Chapter 5

STOCHASTIC COMPROMISE MIXED ALLOCATION IN MULTIVARIATE STRATIFIED SAMPLE SURVEYS
5.1 INTRODUCTION

While estimating the over all population mean $\bar{Y}$ in univariate stratified sampling, in the notations of Cochran (1977), the stratified sample mean $\bar{y}_{sh}$ provides an unbiased estimate of $\bar{Y}$ with a sampling variance

$$V(\bar{y}_{sh}) = \sum_{h=1}^{L} \frac{W_h^2 \sigma_h^2}{n_h} - \sum_{h=1}^{L} \frac{W_h^2 S_h^2}{N_h}$$

(5.1)

The total cost $C$ of the survey is usually expressed as a linear function of the sample allocations $n_h$

$$C = c_0 + \sum_{h=1}^{L} c_h n_h$$

(5.2)

It is assumed that the number of strata $L$ and the strata boundaries are fixed in advance and the sampler has to determine the number of units (allocations) $n_h$ to be selected from each stratum to obtain a stratified sample. The sample allocations $n_h$ that minimize the sampling variance $V(\bar{y}_{sh})$ for fixed cost $C$ or alternatively minimize the cost $C$ for fixed $V(\bar{y}_{sh})$ is known as the ‘Optimum Allocation’ in stratified sampling literature. Another frequently used allocation is the ‘Proportional Allocation’ in which the sample size $n_h$ is proportional to the stratum size $N_h$. Some less frequently used allocations are ‘Equal Allocation’ and ‘Allocation Proportional to $W_h R_h$’ where $R_h$ is the range of the $h^{th}$ stratum. (See Murthy (1967)).

In stratified sampling usually a single allocation is selected and applied to all the strata to select a representative sample of the population. But practical experience suggests that when the complete information about all the strata are not available a single allocation may not be advisable. For example the application of optimum
allocation to $h^{th}$ stratum; $h = 1, 2, \ldots, L$, requires the knowledge of stratum weight $W_h$, stratum variance $S_h^2$ and per unit measurement cost $c_h$. If any of these quantities are not known for a particular stratum, optimum allocation cannot be applied. Similarly if the true values of the strata sizes $N_h$ are not known, proportional allocation can not be applied.

In such situations Ahsan et al. (2005) gave the concept of “Mixed Allocation” in univariate stratified sampling based on the idea of Clark and Steel (2000). Later on Varshney et al. (2011) extended the work of Ahsan et al. (2005) for multivariate stratified populations.

Generally, the true values of $S_h^2$ are unknown and in working out the values of $n_h$, they are replaced by their sample estimates $s_h^2$. Being a sample statistic $s_h^2$ are random variable. Furthermore, the per unit measurement costs $c_h$ may also vary during the course of the survey due to random causes, hence they are also random variable.

In this chapter the problem of finding a compromise mixed allocation in multivariate stratified sampling is viewed as a Stochastic Nonlinear Programming Problem (SNLPP). Where the per unit measurement costs $c_h$ and sample variances $s_h^2$ are random variables with known distributions.

This chapter is based on my research paper accepted for publication in the ‘Journal of Mathematical Modelling and Algorithms in Operations Research’

In the next two section the work of Ahsan et al. (2005) Varshney et al. (2011) are summarized for the sake of continuity.

5.2 MIXED ALLOCATION IN UNIVARIATE STRATIFIED SAMPLING
Consider a stratified population with $L$ strata. Let these $L$ strata be divided into $k$ groups $G_1, G_2, \ldots, G_k$, where the group $G_j; j = 1, 2, \ldots, k$ consists of $L_j$ strata and $\sum_{j=1}^{k} L_j = L$. The strata are grouped according to the available information, that is, full information, partial information and null information (See Kozak (2006(a), 2006(b))).
Without loss of generality we can assume that the first $L_1$ strata constitute the group $G_1$, the next $L_2$ strata constitute the group $G_2$, and so on.

In order to use different type of allocations in different groups Clark and Steel (2000) defined the sample size $n_h$ from the $h^{th}$ stratum as

$$n_h = \alpha_j \beta_h ; h \in I_j ; j = 1, 2, ..., k$$  \hspace{1cm} (5.3)

where $I_j$ is the set of indices of the strata constituting the group $G_j$, $\beta_h ; h \in I_j$ are known constants depending upon the type of allocation to be used in the group $G_j$ and $\alpha_j$ are unknown decision variables to be determined in order to minimize the variance under the $j^{th}$ type of allocation for fixed cost.

