CHAPTER 2

MULTIVARIATE STRATIFIED SAMPLING: THE COMPROMISE ALLOCATION
CHAPTER 2

MULTIVARIATE STRATIFIED SAMPLING: THE COMPROMISE ALLOCATION

2.1 INTRODUCTION

In multivariate surveys where there are more than one characteristics are defined on each unit of the population the problem of allocation becomes more complicated. Because an allocation which is suitable for one characteristic is usually does not suit other characteristics. Hence there is a need for an efficient multivariate sample allocation method.

The problem of obtaining an optimum allocation in multivariate stratified sampling has drawn the attention of researchers since a long time starting apparently with Neyman (1934). It is felt that unless the strata variance for various characteristics are distributed in the same way, the classical optimum allocation based on the variance of a single character is not of much use because an allocation which is optimum for one characteristics may not be acceptable for others. Due to this fact, there is no unique or even widely accepted solution to the problem of optimum allocation in multivariate stratified sampling. One way to resolve this problem is to search for an allocation, which is in some sense optimum for
all the characteristics using some compromise criterion. Such allocations are known as ‘Compromise allocation’.

In the first part of this chapter the compromise allocation in multivariate stratified sample survey using the compromise criterion “Minimizing the weighted sum of squared coefficients of variation of the estimators of the population means” is worked out. The problem is formulated as an All Integer Non Linear Programming Problem (AINLPP) and relaxing some restrictions Lagrange Multipliers Technique is used to obtain a solution.

A numerical example is presented to illustrate the computational details of the proposed method. The solution is compared with the solution obtained by using the optimization software LINGO (2001).


The later part of this chapter has been devoted to the problem of obtaining a compromise allocation in multivariate surveys when the travel cost within stratum to approach the sampled units is substantial. This work is based on my joint research paper entitled “A Multiple Response Stratified Sampling

2.2 COMPROMISE ALLOCATION

Cochran (1977) suggested that when the characteristics are highly correlated characteristic-wise average of the individual optimum allocations may be used as a compromise allocation. If the correlation between characteristics are not high enough the individual optimum allocations may vary widely and there will be no obvious compromise. In such situations the sampler may use an allocation based on some compromise criterion which is optimum for all characteristics in some sense. In other words the sampler have to use a compromise allocations. The researchers who gave new compromise criterion or explored further the already existing criteria are Neyman (1934), Peter & Bucher (undated), Geary (1949), Dahlenius (1953, 1957), Ghosh (1958), Yates (1960), Aoyama (1963), Gren (1964, 1966), Folk & Antle (1965), Hartley (1965), Chatterjee (1967, 1968), Kokan & Khan (1967), Chatterjee (1972), Ahsan (1975-76, 1978), Ahsan & Khan (1977, 1982), Schittkowski (1985, 1986), Bethel (1985, 1989), Chromy (1987), Wywial (1988), Kreienbrock (1993), Jahan et al. (1994, 2001), Jahan & Ahsan (1995), Khan et al. (1997), Bosch & Wildner (2003), Singh (2003), Khan et al. (2003, 2008), Kozak (2006), Diaz-Garcia and Cortez (2006,

2.3 SOME USEFUL COMPROMISE CRITERIA

When the estimation of the $p$-population means is of interest Yates (1960) gave two useful compromise criteria to work out a compromise allocation. He suggested to minimize the weighted sum of the sampling variances of the estimator $\bar{y}_{jst}$ of the population mean $\bar{Y}_j$ of $j$-th characteristics for a fixed cost to obtain a compromise allocation. Yates suggestion may be formulated as the following Nonlinear Programming Problem (NLPP)

\[
\begin{align*}
\text{Minimize} & \quad \sum_{j=1}^{p} a_j V(\bar{y}_{jst}) \\
\text{Subject to} & \quad c_0 + \sum_{h=1}^{L} c_h n_h \leq C \\
& \quad n_h \geq 0; \quad h = 1, 2, ..., L
\end{align*}
\]  \hspace{1cm} (2.1)

where $a_j > 0$ are the weights assigned to $V(\bar{y}_{jst}); j = 1, 2, ..., p$.

In his second approach he suggested to minimize the total cost of the survey for fixed individual precisions of the estimates. Yates second approach is equivalent to solve the NLPP

34
Minimize \[ \sum_{h=1}^{L} c_h n_h \]

Subject to \( V(\tilde{\mu}_{jst}) \leq \nu_j; j = 1,2,\ldots, p \) \hspace{1cm} (2.2)

\[ n_h \geq 0; \ h = 1,2,\ldots, L \]

where \( \nu_j \) are the tolerance limits for the sampling variance \( V(\tilde{\mu}_{jst}); j = 1,2,\ldots, p \).

Algorithms for solving these NLPPs are given by Hartley and Hocking (1963), Chatterjee (1966), Zukhovitsky and Avdeyeva (1966), Kokan and Khan (1967), Huddleston et al. (1970) etc.

Chatterjee (1967) obtained the compromise allocation by minimizing the sum of the relative increases \( E_j \) in the variances of the estimates \( \tilde{\mu}_{jst} \) of the population mean \( \tilde{\mu}_j \) when a compromise allocation is used for fixed cost.

Chatterjee worked out \( E_j \) as

\[ E_j = \frac{1}{C} \sum c_h (n^*_{jth} - n_h)^2 \] \hspace{1cm} (2.3)

where \( n^*_{jth} \) is the individual optimum allocation for fixed budget \( C \) for the \( j \)-th characteristics in the \( h \)-th stratum and \( n_h \) are the Chatterjee's compromise allocation. These are the solutions to the following \( p \)-NLPPs
\[ \text{Minimize } V(\bar{y}_{jst}) = \sum_{h=1}^{L} \frac{W_h^2 S^2_{jh}}{n_h} \]

\[ \text{Subject to } \sum_{h=1}^{L} c_h n_h \leq C_0 \]

\[ n_h \geq 0; \quad h = 1, 2, ..., L \]

(2.4)

where \( S^2_{jh} = \frac{1}{(N_h - 1)} \sum_{i=1}^{N_h} (y_{jhi} - \bar{y}_{jh})^2 \); \( j = 1, 2, ..., p; \quad h = 1, 2, ..., L \)

(2.5)

is the stratum variance with respect to \( j \)-th characteristics and in \( V(\bar{y}_{jst}) \)
finite population correction (fpc) is ignored.

Cochran (1977) suggested to obtain the compromise allocation by
averaging the \( p \)-individual optimum allocations \( n^*_{jh} \) (obtained by NLPP
(2.4)) over the \( p \)-characteristics.

Cochran's compromise allocation is given by

\[ n_h = \frac{1}{p} \sum_{j=1}^{p} n^*_{jh} \]

(2.6)

Sukhatme et al. (1984) suggested the compromise allocation obtained
by minimizing the trace of the variance-covariance matrix. They gave the
compromise allocation as:
\[ n_h = \frac{C_0 W_h \sqrt{\sum_{j=1}^{p} S_{jh}^2 / c_h}}{\sum_{h=1}^{L} W_h \sqrt{c_h \sum_{j=1}^{p} S_{jh}^2}}; \quad h=1, 2, \ldots, L \]  

(2.7)

The following three compromise criteria are used by Holmberg (2002) to obtain compromise allocations when estimation of the population means are of interest.

