal{V}(\bar{y}_{psm}/n_i) \right] = E\left[ \left( \sum_{i=1}^{k} W_{ia} \bar{y}_i \right)^2 \right] \]

\[ = E\left[ \sum_{i=1}^{k} \left( \frac{1}{n_i} - \frac{1}{N_i} \right) W_{ia}^2 S_i^2 \right] \]

\[ = E\left[ \sum_{i=1}^{k} \left( \frac{1}{n_i} - \frac{1}{N_i} \right) \left( \frac{an}{N - an} \frac{N_i}{N} - \left( \frac{an}{N - an} \right) \frac{n_i}{n} \right)^2 S_i^2 \right] \]

\[ = E\left[ \sum_{i=1}^{k} \left( \frac{1}{n_i} - \frac{1}{N_i} \right) \left( (1 + r) \frac{N_i}{N} - r \frac{n_i}{n} \right)^2 S_i^2 \right] \]

Continuing the above final expression, after some algebraic manipulation will be:

\[ E\left[ \mathcal{V}(\bar{y}_{psm}/n_i) \right] \left[ \left( \frac{1}{n} - \frac{1}{N} \right) \sum_{i=1}^{k} W_{i}S_i^2 + \frac{(N-n)}{(N-1)^2} (1 + r)^2 \sum_{i=1}^{k} (1 - W_i)S_i^2 \right. \]

\[ - \frac{(N-n)r^2}{N(N-1)n} \sum_{i=1}^{k} (1 - W_i)S_i^2 \] \quad \ldots \ldots \quad (3.4.6) \]

Further, considering the second component on r.h.s. of expression 3.4.
A NEW ESTIMATION FOR .......

\[
E \left[ \sum_{i=1}^{k} \left( \bar{Y}_{i} \cdot \left( \frac{\bar{Y}_{psm}}{n_{i}} \right) \right) \right] = \sum_{i=1}^{k} \left( \bar{Y}_{i} \cdot \left( \frac{\bar{Y}_{psm}}{n_{i}} \right) \right) \\
= \sum_{i=1}^{k} V(W_{ia}) \bar{Y}_{i}^{2} + \sum_{i=1}^{k} \sum_{j=1}^{N_{i}} \text{cov}(W_{ia}, W_{ja}) \bar{Y}_{i} \bar{Y}_{j}
\]

On continuing, the final expression after some algebraic manipulation;

\[
E \left[ \sum_{i=1}^{k} \left( \bar{Y}_{psm} \cdot \frac{1}{n_{i}} \right) \right] = \frac{\mu^{2}(N-n)}{(N-1)n} \left[ \sum_{i=1}^{k} W_{i}(1-W_{i}) \bar{Y}_{i}^{2} - \sum_{i=1}^{k} \sum_{j=1}^{N_{i}} W_{i}W_{j} \bar{Y}_{i} \bar{Y}_{j} \right]
\]

(3.4.7)

**THEOREM 2.0:**

For the estimator

\[
\bar{Y}_{psm} = \sum_{i=1}^{k} W_{ia} \bar{Y}_{i}
\]

the minimum variance is:

\[
\gamma_{\text{min}}(\bar{Y}_{psm}) = \left( \frac{1}{n} - \frac{1}{N} \right) \sum_{i=1}^{k} W_{i} S_{i}^{2} + \left[ 1 - \frac{f^{2} a_{\text{opt}}}{f a_{\text{opt}} - 1} \right] \frac{N-n}{(N-1)f} \sum_{i=1}^{k} (1-W_{i})^{2} S_{i}^{2}
\]

(3.5.1)

where \( a_{\text{opt}} \) denotes the optimal value of "a" given by:

\[
a_{\text{opt}} = \frac{M_{2} / f}{(1-f)M_{2} - fM_{1}}
\]

(3.5.2)

with \( M_{1} = \frac{1}{N} (S^{2} - \sum_{i=1}^{k} W_{i} S_{i}^{2}) \)

182
\[ M_2 = \frac{1}{(N - 1)n} \sum_{i=1}^{k} (1 - W_i) S_i^2 \]

**PROOF:**

The optimal value of "a", that minimises the variance expression given on (3.5.2) is obtained by differentiating the said expression (3.4.8) with respect to "a" and equating it to zero. It results as

\[
\frac{(N - n)}{(N - 1)n^2} \sum_{i=1}^{k} (1 - W_i) S_i^2 \left[ \frac{\delta}{\delta a} \left\{ 1 + \frac{anf}{N - an} \right\} \right]^2 + \frac{N - n}{N.n} \left( S^2 - \sum_{i=1}^{k} W_i S_i^2 \right) \left[ \frac{\delta}{\delta a} \left\{ \frac{anf}{N - an} \right\} \right]^2 \frac{\delta^2}{\delta a^2} = 0
\]

and after some algebraic manipulation we have

\[
a_{\text{opt}} = \frac{M_2 / f}{(1 - f)M_2 - fM_1}
\]

with \(M_1\) and \(M_2\) as stated above.

