TWO STAGE SAMPLING DESIGN FOR ESTIMATION OF REPRODUCTIVE MORBIDITY RATE: AN APPLICATION TO SLUM DWELLERS OF GUWAHATI CITY

Dilip C. Nath and Aditi Das Deka

ABSTRACT: This paper aims at estimating the Age-Specific Morbidity Rate (ASMR) and the Total Morbidity Rate (TMR) of the married women in the age group 15-59 years residing in the slum pockets of Guwahati city using Two Stage Sampling Technique. The paper also highlights a comparative analysis of the ASMRs and TMR obtained by Simple Random Sampling, and the ones obtained under two stage sampling technique. In our study, the first stage units (fsu) i.e., the slums are selected by Simple Random Sampling (SRS) Technique and the second stage units (ssu) i.e., the households or rather the individuals (married women in the age group 15-59) have been selected using Linear Systematic Sampling Technique. As regards the number of subunits to be selected from each fsu, recourse has been taken to Probability Proportional to Size. The study reveals that the MSEs are higher in case of SRS technique, implying that the sampling error involved in this method is more as compared to two stage sampling technique. It has been thus observed that two stage sampling method as a better technique for estimating TMR.

Keywords: Simple random sampling, Two stage sampling, Morbidity, Age-specific morbidity rate, Total morbidity rate.

1. INTRODUCTION

Women’s reproductive and sexual health had for decades been a neglected area of international research (Campbel et al., 1999, Sen and Snow, 1994). Concern with sexual and reproductive health gained momentum with the International Conference on Population and Development (ICPD, 1994). The Reproductive Health Index developed by Population Foundation of India shows that on a scale of 0 to 100, India scores 43, which shows that there is still much to be achieved on Reproductive Health Index (Nath and Majumdar, 1998). Some studies have been conducted on reproductive health but a comprehensive study is needed to find out the existing knowledge of the women, the prevalent practices and their attitude about reproductive health and to assess their existing health conditions. In the context of India, the impetus to bringing these and related issues into the public domain began with a community based epidemiological study of gynecological morbidity in Maharashtra (Bang et al., 1989). Younis et al., (1993) showed that the majority of women in their study were suffering from at least one gynaecological morbidity. Bang et al., (1989) showed that 92%
of their study participants suffered from gynaecological morbidity. Thapa and Basnet (1998) discerned that an “overwhelming majority” of women in their study suffered from reproductive health problems. In a community based study conducted by Bhatia et al., (1997), 39.8% of women were clinically diagnosed with reproductive morbidity. Bulut et al., (1997) reported that 37% of 918 women interviewed during a contraceptive choice study spontaneously reported a reproductive complaint; when a checklist was given, the percent rose to 81.

Following the ICPD in 1994, and the subsequent demand for information on reproductive morbidity in developing countries, several research studies were undertaken. It became increasingly clear that finding solid evidence on reproductive morbidity presented many obstacles. The most pertinent issue was obvious: how to obtain reliable data in difficult settings. Several studies have been carried out to investigate the various methodological approaches and the ensuing issues in collecting data in developing countries (Bhatia and Cleland, 2000; Khanna, 2001; Koenig and Shepherd, 2001; Younis, et al., 1993). Essentially, study designs are separated into two main categories, community based studies and clinic-based studies.

In clinic-based studies, the sample is usually pulled from patients seeking care for a specific service on their own initiative. In community-based studies, the sample is either all women in a given area or a random representation of women from a given area. Often, these studies include a combination of collection approaches, such as self-reported, medically diagnosed and laboratory tested (Khanna, 2001). Most commonly, these women are interviewed in their home and then receive a physical exam and laboratory work in a clinic setting. Another branch of the community-based study includes the community-based study of specific clinic populations; for example antenatal care patients, contraceptive users, etc. Such studies provide slightly more generalized results than clinic-based studies, but are wrought with complexities and limitations (Bhatia and Cleland 2000; Khanna 2001; Koenig and Sheperd 2001; Younis et al., 1993).