The problem of finding the mixed allocation defined in (5.1) that minimize $V(\bar{Y}_{u})$ given by (5.1), ignoring fpc, for a fixed cost given by (5.2) may be formulated as the following Nonlinear Programming Problem (NLPP):

Minimize $$\sum_{h=1}^{L} \frac{W_h^2 S_h^2}{n_h}$$ \hspace{1cm} (5.4)

subject to $$\sum_{h=1}^{L} c_h n_h \leq C_0$$ \hspace{1cm} (5.5)

$$n_h = \alpha_j \beta_h ; h \in I_j ; j = 1, 2, ..., k$$ \hspace{1cm} (5.6)

and $$2 \leq n_h \leq N_h ; h = 1, 2, ..., L$$ \hspace{1cm} (5.7)

where $C_0 = C - c_0$ is the total cost available for measurements. The restrictions $2 \leq n_h \leq N_h$ are imposed on $n_h$ to have an estimate of the stratum variance and to avoid oversampling.

Using (5.6) $n_h$ can be eliminated from (5.4), (5.5) and (5.7) and NLPP (5.4)-(5.7) may be restated as
Minimize \( F(\alpha_j) = \sum_{j=1}^{k} \sum_{h \in L_j} \frac{W_h^2 S_h^2}{\alpha_j \beta_h} \) \hspace{1cm} (5.8)

subject to \( \sum_{j=1}^{k} \sum_{h \in L_j} \alpha_j c_h \beta_h \leq C_0 \) \hspace{1cm} (5.9)

and \( \frac{2}{\beta_h} \leq \alpha_j \leq \frac{N_h}{\beta_h}; h \in L_j; j = 1, 2, \ldots, k \) \hspace{1cm} (5.10)

where the restrictions in (5.10) are equivalent to the restrictions in (5.7).

Assuming equality in (5.9) and ignoring (5.10), Ahsan et al. (2005) used Lagrange Multipliers Technique to work out the values of \( \alpha_j \) as:

\[
\alpha_j = C_0 \sqrt{\frac{\sum_{h \in L_j} W_h^2 S_h^2 / \beta_h}{\sum_{h \in L_j} c_h \beta_h}}; j = 1, 2, \ldots, k \hspace{1cm} (5.11)
\]

If the values of \( \alpha_j \) given by (5.11) satisfy (5.10) also they will solve the NLPP (5.8)-(5.10). In case (5.10) are violated for some \( j \) then an appropriate Nonlinear Programming Technique may be used.

Substitution of \( \alpha_j \) in (5.3) gives the required mixed allocation. The resulting sampling variance \( V_{mixed} \) of the stratified sample mean \( \tilde{y}_{sl} \) is

\[
V_{mixed} = \frac{1}{C_0} \sqrt{\sum_{j=1}^{k} \frac{\left( \sum_{h \in L_j} W_h^2 S_h^2 / \beta_h \right) \left( \sum_{h \in L_j} c_h \beta_h \right)}{\sum_{h \in L_j} c_h \beta_h}}^2 \hspace{1cm} (5.12)
\]

Section 5.3 summarises the multivariate extension of the work of Ahsan et al. (2005) by Varshney et al. (2011).

5.3 THE COMPROMISE MIXED ALLOCATION

Let in a multivariate stratified population \( p \) characteristics be defined on each unit and the estimation of \( p \) population mean \( \tilde{y}_{l}; l = 1, 2, \ldots, p \) be of interest. Varshney et
al. (2011) formulated the problem of finding a compromise mixed allocation as the following NLPP.

Minimize \[ \sum_{l=1}^{p} a_l \sum_{j=1}^{k} \sum_{h \in I_j} \alpha_j \beta_h \] \[ \text{subject to} \quad \sum_{j=1}^{k} \sum_{h \in I_j} \alpha_j c_h \beta_h \leq C_0 \] \[ \text{and} \quad \frac{2}{\beta_h} \leq \alpha_j \leq \frac{N_h}{\beta_h}; h \in I_j; j = 1, 2, \ldots, k \] where suffix 'l' is introduced to denote the \( l^{th} \) characteristic, and \( a_l > 0; \) \( l = 1, 2, \ldots, p \) are the weights assigned to the variance of the \( l^{th} \) characteristic according to some measure of its relative importance. In (5.14) \( c_h \) denote the cost of measuring all the \( p \) characteristics on a unit selected from the \( h^{th} \) stratum. Note that in NLPP (5.13)-(5.15) is obtained by substituting \( n_h = \alpha_j \beta_h \) in the NLPP formulated by Varshney et al. (2011) and the restrictions (5.15) are equivalent to \( 2 \leq n_h \leq N_h; h = 1, 2, \ldots, L \). The compromise criterion of Varshney et al. (2011) is same as given by Yates (1960), that is: Minimize the weighted sum of variances of the stratified sample means of the \( p \) characteristics under the cost and other restrictions. Taking equality in (5.14) and ignoring restrictions (5.15), Varshney et al. (2011) used Lagrange Multipliers Technique to obtain the optimum values of \( \alpha_j \) as