Now substituting "a_{opt}" in place of "a" in variance expression of optimal minimum variance is:

\[
V_{\text{min}}(y_{pm}) = \left( \frac{1}{N} - \frac{1}{n} \right) \sum_{i=1}^{k} W_i S_i^2 + \frac{(N - n)}{(N - 1)n^2} \left( 1 + \frac{nfa_{\text{opt}}}{N - na_{\text{opt}}} \right)^2 \sum_{i=1}^{k} (1 - W_i) S_i^2
\]

\[
\frac{(N - n)}{N.n} \left[ \frac{nfa_{\text{opt}}}{N - na_{\text{opt}}} \right]^2 \left( S^2 - \sum_{i=1}^{k} W_i S_i^2 \right)
\]

which will further reduce into
A new estimation for $\ldots$ 5 + 1

$$V_{\text{min}}(\bar{Y}_{psm}) = \left(1 - \frac{1}{N}\right)^k \sum_{i=1}^k W_i S_i^2 + \left[1 - \frac{a_{opt}^2}{f_{opt}^2 - 1}\right] \frac{(N - n)}{(N - 1)n^2} \sum_{i=1}^k (1 - W_i) S_i^2$$

4.0 COMPARISON OF $\bar{Y}_{psm}$ WITH $\bar{Y}_{ps}$ AND $\bar{Y}$

The next interest lies around to examine the performance of $\bar{Y}_{psm}$ with customary estimator $\bar{Y}_{ps}$ and $\bar{Y}$ which is the usual sample mean. The variance of $\bar{Y}_{ps}$ to terms of $O(n^{-2})$ is known to be

$$V(\bar{Y}_{ps}) = \left(\frac{1}{n} - \frac{1}{N}\right)^k \sum_{i=1}^k W_i S_i^2 + \frac{(N - n)}{(N - 1)n^2} \sum_{i=1}^k (1 - W_i) S_i^2 \tag{4.1}$$

which as a special case is also obtainable from (3.4.8) by setting 'a' = 0.

Also, the variance of $\bar{Y}$ is

$$V(\bar{Y}) = \left(\frac{1}{n} - \frac{1}{N}\right) S^2 \tag{4.2}$$

Using (3.5.1), (4.1) and (4.2) we can express

$$V_{\text{min}}(\bar{Y}_{psm}) = V(\bar{Y}_{ps}) - T \frac{M_2^2}{M_1 + M_2}$$

and

$$V_{\text{min}}(\bar{Y}_{psm}) = V(\bar{Y}) - T \frac{M_1^2}{M_1 + M_2}$$

where $T = T \frac{N - n}{n}$ which is always a positive quantity along

with the ratio $\frac{M_2^2}{M_1 + M_2}$. Therefore there will be definite gain over

184
6.0 CONCLUSION:

On recaptualisation it is found that

i $\bar{y}_{psm}$ is an unbiased estimator of $\bar{Y}$

ii $\bar{y}_{psm}$ contains the idea of "Variance Proportion" in its design.

iii It is optimum for suitable choice of constant

iv Selection of "a" is easy by prior gusseing or by pilot surveys

v Optimum $\bar{y}_{psm}$ observed efficient than usual post stratified estimator.

7.0 REFERENCES


\( y_{\text{ps}} \) and \( \bar{y} \). It is interesting to watch that the term \( f > 1 \) when \( n < N/2 \). So the proposed estimator will perform better over smaller sample size under the optimum selection of the value of the constant.

4.1 About selection of \( a_{\text{opt}} \)

We have obtained the optimum value of \( 'a' \)

which could be rearranged as,

\[
a_{\text{opt}} = \frac{1}{f(f(1-f) - f(M_1/M_2))}
\]

Here \( M_1/M_2 \) is ratio of two variance terms which may not fluctuate much in successive surveys. It is easy to observe, if the prior knowledge of the ratio \( (M_1/M_2) \) is available or gauged by a pilot survey then the optimum value of \( 'a' \) could be easily obtained. Thus, this ratio is not hard to achieve close to accuracy and so the optimal \( 'a' \).

5.0 Numerical Illustrations

Results obtained above are numerically supported over following data sets

Set 1: [From Singh and Choudhary (1984), PP. 76]

<table>
<thead>
<tr>
<th>Stratum</th>
<th>( N_i )</th>
<th>( \bar{Y}_i )</th>
<th>( S_i^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>22</td>
<td>0.0000</td>
<td>0.000000</td>
</tr>
<tr>
<td>2</td>
<td>25</td>
<td>0.2800</td>
<td>0.001960</td>
</tr>
<tr>
<td>3</td>
<td>18</td>
<td>0.1944</td>
<td>0.011233</td>
</tr>
<tr>
<td>4</td>
<td>26</td>
<td>1.4846</td>
<td>0.151532</td>
</tr>
<tr>
<td>5</td>
<td>20</td>
<td>3.8000</td>
<td>0.0755115</td>
</tr>
<tr>
<td>6</td>
<td>13</td>
<td>6.1544</td>
<td>1.947474</td>
</tr>
</tbody>
</table>

The population parameters are

\( \Pi, \bar{Y}, 16.266, \bar{Y}^2 = 304.82567, n = 30. \)
A new estimation for the comparison of efficiency, we use

\[ \text{Gain in efficiency} = \frac{V(\bar{y}) - V(.)}{V(.)} \times 100 \]

\[ \text{TABLE 1} \]

<table>
<thead>
<tr>
<th>Estimator</th>
<th>Gain in Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\bar{y}_{psm})</td>
<td>86.7 %</td>
</tr>
<tr>
<td>Optimum (\bar{y}_{psm})</td>
<td>88.7 %</td>
</tr>
</tbody>
</table>

Set 2: [From Cochran (1977), P.173]

<table>
<thead>
<tr>
<th>STRATUM</th>
<th>(N_i)</th>
<th>(\bar{y}_i)</th>
<th>(S_i^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>047</td>
<td>8699</td>
<td>69.48</td>
</tr>
<tr>
<td>2</td>
<td>118</td>
<td>4614</td>
<td>43.64</td>
</tr>
<tr>
<td>3</td>
<td>091</td>
<td>7411</td>
<td>66.39</td>
</tr>
</tbody>
</table>