(i) Minimizing the sum of sampling variances of the estimators of the population means of various characteristics.

(ii) Minimizing the sum of squared coefficients of variation of the estimator of the population means over the characteristics.

(iii) Minimizing the sum of efficiency losses for not using the individual optimum allocations.

Khan et al. (2003) used the compromise criteria of Yates (1960) with the weights \( a_j \) defined as

\[ a_j = \frac{\sum_{h=1}^{L} S_{jh}^2}{\sum_{j=1}^{p} \sum_{h=1}^{L} S_{jh}^2}; \quad j = 1, \ldots, p \]  

(2.8)

where \( S_{jh}^2 \) are as defined in (2.5).
Many other compromise criteria were also given by various authors. To cover all of them in this thesis is not possible for want of space.

2.4 COMPROMISE ALLOCATION WITH LINEAR COST

In this section a method is proposed to work out the compromise allocation in a multivariate stratified surveys with linear cost function. The problem is formulated as a Multiobjective Integer Nonlinear Programming Problem. Using the value function technique (see Díaz-García and Ulloa (2008), the problem is expressed as a single objective problem. Lagrange Multipliers Technique (LMT) is then used to obtain a formula for continuous sample sizes.

2.5 THE PROBLEM

Consider a multivariate stratified survey with \( p(\geq 2) \) characteristics. Let the estimation of the \( p \) unknown population means \( \bar{Y}_j; \ j = 1, 2, \ldots, p \) be of interest.

The population mean of the \( j \)-th characteristics is

\[
\bar{Y}_j = \sum_{h=1}^{L} w_h \bar{y}_{jh}; \ j = 1, 2, \ldots, p 
\]  

(2.9)

where \( y_{jhi} \) is the value of the \( j \)-th study variable for of the \( i \)-th unit of the \( h \)-th stratum; \( j = 1, 2, \ldots, p; \ h = 1, 2, \ldots, L \).
\[
\bar{Y}_{jh} = \frac{1}{N_h} \sum_{i=1}^{N_h} y_{jhi}; \ h = 1,2,\ldots,L \quad \text{are the stratum means.}
\]

The stratified sampling mean
\[
\bar{y}_{jst} = \sum_{h=1}^{L} W_h \bar{Y}_{jh}; \ j = 1,2,\ldots,p \quad (2.10)
\]
serves as an unbiased estimate of population mean \( \bar{Y}_{j} \), with a sampling variance
\[
V(\bar{y}_{jst}) = \sum_{h=1}^{L} W_h^2 \left( 1 - \frac{n_h}{N_h} \right) \frac{S_{jh}^2}{n_h}; \ j = 1,2,\ldots,p \quad (2.11)
\]
where \( S_{jh}^2 \) are as defined in (2.5).

A compromise allocation which minimizes all the \( p \) variances (given by (2.11)), simultaneously will be the best suited for a multivariate sample survey. Thus for obtaining the optimum compromise allocation we need to solve the following Multi-objective Integer Nonlinear Programming Problem (MINLPP).
Minimize \[ \left\{ \begin{array}{l} V(\bar{y}_{1st}) \\ \vdots \\ V(\bar{y}_{pst}) \end{array} \right. \]

Subject to \[ \sum_{h=1}^{L} c_h n_h \leq C_0 \]
\[ 2 \leq n_h \leq N_h \]

and \( n_h \) are integers; \( h = 1, 2, ..., L \)

(2.12)

Where the restrictions \( 2 \leq n_h \leq N_h; h = 1, 2, ..., L \) are imposed on the sample sizes to have an estimate of the stratum variances and to avoid the problem of over sampling. The integer values of the sample sizes are required for their practical implication.

Some of them who studied the multiobjective optimization are Steuer (1986), Rios et al. (1989) and Miettinen (1999) etcetera. Using the value function technique Díaz-García and Ulloa (2008) expressed this problem as

Minimize \[ \phi \left[ V \left( \bar{y}_{jst} \right) \right] \]

Subject to \[ \sum_{h=1}^{L} c_h n_h \leq C_0 \]
\[ 2 \leq n_h \leq N_h \]

and \( n_h \) are integers; \( h = 1, 2, ..., L \)

(2.13)

In (2.13) \( \phi() \) is a scalar function that characterizes the importance of the \( p \)-variances. One way to define \( \phi \) is
\[
\phi = \sum_{j=1}^{p} w_j V(\bar{y}_{jst})
\]  
(2.14)

which is the weighted sum of the \( p \) variances.

Where \( \sum_{j=1}^{p} w_j = 1 \), \( w_j \geq 0 \), \( j = 1, 2, \ldots, p \), \( w_j \) are the weights assigned to the variances according to their relative importance (See Yates (1960)).

An objection on the above definition of \( \phi \) is that since the variances are measured in different units therefore they cannot be added. To avoid this the squared coefficients of variation that are unit free and positive are used.

A problem parallel to (2.13) may thus be expressed in terms of the squared coefficients of variance as

\[
\text{Minimize} \quad Z = \sum_{j=1}^{p} w_j (CV)^2_j
\]

Subject to \( \sum_{h=1}^{L} c_h n_h \leq C_0 \)

\( 2 \leq n_h \leq N_h \)

and \( n_h \) are integers; \( h = 1, 2, \ldots, L \)

\[
(CV)^2_j = \left[ \sum_{h=1}^{L} W_h^2 \left( 1 - \frac{n_h}{N_h} \right) \frac{S_{jh}^2}{n_h} \right] \left/ \bar{y}_{jst}^2 \right. ; \quad j = 1, 2, \ldots, p
\]  
(2.15)
where $\bar{y}_j, \bar{y}_{jst}$ and $V(\bar{y}_{jst})$ are as defined in (2.9), (2.10) and (2.11) respectively and $\left(CV\right)_j = \frac{SD(\bar{y}_{jst})}{\bar{y}_j}$ denote the $j$-th coefficient of variation.

AINLPP (2.15) may now be restated as:

$$
\text{Minimize } Z = \sum_{j=1}^{p} \frac{1}{\bar{y}_j^2} \left( \sum_{h=1}^{L} w_h^2 \left( 1 - \frac{n_h}{N_h} \right) \frac{S_{jh}^2}{n_h} \right)
$$

Subject to

$$
\begin{align*}
\sum_{h=1}^{L} c_h n_h & \leq C_0 \\
2 & \leq n_h \leq N_h \\
\text{and } n_h & \text{ are integers; } h = 1, 2, ..., L
\end{align*}
$$

Where the value of $(CV)^2_j$ is substituted from (2.16), the $p$ weights are assumed to be equal, that is, $w_1 = w_2 = \cdots = w_p = \frac{1}{p}$, and the common factor $\frac{1}{p}$ is dropped from the objective function.