\[ \alpha_j = C_0 \sqrt{\frac{\sum_{h \in I_j} \beta_h \left( \sum_{h \in I_j} \frac{A_h}{\beta_h} \right)}{\left( \sum_{h \in I_j} \beta_h \right)^2}} ; j = 1, 2, \ldots, k \] \[ \text{where} \quad A_h = \sum_{l=1}^{p} a_l S_{ih}^2 ; h = 1, 2, \ldots, L \]
Once \( \alpha_j \) are obtained (5.3) will give the required compromise mixed allocation. The sampling variance \( V_{(cm)} \) of the stratified sample mean under this allocation is

\[
V_{(cm)} = \frac{1}{C_0} \left[ \sum_{j=1}^k \sqrt{ \sum_{h \in I_j} \frac{A_h}{\beta_h} \left( \sum_{h \in I_j} c_h \beta_h \right) } \right]^2
\]  

(5.18)

where \( (cm) \) stands for ‘Compromise Mixed’ allocation.

In the next section the problem of obtaining the compromise mixed allocation when the parameters of the NLPP (5.13)-(5.15) are independent random variables with known distributions is discussed as an SNLPP.

### 5.4 THE STOCHASTIC COMPROMISE MIXED ALLOCATION

The general mathematical model of an \( m \times n \) Mathematical Programming Problem (MPP) may be given as

Minimize (or Maximize) \( f(x) \)

subject to \( g_i(x) \) \( \{ \leq = \geq \} b_i; i = 1, 2, ..., m \)

and \( x \geq 0 \)

where \( f(x) \) and \( g_i(x); i = 1, 2, ..., m \) are functions of decision variables \( x = (x_1, x_2, ..., x_n) \) and one and only one of the symbols \( \leq, =, \geq \) holds for each \( i = 1, 2, ..., m \). The coefficients of the decision variables in the objective function \( f(x) \), constraint functions \( g_i(x) \) and the RHS \( b_i; i = 1, 2, ..., m \) are called parameters of the MPP. If these parameters are known constants the MPP is termed as deterministic. On the other hand if the parameters are random variables the MPP becomes a Stochastic Programming Problem. In this chapter a more general as well as practical instance of the compromise mixed allocation formulated as NLPP (5.13)-(5.15) is considered. The true values of the strata variances \( S_h^2 \) are usually unknown and they are replaced by their sample estimates \( s_h^2 \) which are random variables.

Furthermore, the measurement costs \( c_h \) may also be a random variable because
during the course of survey they may vary due to the random causes. Thus we have the following SNLPP to solve.

\[
\begin{align*}
\text{Minimize} & \quad \sum_{i=1}^{p} a_i \sum_{j=1}^{k} \sum_{h_{i,j}} \frac{W_{j,h}^2 s_{ih}^2}{\alpha_j \beta_h} = \sum_{i=1}^{p} a_i F(s_{ih}^2) \quad (5.19) \\
\text{subject to} & \quad \sum_{j=1}^{k} \sum_{h_{i,j}} \alpha_j c_h \beta_h \leq C_0 \quad (5.20) \\
& \quad \frac{2}{\beta_h} \leq \alpha_j \leq \frac{N_h}{\beta_h} ; h \in I_j \quad (5.21) \\
& \quad \alpha_j \geq 0 ; j = 1, 2, \ldots, k \quad (5.22)
\end{align*}
\]

where \( F(s_{ih}^2) = \sum_{j=1}^{k} \sum_{h_{i,j}} \frac{W_{j,h}^2 s_{ih}^2}{\alpha_j \beta_h} \), \( s_{ih}^2 \) and \( c_h ; h = 1, 2, \ldots, L \) are random variables.

The next section proposes a method to solve the SNLPP (5.19)-(5.22) by converting it into its deterministic equivalent.