With population parameters:

\[ N = 254, \bar{y} = 56.47, S^2 = 6444.38, n = 100 \]

\[ \text{TABLE 2} \]

<table>
<thead>
<tr>
<th>Estimator</th>
<th>Gain in Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\bar{y}_{psm})</td>
<td>1.52 %</td>
</tr>
<tr>
<td>Optimum (\bar{y}_{psm})</td>
<td>1.02 %</td>
</tr>
</tbody>
</table>
Appendices C
Mean estimation in $2 \times 2$ classified dichotomous population under poststratification

by D. Shukla & Manish Trivedi, Department of Mathematics and Statistics,
Dr. Hari Singh Gour University, Sagar-470003
(Received May 27, 2000)

Abstract

This paper deals with the problem of mean estimation for $2 \times 2$ classified dichotomous population and an estimator is constructed for this purpose. An approximate expression of m.s.e. is derived.

1.0 Notations

Let $Y_{ijk}$ be the $k^{th}$ value of a variable $Y$ under study in a $(i, j)^{th}$ class ($i, j = 1, 2$) of a $2 \times 2$ classified dichotomous population of size $N$ with respect to attributes, say $A$ and $B$ having only two options, possessing or not possessing to trait. Further, let $N_{ij}$ be the population size of $(i, j)^{th}$ strata with population mean $\bar{Y}_{ij}$ and population mean square $S^2_{ij}$. Symbols, $\bar{Y}$ and $S^2$ are as usual and $N_i$ and $N_j$ are row and column size totals respectively. Let a sample of size $n$ is drawn from $N$ by SRSWOR and post-stratified such that $n_{ij}$ units represent to $N_{ij}$ with sample means $\bar{y}_{ij}$, and $\bar{y}$ be mean on $n$.

2.0 An estimation strategy

We suppose the frame of $N$ units of population is available but when the same population is $2 \times 2$ classified, the sizes $N_{ij}$ and list of units of these classes are not known. Moreover row-column totals $N_i$ and $N_j$ are known.

To estimate $\bar{Y}$, proposed, "Classified Poststratified" estimator with constant $\alpha$ is

$$\bar{Y}_{cps} = \frac{2}{\Sigma_{i=1}^{2} \Sigma_{j=1}^{2} W_{aij}}$$

$$W_{aij} = \left\{ \frac{a}{2} \left\{ \frac{n_i}{n} + \frac{N_i}{N} \right\} + \left( \frac{1-a}{2} \right) \left\{ \frac{n_j}{n} + \frac{N_j}{N} \right\} \right\}$$

(2.1)

The logic of $W_{aij}$ is supported by Agrawal and Panda (1993).
2.1 Approximation used

Let \( P_{ij} = N_{ij} / N, \) \( P_{ij} = n_{ij} / n \) and assume \( n \) large enough, approximations are:

\[
\rho_{ij} = P_{ij} (1 + \epsilon_{ij}), \quad \rho_i' = P_{i'} (1 + \epsilon_i'), \quad \rho_{ij}' = P_{ij} (1 + \epsilon_{ij}')
\]

where, \( E[\epsilon_{ij}] = E[\epsilon_i'] = E[\epsilon_{ij}'] = 0 \)

**Theorem 2.1:**

Using (2.2) and avoiding terms of higher order, approximate result, for \( i \neq i', j \neq j' \) are

\[
(a) \quad E \left[ \frac{\rho_{ij}'}{\rho_{ij}} \right] = \frac{P_{ij}'}{P_{ij}} \left[ 1 + \frac{\operatorname{Var}(\rho_{ij})}{p_{ij}^2} - \frac{\operatorname{Cov}(\rho_{ij}, \rho_{ij}')}{P_{ij} P_{ij}'} \right]
\]

\[
(b) \quad E \left[ \frac{\rho_{ij}'}{\rho_{ij}} \right] = \frac{P_{ij}'}{P_{ij}} \left[ 1 + \frac{\operatorname{Var}(\rho_{ij})}{p_{ij}^2} + \frac{\operatorname{Var}(\rho_{ij}')}{{p_{ij}'}^2} - \frac{2\operatorname{Cov}(\rho_{ij}, \rho_{ij}')}{{P_{ij} P_{ij}'}} \right]
\]

\[
(c) \quad E \left[ \frac{\rho_{ij} \rho_{ij}'}{\rho_{ij} P_{ij}'} \right] = \frac{P_{ij} P_{ij}'}{P_{ij}} \left[ 1 + \frac{\operatorname{Var}(\rho_{ij})}{p_{ij}^2} + \frac{\operatorname{Cov}(\rho_i', \rho_{ij}')}{{P_{ij}'} P_{ij}'} - \frac{\operatorname{Cov}(\rho_{ij}, \rho_{ij}')}{{P_{ij} P_{ij}'}} - \frac{\operatorname{Cov}(\rho_{ij}, \rho_{ij}')}{{P_{ij} P_{ij}'}} \right]
\]

3.0 Bias and mean square error

**Theorem 3.1:** The estimator \( \tilde{y}_{\text{cps}} \) is biased for \( \bar{y} \).

**Proof:** Let \( E[\{\cdot\} / n_{ij}] \) denotes conditional expectation given \( n_{ij} \).

\[
E(\tilde{y}_{\text{cps}}) = E[E[\{\tilde{y}_{\text{cps}}\} / n_{ij}] = \frac{2}{\Sigma_{i=1}^{2} \Sigma_{j=1}^{2}} \left[ \alpha \left(N_{i\cdot} / N \right) + (1 - \alpha) \left(N_{j\cdot} / N \right) \right] \tilde{y}_{ij}
\]

\[
= [\bar{y} + \alpha V_1 + (1 - \alpha) V_2] = \bar{y} + \text{Bias}(\tilde{y}_{\text{cps}})
\]

where \( V_1 = \Sigma_{i=1}^{2} \Sigma_{j=1}^{2} W_{ij} \tilde{y}_{ij} \), \( V_2 = \Sigma_{i=1}^{2} \Sigma_{j=1}^{2} W_{ij} \tilde{y}_{ij}' \)