2.6 THE SOLUTION

With some relaxations in the restrictions Lagrange Multipliers Technique may be used to solve NLPP (2.17). If equality is assumed in the cost constraint and the integer restrictions and the restrictions $2 \leq n_h \leq N_h; h = 1, 2, ..., L$ are ignored, Lagrange Multipliers Technique can be applied to NLPP (2.17). The objective function in NLPP (2.17) is convex.
in $n_h$ and the cost constraint is linear this ensures that the continuous solution of the NLPP (2.17) with relaxations will be a boundary point of the convex feasible region on the hyperplane $\sum_{h=1}^{L} c_h n_h = C_0$. Thus the above relaxations are without loss of generality.

With $\lambda$ as Lagrange Multiplier the Lagrangian function $L(n_h, \lambda)$ may be defined as

$$L(n_h, \lambda) = \sum_{h=1}^{L} \sum_{j=1}^{p} \frac{W^2_h S^2_{jh}}{\overline{Y}_j^2 n_h} \left(1 - \frac{n_h}{N_h}\right) + \lambda \left(\sum_{h=1}^{L} c_h n_h - C_0\right)$$  (2.18)

Differentiating (2.18) with respect to $n_h$; $h=1,2,\ldots,L$ and $\lambda$ partially and equating to zero we get the following set of $L+1$ simultaneous equations in $n_h$ and $\lambda$.

$$\frac{\partial L(n_h, \lambda)}{\partial n_h} = -\sum_{j=1}^{p} \frac{W^2_h S^2_{jh}}{\overline{Y}_j^2 n_h^2} + \lambda c_h = 0; \ h=1,2,\ldots,L$$  (2.19)

and

$$\frac{\partial L(n_h, \lambda)}{\partial \lambda} = \sum_{h=1}^{L} c_h n_h - C_0 = 0$$  (2.20)

From (2.19)

$$n_h = \sqrt{\frac{W^2_h}{\lambda c_h} \sum_{j=1}^{p} \frac{S^2_{jh}}{\overline{Y}_j^2}}$$  (2.21)
Substituting this value of \( n_h \) in (2.20) we get

\[
\frac{1}{\sqrt{\lambda}} = \frac{C_0}{\left\{ \sum_{h=1}^{L} c_h W_h^2 \sum_{j=1}^{P} \frac{S_{jh}^2}{\bar{Y}_j^2} \right\}^{\frac{1}{2}}}
\]  

(2.22)

From (2.21) and (2.22)

\[
n_h = \frac{C_0 W_h \sqrt{\frac{\sum_{j=1}^{P} S_{jh}^2}{\bar{Y}_j^2}}}{\sum_{h=1}^{L} c_h}
\]  

(2.23)

or

\[
n_h = \frac{C_0 W_h \sqrt{A_h / c_h}}{\sum_{h=1}^{L} W_h \sqrt{A_h c_h}}
\]  

(2.24)

where

\[ A_h = \sum_{j=1}^{P} \frac{S_{jh}^2}{\bar{Y}_j^2}, \quad h = 1, 2, \ldots, L. \]

If the restrictions \( 2 \leq n_h \leq N_h \) are satisfied by \( n_h; \ h = 1, 2, \ldots, L \) given by (2.24) then a continuous solution of NLPP (2.17) is obtained. This continuous solution, rounded off to the nearest integer values, will provide an integer solution. After rounding off we must check that the rounded off
values satisfy the cost constraint \[ \sum_{h=1}^{L} c_h n_h \leq C_0 \] or not. If not the rounding off may be adjusted accordingly to get an approximate integer solution.

For large \( N_h \), usually \( n_h \) given by (2.24) satisfy the restrictions \( 2 \leq n_h \leq N_h \); \( h = 1, 2, \ldots, L \) automatically.

However if any \( n_h > N_h \) it can be taken as equal to \( N_h \) and for the remaining \( (L - 1) \) strata \( n_h \) are obtained afresh. Similarly if any \( n_h < 2 \) it can be put equal to 2 and the remaining \( n_h \) are obtained afresh.

The restriction \( 2 \leq n_h \) may also be taken care by reducing the total sample size by \( 2L \) assuming that 2 units are already allocated to each strata. This is equivalent to substitute \( n'_h = n_h + 2 \) or \( n'_h = n'_h - 2 \), in (2.17).

As an alternative the constraints \( 2 \leq n_h \leq N_h \); \( h = 1, 2, \ldots, L \) may also be included and the AINLPP (2.17) may be solved by some integer nonlinear programming technique. Softwares are also available to solve AINLLP. One such software is LINGO mentioned in chapter 1 of this thesis.

2.7 A NUMERICAL ILLUSTRATION

State are divided into four strata. The relevant data with respect to two characteristics (i) the quantity of corn harvested $Y_1$, and (ii) the quantity of oats harvested $Y_2$ are given in Table 2.1.

**Table 2.1**
Data for 4 strata and 2 characteristics

<table>
<thead>
<tr>
<th>$h$</th>
<th>$N_h$</th>
<th>$W_h$</th>
<th>$S_{1h}^2$</th>
<th>$S_{2h}^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8</td>
<td>0.0808</td>
<td>21601503189.8</td>
<td>1154134.2</td>
</tr>
<tr>
<td>2</td>
<td>34</td>
<td>0.3434</td>
<td>19734615816.7</td>
<td>7056074.8</td>
</tr>
<tr>
<td>3</td>
<td>45</td>
<td>0.4545</td>
<td>27129658750.0</td>
<td>2082871.3</td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>0.1212</td>
<td>17258237358.5</td>
<td>732004.9</td>
</tr>
</tbody>
</table>

The population means $\bar{Y}_1$ and $\bar{Y}_2$ are assumed to be known as $\bar{Y}_1 = 405654.19$ and $\bar{Y}_2 = 2116.70$. However, they are unknown in real surveys, and either some approximations of their values are used or values known from a recent or preliminary survey are used (Kozak (2006)).

The costs are assumed as follows.

Total cost $C = 350$ units

Overhead cost $c_0 = 70$ units

Per unit measurement cost $c_h$ in the four strata $c_1 = 10, c_2 = 5, c_3 = 3$ and $c_4 = 7$ units respectively.