2. RATIONALE OF THE STUDY AND ITS OBJECTIVES
The world community, today, views health as a burning issue which needs international concern and utmost priority. The health of a country’s female population has profound implications on the health, education of children and the economic well-being of households as well as for women themselves. As the effects of pervasive ill health extend beyond the woman herself, it has become highly necessary to understand the roots of the morbid conditions in which the female folk has been trapped and take immediate steps with pre-determined objectives so as to attain the ‘good health for all’ target of the World Health Organisation (WHO). Investing into women would save human folk; otherwise survival of whole generation would be at risk.

Direct selection of sampling units from study populations, like simple random sampling (SRS), is expensive and the findings are also not cent percent acceptable. Even cluster
Two Stage Sampling Design for Estimation of Reproductive Morbidity Rate:  

sampling is not a favourable alternative in such situations as it may not give rise to a representative sample. A compromise between direct sampling and cluster sampling can be achieved by selecting a sample of clusters and surveying only a sample of units in each sample cluster instead of completely enumerating all the units in the sample clusters (Murthy, 1977). Such sampling procedure is referred to as two stage sampling. This sampling technique is very much in use in demographic surveys because of its practical advantage of providing a feasible procedure for framing of a satisfactory Sampling Frame. The clusters may be called the first stage units (fsu) and the observational units may be called the second stage units (ssu).

Again to ascertain and enhance precision of estimates, in case of two stage sampling schemes, it is beneficial to adopt probability proportional to cluster sizes for selection of fsu. And for selection of ssu, emphasis should be given on such sampling procedures which can ensure wider spread of ssu’s in the fsu’s (Singh, 2003).

The basic objective of this study is to identify the obstetrics and gynecological diseases prevailing among the slum dwellers of Guwahati city and hence obtain a comparative analysis of the Age Specific Morbidity Rates (ASMR) and the Total Morbidity Rate (TMR), as far as reproductive health is concerned, under the SRS technique and two stage sampling design.

3. STUDY POPULATION AND SAMPLING TECHNIQUE

There are 25 slum pockets in the Greater Guwahati area, the total number of households in these pockets being recorded as 24,603 with a total population of 15,6906 out of which 7,1995 are females (Guwahati Municipality Corporation Slum Survey Report, 2003). From these 25 slum pockets, 5 slum pockets have been selected for carrying out the survey. The female respondents who were the residents of the said slums and who were married and were in the age group 15-59 years were interviewed personally and necessary information were collected.

For our study, we have resorted to two stage sampling procedure, so as to obtain a better estimate of fertility for the population under consideration. In our study, the first stage units (fsu) i.e., the slums are selected by SRS technique and the second stage units (ssu) i.e., the households or rather the individuals (married women in the age group 15-59) have been selected using linear systematic sampling technique. As regards the number of subunits to be selected from each fsu, recourse has been taken to Probability Proportional to Size.

In the first stage, with application of SRS technique 5 slum pockets have been selected for carrying out the survey. The selected slums are:

1. Islampur  
2. Rajabari  
3. Athgaon  
4. Hatigaon (Sijubari)  
5. Santipur (East)

In the second stage, for selection of households or rather the individuals (i.e., ssu) from each of the selected slum pockets we have made use of Linear Systematic Sampling
Technique. This technique requires that the ratio of the number of ssu’s in the \(i^{th}\) fsu \((M_i)\) to the selected number of ssu \((m_i)\), is an integer and not a fractional number i.e.,

\[
M_i / m_i = k.
\]

Although, in our case \(k\) is a fractional number, yet we have resorted to Linear Systematic Sampling owing to the fact that we have approximated the fractional numbers to its nearest integer.

