2.1 Introduction

Non-response refers to the failure to obtain information, for some reason or the other, from some of the population units that are selected in the sample. Hansen and Hurwitz (1946) presented the classical non-response theory for eliciting responses from a sub sample of the non-respondents. The technique was first developed for the surveys in which the first attempt was made by mailing the questionnaires and a second attempt was made by personal interviews to a sub sample of the non-respondents. They constructed the estimator for the population mean and derived the expression for its variance. Hansen and Hurwitz’s technique was further extended by El-Badry (1956) by sending waves of questionnaires to the non-respondents units to increase the response rate. Foradari (1961) generalized El-Badry’s approach for different sampling designs. Srinath (1971) suggested the selection of sub samples by making several attempts. Khare (1987) investigated the problem of optimum allocation in stratified sampling in presence of non-response for fixed cost as well as for fixed precision of the estimate.

The problem of optimum allocation in stratified random sampling for a univariate population is well known in sampling literature; (Cochran (1977) and Sukhatme et al. (1984)). When we study more than one characteristic, it is not possible to use the individual optimum allocations to various strata because an allocation, which is optimum for one characteristic may not be optimum for other characteristics. Moreover, in the absence of a strong positive correlation between the characteristics under study the individual optimum allocations may differ a lot and there may be no obvious compromise. In such situations, some criteria is needed to work out an
allocation which is optimum, in some sense, for all characteristics. Methods for solving the problem of optimum allocation in multivariate stratified sampling are proposed by various authors. Dalenius (1953), Yates (1960), Folks and Antle (1965), Hartley (1965), Kish (1988), Khan at al. (1997), worked out the multivariate allocation by minimizing the weighted average of different characters. The second approach of minimizing the total cost of the survey when variances are subjected to fixed tolerance limits is discussed by Dalenius (1957), Yates (1960), Kokan (1963), Kokan and Khan (1967), Chatterjee (1968), Haddleston et al. (1970), Chatterjee (1972), Bethal (1985), Chromy (1987), Bethal (1989) etc. Sarndal et al. (1992) formulated the generalized optimization problem for model based sampling that is of interest for several specified allocation problems. Zayatz and Sigman (1994) studied the feasibility of the use of Chromy’s algorithm in a practical situation related to the annual sample survey of manufactures.

In this chapter we have used the Khare’s (1987), problem of determining the sample sizes for the fixed total sample size to various strata in presence of non-response and optimum sub sample sizes among the non-respondents in stratified sampling. The problem is first formulated as a Mathematical Programming Problem (MPP), and then solved by using Lagrange’s multiplier technique. To find the integer solution we have used Branch and Bound technique. A numerical example is also presented to illustrate the computational details.

2.2 Formulation of the problem

In stratified sampling where population of size $N$ is divided into $L$ strata. Let $N_i$, $\bar{Y}_i$, $S_i^2$ and $p_i = \frac{N_i}{N}$ denote the stratum size, stratum mean, stratum variance and stratum weight of $i^{th}$ stratum. The population of each stratum is divided into two classes, those who will response at first attempt and those who will not response, hence
creates the problem of non response. Select a random sample of size $n_i$ from $i^{th}$ stratum in which $n_{i1}$ units who will response and $n_{i2}$ units who will not response. These are known as sizes of respondent sample and non-respondent sample respectively. We select a sub sample of size $r_i$ units out of a non-respondent sample of size $n_{i2}$ units, such that

$$n_{i2} = K_i r_i \ (K_i > 1)$$

where $\frac{1}{K_i}$ is known as the sampling fraction among non-respondents in the $i^{th}$ stratum.

Mathematical programming Problem of Khare (1987) for fixed sample size can be written as

$$\begin{align*}
\text{Min } V(\bar{y}_w) &= \sum_{i=1}^{L} \left\{ \left( \frac{N_i-n_i}{N_i} \right) + \frac{(K_i-1)}{n_i} W_{i2} \right\} p_i^2 S_i^2 \ (i) \\
\text{subject to } &\sum_{i=1}^{L} n_i = n \quad (ii) \\
\text{and } &2 \leq n_i \leq N_i \quad (iii)
\end{align*}$$

(2.2.1)

where $\bar{y}_w = \sum_{i=1}^{L} \frac{n_i}{n_i} (n_{i1} \bar{y}_{i1} + n_{i2} \bar{y}_{i2} r_i)$ and $\bar{y}_{i1}, \bar{y}_{i2}$ be the sample means of the respondents group in the $i^{th}$ stratum. The subscript $r_i$ is introduced as a reminder that the sample in the second group is of size $r_i$, in the $i^{th}$ stratum.