The amount available for measurement $C_0 = C - c_0 = 350 - 70 = 280$ units.
**Table 2.2**  
The compromise allocation for four strata

<p>| | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
<td>(7)</td>
<td>(8)</td>
<td>(9)</td>
<td>(10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c_h</td>
<td>(S^2_{h}/F^2_1)</td>
<td>(S^2_{2h}/F^2_2)</td>
<td>(A_h = (4) + (5))</td>
<td>(C_0W_h\sqrt{A_h/c_h})</td>
<td>(W_h\sqrt{c_hA_h})</td>
<td>(n_h = (7)/\sum(8))</td>
<td>Rounded off (n_h)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.131271982</td>
<td>0.257595</td>
<td>0.388866982</td>
<td>4.461387</td>
<td>0.159335</td>
<td>2.288161</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.119926938</td>
<td>1.574868</td>
<td>1.694795938</td>
<td>55.979888</td>
<td>0.999641</td>
<td>28.711021</td>
<td>29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.164866493</td>
<td>0.464883</td>
<td>0.629749493</td>
<td>65.806262</td>
<td>0.624710</td>
<td>29.904174</td>
<td>30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.104878026</td>
<td>0.163379</td>
<td>0.268257026</td>
<td>6.643350</td>
<td>0.166084</td>
<td>3.407248</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(\sum = 1.94977)</td>
<td></td>
<td></td>
<td>(\sum = 64)</td>
</tr>
</tbody>
</table>
Based on the data of Table 2.1 the compromise allocation is worked out and is given in column 10 of Table 2.2. The value of the objective function at this solution is \( Z = 0.004 \).

### 2.8 COMPROMISE ALLOCATION USING LINGO

Using the relevant data from Table 2.1 and Table 2.2 the instance of the AINLPP (2.17) is

\[
\begin{align*}
\text{Minimize} \quad Z &= \{(0.000857027)(\frac{1}{n_1} - \frac{1}{8}) + (0.014142211)(\frac{1}{n_2} - \frac{1}{34}) \\
& \quad + (0.034056512)(\frac{1}{n_3} - \frac{1}{45}) + (0.001540599)(\frac{1}{n_4} - \frac{1}{12})\} \\
& \quad + \{(0.0019681745)(\frac{1}{n_1} - \frac{1}{8}) + (0.185714041)(\frac{1}{n_2} - \frac{1}{34}) \\
& \quad + (0.096030997)(\frac{1}{n_3} - \frac{1}{45}) + (0.002399946)(\frac{1}{n_4} - \frac{1}{34})\} \\
\text{Subject to} \quad & 10n_1 + 5n_2 + 3n_3 + 7n_4 \leq 280 \\
& 2 \leq n_1 \leq 8 \\
& 2 \leq n_2 \leq 34 \\
& 2 \leq n_3 \leq 45 \\
& 2 \leq n_4 \leq 12 \\
\text{and} \quad & n_h \text{ are integers; } h = 1, 2, 3, 4
\end{align*}
\]

Using optimization software LINGO the following compromise allocation is obtained as a solution to (2.25) \( n_1 = 2, n_2 = 29, n_3 = 29, n_4 = 4 \). The minimum value of the objective function is obtained as \( Z = 0.004325 \). This allocation differs only by one unit in third and fourth strata from the
allocation obtained by the formula (2.24). The difference in the values of the
objective function is also negligible.

This shows that the formula (2.24) gives a near optimum solution to
the AINLPP (2.17) and can be used without any significant loss in precision.

2.9 COMPROMISE ALLOCATION WITH TRAVEL COST

To collect the information from the units selected in the sample from a
particular stratum the investigator has to travel from unit to unit. If the
stratum consist of large geographical and difficult to travel area it may be
costly to travel between the selected units. In this situation the linear cost
function given in (2.1) will not be an adequate approximation to the actual
cost incurred. The investigator will have to spend a significant amount on
travel between the selected units.

Beardwood et al. (1959) suggested that the cost of visiting the $n_h$
selected units in the $h$-th stratum may be taken as $t_h \sqrt{n_h}$ ; $h = 1, 2, ..., L$
approximately, where $t_h$ is the travel cost per unit in the $h$-th stratum. This
conjecture is based on the fact that the distance between $k$ randomly
scattered points is proportional to $\sqrt{k}$.

Under the above situation the total cost of a stratified sample survey will be
the sum of (i) the overhead cost, (ii) the measurement cost, and (iii) the
travel cost.
This gives the total cost $C$ as

$$C = c_o + \sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} t_h \sqrt{n_h}$$

(2.26)

which is quadratic in $\sqrt{n_h}$.

When the travel cost is significant and varies from stratum to stratum, that is the cost function is as given in (2.26) the problem of finding the optimum allocation may be given as the following Nonlinear Programming Problem (NLPP):

$$\text{Minimize } V(\bar{y}_{jsl}) = \sum_{h=1}^{L} \left( \frac{1}{n_h} - \frac{1}{N_h} \right) W_h^2 S_h^2$$

Subject to

$$\frac{\sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} t_h \sqrt{n_h}}{C_0} \leq C_0$$

(2.27)

and $n_h \geq 0; \ h = 1, 2, ..., L$

where $C_0 = C - c_o$, is the amount available to meet the travel and measurement expenses.

For solving the NLPP (2.27) if Lagrange Multipliers Technique is to be used one has to take the cost constraint as an equation and the non-negativity restrictions are to be ignored. The Lagrangian function is defined as
\[ L(n_h, \lambda) = \sum_{h=1}^{L} \frac{W_{h}^2 S_{h}^2}{n_h} + \lambda \left( \sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} t_h \sqrt{n_h} - C_0 \right) \] 

(2.28)

where \( \lambda \) is the Lagrange Multiplier.

Differentiating (2.28) with respect to \( n_h \); \( h = 1, 2, ..., L \) and \( \lambda \) partially and equating to zero, we get the following \((L+1)\) equations

\[
\frac{\partial L(n_h, \lambda)}{\partial n_h} = -\frac{W_{h}^2 S_{h}^2}{n_h^2} + \lambda \left( c_h + \frac{t_h}{2 \sqrt{n_h}} \right) = 0; \quad h = 1, 2, ..., L
\] 

(2.29)

\[
\frac{\partial L(n_h, \lambda)}{\partial \lambda} = \left( \sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} t_h \sqrt{n_h} - C_0 \right) = 0
\] 

(2.30)

Equations (2.29) are implicit equations in \( n_h \). An exact solution of the system is very tedious if not impossible. However, an approximate solution may be obtained. In the absence of an explicit solution, if the numerical values of \( W_h, S_h, c_h, t_h, c_o \) and \( C \) are available, the software package LINGO can be used to solve the NLPP (2.27).

It is assumed that the cost of traveling \( (t_h) \) with in stratum to contact the selected units is significant and the cost function is of the form as given in (2.26), that is, quadratic in \( \sqrt{n_h} \). In the next section the problem of allocation in multivariate stratified sample surveys with \( p \) -independent characteristics is formulated as a multiobjective NLPP. The \( 'p' \) objectives
are to minimize the individual variances of the estimates of the population means of \( p \)-characteristics simultaneously, subject to the cost constraint and the restrictions on the sample sizes. The formulated multiobjective NLPP is then solved by the “Goal Programming Technique” using software package LINGO.