### Table 1

**Showing the fsu Size, No. of ssu in the Sample**

<table>
<thead>
<tr>
<th>Sl. No.</th>
<th>Name of slum pocket</th>
<th>fsu Size ((M_i))</th>
<th>No. of selected ssu ((m_i))</th>
<th>Ratio ((M_i / m_i))</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Islampur</td>
<td>1896</td>
<td>315</td>
<td>6.01</td>
</tr>
<tr>
<td>2.</td>
<td>Rajabari</td>
<td>520</td>
<td>95</td>
<td>6.11</td>
</tr>
<tr>
<td>3.</td>
<td>Athgaon</td>
<td>1275</td>
<td>210</td>
<td>6.07</td>
</tr>
<tr>
<td>4.</td>
<td>Hatigaon (Sijubari)</td>
<td>3492</td>
<td>576</td>
<td>6.06</td>
</tr>
<tr>
<td>5.</td>
<td>Santipur (East)</td>
<td>995</td>
<td>154</td>
<td>6.06</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>8178</strong></td>
<td><strong>1350</strong></td>
<td></td>
</tr>
</tbody>
</table>

Since the integer \(k\) is approximately equal to 6 in all the cases, the possible systematic samples were drawn in such a manner that every 6\(^{th}\) household is selected until the desired numbers of individuals (ssu) are selected.

### 4. ESTIMATION OF ASMR AND TMR

#### 4.1 SRS Technique

ASMR of the age group \((x, x + h)\) is defined as the ratio of females in the said age group suffering from any of the morbid conditions under consideration to the total number of women in that age group. Mathematically,

\[
ASMR = \frac{m_x}{f_x} = \frac{d_x}{f_x}
\]

where \(d_x\) is the number of women in the age group \((x, x + h)\) living under any one of the morbid conditions mentioned above and \(f_x\) is the number of resident women in that age group. The sampling technique used in such conventional set up is usually SRS Technique.

We shall also estimate the TMR by cumulating the ASMRs of female residents in the reproductive age group (15-59), residing in the slums of Guwahati, multiplied by \(h\), the age interval usually taken as 5 years, by the following formula:

\[
TMR = \left(\sum ASMR\right) \times h
\]

#### 4.2 Two Stage Sampling Design

We shall now attempt to calculate the age-specific morbidity rate of the cohort of 1350 females, the age groups being taken as 15-19, 20-24, …, 55-59, under the proposed sampling frame.
We define

\[ N = \text{number of } fsu \text{ in the population.} \]

\[ M_i = \text{number of households in the } i^{th} \text{ } fsu. \]

\[ n = \text{number of } fsu \text{ drawn out of } N \text{ under SRSWR.} \]

\[ m_i = \text{number of households drawn from } M_i \text{ households in the } i^{th} \text{ selected } fsu, \text{ which are selected using systematic sampling and whose sizes are determined by PPS.} \]

\[ hF_x = \text{total number of women in the reproductive age range i.e., } 15-59 \text{ years.} \]

\[ hD_x = \text{total number of women living under morbid conditions in the reference } 2 \text{ year period preceding the survey.} \]

\[ hf_x = \text{total number of women in the age group } (x, x+h) \text{ in the sample.} \]

\[ sd_x = \text{total number of women living under morbid conditions in the age group } (x, x+h) \text{ in the sample.} \]

\[ hF_{xij} = \text{total number of women in the age group } (x, x+h) \text{ in the } j^{th} \text{ household of the } i^{th} \text{ } fsu. \]

\[ hd_{xij} = \text{total number of women living under morbid conditions in the age group } (x, x+h) \text{ in the } j^{th} \text{ household of the } i^{th} \text{ } fsu. \]

\[ P_i = \text{inclusion probability of the } i^{th} \text{ } fsu \text{ in the sample.} \]

The ASMR for the study population of age group \((x, x+h)\) is defined as:

\[ hM_x = hD_x / hF_x \quad (4.2.1) \]