Ignoring the terms independent of $n_i$’s the objective function (2.2.1)(i) can be expressed as

$$Z(n_1, n_2, \ldots, n_L) = \sum_{i=1}^{L} \left\{ \frac{1+(K_i-1) W_{i2}}{n_i} \right\} p_i^2 S_i^2$$

$$= \sum_{i=1}^{L} \frac{b_i}{n_i}$$

where $b_i = \{1 + (K_i - 1) W_{i2}\} p_i^2 S_i^2$.

Thus the problem (2.2.1) is simplified as

$$\begin{align*}
\text{Min } Z(n_1, n_2, \ldots, n_L) &= \sum_{i=1}^{L} \frac{b_i}{n_i} \ (i) \\
\text{subject to } &\sum_{i=1}^{L} n_i = n \quad (ii) \\
\text{and } &2 \leq n_i \leq N_i \quad (iii)
\end{align*}$$

(2.2.2)
The restriction \((2.2.2)(ii)\) are imposed to avoid over sampling, that is, the situation where \(n_i \geq N_i\) and to have the representation of every stratum in the sample.

To determine the optimum values of \(n_i\), we consider the function:

\[
\phi(n_i, \lambda) = \sum_{i=1}^{L} \left( \frac{b_i}{n_i} \right) + \lambda \left( \sum_{i=1}^{L} n_i - n \right)
\]

where \(\lambda\) is Lagrange’s multiplier.

Differentiating \(\phi\) partially w. r. t. \(n_i\) and equating to zero, we get

\[
\frac{\partial \phi}{\partial n_i} = -\frac{b_i}{n_i^2} + \lambda = 0, \quad i = 1, 2, \ldots, L
\]

and finally

\[
n_i = \frac{n \sqrt{b_i}}{\sum_{i=1}^{L} \sqrt{b_i}} \quad (2.2.3)
\]

If the above values of \(n_i\) satisfies \((2.2.2)\) then the non-linear integer programming problem \((\text{NLIPP}) (2.2.2)\) is solved and \((2.2.3)\) will give the required optimum allocation. If we do not find the integer solution then we use the following Branch and Bound technique to find the integer solution.

### 2.3 The computation procedure based on Branch and Bound technique

To explain the branching process which will be used, let the notation \([t]\) represent the greatest integer less than or equal to a real no. \(t\). For example

\([4.2] = 4, \quad [-2.1] = -3\)

Now, suppose that we first solve the linear relaxation. If this yields an all integer solution then the problem is solved, otherwise, some integer variable will be selected, say \(x_k = \beta_k\), where \(\beta_k\) is the current non integer value of \(x_k\). A partitioning of the problem will then be created utilizing the conditions.

\[
x_k \leq [\beta_k]\n\]

\[
x_k \geq [\beta_k] + 1
\]
For example if $x_k = 5.82$, then the next branching will be defined by $x_k \leq \lfloor 5.82 \rfloor = 5$, and $x_k \geq \lfloor 5.82 \rfloor + 1 = 5 + 1 = 6$

### 2.4 Numerical Illustration

Table 1 & 2 gives the population parameters obtained from the data as given in Okafor (1994)

**Table 1: Overall stratum population parameter**

<table>
<thead>
<tr>
<th>Stratum</th>
<th>$p_i$</th>
<th>$S_i$</th>
<th>$K_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 – 930</td>
<td>0.336</td>
<td>199.9370</td>
<td>2</td>
</tr>
<tr>
<td>931 – 1700</td>
<td>0.352</td>
<td>247.9022</td>
<td>1.67</td>
</tr>
<tr>
<td>1701 – 4300</td>
<td>0.313</td>
<td>415.2409</td>
<td>1.43</td>
</tr>
</tbody>
</table>