### 3.1 Some symbols

\[
A_{ij}(j') = E \left[ \frac{\rho_{ij}'}{\rho_{ij}} \right]; \quad B_{ij}(j') = E \left[ \frac{\rho_{ij}'}{\rho_{ij}} \right]; \quad C_{ij}(i'j') = E \left[ \frac{\rho_{ij} \rho_{i'j'}}{\rho_{ij}'} \right]
\]

\[
M_i = \Sigma_{j=1}^{2} \tilde{y}_{ij}; \quad M_j = \Sigma_{i=1}^{2} \tilde{y}_{ij} \quad W
\]

\[
W_{ij} = P_{ij} = (N_{ij} / N) \quad W_i = \Sigma_{j=1}^{2} W_{ij}; \quad W_j = \Sigma_{i=1}^{2} W_{ij}
\]
\[ F_i = \frac{1}{N_{ij}} \left\{ \frac{(N-n)}{(N-1)} \frac{W_i(1-W_{i,j})}{n} + W_i^2 \right\} ; F_{ij} = \frac{1}{N_{ij}} \left\{ \frac{(N-n)}{(N-1)} \frac{W_j(1-W_{i,j})}{n} + W_j^2 \right\} \]

\[ F_{ij} = E \left[ \frac{\rho_{i,j}^2}{N_{ij}} \right] = (1/N_{ij}) \{ \text{Cov}(\rho_{i,j}) + E(\rho_{i,j})E(\rho_{j}) \}; \]

\[ D_{ij} = E \left[ \frac{1}{n_{ij}} \right] = \left[ \frac{1}{nW_{ij}} + \frac{(N-n)(1-W_{ij})}{(N-1)n^2W_{ij}^2} \right] \]

**Theorem 3.2**: The mean square error of \( \hat{y}_{\text{cps}} \) is

\[
\text{MSE} \left( \hat{y}_{\text{cps}} \right) = \frac{1}{4} \left[ U_1 + a^2 (R_1 + R_2 + 4V_1^2) + (1-a)^2 (S_1 + S_2 + 4V_2^2) \right] + 2a(1-a) \left( T_1 + T_2 + 4V_1 V_2 \right)
\]

Where, \( U_1 = \frac{2}{7/n} - \frac{9}{N/N} \), \( \Sigma \)

\[ \begin{align*}
R_1 &= \sum_{i=1}^{2} \sum_{j=1}^{2} \frac{(1/n)}{B_{ij}(j')} S_{ij}^2 + \sum_{i=1}^{2} \sum_{j=1}^{2} \frac{W_i^2}{D_{ij} S_{ij}^2} \\
R_2 &= \left( \frac{(N-n)}{(N-1)n} \right) \left[ \frac{2}{i=1} W_i(1-W_{i,j}) M_i^2 - \frac{2}{i=1} \sum_{i=1}^{2} M_i M_i' W_i W_i' \right] \\
S_1 &= \sum_{i=1}^{2} \sum_{j=1}^{2} \frac{(1/n)}{B_{ij}(j')} S_{ij}^2 + \sum_{i=1}^{2} \sum_{j=1}^{2} \frac{W_j^2}{D_{ij} S_{ij}^2} \\
S_2 &= \left( \frac{(N-n)}{(N-1)n} \right) \left[ \frac{2}{i=1} W_j(1-W_{i,j}) M_j^2 - \frac{2}{j=1} \sum_{j=1}^{2} M_j M_j' W_j W_j' \right] \\
T_1 &= \sum_{i=1}^{2} \sum_{j=1}^{2} \frac{(1/n)}{C_{ij}(i')} S_{ij}^2 + \sum_{i=1}^{2} \sum_{j=1}^{2} \frac{W_i}{D_{ij} S_{ij}^2} \\
T_2 &= \left( \frac{(N-n)}{(N-1)n} \right) \left[ \frac{2}{j=1} W_j(1-W_{i,j}) M_j^2 - \frac{2}{j=1} \sum_{j=1}^{2} M_j M_j' W_j W_j' \right] 
\end{align*} \]
\begin{align*}
&\frac{1}{(n)} \sum_{i=1}^{2} \sum_{j \neq j'}^{2} \left( W_{i} + W_{j} \right) A_{ij} \quad S_{ij}^2 \\
&- \frac{2}{\sum_{i=1}^{2}} \sum_{j=1}^{2} F_{ij} S_{ij}^2 \quad \left( 3/N \right) \quad \sum_{i=1}^{2} \sum_{j \neq j'}^{2} \left\{ \left( W_{i} W_{j} \right) S_{ij}^2 \right\} / W_{ij} \\
T_2 &= -\left\{ (N-n) / (N-1) \right\} \n \sum_{i=1}^{2} W_i, W_j, M_i, M_j
\end{align*}

Proof: MSE \( \bar{y}_{\text{cps}} \) = \( E[V(\bar{y}_{\text{cps}}) / \mu_i] + V[E(\bar{y}_{\text{cps}}) / \mu_i] + [\text{Bias}(\bar{y}_{\text{cps}})]^2 \)