Then under the proposed sampling design we have,

\[ d_x = \sum_i \{M_i \cdot hF_{xij} \} / P_i \quad i = 1, 2, \ldots, n \quad (4.2.2) \]

where \(d_x\) is an unbiased estimate of \(hD_x\) and where

\[ D_x = \sum_i \{M_i \cdot hF_{xij} \} / P_i \quad i = 1, 2, \ldots, n \quad (4.2.3) \]

\[ D_{xij} = \{1/M_i\} \sum_i d_{xij} \quad i = 1, 2, \ldots, M_i \quad (4.2.4) \]

\[ d_{xij} = \{1/m_i\} \sum_i b_{xij} \quad i = 1, 2, \ldots, m \quad (4.2.5) \]

and

\[ f_x = \sum_i \{M_i \cdot hF_{xij} \} / P_i \quad i = 1, 2, \ldots, n \quad (4.2.6) \]

where \(f_x\) is an unbiased estimate of \(hF_x\) and where

\[ F_x = \sum_i \{M_i \cdot hF_{xij} \} / P_i \quad i = 1, 2, \ldots, N \quad (4.2.7) \]

\[ F_{xij} = \{1/M_i\} \sum_i f_{xij} \quad i = 1, 2, \ldots, M_i \quad (4.2.8) \]
\[ \mu_f(x) = \{1/m_i\} \sum_{j} m_if_{ij}, \quad i = 1, 2, \ldots, m_i \quad (4.2.9) \]

Then the sample estimates of \( ASMR \) is defined as:

\[ \hat{m}_i = h \hat{d}_i / \hat{f}_x \quad (4.2.10) \]

Then the sample estimates of \( ASMR \) and \( TMR \) are defined as:

\[ \hat{m}_i = h \hat{d}_i / \hat{f}_x \quad (4.2.11) \]

\[ TMR(\text{est.}) = h \sum \hat{m}_x \quad (4.2.12) \]

### 5. ESTIMATION OF VARIANCE OF \( TMR \)

#### 5.1 SRS Technique

We know that under SRS sample mean is an unbiased estimate of population mean. Thus, we have,

\[ E(\hat{\mu}_x) = \mu_x \quad E(\hat{\mu}_d) = \mu_d, \]

where \( \hat{\mu}_x \) and \( \hat{\mu}_d \) are the sample estimates of \( \mu_x \) and \( \mu_d \) respectively.

Using \( \Delta \)-method of expansion for ratio estimator (Murthy, 1977), \( O(n^{-2}) \) approximations to the bias and the mean square of \( \hat{\mu}_x \), we have,

\[ B = M_x \{(V(\hat{\mu}_x) / \hat{F}_x^2)\} - \text{Cov}(\hat{\mu}_x, \hat{d}_x) / \hat{F}_x \]

\[ = M_x \{(CV(\hat{\mu}_x))^2 - \rho_{\hat{\mu}_x, \hat{d}_x} CV(\hat{\mu}_x) \cdot CV(\hat{d}_x)\} \quad (5.1.1) \]

and,

\[ \text{MSE}(\hat{m}_x) = M_x^2 \{ (CV(\hat{\mu}_x))^2 + (CV(\hat{d}_x))^2 - 2 \rho_{\hat{\mu}_x, \hat{d}_x} CV(\hat{\mu}_x) \cdot CV(\hat{d}_x)\} \quad (5.1.2) \]

Where \( \rho_{\hat{\mu}_x, \hat{d}_x} \) is the correlation co-efficient between \( \hat{\mu}_x \) and \( \hat{d}_x \) and \( CV(\hat{\mu}_x) \) and \( CV(\hat{d}_x) \) are the respective coefficient of variation.