5.5 CONVERSION OF THE SNLPP INTO ITS DETERMINISTIC EQUIVALENT

Diaz-García and Gary-Tapia (2007) using results of Melaku (1986) showed that in stratified sampling the sample variances \( s_h^2 \) have an asymptotic normal distribution with mean \( E(\xi_h) \) and variance \( V(\xi_h) ; h = 1, 2, \ldots, L \), where

\[
\xi_h = \frac{1}{n_h - 1} \sum_{i=1}^{n_h} (y_{ih} - \bar{Y}_h)^2 \quad (5.23)
\]

\[
E(\xi_h) = \frac{n_h}{n_h - 1} S_h^2 \quad (5.24)
\]

\[
V(\xi_h) = \frac{n_h}{(n_h - 1)^2} \left[ C_{yh}^4 - (S_h^2)^2 \right] \quad (5.25)
\]

and

\[
C_{yh}^4 = \frac{1}{N_h} \sum_{i=1}^{N_h} (y_{ih} - \bar{Y}_h)^4 \quad (5.26)
\]
where $C_{vh}^4$ are the fourth moments about stratum means.

The above results can be applied to $s_{ih}^2; l = 1,2,...,p; h = 1,2,...,L$, for multivariate stratified sampling when the characteristics are independent. The objective function (5.19) is a linear function of $s_{ih}^2; l = 1,2,...,p$ therefore, it will also follow a normal distribution.

Using (5.24) we can substitute $E(s_{ih}^2) = \frac{n_h}{n_h - 1} S_{ih}^2$ in the objective function (5.19) to get

$$E\left(\sum_{i=1}^{p} a_i F(s_{ih}^2)\right) = \sum_{i=1}^{p} a_i \sum_{j=1}^{k} \frac{W_{h}^2 s_{ijh}^2}{n_h \beta_h (n_a - 1)}$$

and

$$V\left(\sum_{i=1}^{p} a_i F(s_{ih}^2)\right) = \sum_{i=1}^{p} a_i^2 V(F(s_{ih}^2))$$

$$= \sum_{i=1}^{p} a_i^2 \left(\sum_{j=1}^{k} \frac{W_{h}^2 s_{ijh}^2}{n_h \beta_h (n_a - 1)}\right)$$

$$= \sum_{i=1}^{p} a_i^2 \sum_{j=1}^{k} \frac{W_{h}^4}{n_h \beta_h^2} V(s_{ih}^2)$$

$$= \sum_{i=1}^{p} a_i^2 \sum_{j=1}^{k} \frac{W_{h}^4}{n_h \beta_h^2} V(s_{ih}^2)$$

$$= \sum_{i=1}^{p} a_i^2 \sum_{j=1}^{k} \frac{W_{h}^4}{n_h \beta_h^2} V(s_{ih}^2)$$

$$= \sum_{i=1}^{p} a_i^2 \sum_{j=1}^{k} \frac{W_{h}^4}{n_h \beta_h^2} \left(\frac{n_h}{n_h - 1}\right)^2 \left[C_{vh}^4 \left(-s_{ih}^2\right)^2\right]$$

where $C_{vh}^4$ denote the fourth moments with respect to the $l^{th}$ characteristic.

Thus the deterministic equivalent of the objective function (5.19) may be given as
\[ f(n) = k_1 E\left( \sum_{i=1}^{g} a_i F(s_{ih})^2 \right) + k_2 \sqrt{V\left( \sum_{i=1}^{g} a_i F(s_{ih})^2 \right)} \]  \hspace{1cm} (5.29)

where \( k_1 \) and \( k_2 \geq 0 \) are constants fixed according to the relative importance of the
\[ E\left( \sum_{i=1}^{g} a_i F(s_{ih})^2 \right) \]
and the standard deviation \( \sqrt{V\left( \sum_{i=1}^{g} a_i F(s_{ih})^2 \right)} \). (See Rao, S. S. (1977)).

Without loss of generality we can assume that \( k_1 + k_2 = 1 \).

Now consider the cost constraint (5.20), that is,
\[ \sum_{j=1}^{k} \sum_{h \in I_j} C_j \alpha_j \beta_h \leq C_0 \]

where the costs \( C_h \) are independently and normally distributed as \( N(\mu_{C_h}, \sigma^2_{C_h}); h = 1, 2, ..., L \). Being a linear combination of independent normal variates the total cost \( C_0 \) in (5.20) will also follow a normal distribution.