Where \( V(\cdot)/\mu_i \) is conditional variance for given \( \mu_i \). Now, picking first component

\begin{align*}
E[V(\bar{y}_{\text{cps}}) / \mu_i] &= E\left[ \sum_{i=1}^{2} \sum_{j=1}^{2} \left( 1/\mu_i \right) W_{aij}^2 S_{ij}^2 \right] - E\left[ \sum_{i=1}^{2} \sum_{j=1}^{2} \left( 1/\mu_i \right) W_{aij}^2 S_{ij}^2 \right] \\
&= \left( \frac{7}{4n} - \frac{9}{4N} \right) \sum_{i=1}^{2} \sum_{j=1}^{2} W_{ij} S_{ij}^2 + (\alpha^2/4) \left[ \left( \frac{1}{n} \right) \sum_{i=1}^{2} \sum_{j \neq j'}^{2} B_{ij} \right] \\
&+ \sum_{i=1}^{2} \sum_{j=1}^{2} W_{i}^2 D_{ij} S_{ij}^2 - \sum_{i=1}^{2} \sum_{j=1}^{2} F_{ij} S_{ij}^2 + (2n) \sum_{i=1}^{2} \sum_{j \neq j'}^{2} W_{i} A_{ij(j')} S_{ij}^2 \\
&- \sum_{i=1}^{2} \sum_{j \neq j'}^{2} \left( W_{i} W_{j} \right) S_{ij}^2 \right] \right\} + \left( 1-\alpha \right)^2 / 4 \left[ \left( \frac{1}{n} \right) \sum_{i=1}^{2} \sum_{j \neq j'}^{2} B_{ij(j')} S_{ij}^2 \\
&+ \sum_{i=1}^{2} \sum_{j=1}^{2} W_{i}^2 D_{ij} S_{ij}^2 + (2n) \sum_{i=1}^{2} \sum_{j \neq j'}^{2} W_{i} A_{ij(j')} S_{ij}^2 - \sum_{i=1}^{2} \sum_{j=1}^{2} F_{ij} S_{ij}^2 \\
&- \sum_{i=1}^{2} \sum_{j \neq j'}^{2} \left( W_{i} W_{j} \right) S_{ij}^2 \right] \right\} + \left( 1-\alpha \right)/ 2 \left[ \left( \frac{1}{n} \right) \sum_{i \neq i'}^{2} \sum_{j \neq j'}^{2} C_{ij(i') S_{ij}^2} \right] \\
&+ \sum_{i=1}^{2} \sum_{j=1}^{2} W_{i} W_{j} D_{ij} S_{ij}^2 + (1/n) \sum_{i=1}^{2} \sum_{j \neq j'}^{2} \left( W_{i} + W_{j} \right) A_{ij(j')} S_{ij}^2 \\
&- \sum_{i=1}^{2} \sum_{j \neq j'}^{2} F_{ij} S_{ij}^2 \right\} + \left( 3/N \right) \sum_{i \neq i'}^{2} \sum_{j \neq j'}^{2} \left\{ \left( W_{i} W_{j} \right) S_{ij}^2 \right\} \\
&= (1/4) U_1 + (\alpha^2/4) R_1 + \left( 1-\alpha \right)^2 S_1 + \left( \alpha(1-\alpha)/2 \right) T_1
\end{align*}

\begin{align*}
V(E(\bar{y}_{\text{cps}}) / \mu_i) &= V\left[ \sum_{i=1}^{2} \sum_{j=1}^{2} W_{aij} T_{ij} \right] - (\alpha^2/4) \left[ \sum_{i=1}^{2} M_i^2 V(\rho_{i}) \right] \\
&+ \sum_{i=1}^{2} \sum_{i \neq i'}^{2} M_i M_{i'} \text{Cov}(\rho_{i}, \rho_{i'}) \right\} + \left( 1-\alpha \right)^2 / 4 \left[ \sum_{j=1}^{2} M_j^2 V(\rho_{j}) \right]
\end{align*}
\[ + \sum_{j=1}^{2} \sum_{j' \neq j} M_{ij} M_{ij'} \text{Cov}(\varepsilon_{ij}, \varepsilon_{ij'}) + \left( (a-1)/2 \right) \sum_{i=1}^{2} M_{ij} \text{Cov}(\varepsilon_{ij}, \varepsilon_{ij'}) \]

\[ = (a^2/4) R + (1-a^2/4) S + (a-1)/2 T \]

And, \[ \text{Bias}(\bar{y}_{\text{cns}}) \] = \[ \left[ a^2 V_1^2 + (1-a)^2 V_2^2 + 2a(1-a) V_1 V_2 \right] \]

On adding all (3.1), (3.2) and (3.3), proof of the theorem holds.

\section*{4.0 Optimum estimator}

On differentiating \( \text{MSE}(\bar{y}_{\text{cns}}) \) with respect to \( a \) and equating to zero, one can easily obtain

\[ \alpha_{\text{opt}} = \left[ \frac{(S_1+S_2+4V_2^2) - (T_1+T_2+4V_1 V_2)}{(R_1+R_2+4V_1^2) + (S_1+S_2+4V_2^2) - 2(T_1+T_2+4V_1 V_2)} \right] \]

Substituting \( \alpha_{\text{opt}} \), an optimal estimator of \( \bar{Y} \) is \( \left( \bar{y}_{\text{cns}} \right)_{\text{opt}} \) with optimum m.s.e.

\[ \text{MSE} \left( \left( \bar{y}_{\text{cns}} \right)_{\text{opt}} \right) = (1/4) \left[ U_1 + \left\{ \frac{((R_1+R_2+V_1^2)(S_1+S_2+V_2^2) - (T_1+T_2+4V_1 V_2)^2)}{((R_1+R_2+V_1^2)(S_1+S_2+V_2^2) - 2(T_1+T_2+4V_1 V_2))^2} \right\} \right] \]

\section*{References}


Appendices D
'This is to certify that Prof./Dr./Shri/Ku./Smt. Manish Trivedi took part in the deliberations and presented a paper on Efficient Estimation in Post-stratification Using Prior Information in the seminar.