Thus, we have,

\[ E(hm_x - \hat{M}_x)(hm_y - \hat{M}_y) = M_x M_y \{\rho_{\hat{\mu}_x, \hat{d}_x} CV(\hat{\mu}_x) \cdot CV(\hat{d}_x) + \rho_{\hat{\mu}_y, \hat{d}_y} CV(\hat{\mu}_y) \cdot CV(\hat{d}_y) - \rho_{\hat{\mu}_x, \hat{d}_y} CV(\hat{\mu}_x) \cdot CV(\hat{d}_y)\} \]

\[ \quad - \rho_{\hat{\mu}_x, \hat{d}_y} CV(\hat{\mu}_x) \cdot CV(\hat{d}_y) \quad (5.1.3) \]

Thus, the estimated adjusted \( TMR \) and its \( MSE \) are defined by:

\[ TMR(\text{adj}) = \sum_i \{hm_x - \text{est. } B(hm_x)\} = TMR_{\text{SRS}} - h \sum \text{est. } B(hm_x) \quad (5.1.4) \]

Est. MSE of \( TMR(\text{adj}) \) is:

\[ = h^2 \{ \sum_i \text{est. MSE}(hm_x) + \sum_i \sum_j \text{est. } E(hm_x - \hat{M}_x)(hm_y - \hat{M}_y) \} \]

\[ - \{ \sum_i \text{est. } B(hm_x) \}^2 \quad (5.1.5) \]

Estimators \( \text{est. } B(hm_x), \text{est. MSE}(hm_x) \) and \( \text{est. } E(hm_x - \hat{M}_x)(hm_y - \hat{M}_y) \) of \( \hat{B}(hm_x), \text{MSE}(hm_x) \) and \( E(hm_x - \hat{M}_x)(hm_y - \hat{M}_y) \) respectively can be obtained replacing population parameters by the corresponding sample unbiased estimators in their respective equations.
5.2 Two Stage Sampling Technique

Let \( E_i, \quad i = 1, 2 \) denote the expectation over the \( i^{th} \) stage sampling and
\[
d_i = 1, \text{ if } i^{th} \text{ fsu is included in the sample}
\]
\[
d_i = 0, \text{ otherwise.}
\]

It is to be noted that \( E_2[\frac{\bar{x} \cdot d_i}{m_i}] = \frac{\bar{D}_x}{n} \), as systematic sample mean is an unbiased estimator of population mean.

As a consequence, we get,
\[
E(\bar{x} \cdot d_i) = E_i[\sum M_i E_2(\frac{\bar{x} \cdot d_i}{m_i})]/P_i = \frac{\bar{D}_x}{n} \text{ showing that } b_x \text{ is an unbiased estimate of } \bar{D}_x.
\]

(see appendix A1)

Further, taking \( V_i, \quad i = 1, 2 \) as the variance for the \( i^{th} \) stage sampling, by definition, we have,
\[
V(\bar{x} \cdot d_i) = V_i E_2(\bar{x} \cdot d_i) + E_i V_2(\bar{x} \cdot d_i)
\]
\[
= \sum \sum (P_i P_j - P_{ij}) \left[ \left( M_i \frac{\bar{D}_x}{P_i} - M_j \frac{\bar{D}_x}{P_j} \right) \right]^2
\]
\[
+ \sum \left[ \left( M_i^2 / P_i \right) S_{d_{ij}}^2 / m_i \right] \left[ 1 / \left( m_i - 1 \right) \right] S_{d_{ij}}^2
\]
\[
\left( \text{see appendix A2} \right)
\]

where \( p_{d_{ij}} \) is the intra class correlation co-eff. of births to the women in the age group \((x, x + h)\) in the \( i^{th} \) fsu and \( S_{d_{ij}}^2 \) is the population variance of the \( i^{th} \) fsu.