$p_i = \frac{N_i}{N}, i = 1, 2, \text{and } 3 N_1 = 67.2, N_2 = 70.40 \text{ and } N_3 = 62.60$
Table 2: Class stratum population parameter

<table>
<thead>
<tr>
<th>Stratum</th>
<th>Class</th>
<th>Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 – 930</td>
<td>Respondent</td>
<td>0.188</td>
</tr>
<tr>
<td></td>
<td>Non – Respondent</td>
<td>0.148</td>
</tr>
<tr>
<td>931 – 1700</td>
<td>Respondent</td>
<td>0.219</td>
</tr>
<tr>
<td></td>
<td>Non – Respondent</td>
<td>0.133</td>
</tr>
<tr>
<td>1701 – 4300</td>
<td>Respondent</td>
<td>0.188</td>
</tr>
<tr>
<td></td>
<td>Non – Respondent</td>
<td>0.125</td>
</tr>
</tbody>
</table>

It is assumed that $N = 200, n = 50, N_1 = 67, N_2 = 70$ and $N_3 = 63$

Table 3: Calculation of $n_i$ using formula (2.2.3)

<table>
<thead>
<tr>
<th>$i$</th>
<th>$b_i$</th>
<th>$\sqrt{b_i}$</th>
<th>$n\sqrt{b_i}$</th>
<th>$n_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5180.9188</td>
<td>71.9787</td>
<td>3598.9300</td>
<td>12.1396 ≈ 12</td>
</tr>
<tr>
<td>2</td>
<td>8293.1178</td>
<td>91.0666</td>
<td>4553.3278</td>
<td>15.3588 ≈ 15</td>
</tr>
<tr>
<td>3</td>
<td>17800.2710</td>
<td>133.4177</td>
<td>6670.8828</td>
<td>22.5016 ≈ 23</td>
</tr>
</tbody>
</table>

The problem (2.2.2) can be rewritten as

$$
\begin{align*}
\text{Min } Z &= \frac{b_1}{n_1} + \frac{b_2}{n_2} + \frac{b_3}{n_3} (i) \\
\text{subject to } &n_1 + n_2 + n_3 = 50 \quad (ii) \\
&2 \leq n_1 \leq 67 \quad (iii) \\
&2 \leq n_2 \leq 70 \\
&2 \leq n_3 \leq 63 \\
\end{align*}
$$

and $n_i$'s are integers, $i = 1,2,3$  (iv)

The values of $n_i$'s from table 3 satisfy 2.4.1(iii) and the optimal value of the objective function is $\text{Min } Z = 1757.804$. 

118
Branch and Bound Method to find out the integer optimum of the problem given by 2.4.1 and the optimum solution of the given problem from figure 1 is \( n_1 = 12, n_2 = 15 \) and \( n_3 = 23 \) with objective function \( \text{Min } Z = 1758.543 \)

\[
\begin{align*}
\text{Min } Z & = 1757.60 \\
n_1 & = 12.14 \\
n_2 & = 15.36 \\
n_3 & = 22.50 \\
\text{Min } Z & = 1758.51 \\
n_1 & = 11.92 \\
n_2 & = 15.08 \\
n_3 & = 23 \\
\text{Min } Z & = 1758.52 \\
n_1 & = 12.36 \\
n_2 & = 15.64 \\
n_3 & = 22 \\
\text{Min } Z & = 1763.24 \\
n_1 & = 11 \\
n_2 & = 16 \\
n_3 & = 23 \\
\text{Min } Z & = 1758.54 \\
n_1 & = 12 \\
n_2 & = 15 \\
n_3 & = 23 \\
\end{align*}
\]

\( n_2 \leq 22 \) \( n_2 \geq 23 \) \( n_2 \leq 15 \) \( n_2 \geq 16 \) \( n_1 \leq 11 \) \( n_1 \geq 12 \)

\[ \text{Fathomed Fathomed Fathomed Optimum Integer Solution} \]

**Figure (1)**

2.5 **Determination for sub sample sizes of non-response \( r_i \)'s for fixed cost**

Consider the variance of the given mean \( \bar{y}_w \) i.e.