K.K.N. Sharma
Organising Secretary
XV M.P. YOUNG SCIENTIST CONGRESS
Sponsored By: M.P. Council of Science & Technology, Bhopal.
Organised By: Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal

CERTIFICATE

Dr./Mr./Ms. Manish Trivedi
of H.S. Gour Univ., Sagar, presented his/her research paper in
Maths discipline at the XV M.P. Young Scientist Congress
Organised by Rajiv Gandhi Proudyogiki Vishwavidyalaya (University of Technology

Prof. U.S. Vijaywardiya
Registrar
Rajeev Gandhi Proudyogiki Vishwavidyalaya, Bhopal.

Dr. S.N. Dwivedi
Director General
M.P. Council of Science and Technology, Bhopal.

Prof. P.B. Singh
Vice Chancellor
Rajeev Gandhi Proudyogiki Vishwavidyalaya, Bhopal.
INdIAN SCIENCIe CONGRESS
88th Session
3-7, January 2001

Sh. Manish Trivedi

participated in the 88th Session of the Indian Science Congress held at
Indian Agricultural Research Institute, New Delhi. Presented a paper
for Young Scientist Award.

Dr. Anupam Verma
Organising Secretary
THE XX ANNUAL CONFERENCE OF INDIAN SOCIETY
FOR PROBABILITY AND STATISTICS
(ISPS)
AND
NATIONAL SEMINAR ON
INDUSTRIES ORIENTED STATISTICAL RESEARCH
AND
ANNUAL CONFERENCE OF BAYESIAN SOCIETY OF INDIA
(FEBRUARY 19-21, 2001)

Certificate

This is to certify that Prof./Dr./Sri/Ms. MANISH TRIVEDI

of DEPT. OF MATHEMATICS & STATISTICS, Sagar,
Participated/Presented Paper/Delivered Invited Talk/Chaired the Session in the
Conference and Seminar held at the S.O.S. in Statistics, Pt. Ravishankar Shukla
University, Raipur (C.G.) from February 19-21, 2001
Title: Mean Estimation Under Set-up of Post-stratified
Cluster Sampling

Raipur
21-02-2001

Dr. S. K. Singh
Organizing Secretary
XVI M.P. YOUNG SCIENTISTS CONGRESS
Sponsored by: M.P. Council of Science and Technology, Bhopa
Organised by: Devi Ahilya Vishwavidyalaya, Indore

Certificate

Dr. / Mr. / As. MANISH TRIVEDI

of Dy. H.S. Gaur Univ. SAGAR

presented his research paper in the discipline of MATHEMATICS & STATISTICS at the XVI M.P. Young Scientists Congress organised by Devi Ahilya Vishwavidyalaya, Indore, on 6 & 7 September, 2001.

Prof. B.C. Chhaparwal
Vice-Chancellor
Devi Ahilya Vishwavidyalaya, Indore
ABSTRACTS

SPONSORED BY
UNIVERSITY GRANT COMMISSION, NEW DELHI

ORGANISED BY
FACULTIES OF SCIENCE, LIFE SCIENCES, ENGINEERING AND TECHNOLOGY

Dr. Harisingh Gour Vishwavidyalaya
(FORMERLY: UNIVERSITY OF SAGAR)
SAGAR

27 - 28 February 2001

198
USE OF INTERNET IN EDUCATION

Tripti Gehlot, Vijay Athavale & Tara Singh
Barkatullah University, Bhopal (MP)

The paper incorporates the ways in which Internet can be used as a better alternative to traditional classroom teaching by looking at various delivery methods. The authors have argued that the judicious mix of these methods with other contemporary methods of instructions can help the education system immensely by providing better understanding and comprehension of the subject. Advantages and disadvantages of using Internet in education are also discussed in the paper.

AN ESTIMATOR UNDER: POST-STRATIFIED CLUSTER SAMPLING

D. Shukla and Manish Trivedi,
Department of Mathematics and Statistics
Dr. H. S. Gour University, Sagar-470 003 M.P.

Suppose a population contains clusters of unequal size and stratified into several strata. The post stratification is useful when frame of each stratum is unknown in the set-up of a stratified population. This paper presents an estimation strategy for population mean in the set up of post stratification assuming that a sample of unequal cluster is drawn randomly from the population. The sample of clusters is post stratified and a weighting system is proposed. The expression of optimum strategy is derived and results are numerically supported.
PLATINUM JUBILEE YEAR 2001

NATIONAL SEMINAR

ON

ADVANCES IN MATHEMATICAL, STATISTICAL AND COMPUTATIONAL METHODS IN SCIENCE AND TECHNOLOGY

November 29 - 30, 2001

Abstract of papers

Organised by

DEPARTMENT OF APPLIED MATHEMATICS INDIAN SCHOOL OF MINES, DHANBAD

Sponsored by

All India Council for Technical Education, New Delhi
Council of Scientific & Industrial Research, New Delhi
Deptt. of Science & Technology, New Delhi
Indian School of Mines, Dhanbad
study is carried out to demonstrate the performance of constructed estimator over other estimators.

Keywords: Finite population mean, Class of unbiased ratio-type estimators, Negatively correlated variables, Variance.