To obtain an unbiased estimator of \( V(\bar{x} \cdot d_i) \), we note that
\[
E(\sum \left( P_i P_j - P_{ij} \right) (M_i \frac{\bar{D}_x}{P_i} - M_j \frac{\bar{D}_x}{P_j})^2 )
\]
\[
= V(\bar{x} \cdot d_i) - E \left[ \sum M_i^2 / P_i \text{ est. } V_2(\bar{x} \cdot d_i) \right] m_i
\]
\[
\left( \text{see appendix A3} \right)
\]

Therefore, an unbiased estimator of \( V(\bar{x} \cdot d_i) \) is of the form:
\[
V(\bar{x} \cdot d_i) = \sum \left( P_i P_j - P_{ij} \right) (M_i \frac{\bar{D}_x}{P_i} - M_j \frac{\bar{D}_x}{P_j})^2
\]
\[
+ \sum M_i^2 / P_i \left[ 1 / \left( m_i (m_i - 1) \right) \right] \sum \left[ d_{ij}^2 - 2 d_{ij} \right]
\]
\[
\left( \text{see appendix A3} \right)
\]

Similar results for women of age group \((x, x + h)\) can be obtained.

As in the usual ratio estimator \( m \) is a biased estimator of \( M_x \). Thus, using the delta method of expansion and following Sukhatme et al., (1984), the first order approximation to the bias and mean square error of \( m \) are obtained as:
\[
B(\bar{x}, m) = \frac{\bar{D}_x}{n} \left\{ CV(\bar{x}) \right\}^2 - p_{d_{ij}} \bar{x} CV(\bar{x}) \cdot CV(\bar{x} \cdot d_i).
\]
\[
\left( \text{5.2.5} \right)
\]
\[
MSE(m_x) = h M_x^2 \left[ \left\{ CV(f_x) \right\}^2 + \left\{ CV(d_x) \right\}^2 - 2p_{h(f_x), h(d_x)} CV(f_x) \cdot CV(d_x) \right]
\]  \hspace{1cm} (5.2.6)

where \(p_{h(f_x), h(d_x)}\) is the correlation co-eff. between \(f_x\) and \(d_x\) and \(CV(f_x)\) and \(CV(d_x)\) are the coeff. of variation of \(f_x\) and \(d_x\) respectively.

Equivalently,
\[
E(h m_x - M_x)(h m_y - M_y) = M_x M_y \left[ p_{h(f_x), h(d_x)} CV(f_x) \cdot CV(d_x) \right] + p_{h(d_x), h(d_y)} CV(d_x) \cdot CV(d_y)
\]
\[
- p_{h(f_x), h(d_y)} CV(f_x) \cdot CV(d_y) - p_{h(d_x), h(f_y)} CV(d_x) \cdot CV(f_y).
\]

Estimators \(\text{est} B(h m_x), \text{est} MSE(h m_x)\) and \(\text{est} E(h m_x - M_x)(h m_y - M_y)\) of \(B(h m_x), MSE(h m_x)\) and \(E(h m_x - M_x)(h m_y - M_y)\) respectively can be obtained replacing population parameters by the corresponding sample unbiased estimators in their respective equations.

Thus, the estimated adjusted \(TMR\) and its \(MSE\) are defined by:
\[
TMR(\text{adj}) = \sum [h m_x - \text{est} B(h m_x)]
\]  \hspace{1cm} (5.2.7)

\[
\text{Est. MSE of TMR(\text{adj})} = \sum \text{Est. MSE}(h m_x) + \sum \sum \text{Est.} E(h m_x - M_x)(h m_y - M_y)
\]
\[
- (\sum \text{est} B(h m_x))^2
\]  \hspace{1cm} (5.2.8)

6. RESULTS AND DISCUSSION

The various estimates of \(ASMRs\) for the different age groups and the \(TMRs\) as obtained under the two methods of sampling along with the adjusted \(TMR\) and the mean square error of the estimated \(TMR\) are shown in Table 2.