\[
V(\bar{y}_w) = \sum_{i=1}^{L} \left\{ \left( \frac{N_i-n_i}{N_i n_i} \right) + \frac{(K_i-1)}{n_i} W_i \right\} p_i^2 s_i^2
\]  \hspace{1cm} (2.5.1)

Cost function is of the form

\[
C = \sum_{i=1}^{L} C_{i0} n_i + \sum_{i=1}^{L} C_{i1} n_{i1} + \sum_{i=1}^{L} C_{i2} r_i
\]  \hspace{1cm} (2.5.2)

where \( C_{i0} \) is the cost of making the first attempt while \( C_{i1} \) and \( C_{i2} \) are the cost of getting, editing and processing information per unit in the response and non-response groups respectively in the \( i^{th} \) stratum.

\[
n_{i2} = K_i r_i \Rightarrow K_i = \frac{n_{i2}}{r_i}
\]
ignoring the terms Independent of \( r_i \) in the R.H.S. of (2.5.1) and putting

\[ K_i = \frac{n_{iz}}{r_i} \]

The problem becomes

\[
\begin{aligned}
\text{Min } Z(r_1, r_2, \ldots, r_L) &= \frac{n_{iz}w_{iz}}{r_i n_i} p_i^2 S_i^2 \quad (i) \\
\text{subject to} \\
\sum_{i=1}^{L} c_{i2} r_i &\leq C_0 \quad (ii) \\
2 \leq r_i &\leq n_{i2} \quad (iii) \\
\text{and } r_i \text{'s are integers, } i = 1, 2, \ldots, L \quad (iv)
\end{aligned}
\]

where \( C_0 = \sum_{i=1}^{L} c_{i0} n_i + \sum_{i=1}^{L} c_{i1} n_{i1} \)

Let \( d_i = \frac{n_{iz}w_{iz} p_i^2 S_i^2}{n_i} \quad (2.5.4) \)

The AINLPP (2.5.3) can be redefined as

\[
\begin{aligned}
\text{Min } Z(r_1, r_2, \ldots, r_L) &= \sum_{i=1}^{L} \frac{d_i}{r_i} \quad (i) \\
\text{subject to} \\
\sum_{i=1}^{L} c_{i2} r_i &\leq C_0 \quad (ii) \\
2 \leq r_i &\leq n_{i2} \quad (iii) \\
\text{and } r_i \text{'s are integers, } i = 1, 2, \ldots, L \quad (iv)
\end{aligned}
\]

Applying Lagrange’s multiplier technique, with equality in (2.5.5) (ii) and ignoring (2.5.5)(iii-iv), we get

\[
\phi(r_i, \lambda) = \sum_{i=1}^{L} \frac{d_i}{r_i} + \lambda (\sum_{i=1}^{L} c_{i2} r_i - C_0) \quad (2.5.6)
\]

Differentiating \( \phi \) with respect to \( r_i \) and \( \lambda \) and equating to zero, we get

\[
\frac{\partial \phi}{\partial r_i} = -\frac{d_i}{r_i^2} + \lambda c_{i2} = 0 \quad (2.5.7)
\]

\[
\frac{\partial \phi}{\partial \lambda} = \sum_{i=1}^{L} c_{i2} r_i - C_0 = 0 \quad (2.5.8)
\]

Solving equation (2.5.7) and (2.5.8), we get the optimum value of \( r_i \) as :-

\[
r_i = \frac{c_{i0} \sqrt{d_i}}{\sum_{i=1}^{L} c_{i2} \sqrt{d_i c_{i2}}} \quad (2.5.9)
\]
From the example given in the section (2.4) and letting \( C_{i2} = 12,14,10 \) for \( i = 1,2 \) and 3 respectively and \( C_0 = 200 \), we have the following table. Since \( W_{i1} \) and \( W_{i2} \) are known for \( i = 1,2 \) and 3. These values are used to work out for the expected value of \( n_{i2} \) as 

\[
n_{i2} = \frac{n_i W_{i2}}{W_{i1} + W_{i2}}
\]