62. Mean estimation under post – stratified cluster sampling scheme

D. Shukla and Manish Trivedi
Department of Mathematics and Statistics, Dr. Hari Singh Gour Vishwavidyalaya, Sagar (M.P.) India.

Abstract

Post stratification is used when the size of each strata the frame of each is unknown. Let strata contain clusters of size and a random sample of 'n' clusters post – stratified a stratum of the population. This paper considers the compare several post-stratified cluster estimators alongwith the get a class of estimators. A modified weight structure is proper to combine different cluster means. An attempt has been made optimum variance of the proposed class. The efficiency comp. Estimators is numerically supported by a database study.

63. A new approach to the estimation of population mean in two-phase sampling

S. C. Senapati¹, L. N. Sahov¹, G. N. Singh² and L. N. Upadhyaya²
1. Department of Statistics, Utkal University, Bhubaneswar-751004.
2. Department of Applied Mathematics, Indian School of Mines, Dhanbad-826 004.

Abstract

This paper considers the problem of estimating a finite population mean by adopting a two-phase sampling mechanism when the unknown mean of the auxiliary variable x is estimated with the help of an additional auxiliary variable z. Developing a new concept, we compose a general class of estimators that is superior to some of the previously studied classes of estimators under the minimum variance bound criterion.

Key Words: Asymptotic variance, auxiliary variable, minimum variance bound, two-phase sampling.

64. Unbiased product-type estimators under a linear model

Arun K. Singh and L. N. Upadhyaya
S.S.R.D, Nagaland University, Medishpema – 797 106

Abstract

In this paper, a general class of unbiased product-type estimators for population mean Y of the study character y has been obtained. Two new unbiased product type estimators are generated. The exact expressions for their variances are obtained and their merits
Appendices

E
To,

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

Dear Mr. Trivedi,

It is my pleasure to inform you that your research paper entitled "Efficient Estimation in Post Stratification using Prior Information" (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.)
File No. 1(10)/99

Dr. Diwakar Shukla
Reader in Statistics
Department of Mathematics & Statistics
Dr. H.S. Gour University
SAGAR – 470 003.

Dear Dr. Shukla,

On behalf of the Editorial Board, I am glad to inform you that the paper entitled MEAN ESTIMATION IN DEEPLY STRATIFIED POPULATION UNDER POST-STRATIFICATION by D. Shukla and Manish Trivedi has been ACCEPTED for publication in the Journal of the Indian Society of Agricultural Statistics.

With kind regards,

Yours sincerely,

(V.K. BHATIA)
File No. 1(25)/01

Subject: POST-STRATIFICATION IN TWO-WAY STRATIFIED POPULATION

By D. Shukla and Manish Trivedi

Dear Dr. Shukla,

I acknowledge with thanks the receipt of your article sent with your letter dated 01.08.2001 which is receiving attention. Please quote the reference number which is given above for further correspondence.

With best wishes,

Yours sincerely,

(V.K. Bhatia)

Dr. Diwakar Shukla
Reader in Statistics
Department of Mathematics & Statistics
Dr. H.S. Gour University of Sagar
SAGAR – 470 003.

August 06, 2001
Dear Dr. Shukla,

I am in receipt of three copies of your paper entitled "Efficient estimation using deep-post-stratification under a two-way r x r set-up" submitted for publication in the Aligarh Journal of Statistics. Our referees comments and the decision of the Editorial Board will be communicated to you in due course.

With thanks and regards,

Sincerely yours,

[Signature]

Prof. M.Z. Khan
Editor
File No. 1(3)/99

Subject: A New Estimator for Poststratification sampling scheme

By D. Shukla and Manish Trivedi

Dear Dr. Shukla,

I acknowledge with thanks the receipt of your article sent with your letter dated 27.03.1999 which is receiving attention. Please quote the reference number which is given above for further correspondence.

With kind regards,

Yours sincerely,

(V.K. Bhatia)

Dr. Diwakar Shukla
Reader in Statistics
Department of Mathematics & Statistics
Dr. Hari Singh Gour University
RAGAM - 476 003

08.04.1999
Subject: USE OF STABILITY RATIO $M_1/M_2$ IN POST-STRATIFICATION

By D. Shukla and Manish Trivedi

Dear Dr. Shukla,

I acknowledge with thanks the receipt of your article sent with your letter dated August 09, 2000 which is receiving attention. Please quote the reference number which is given above for further correspondence.

With kind regards,

Yours sincerely,

(V.K. BHATIA)

Dr. Dwakar Shukla
Reader in Statistics
Department of Mathematics & Statistics
Dr. H.S. Gour University
Sagar 470 003 (M.P.)
Dear Dr. Shukla,

I am sending herewith the comments on the paper entitled **A NEW ESTIMATOR FOR POSTSTRATIFICATION SAMPLING SCHEME** by D. Shukla and Manish Tripathi which was received for possible publication in the Journal of the Indian Society of Agricultural Statistics. You may please revise the paper in the light of the comments and send three copies of the same for considering it as a fresh paper.

Please give 'key-words' as well as a 'short title' of your paper.

With kind regards,

Yours sincerely,

(V.K. BHATIA)
Dr. S.D. Sharma
Secretary

Dr. V.K. Bhatla
Jt. Secretary & Member, Editorial Board

File No. 1(15)/98

Dr. Diwakar Shukla
Reader in Statistics
Department of Mathematics & Statistics
Dr. H.S. Gour University of Sagar
Sagar - 470 003.