2.10 THE PROBLEM

The Multiobjective Non-linear Programming Problem (MNLPP) discussed in the previous section may be expressed as:

\[
\begin{align*}
\text{Minimize} & \quad \begin{pmatrix} V(\bar{y}_{1st}) \\ V(\bar{y}_{2st}) \\ \vdots \\ V(\bar{y}_{pst}) \end{pmatrix} \\
\text{Subject to} & \quad \sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} t_h \sqrt{n_h} \leq C_0 \\
\text{and} & \quad 2 \leq n_h \leq N_h; \quad h = 1, 2, ..., L 
\end{align*}
\]  

(2.31)

where \( V(\bar{y}_{jst}) = \sum_{h=1}^{L} \frac{W_h^2 S_{jh}^2}{n_h} \); \( j = 1, 2, ..., p \)  

(2.32)

denote the sampling variance (ignoring fpc) of the estimate

\[
\bar{y}_{jst} = \sum_{h=1}^{L} W_h \bar{y}_{jh}
\]  

(2.33)
of the overall population mean \( \bar{Y}_j \); \( j=1,2,\ldots,p \) of the \( j \)-th characteristic.

\[
\bar{y}_{jh} = \frac{1}{n_h} \sum_{k=1}^{n_h} y_{jhk}
\]

is the sample mean from the \( h \)-th stratum for the \( j \)-th characteristic and \( y_{jhk} \) is the value of the \( k \)-th selected unit of the sample from the \( h \)-th stratum for the \( j \)-th characteristic; \( k = 1,2,\ldots,n_h; h = 1,2,\ldots,L; j = 1,2,\ldots,p \).

It is assumed that the true values of \( S_{jh}^2 \) are known. If not known, approximation of these parameters may be obtained from some recent or preliminary survey, may be used.

2.11 THE GOAL PROGRAMMING APPROACH

To solve the problem (2.31) using goal programming, we first solve the following \( p \) Non Linear Programming Problems (NLPPs) for all the ‘\( p \)’ characteristics separately.

\[
\begin{align*}
\text{Minimize} & \quad \sum_{h=1}^{L} \frac{W_h S_{jh}^2}{n_h} \\
\text{Subject to} & \quad \sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} l_h \sqrt{n_h} \leq C_0 \quad j = 1,2,\ldots,p \\
\text{and} & \quad 2 \leq n_h \leq N_h; \quad h = 1,2,\ldots,L
\end{align*}
\]
Let \( n_j^* = (n_{j1}^*, n_{j2}^*, ..., n_{jL}^*) \) denote the solution to the \( j \)-th NLPP in (2.34) with \( V_j^* \) as the value of the objective function given by

\[
V_j^* = \sum_{h=1}^{L} \frac{W_h^2 S_{jh}^2}{n_{jh}^*}; \quad j = 1, 2, ..., p
\]  

(2.35)

Further, let \( n_c^* = (n_{1c}^*, n_{2c}^*, ..., n_{Lc}^*) \) be the vector of optimum compromise allocations with

\[
V_{cj}^* = \sum_{h=1}^{L} \frac{W_h^2 S_{jh}^2}{n_{hc}^*}
\]

as the optimal value of the objective function for \( j \)-th characteristics under this allocation.

Obviously, \( V_{cj}^* \geq V_j^* \) or \( V_{cj}^* - V_j^* \geq 0; \quad j = 1, 2, ..., p \)  

(2.36)

A reasonable criterion to workout a compromise allocation may be to “Minimize the sum of increases in the variances \( V_j \); \quad j = 1, 2, ..., p \) due to the use of the compromise allocation”. We may express the multiobjective NLPP (2.31) using (2.36) with the above compromise criterion as the following single objective NLPP
**Multivariate Stratified...**

Minimize \[ \sum_{j=1}^{p} x_j \]

Subject to \[ V_{cj} - V_{j}^{*} \leq x_j \]
\[ \sum_{h=1}^{L} c_{h n_{h}} + \sum_{h=1}^{L} t_{h} \sqrt{n_{h}} \leq C_0 \]
\[ x_j \geq 0 \]
\[ j = 1, 2, ..., p \]  \( (2.37) \)

and \[ 2 \leq n_{h} \leq N_{h} \quad ; \quad h = 1, 2, ..., L_j \]

where \( n_c = (n_{1c}, n_{2c}, ..., n_{Lc}) \) denotes a compromise allocation with variance

\[ V_{cj} = \sum_{h=1}^{L} \frac{W_{h}^{2} S_{j}^{2}}{n_{ch}} ; \quad j = 1, 2, ..., p \]  \( (2.38) \)

and \( x_j \geq 0 \); \( j = 1, 2, ..., p \) are called goal variables whose values are to be determined.

The ‘Goal’ is to “Find the compromise allocation \( n_c^{*} = (n_{1c}^{*}, n_{2c}^{*}, ..., n_{Lc}^{*}) \) such that the increases in the \( j \)-th variance due to the use of compromise allocation should not exceed \( x_j \); \( j = 1, 2, ..., p \) and

\[ \sum_{j=1}^{p} x_j \] is minimum”. NLPP \( (2.37) \) may be restated as:

55
minimize \[ \sum_{j=1}^{p} x_j \]

subject to \[ \sum_{h=1}^{L} \frac{W_h^2 S_j^2}{n_h} - x_j \leq V_j^* \]

\[ \sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} t_h \sqrt{n_h} \leq C_0 \]

\[ x_j \geq 0 \]

and \[ 2 \leq n_h \leq N_h; \quad h=1,2,\ldots,L \]

where the value of \( V_{ej} \) is substituted from (2.38) and the compromise allocation \( n_{hc}; \quad h=1,2,\ldots,L \) is replaced by \( n_h \) for simplicity. The optimal solution to the NLPP (2.39) will be the required optimum compromise allocation \( n^*_c \) that minimizes the sum of deviations of the variances from their optimum values. NLPP (2.39) may be solved by using software package LINGO.

2.12 SOME OTHER COMPROMISE ALLOCATIONS WITH TRAVEL COST

In this section three other compromise allocations are discussed for the sake of comparison with the proposed allocation.

2.12.1 The Proportional Allocation

Because of its simplicity the proportional allocation is the most commonly used allocation in stratified sample surveys. In the proportional
allocation the sample size from the $h$-th stratum is proportional to its stratum weights, that is,

$$n_h \propto W_h; \quad h = 1, 2, ..., L$$

or

$$n_h = k W_h; \quad h = 1, 2, ..., L$$

where $k > 0$ is the constant of proportionality.

From (2.40)

$$\sum_{h=1}^{L} n_h = k \sum_{h=1}^{L} W_h$$

or, $n = k$, where $n$ is the total sample size. Thus (2.40) gives

$$n_h = n W_h; \quad h = 1, 2, ..., L.$$  (2.41)

To work out the value of the total sample size $n$ for fixed cost we proceed as follows. Substitution of the values of $n_h$ from (2.41) in the cost function (2.26) with $C_0 = C - c_o$

we get

$$C_0 = n \sum_{h=1}^{L} c_h W_h + \sqrt{n} \sum_{h=1}^{L} t_h \sqrt{W_h}$$  (2.42)

The RHS of (2.42) is quadratic in $\sqrt{n}$ which is equivalent to

$$AX^2 + BX - C_0 = 0$$  (2.43)

where $\sum_{h=1}^{L} c_h W_h = A > 0$, $\sum_{h=1}^{L} t_h \sqrt{W_h} = B > 0$ and $\sqrt{n} = X > 0$.