**Table 4: Calculation of \( n_{i2} \)**

<table>
<thead>
<tr>
<th>( i )</th>
<th>( W_{i1} )</th>
<th>( W_{i2} )</th>
<th>( S_i )</th>
<th>( p_i )</th>
<th>( n_i )</th>
<th>( C_{i2} )</th>
<th>( n_{i2} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.188</td>
<td>0.148</td>
<td>199.9370</td>
<td>0.336</td>
<td>12</td>
<td>12</td>
<td>5.2857 ( \approx 5 )</td>
</tr>
<tr>
<td>2</td>
<td>0.219</td>
<td>0.133</td>
<td>247.9022</td>
<td>0.352</td>
<td>15</td>
<td>14</td>
<td>5.6676 ( \approx 6 )</td>
</tr>
<tr>
<td>3</td>
<td>0.188</td>
<td>0.125</td>
<td>415.2410</td>
<td>0.313</td>
<td>23</td>
<td>10</td>
<td>9.1853 ( \approx 9 )</td>
</tr>
</tbody>
</table>

**Table 5: Calculation of \( r_i \) using (2.5.9)**

<table>
<thead>
<tr>
<th>( i )</th>
<th>( d_i )</th>
<th>( \sqrt{d_i} )</th>
<th>( \sqrt{d_i / C_{i2}} )</th>
<th>( \sqrt{d_i \times C_{i2}} )</th>
<th>( r_i )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>294.2035</td>
<td>17.1524</td>
<td>4.9515</td>
<td>59.4175</td>
<td>4.4123 ( \approx 4 )</td>
</tr>
<tr>
<td>2</td>
<td>382.6535</td>
<td>19.5615</td>
<td>5.2280</td>
<td>73.19256</td>
<td>4.6588 ( \approx 5 )</td>
</tr>
<tr>
<td>3</td>
<td>843.2659</td>
<td>29.0390</td>
<td>9.1830</td>
<td>91.8295</td>
<td>8.1830 ( \approx 8 )</td>
</tr>
</tbody>
</table>

For \( L = 3 \) the problem (2.5.5) can be restated as

\[
\begin{align*}
\text{Min } Z &= \sum_{i=1}^{3} \frac{d_i}{r_i} \\
\text{subject to} & \sum_{i=1}^{3} C_{i2} r_i \leq C_0 \\
& 2 \leq r_i \leq n_{i2} \\
\text{and } r_i's \text{ are integers, } i = 1,2,\ldots,L
\end{align*}
\]

(2.5.10)

The problem (2.5.10) can be rewritten as
If we solve these equations by adopted method we find the solution as \( r_1 = 4, r_2 = 5 \) and \( r_3 = 8 \) with optimum value of objective function \( Z^* = 255.4898 \).

Branch and Bound method to find out the integer values of sub sample of size \( r_i \)'s of non respondents of the problem (2.5.11) as

\[
Min Z = \frac{d_1}{r_1} + \frac{d_2}{r_2} + \frac{d_3}{r_3} \quad (i)
\]

subject to

\[
12r_1 + 14r_2 + 10r_3 \leq 200 \quad (ii)
\]

\[
2 \leq r_1 \leq 5
\]

\[
2 \leq r_2 \leq 6
\]

\[
2 \leq r_3 \leq 9
\]

where \( r_1, r_2 \) and \( r_3 \) are integers \( (iv) \)

The optimum solution of the given problem from figure 2 is \( r_1 = 4, r_2 = 5 \) and \( r_3 = 8 \) with objective function \( Min Z = 255.4898 \).
2.6 Computer Program