Usually the parameter \( \mu_{C_h} \) and \( \sigma^2_{C_h} \) are unknown but they can be estimated from a pilot survey or their values at some previous occasion may be used. If \( \hat{\mu}_{C_h} \) and \( \hat{\sigma}^2_{C_h} \) are the estimated value of \( \mu_{C_h} \) and \( \sigma^2_{C_h} \) respectively, then the usual deterministic equivalent of the LHS of (5.20) will be
\[ k_1' \left( \sum_{j=1}^{k} \sum_{h \in I_j} \hat{\mu}_{C_h} \hat{\alpha}_j \hat{\beta}_h \right) + k_2' \sqrt{\sum_{j=1}^{k} \sum_{h \in I_j} (\alpha_j \beta_h)^2 \hat{\sigma}^2_{C_h}} \]  \hspace{1cm} (5.30)

where \( k_1' \) and \( k_2' \geq 0 \) are known constants representing the relative importance of the terms,
\[ E\left( \sum_{j=1}^{k} \sum_{h \in I_j} \alpha_j C_h \beta_h \right) \]
and \( \sqrt{V\left( \sum_{j=1}^{k} \sum_{h \in I_j} \alpha_j C_h \beta_h \right)} \) in the cost constraint. Without loss of generality we can assume that \( k_1' + k_2' = 1 \). Since sampling is independent in
each stratum the cross product terms in \( \nu \left( \sum_{l=1}^{p} a_l F(s^2_{lh}) \right) \) vanish. Thus the
deterministic equivalent of the cost constraint (5.20) will be

\[
k_1 \left( \sum_{j=1}^{k} \sum_{h \in I_j} \alpha_j \beta_h \hat{c}_h \right) + k_2 \sqrt{\sum_{j=1}^{k} \sum_{h \in I_j} (\alpha_j \beta_h)^2 \hat{\sigma}^2_{\epsilon_h}} \leq C_0 \quad (5.31)
\]

The expressions (5.27), (5.28), (5.29) and (5.31) give the deterministic equivalent of
the SNLPP (5.19)-(5.21) as the following NLPP:

Minimize \( \phi(\alpha) = k_1 \left[ \sum_{l=1}^{p} a_l \sum_{j=1}^{k} \sum_{h \in I_j} \frac{W_h^2 s^2_{lh}}{(\alpha_j \beta_h - 1)^2} \right] \)

\[
+ k_2 \sqrt{\sum_{l=1}^{p} a_l^2 \sum_{j=1}^{k} \sum_{h \in I_j} \frac{W_h^4}{\alpha_j \beta_h (\alpha_j \beta_h - 1)^2} \left[ C^4_{ylh} - (s^2_{lh})^2 \right]}
\]

subject to \( k_1 \left( \sum_{j=1}^{k} \sum_{h \in I_j} \hat{c}_h \alpha_j \beta_h \right) + k_2 \sqrt{\sum_{j=1}^{k} \sum_{h \in I_j} (\alpha_j \beta_h)^2 \hat{\sigma}^2_{\epsilon_h}} \leq C_0 \quad (5.33) \)

and \( \frac{2}{\beta_h} \leq \alpha_j \leq 1; \alpha_j \leq \frac{N_h}{\beta_h}; h \in I_j; j = 1, 2, \ldots, k \quad (5.34) \)

where \( \alpha = (\alpha_1, \alpha_2, \ldots, \alpha_k) \) is the vector of decision variables.

Note that in the objective function (5.32) \( \alpha_j \beta_h \) is substituted for \( n_h \) using (5.3).

When the numerical data are available the NLPP (5.32)-(5.34) may be solved using a
suitable Nonlinear Programming Technique. In the next section an application of the
proposed formulation is given for an artificial data. The solution to the NLPP (5.32)-(5.34) is obtained by the optimization software LINGO (2001).
5.6 APPLICATION OF THE PROPOSED METHOD

In the following two different situation are studied. The first case deals with the situation when the variations in the objective and constraint functions are assumed to be insignificant, that is, $k_1 = 1$, $k_2 = 0$ and $k_1' = 1$, $k_2' = 0$. The second case deals with the situation, when $k_1$, $k_2$ and $k_1'$, $k_2'$ are all > 0.