September 25, 2001

Dear Dr. Shukla,

I am sending herewith the comments on the paper entitled "ESTIMATION IN POST STRATIFICATION USING PRIOR INFORMATION AND GROUPING STRATEGY" by D. Shukla, Ajay Bankey and Manish Trivedi which was received for possible publication in the Journal of the Indian Society of Agricultural Statistics. You may please revise the paper in the light of the comments and send three copies of the same after typing afresh along with the original marked manuscript.

Please give 'key-words' as well as a 'short title' of your paper.

With best wishes,

Yours sincerely,

(V.K. Bhatla)

(209)
THE XX ANNUAL CONFERENCE OF INDIAN SOCIETY FOR PROBABILITY
AND STATISTICS
AND NATIONAL SEMINAR ON INDUSTRIES ORIENTED STATISTICAL RESEARCH
AND CONFERENCE OF BAYESIAN SOCIETY OF INDIA
(FEBRUARY 19-21,2001)

To

Manish Trivedi
Department of Maths & Statistics
Dr.H.S.Gour University
Sagar (M.P.) - 470 003

ACCEPTANCE LETTER

I am happy to inform you that your paper titled mention below has been accepted for presentation at the "THE XX ANNUAL CONFERENCE OF INDIAN SOCIETY FOR PROBABILITY AND STATISTICS AND NATIONAL SEMINAR ON INDUSTRIES ORIENTED STATISTICAL RESEARCH AND CONFERENCE OF BAYESIAN SOCIETY OF INDIA " to be held at the School of Studies in Statistics ,Pt. Ravishankar Shukla University,Raipur,Chhattisgarh Pradesh during Feb. 19-21,2001.

Topic of the paper : Mean estimation under set up of post startified cluster sampling
Co-author : D.Shukla
Fees Paid : No.

(Dr.S.K.Singh)
Organizing Secretary
ISPS Conference.
Dear Prof. / Dr. Mr. Manish K., Dr. Shrivastava,

It is our pleasure to inform you that your paper "Data Efficient Estimation in Post
Stabilization Using Prior Information" has been accepted for the presentation in the seminar on "Role of Science in Sustainable Development", which is going to be organised on the occasion of national science day (Feb. 28, 2000). Outstation participants are entitled to get second class train (ordinary) bus fare. During the seminar lunch and refreshment will be provided. The programme of the seminar is as below:

**February 28, 2000.**

<table>
<thead>
<tr>
<th>Programmes</th>
<th>Time</th>
<th>Venue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Registration (Fess Rs. 60/)</td>
<td>08.30 AM to 09.30 AM</td>
<td>Science Lecture Hall</td>
</tr>
<tr>
<td>Inauguration</td>
<td>10.00 AM to 11.30 AM</td>
<td>Jawahar Lal Nehru Library Hall</td>
</tr>
<tr>
<td>Refreshment</td>
<td>11.30 AM to 12.00 AM</td>
<td></td>
</tr>
<tr>
<td>First Scientific Session</td>
<td>12.10 PM to 01.30 PM</td>
<td>Science Lecture Hall</td>
</tr>
<tr>
<td>Lunch</td>
<td>01.30 PM to 02.00 PM</td>
<td></td>
</tr>
<tr>
<td>Second Scientific Session</td>
<td>02.00 PM to 03.45 PM</td>
<td>Science Lecture Hall</td>
</tr>
<tr>
<td>Valedictory</td>
<td>04.00 PM to 05.00 PM</td>
<td>Science Lecture Hall</td>
</tr>
</tbody>
</table>

Your participation will be highly appreciated. It is necessary that before presentation of paper participants should submit their full length paper for the publication.

Thanking you.

V.K. Saxena

Phone : Res 23689 (07582)

Dr. K.K.N. Sharma

Phone : Res. 30620 (07582)
No. 337C /30/2000-2001
November 5 , 2000.

Dear Sir/Madam,

I am glad to inform you that your paper has been accepted for presentation under ISCA Young Scientists Award programme in the section at the 88th Indian Science Congress to be held at Indian Agricultural Research Institute, New Delhi from January 3 to 7, 2000.

You are requested to kindly present your full paper membership card, certificate of age (original) and a copy of your biodata at the Indian Science Congress Association Camp Office, at New Delhi, on 3-5 January, 2001.

T.A./D.A. will be paid by ISCA at the venue by cheque (maximum of first class/A.C.III Tier, which is available, train fare by the convenient shortest route to and from residence/institute to venue and back and D.A. as per ISCA rules not more than 9 days). You will be required to submit xerox copy of the ticket for reimbursement of your T.A.

Thanking you,

Yours faithfully,

(A.B. Banerjee)
General Secretary (Hqrs.)

C:\ns\A\B\T\BOB\SGC
Dear Young Scientist,


With best wishes,

Yours sincerely,

Prof. U.S. Vijayawargiya
Co-ordinator
15th M.P. Young Scientist Congress

Encl: As above

To,

Ms. Monisha Bhawani
H.O. (Officer in Charge)
No AU/513/03/98
Date : November 20, 98

Prof. D. Shukla
School of Statistics,
Devki Ahilya Vishwavidyalaya,

Dear Sir,

It gives us great pleasure to inform you that the abstract of your research paper entitled **A NEW ESTIMATOR FOR POSTSTRATIFICATION SAMPLING SCHEME** jointly with Prof. Manish Trivedi has been received and is accepted for presentation in the seminar.

You may please send the registration fee of Rs. 400=00 (Rupees four hundred only) by demand draft in favour of the Registrar, Amravati University, payable at Amravati.

Please intimate your itinerary well in advance so as to enable us to make the necessary arrangement for your local hospitality. Kindly inform your difficulties, we shall be happy to solve.

Looking forward to see you at Amravati,

Yours sincerely,

(R. SINGH)