The computer program for section 2.4 and section 2.5 as follows:

```cpp
#include<iostream.h>
#include<conio.h>
#include<math.h>
#include<iomanip.h>
class data
{
private:
float pi[10],si[10],ki[10],wi1[10],wi2[10],bi[10];
float ni[10],Ni[10],Ci2[10],di[10],ri[10],ni2[10];
int stratum;
int N;
int n;
int C0;
public:
void read();
void display();
void samplesize(); //ni
void stratumsize(); //Ni
void minimize1();
void minimize2();
}d;
//function defination************************************************************************
void data::read()
{
    cout<<"how many stratum you want to enter\n";
    cin>>stratum;
    for(int i=0;i<stratum;i++)
    {
        cout<<"enter the value of pi\n";
        cin>>pi[i];
        cout<<"enter the value Si\n";
        cin>>si[i];
        cout<<"enter the value of ki\n";
        cin>>ki[i];
        cout<<"enter the value of wi1\n";
        cin>>wi1[i];
        cout<<"enter tthe value of wi2\n";
        cin>>wi2[i];
        cout<<"enter the value of Ci2\n";
        cin>>Ci2[i];
    }
    cout<<"***************888888888888\n";
    cout<<"enter tthe value of C0\n";
    cin>>C0;
    cout<<"enter tthe value of N\n";
    cin>>N;
```
cout<<"enter the value of n\n";
cin>>n;
void data::display()
{
    cout<<"pi="<<setw(5);
    for(int i=0;i<stratum;i++)
    {
        cout<<setw(6)<<pi[i];
    }
    cout<<"\n";
    cout<<"si="<<setw(5);
    for( i=0;i<stratum;i++)
    {
        cout<<setw(6)<<si[i];
    }
    cout<<"\n";
    cout<<"ki="<<setw(5);
    for( i=0;i<stratum;i++)
    {
        cout<<setw(5)<<ki[i];
    }
    cout<<"\n";
    cout<<"wi1="<<setw(5);
    for( i=0;i<stratum;i++)
    {
        cout<<setw(5)<<wi1[i];
    }
    cout<<"\n";
    cout<<"wi2="<<setw(5);
    for( i=0;i<stratum;i++)
    {
        cout<<setw(5)<<wi2[i];
    }
    cout<<"\n";
    cout<<"Ci2="<<setw(5);
    for( i=0;i<stratum;i++)
    {
        cout<<setw(5)<<Ci2[i];
    }
    cout<<"\n******************************\n";
    cout<<N<<endl;
    cout<<n<<endl;
    cout<<C0<<endl;
}
void data::stratumsize()
{
    for(int i=0;i<stratum;i++)
    {
        Ni[i]=pi[i]*N;
cout<<"********the value of Ni**********\n";
for(i=0;i<stratum;i++)
    cout<<Ni[i]<<endl;
}
void data::samplesize()
{
    float b=0;
    for(int i=0;i<stratum;i++)
    {
        bi[i]=(1+(ki[i]-1)*wi2[i])*pow(pi[i],2)*pow(si[i],2);
        b=b+sqrt(bi[i]);
    }
    for(i=0;i<stratum;i++)
    {
        ni[i]=(n*sqrt(bi[i]))/b;
        cout<<"ni="<<ni[i]<<endl;
    }
}
void data::minimize1()
{
    float z1=0;
    for(int i=0;i<stratum;i++)
    {
        z1=z1+bi[i]/ni[i];
    }
    cout<<"n the value of minimized z1 is="<<z1<<endl;
}
void data::minimize2()
{
    float y=0,z2=0;
    for(int i=0;i<stratum;i++)
    {
        ni2[i]=(ni[i]*wi2[i])/(wi1[i]+wi2[i]);
        di[i]=(ni2[i]*wi2[i]*pow(pi[i],2)*pow(si[i],2))/ni[i];
        cout<<"nthe ni2=="<<ni2[i]<<endl;
        cout<<"nthe di=="<<di[i]<<endl;
        y=y+sqrt(di[i]*Ci2[i]);
    }
    for(i=0;i<stratum;i++)
    {
        ri[i]=C0*sqrt(di[i]/Ci2[i])/y;
        cout<<"n ri="<<ri[i]<<endl;
        z2=z2+di[i]/ri[i];
    }
    cout<<"n********the value of minimize z2="<<z2;
}
void main()
{
    clrscr();
}
The Computer Program (in C++ language) of the procedure given in section 2.4 for solving the AINLPP (2.4.1) and of the procedure in section 2.5 for solving the AINLPP (2.5.10) gives the following results respectively. $n_1 = 12$, $n_2 = 15$ and $n_3 = 23$ with objective function $Min Z_1 = 1758.543$ and $r_1 = 4, r_2 = 5$ and $r_3 = 8$ with objective function $Min Z_2 = 255.4898$. 