5.6.1 The First Case

When $k_1 = 1$, $k_2 = 0$ and $k_1' = 1$, $k_2' = 0$ the NLPP (5.32)-(5.34) takes a simpler form as:

\[
\text{Minimize } \phi(\alpha) = \left[ \sum_{l=1}^{p} \alpha_l \sum_{j=1}^{k} \sum_{h \in I_j} \frac{W_{h}^2 s_{lh}^2}{(\alpha_j \beta_h - 1)} \right] \tag{5.35}
\]

subject to

\[
\sum_{j=1}^{k} \sum_{h \in I_j} \hat{c}_h \alpha_j \beta_h \leq C_0 \tag{5.36}
\]

and

\[
\frac{2}{\beta_h} \leq \alpha_j \leq \frac{N_h}{\beta_h} ; h \in I_j ; j = 1, 2, \ldots, k \tag{5.37}
\]

It can be seen that NLPP (5.35)-(5.37) is same as formulated by Varshney et al. (2011) except for the change that in the proposed method the costs $c_h$ and the sample variance $s_{lh}^2$ are considered as random variable and their estimates are substituted for them.

A Numerical Illustration

The following data are from Varshney et al. (2011). The values of $N_h$, $W_h$, $s_{lh}$, $s_{2h}$ and $c_h$ for seven strata and two characteristics are given in Table 5.1.
Let the total budget of the survey be 4500 units that includes an overhead cost of 500 units. So that the $C - c_0 = 4500 - 500 = 4000 = C_0$ units are available for measurement.

Both the characteristics are assumed to be of equal importance, that is, $a_1 = a_2 = 0.5$.

The groupings of strata are also same as Varshney et al. (2011), that is,

(i) Strata 1, 2 and 3 constitute group $G_1$ in which equal allocation is to be used, that is,

$$\beta_h = 1; h \in I_1 = \{1, 2, 3\}$$  \hspace{1cm} (5.38)

(ii) Strata 4 and 5 constitute group $G_2$ in which proportional allocation is to be used, that is,

$$\beta_h = W_h; h \in I_2 = \{4, 5\}$$  \hspace{1cm} (5.39)

(iii) Strata 6 and 7 constitute group $G_3$ in which optimum allocation is to be used, that is,

$$\beta_h = \sqrt{\frac{W_h^2 \sum_{l=1}^n a_l s_{l}^2}{\hat{c}_h}} h \in I_3 = \{6, 7\}$$  \hspace{1cm} (5.40)

It can be seen that $I_j; j = 1, 2, 3$ are mutually exclusive and exhaustive.

Using (5.38), (5.39) and (5.40) for the data given in Table 5.1 the values of $\beta_h; h = 1, 2, ..., 7$ are obtained as $\beta_1 = \beta_2 = \beta_3 = 1, \beta_4 = 0.087, \beta_5 = 0.093, \beta_6 = 0.699$ and $\beta_7 = 1.172$.  

<table>
<thead>
<tr>
<th>$h$</th>
<th>$N_h$</th>
<th>$W_h$</th>
<th>$s_{1h}$</th>
<th>$s_{2h}$</th>
<th>$c_h$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>472</td>
<td>0.189</td>
<td>05.237</td>
<td>07.815</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>559</td>
<td>0.224</td>
<td>05.821</td>
<td>07.949</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>425</td>
<td>0.170</td>
<td>05.238</td>
<td>07.725</td>
<td>7</td>
</tr>
<tr>
<td>4</td>
<td>218</td>
<td>0.087</td>
<td>25.528</td>
<td>30.125</td>
<td>12</td>
</tr>
<tr>
<td>5</td>
<td>233</td>
<td>0.093</td>
<td>22.232</td>
<td>32.231</td>
<td>11</td>
</tr>
<tr>
<td>6</td>
<td>328</td>
<td>0.131</td>
<td>15.129</td>
<td>18.455</td>
<td>10</td>
</tr>
<tr>
<td>7</td>
<td>265</td>
<td>0.106</td>
<td>40.125</td>
<td>45.358</td>
<td>15</td>
</tr>
</tbody>
</table>
Thus we get NLPP (5.35)-(5.37) as

\[
\begin{align*}
\text{Minimize } & \phi(\alpha) = \left[ \frac{10.5495087}{(\alpha_1 - 1)} + \frac{11.80154696}{(0.087\alpha_2 - 1)} + \frac{13.25977325}{(0.093\alpha_2 - 1)} \right. \\
& \left. + \frac{9.772743314}{(0.699\alpha_3 - 1)} + \frac{41.20650191}{(1.172\alpha_3 - 1)} \right] \\
\text{subject to } & 21\alpha_1 + 2.067\alpha_2 + 24.57\alpha_3 \leq 4000 \\
& 2 \leq \alpha_1 \leq 425 \\
& 23 \leq \alpha_2 \leq 2505 \\
& 3 \leq \alpha_3 \leq 226
\end{align*}
\] (5.41)