The roots of (2.43) are
\[ X = \frac{-B \pm \sqrt{B^2 + 4AC_0}}{2A} \]

We have the only usable root as

\[ X = \frac{-B + \sqrt{B^2 + 4AC_0}}{2A} > 0 \quad (2.44) \]

as \( \sqrt{B^2 + 4AC_0} > B \) because \( 4AC_0 > 0 \).

When the numerical values of \( A, B \) and \( C_0 \) are available we can easily compute the value \( X \) and \( X^2 \) will give the total sample size \( n \). Substitution of the value of \( n \) in (2.41) gives the proportional allocation.

### 2.12.2 Cochran’s Compromise Allocation

Cochran (1977) gave the compromise criteria by averaging the individual optimum allocations \( n_{jh}^* \) that are solutions to NLPP (2.34) for \( j = 1, 2, \ldots, p \), over the characteristics.

Cochran’s compromise allocation is given by

\[ n_h = \frac{1}{p} \sum_{j=1}^{p} n_{jh}^* \quad (2.45) \]
2.12.3 Minimizing Weighted Sum of Variances

To work out a compromise allocation Khan et al. (2003) used the compromise criteria as “Minimize $\sum_{j=1}^{p} a_j V_j$, where $a_j > 0$ is the weights assigned to $V_j$.

The above compromise criterion was first used by Yates (1960).

Khan et al. (2003) conjectured that $a_j = \frac{\sum_{h=1}^{L} S_{jh}^2}{\sum_{j=1}^{p} \sum_{h=1}^{L} S_{jh}^2}$. It can be seen that $\sum_{j=1}^{p} a_j = 1$.

To compare their allocation with the proposed allocation the Khan et al. (2003) compromise allocation is worked out with a quadratic cost function in the following.

The objective function of this problem may be expressed as:
\[ Z(n_1, n_2, \ldots, n_L) = \sum_{j=1}^{p} a_j V_j \]
\[ = \sum_{j=1}^{p} a_j \left( \sum_{h=1}^{L} \frac{W_h^2 S_{jh}^2}{n_h} \right) \]
\[ = \sum_{h=1}^{L} \sum_{j=1}^{p} \frac{W_h^2 \left( a_j S_{jh}^2 \right)}{n_h} \]
\[ = \sum_{h=1}^{L} \frac{W_h^2 A_h^2}{n_h}, \]

where \( A_h^2 = \sum_{j=1}^{p} a_j S_{jh}^2; \ h = 1, 2, \ldots, L. \)

The NLPP for finding the optimum compromise allocation according to \textit{Khan et al.} (2003) compromise criterion may be given as

\[
\begin{align*}
\text{Minimize } Z &= \sum_{h=1}^{L} \frac{W_h^2 A_h^2}{n_h} \\
\text{Subject to } & \sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} t_h \sqrt{n_h} \leq C_0 \quad j = 1, 2, \ldots, p. \\
\text{and } & 2 \leq n_h \leq N_h; \quad h = 1, 2, \ldots, L
\end{align*}
\]  

(2.46)

When the numerical values of \( W_h, A_h, c_h, t_h, C_0 \) and \( N_h \) are available, LINGO optimization software may be used to obtain a solution.
2.13 A NUMERICAL ILLUSTRATION

In the table below the stratum sizes, stratum weights, stratum standard deviations, measurement costs, and the travel costs within stratum are given for four different characteristics under study in a population stratified in five strata. The data are mainly from Chatterjee (1968). The values of strata sizes are added assuming the population size $N = 6000$. The traveling cost $t_h$ is also assumed for the five strata by the authors.

<table>
<thead>
<tr>
<th>$h$</th>
<th>$N_h$</th>
<th>$W_h$</th>
<th>$c_h$</th>
<th>$t_h$</th>
<th>$S_{jh}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$S_{1h}$</td>
</tr>
<tr>
<td>1</td>
<td>1500</td>
<td>0.25</td>
<td>1</td>
<td>0.5</td>
<td>28</td>
</tr>
<tr>
<td>2</td>
<td>1920</td>
<td>0.32</td>
<td>1</td>
<td>0.5</td>
<td>24</td>
</tr>
<tr>
<td>3</td>
<td>1260</td>
<td>0.21</td>
<td>1.5</td>
<td>1</td>
<td>32</td>
</tr>
<tr>
<td>4</td>
<td>480</td>
<td>0.08</td>
<td>1.5</td>
<td>1</td>
<td>54</td>
</tr>
<tr>
<td>5</td>
<td>840</td>
<td>0.14</td>
<td>2</td>
<td>1.5</td>
<td>67</td>
</tr>
</tbody>
</table>

The total budget of the survey is assumed to be 1500 units with an overhead cost $c_o = 300$ units. Thus $C_0 = C - c_o = 1500 - 300 = 1200$ units are available for measurement and travel within stratum for approaching the selected units for measurement.
2.13.1 The Proposed Compromise Allocation

Using the values given in Table 2.3 the NLPP (2.34) and their optimal solutions \( n_j^* ; j = 1, 2, 3, 4 \) with the corresponding values of \( V_j^* \) are given below. These values are obtained by software LINGO.

For \( j = 1 \)

\[
\begin{align*}
\text{Minimize} & \quad \frac{49}{n_1} + \frac{58.9824}{n_2} + \frac{45.1584}{n_3} + \frac{18.6624}{n_4} + \frac{87.9844}{n_5} \\
\text{Subject to} & \quad \ln_1 + \ln_2 + 1.5n_3 + 1.5n_4 + 2n_5 + 0.5\sqrt{n_1} + 0.5\sqrt{n_2} + 1\sqrt{n_3} + 1\sqrt{n_4} + 1.5\sqrt{n_5} \leq 1200 \\
& \quad 2 \leq n_1 \leq 1500 \\
& \quad 2 \leq n_2 \leq 1920 \\
& \quad 2 \leq n_3 \leq 1260 \\
& \quad 2 \leq n_4 \leq 480 \\
\text{and} & \quad 2 \leq n_5 \leq 840 \\
\end{align*}
\]

(2.47)

The optimum allocation \( n_1^* = (n_1^*, n_2^*, n_3^*, n_4^*, n_5^*) \) is

\( n_{11}^* = 193.6535, \quad n_{12}^* = 212.5509, \quad n_{13}^* = 151.1150, \quad n_{14}^* = 96.82734, \)

\( n_{15}^* = 182.6156. \) The corresponding value of the variance ignoring fpc is \( V_1^* = 1.503902 \).
For $j = 2$

Minimize \[ \frac{2652.25}{n_1} + \frac{1811.3536}{n_2} + \frac{101.6064}{n_3} + \frac{8.7616}{n_4} + \frac{1.5876}{n_5} \]

Subject to \[ l n_1 + 1 n_2 + 1.5 n_3 + 1.5 n_4 + 2 n_5 + \\
0.5 \sqrt{n_1} + 0.5 \sqrt{n_2} + 1 \sqrt{n_3} + 1 \sqrt{n_4} + 1.5 \sqrt{n_5} \leq 1200 \]

and

\[ \begin{aligned}
2 & \leq n_1 \leq 1500 \\
2 & \leq n_2 \leq 1920 \\
2 & \leq n_3 \leq 1260 \\
2 & \leq n_4 \leq 480 \\
2 & \leq n_5 \leq 840
\end{aligned} \] \hfill (2.48)

The optimum allocation \( n_2^* = (n_{21}^*, n_{22}^*, n_{23}^*, n_{24}^*, n_{25}^*) \) is

\[ n_{21}^* = 535.7324, \quad n_{22}^* = 442.4963, \quad n_{23}^* = 84.55834, \quad n_{24}^* = 24.46372, \]

\[ n_{25}^* = 8.779119. \] The corresponding value of the variance ignoring fpc is \( V_2^* = 10.78476 \).