<table>
<thead>
<tr>
<th>Age group</th>
<th>SRS</th>
<th>Under two stage sampling</th>
</tr>
</thead>
<tbody>
<tr>
<td>15-19</td>
<td>0.6974</td>
<td>0.2679</td>
</tr>
<tr>
<td>20-24</td>
<td>0.6495</td>
<td>0.6440</td>
</tr>
<tr>
<td>25-29</td>
<td>0.6379</td>
<td>0.6848</td>
</tr>
<tr>
<td>30-34</td>
<td>0.6302</td>
<td>0.6587</td>
</tr>
<tr>
<td>35-39</td>
<td>0.6667</td>
<td>0.6588</td>
</tr>
<tr>
<td>40-44</td>
<td>0.8257</td>
<td>0.8407</td>
</tr>
<tr>
<td>45-49</td>
<td>0.9121</td>
<td>0.9389</td>
</tr>
<tr>
<td>50-54</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>55-59</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>TMR</td>
<td>30.54</td>
<td>30.05</td>
</tr>
<tr>
<td>Adjusted TMR</td>
<td>30.556</td>
<td>30.014</td>
</tr>
<tr>
<td>MSE of TMR</td>
<td>0.7954</td>
<td>0.4893</td>
</tr>
</tbody>
</table>
Two Stage Sampling Design for Estimation of Reproductive Morbidity Rate:

The study reveals that the age specific morbidity rate (ASMR) calculated directly as a ratio between the total number of women living under morbid conditions in the age group \((x, x + h)\) in the sample and the total number of women in the age group \((x, x + h)\) in the sample and under the proposed two-stage sampling design, gives almost the same results.

However, a difference has been seen in the ASMR for the age interval 15-19 in the two methods. It might be a reflection of the fact that in this particular age group, in the Rajabari area none of the women were reported to live under morbid conditions.

ASMR shows an increasing trend from age 34 onwards. It has also been observed that there is a sudden skip in the ASMR figures as the women enters the age group 40-44, implying a high risk of morbidity during this age interval. It seems to be the most crucial stage as far as women’s reproductive health is concerned. It is due to the fact that there occurs hormonal changes in her body and she is in her pre-menopausal stage. It is noteworthy to mention that for the age groups 50-54 and 55-59, the ASMR figures are as high as 1, implying that all women residing in the various slums are living under morbid conditions as far as their reproductive health is concerned.

As far as the TMRs are concerned, it has been found that the TMR calculated by conventional method gives slightly higher values than the one obtained under two stage sampling method, implying over estimation of the TMR by the conventional technique. Moreover, it has been seen that the adjusted TMRs derived after adjusting the biases marginally increases in case of SRS technique and decreases in case of two stage sampling technique. The MSEs are found to be 0.7954 and 0.4893 respectively, implying that the sampling error involved in the conventional method is more as compared to two stage sampling technique. It is thus observed that two stage sampling method is a better technique for estimating TMR.

REFERENCES


**Dilip C. Nath**  
Professor, Dept. of Statistics,  
Gauhati University, Guwahati-14, Assam, India.  
E-mail: dcnath@rediffmail.com

**Aditi Das Deka**  
Research Scholar, Dept. of Statistics,  
Gauhati University, Guwahati-14, Assam, India.  
E-mail: aditidasdeka@gmail.com
APPENDIX

A1:

\[ E(d) = E_i \left[ \sum M_i E_j \left( \tilde{d}_{i,j} \right) / m \right] / P_i \]

\[ = \sum E_i \left( \tilde{d} \right) \left( M_i \tilde{D} \right) / P_i \]

\[ = \sum \left( M_i \tilde{D} \right) / P_i \]

\[ = D, \text{ showing that } d \text{ is an unbiased estimate of } D \]

A2:

\[ V_2(d) = \sum M_i^2 / P_i^2 V_2 \left( \frac{\tilde{d}_{i,j}}{m_{ij}} \right) \]

\[ = \sum M_i^2 / P_i^2 S_{i,j}^2 / m_{ij} \left[ 1 + (m_i - 1) p_{i,j} \right] \]