The optimization software LINGO gives the optimal solution to the NLPP (5.41)-(5.43) as:

\[\alpha_1 = 38.84914, \alpha_2 = 630.6696 \text{ and } \alpha_3 = 76.53944.\]

With these values of \(\alpha_j; j = 1, 2, 3\), the compromise mixed allocation \(n_{h(m)}; h = 1, 2, ..., 7\) given by (5.3) is

For \(j = 1\) \(n_{1(m)} = n_{2(m)} = n_{3(m)} = \alpha_1\beta_1 = 38.84914 \times 1 = 39\) since \(\beta_1 = \beta_2 = \beta_3 = 1.\)

For \(j = 2\) \(n_{4(m)} = \alpha_2\beta_4 = 630.6696 \times 0.087 = 54.8682552 \equiv 55\)

\(n_{5(m)} = \alpha_2\beta_5 = 630.6696 \times 0.093 = 58.6522728 \equiv 59\)

For \(j = 3\) \(n_{6(m)} = \alpha_3\beta_6 = 76.53944 \times 0.699 = 53.50106856 \equiv 53\)

\(n_{7(m)} = \alpha_3\beta_7 = 76.53944 \times 1.172 = 89.70422368 \equiv 90\).

with a total sample size \(n = \sum_{h=1}^{7} n_{h(m)} = 374\).

The estimated variances \(v(\bar{y}_{1st})\) (ignoring fpc) for the two characteristics under the proposed compromise mixed allocation are

\(v(\bar{y}_{1st}) = 0.525015287\) and \(v(\bar{y}_{2st}) = 0.825759789\) with a trace value of 1.350775076.
5.6.2 The Second Case

In the second case, when the variations in the cost and variance functions are also important, that is, \( k_1, k_2 \) and \( k_1', k_2' \) are all strictly greater than zero.

A Numerical Illustration

The data and the grouping scheme are same as that of the numerical illustration presented in section 5.6.1. The amount available for measurements \( C_0 = C - c_0 = 4500 - 500 = 4000 \). This amount is now bifurcated in proportion to

\[
\sum_{h=1}^{7} c_{1h} \quad \text{and} \quad \sum_{h=1}^{7} c_{2h} \quad (\text{approximately}) \quad \text{as} \quad C_{01} = 2300 \quad \text{and} \quad C_{02} = 1700
\]

for measuring the first \((l = 1)\) and second \((l = 2)\) characteristics respectively. (See Jahan et al. (1994)).

Where \( c_{1h} \) and \( c_{2h} \) are per unit expected costs for measuring the first and second characteristics respectively. Note that \( \hat{c}_h = \hat{c}_{1h} + \hat{c}_{2h} \).

In addition we assume that the fourth moments about strata means for the two characteristics and the expected measurement costs \( \hat{c}_{1h}, \hat{c}_{2h}, \hat{c}_h \) and the estimates of the variances of \( \hat{c}_h; h = 1, 2, \ldots, 7 \) are as given in Table 5.2.

Table 5.2: Fourth moments about strata means and expected cost with their estimated variances for the two characteristics

<table>
<thead>
<tr>
<th>( h )</th>
<th>( C_{1h}^4 )</th>
<th>( C_{2h}^4 )</th>
<th>( \hat{c}_{1h} )</th>
<th>( \hat{c}_{2h} )</th>
<th>( \hat{c}<em>h = \hat{c}</em>{1h} + \hat{c}_{2h} )</th>
<th>( \hat{c}_{2h}^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3999.737</td>
<td>10245.520</td>
<td>3</td>
<td>3</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>3981.273</td>
<td>10444.760</td>
<td>5</td>
<td>3</td>
<td>8</td>
<td>12</td>
</tr>
<tr>
<td>3</td>
<td>1637.434</td>
<td>7284.730</td>
<td>4</td>
<td>3</td>
<td>7</td>
<td>15</td>
</tr>
<tr>
<td>4</td>
<td>1370394.000</td>
<td>2651515.000</td>
<td>7</td>
<td>5</td>
<td>12</td>
<td>20</td>
</tr>
<tr>
<td>5</td>
<td>912987.900</td>
<td>2740463.000</td>
<td>6</td>
<td>5</td>
<td>11</td>
<td>22</td>
</tr>
<tr>
<td>6</td>
<td>218262.200</td>
<td>303569.500</td>
<td>6</td>
<td>4</td>
<td>10</td>
<td>27</td>
</tr>
<tr>
<td>7</td>
<td>6829099.000</td>
<td>16474685.000</td>
<td>9</td>
<td>6</td>
<td>15</td>
<td>38</td>
</tr>
</tbody>
</table>
For simplicity we take \( k_1 = k_2 = 0.5 \) and \( k'_1 = k'_2 = 0.5 \). Also the two characteristics are assumed to be of equal importance, that is, \( a_l = 0.5 \) for \( l = 1, 2 \).