For $j = 3$

Minimize \[ \frac{90.25}{n_1} + \frac{69.2224}{n_2} + \frac{85.3776}{n_3} + \frac{38.9376}{n_4} + \frac{113.2096}{n_5} \]

Subject to \[ l n_1 + 1 n_2 + 1.5 n_3 + 1.5 n_4 + 2 n_5 + \\
0.5 \sqrt{n_1} + 0.5 \sqrt{n_2} + 1 \sqrt{n_3} + 1 \sqrt{n_4} + 1.5 \sqrt{n_5} \leq 1200 \]

and

\[ \begin{aligned}
2 & \leq n_1 \leq 1500 \\
2 & \leq n_2 \leq 1920 \\
2 & \leq n_3 \leq 1260 \\
2 & \leq n_4 \leq 480 \\
2 & \leq n_5 \leq 840
\end{aligned} \] \hfill (2.49)

The optimum allocation \( n_3^* = (n_{31}^*, n_{32}^*, n_{33}^*, n_{34}^*, n_{35}^*) \) is
\[ n_{31}^* = 210.3325, \quad n_{32}^* = 184.0994, \quad n_{33}^* = 166.3332, \quad n_{34}^* = 112.0209, \]
\[ n_{35}^* = 165.6085. \] The corresponding value of the variance ignoring fpc is
\[ V_3^* = 2.349571. \]

For \( j = 4 \)

\[
\begin{aligned}
\text{Minimize} & \quad \frac{900}{n_1} + \frac{3466.8544}{n_2} + \frac{1319.8689}{n_3} + \frac{54.1696}{n_4} + \frac{268.3044}{n_5} \\
\text{Subject to} & \quad 1n_1 + 2.5n_2 + 1.5n_3 + 1.5n_4 + 2n_5 + \\
& \quad 0.5\sqrt{n_1} + 0.5\sqrt{n_2} + 1\sqrt{n_3} + 1\sqrt{n_4} + 1.5\sqrt{n_5} \leq 1200 \\
\text{and} & \quad 2 \leq n_1 \leq 1500 \\
& \quad 2 \leq n_2 \leq 1920 \\
& \quad 2 \leq n_3 \leq 1260 \\
& \quad 2 \leq n_4 \leq 480 \\
& \quad 2 \leq n_5 \leq 840 \\
\end{aligned}
\]

(2.50)

The optimum allocation \( n_4^* = (n_{41}^*, n_{42}^*, n_{43}^*, n_{44}^*, n_{45}^*) \) is
\[ n_{41}^* = 208.6400, \quad n_{42}^* = 410.4944, \quad n_{43}^* = 205.6995, \quad n_{44}^* = 41.09868, \]
\[ n_{45}^* = 79.59035. \] The corresponding value of the variance ignoring fpc is
\[ V_4^* = 23.86480. \]

Using the computed values of \( V_j^*; \ j = 1, 2, 3, 4 \) and the compromise criterion conjectured in section 3, the Goal Programming Problem given in (2.39) may be expressed as
Minimize \[ \sum_{j=1}^{4} x_j \]

Subject to
\[ \frac{49}{n_1} + \frac{58.9824}{n_2} + \frac{45.1584}{n_3} + \frac{18.6624}{n_4} + \frac{87.9844}{n_5} - x_1 \leq 1.503902 \]
\[ \frac{2652.25}{n_1} + \frac{1811.3536}{n_2} + \frac{101.6064}{n_3} + \frac{8.7616}{n_4} + \frac{1.5876}{n_5} - x_2 \leq 10.78476 \]
\[ \frac{90.25}{n_1} + \frac{69.2224}{n_2} + \frac{85.3776}{n_3} + \frac{38.9376}{n_4} + \frac{113.2096}{n_5} - x_3 \leq 2.349571 \]
\[ \frac{900}{n_1} + \frac{3466.8544}{n_2} + \frac{1319.8689}{n_3} + \frac{54.1696}{n_4} + \frac{268.3044}{n_5} - x_4 \leq 23.86480 \]
\[ \ln(n_1) + \ln(n_2) + 1.5n_3 + 1.5n_4 + 2n_5 + \]
\[ 0.5\sqrt{n_1} + 0.5\sqrt{n_2} + 1\sqrt{n_3} + 1\sqrt{n_4} + 1.5\sqrt{n_5} \leq 1200 \]
\[ 2 \leq n_1 \leq 1500 \]
\[ 2 \leq n_2 \leq 1920 \]
\[ 2 \leq n_3 \leq 1260 \]
\[ 2 \leq n_4 \leq 480 \]
\[ 2 \leq n_5 \leq 840 \]

and \[ x_j \geq 0 \; ; j = 1, 2, 3, 4 \]

(2.51)

The proposed compromise allocation which is the solution to the NLPP (2.51) given by the optimization software LINGO is:
\[ n_{1c}^* = 309.1612, \quad n_{2c}^* = 374.3831, \quad n_{3c}^* = 162.7233, \quad n_{4c}^* = 44.8336, \quad n_{5c}^* = 77.01911. \]

After rounding off to the nearest integer value we get the optimum compromise allocation as

\[ n_{1c}^* = 309, \quad n_{2c}^* = 374, \quad n_{3c}^* = 162, \quad n_{4c}^* = 45, \quad n_{5c}^* = 77. \]

The variances \( \nu(\bar{y}_{fst}) \) under compromise allocations denoted by \( \nu(\bar{y}_{fst})_{comp.} \) are:

\[ \nu(\bar{y}_{1st})_{comp.} = 2.152413105, \]
\[ \nu(\bar{y}_{2st})_{comp.} = 14.26904518, \]
\[ \nu(\bar{y}_{3st})_{comp.} = 3.339714596 \]
and
\[ \nu(\bar{y}_{4st})_{comp.} = 25.01786604 \]

with increases in the variances for the individual characteristics as:

\[ x_1 = 0.6482838, \quad x_2 = 3.472783, \quad x_3 = 0.9903062 \quad and \quad x_4 = 1.109451 \]

respectively. Where ‘comp’ stands for the proposed compromise allocation.