Thus,

\[ E_1 V_2(d) = \sum \left( \frac{M_i^2}{P_i^2} \right) S_{i,j}^2 / m_{ij} \left[ 1 + (m_i - 1) p_{i,j} \right] \]

And,

\[ E_2(d) = \left[ \sum M_i E_j \left( \tilde{d}_{i,j} \right) / m \right] / P_i \]

\[ = \sum \left( M_i \tilde{D} \right) / P_i \]

Then,

\[ V_1 E_2(d) = \sum \sum \left( P_i P_j - P_o \right) \left[ \left( M_i \tilde{D} \right) / P_i - \left( M_j \tilde{D} \right) / P_j \right]^2 \]

Combining the results we have

\[ V(d) = \sum \sum \left( P_i P_j - P_o \right) \left[ \left( M_i \tilde{D} \right) / P_i - \left( M_j \tilde{D} \right) / P_j \right]^2 \]

\[ + \sum \left( M_i^2 / P_i^2 \right) S_{i,j}^2 / m_{ij} \left[ 1 + (m_i - 1) p_{i,j} \right] \]

A3:

\[ E \left[ \sum \sum \left( P_i P_j - P_o \right) \left( \tilde{D} \right) / P_i - \left( D \right) / P_j \right]^2 \]

\[ = \sum \sum \left( P_i P_j - P_o \right) \left[ \left( M_i^2 / P_i^2 \right) E_j \left( \tilde{d}_{i,j}^2 \right) / m_{ij} \right] - \left( M_j^2 / P_j^2 \right) E_j \left( \tilde{d}_{i,j}^2 \right) / m_{ij} \]

\[ - 2 \left( M_i M_j / P_i P_j \right) E_j \left( \tilde{d}_{i,j} \right) / m_{ij} \]

\[ = \sum \sum \left( P_i P_j - P_o \right) \left[ \left( M_i^2 / P_i^2 \right) \tilde{D}_{i,j}^2 + V_2 \left( \tilde{d}_{i,j} \right) / m_{ij} \right] \]

\[ + \left[ \left( M_j^2 / P_j^2 \right) \tilde{D}_{i,j}^2 + V_2 \left( \tilde{d}_{i,j} \right) / m_{ij} \right] - 2 \left( M_i M_j / P_i P_j \right) \left\{ \tilde{D}_{i,j} \right\} / m_{ij} \]

\[ = \sum \sum \left( P_i P_j - P_o \right) \left[ \left( M_i \tilde{D} \right) / P_i - \left( M_j \tilde{D} \right) / P_j \right]^2 \]

\[ + \sum \sum \left( P_i P_j - P_o \right) \left[ \left( M_i \tilde{D} \right) / P_i - \left( M_j \tilde{D} \right) / P_j \right] \]
\[ \begin{align*}
&= \sum \sum (P_i P_j - P_{ij}) \left[ \left\{ M_{ij} D_{ij} \right\} / P_i - M_{ij} D_{ij} / P_j \right]^2 \\
&\quad + \sum M_i^2 / P_i \} S_{\mu_i}^2 / m_i \left\{ 1 + (m_i - 1) p_{\mu_i} - \sum M_i^2 S_{\mu_i}^2 / m_i \left\{ 1 + (m_i - 1) p_{\mu_i} \right\} \right\} \\
&= V(\bar{d}_x) - \sum M_i^2 S_{\mu_i}^2 / m_i \left\{ 1 + (m_i - 1) p_{\mu_i} \right\} \\
&= V(\bar{d}_x) - \sum M_i^2 \{ V_2(\bar{d}_x) \} | m_i \} \\
&= V(\bar{d}_x) - E \left\{ \sum M_i^2 / P_i \text{ est.} \{ V_2(\bar{d}_x) \} | m_i \} \right\}
\end{align*} \]