2.13.2 Proportional Allocation

Using the values of \( c_h, t_h \) and \( W_h \) as given in Table 2.3 with \( C_0 = 1200 \) the numerical values of \( A \) and \( B \) are obtained as

\[ A = 1.2850 \quad and \quad B = 1.8353. \] This gives
\[ X = \sqrt{n} = 29.8532. \]

or \[ n = X^2 = 891.2136. \]

Substituting this value of \( n \) in (2.40) the proportional allocation is obtained as:

\[ n_1 = 222.8034, \quad n_2 = 285.1884, \quad n_3 = 187.1549, \quad n_4 = 71.2971, \quad n_5 = 124.7699. \]

After rounding off to the nearest integer value we get

\[ n_1 = 223, \quad n_2 = 285, \quad n_3 = 187, \quad n_4 = 71, \quad n_5 = 125. \]

The variances of \( V(\bar{y}_{jst}) \) under proportional allocation (ignoring fpc) may be obtained by substituting the above values of \( n_h \) in variance formula

\[ V(\bar{y}_{jst}) = \sum_{h=1}^{L} \frac{W_h^2 S_{jh}^2}{n_h}, \quad j = 1, 2, \ldots, 4 \quad (2.52) \]

which gives \( V(\bar{y}_{jst}) \) for \( j = 1, 2, \ldots, 4 \) as:

\[ V_1 = V(\bar{y}_{1st})_{\text{prop}} = 1.6350, \]

\[ V_2 = V(\bar{y}_{2st})_{\text{prop}} = 18.9285, \]

\[ V_3 = V(\bar{y}_{3st})_{\text{prop}} = 2.5583 \quad \text{and} \]

\[ V_4 = V(\bar{y}_{4st})_{\text{prop}} = 26.1678. \]

Where 'prop' stands for the proportional allocation.
2.13.3 Cochran's Compromise Allocation

For the present example Cochran's compromise allocation given by (2.45) is

\[ n_1 = 287.0896, \quad n_2 = 312.4103, \quad n_3 = 151.9265, \quad n_4 = 68.6031, \]
\[ n_5 = 109.1480. \]

After rounding off to the nearest integer value we get

\[ n_1 = 287, \quad n_2 = 312, \quad n_3 = 152, \quad n_4 = 69, \quad n_5 = 109. \]

The variances \( \nu \{(\bar{y}_{jst})_C \} \) under the Cochran's compromise allocation are

\[ \nu_1C = \nu \{(\bar{y}_{1st})_C \} = 1.7345, \]
\[ \nu_2C = \nu \{(\bar{y}_{2st})_C \} = 15.8570, \]
\[ \nu_3C = \nu \{(\bar{y}_{3st})_C \} = 2.7010 \quad \text{and} \]
\[ \nu_4C = \nu \{(\bar{y}_{4st})_C \} = 26.1775. \]

Where 'C' stands for the Cochran's compromise allocation.

2.13.4 Khan's Compromise Allocation

Using the values given in Table 2.1 the NLPP (2.46) becomes
Minimize \[ Z = \frac{1377.5081}{n_1} + \frac{2449.5358}{n_2} + \frac{740.9373}{n_3} + \frac{35.7884}{n_4} + \frac{56.2973}{n_5} \]

Subject to \[ \begin{align*}
& n_1 + 1n_2 + 1.5n_3 + 1.5n_4 + 2n_5 + \\
& 0.5\sqrt{n_1} + 0.5\sqrt{n_2} + 1\sqrt{n_3} + 1\sqrt{n_4} + 1.5\sqrt{n_5} \leq 1200
\end{align*} \]

\[ \begin{align*}
& 2 \leq n_1 \leq 1500 \\
& 2 \leq n_2 \leq 1920 \\
& 2 \leq n_3 \leq 1260 \\
& 2 \leq n_4 \leq 480 \\
& 2 \leq n_5 \leq 840
\end{align*} \]

(2.53)

The optimal solution to NLPP (2.53) using optimization software LINGO is obtained as:

\[ n_1 = 310.9282, \quad n_2 = 415.0147, \quad n_3 = 185.2513, \quad n_4 = 40.1666, \quad n_5 = 43.5408. \]

After rounding off to the nearest integer value we get

\[ n_1 = 311, \quad n_2 = 415, \quad n_3 = 185, \quad n_4 = 40, \quad n_5 = 44. \]

The variances \( V_{(\overline{y}_{jst})_K} \) using the Khan et al. (2003) compromise criterion are

\[ V_{1K} = 3.0100, \quad V_{2K} = 13.6979, \quad V_{3K} = 4.4648 \quad \text{and} \quad V_{4K} = 25.8342. \]

Where ‘\( K \)’ stands for Khan’s compromise allocation.
2.14 DISCUSSION

In this section a comparative study of the four compromise allocations discussed in Section 2.3 has been made. The basis of comparison is the traces of the variance-covariance matrices of the estimates under various compromise allocations. Since the characteristics under study are assumed as independent, the covariances are zero. The traces are the sum of the diagonal elements of the variances-covariance matrices that are the variances of the estimates of the population means of the different characteristics. Sukhatme et al. (1984) define the relative efficiency (R. E.) of a compromise allocation with respect to proportional allocation as

\[
\text{R.E.} = \frac{T_{\text{Prop.}}}{T_{\text{Comp.}}} \quad (2.54)
\]

where \( T_{\text{Prop.}} = \) Sum of the variances under proportional allocation. and \( T_{\text{Comp.}} = \) Sum of the variances under the given compromise allocations.

Column (8) of Table 2.4 gives the R.E. of the three compromise allocation discussed in this article as compared to the proportional allocation.

2.15 CONCLUSION

The results are summarized in Table 2.4.
Table 2.4
R. E. as Compared to Proportional Allocation

<table>
<thead>
<tr>
<th>Compromise allocations (2)</th>
<th>Values of ( V_j, j = 1,2,3 &amp; 4 ) under various compromise allocation</th>
<th>Trace ((7)) = (3) + (4) + (5) + (6)</th>
<th>R.E. ((8) = \frac{T_{Prop.}}{T_{Comp.}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportional</td>
<td>1.6350 18.9285 2.5583 26.1678</td>
<td>49.2896</td>
<td>1.00</td>
</tr>
<tr>
<td>Cochran’s Compromise allocation</td>
<td>1.7345 15.8570 2.7010 26.1775</td>
<td>46.4700</td>
<td>1.06</td>
</tr>
<tr>
<td>Khan’s Compromise allocation</td>
<td>3.0100 13.6979 4.4648 25.8342</td>
<td>47.0061</td>
<td>1.05</td>
</tr>
<tr>
<td>Proposed</td>
<td>2.1524 14.2690 3.3397 25.0179</td>
<td>44.7790</td>
<td>1.10</td>
</tr>
</tbody>
</table>
An observation of the column 10 of the Table 2.4 reveals that the proposed compromise allocation is the most efficient. Thus it is concluded that the proposed compromise allocation compares favorably with the other studied allocations when the travel cost within the strata are significant.