Substituting the estimated values of the parameters from Tables 5.1 and 5.2 we get the NLPP (5.32)-(5.34) as

Minimize \( \phi(\alpha) = 
\begin{align*}
0.5 \left[ 0.5 \left( \frac{10.5495087}{\alpha_1 - 1} + \frac{11.80154696}{0.087\alpha_2 - 1} + \frac{13.259777325}{0.093\alpha_2 - 1} + \frac{9.772743314}{0.699\alpha_3 - 1} \right) 
+ \frac{41.20650191}{1.172\alpha_3 - 1} \right] \\
0.5 \left[ (0.5)^2 \left( \frac{39.68329591}{\alpha_1 (\alpha_1 - 1)^2} + \frac{158.9010691}{0.087\alpha_2 (0.087\alpha_2 - 1)^2} + \frac{174.2941481}{0.093\alpha_2 (0.093\alpha_2 - 1)^2} \\
+ \frac{104.0889169}{0.699\alpha_3 (0.699\alpha_3 - 1)^2} + \frac{2080.429683}{1.172\alpha_3 (1.172\alpha_3 - 1)^2} \right) \right]^{1/2} 
\end{align*} 
\tag{5.44}
\]

subject to
\begin{align*}
0.5 (21\alpha_1 + 2.067\alpha_2 + 24.57\alpha_3 ) \\
+ 0.5 \sqrt{(37\alpha_1^2 + 0.341658\alpha_2^2 + 65.388419\alpha_3^2 )} \leq 4000 
\tag{5.45}
\end{align*}
\begin{align*}
2 \leq \alpha_1 \leq 425 \\
23 \leq \alpha_2 \leq 2505 \\
3 \leq \alpha_3 \leq 226 
\tag{5.46}
\end{align*}

Using optimization software LINGO the solution to the NLPP (5.44)-(5.46) is obtained as:
\[ \alpha_1 = 51.44602, \alpha_2 = 617.1024 \text{ and } \alpha_3 = 73.39763. \]

With these values of \( \alpha_j; j = 1, 2, 3 \), the compromise mixed allocation given by (5.3) is

For \( j = 1 \) \( n_{1(m)} = n_{2(m)} = n_{3(m)} = \alpha_1 \beta_1 = 51.44602 \times 1 \equiv 51 \)

For \( j = 2 \) \( n_{4(m)} = \alpha_2 \beta_4 = 617.1024 \times 0.087 = 53.6879088 \equiv 54 \)

\( n_{5(m)} = \alpha_2 \beta_5 = 617.1024 \times 0.093 = 57.3905232 \equiv 57 \)
For $j = 3 \ n_{6(m)} = \alpha_3 \beta_6 = 73.39763 \times 0.699 = 51.30494337 \cong 51$

$n_{7(m)} = \alpha_3 \beta_7 = 73.39763 \times 1.172 = 86.02202236 \cong 86$.

The total sample size $n$ is $\sum_{h=1}^{7} n_{h(m)} = 401$.

The estimated variances $\nu(\bar{y}_{1st})$ (ignoring fpc) under the compromise mixed allocation for the two characteristics are

$\nu(\bar{y}_{1st}) = 0.521803789$ and $\nu(\bar{y}_{2st}) = 0.806991877$ respectively with a trace value of $1.3287956686$.

5.7 A COMPARATIVE STUDY WITH SOME OTHER ALLOCATIONS

In this section a comparative study of the proposed allocation is made with proportional and some other compromise allocations.

The allocations compared are

1. Proportional allocation
2. Cochran’s average compromise allocation
3. Chatterjee’s compromise allocation
4. Sukhatme’s compromise allocation
5. Proposed compromise mixed allocation

5.7.1 Proportional Allocation

The proportional allocation is given by

$$n_h = nW_h; h = 1, 2, \ldots, L$$ (5.47)

where $n = \sum_{h=1}^{L} n_h$ is the total sample size.

Taking $n = 401$, the rounded off proportional allocation is obtained as:

$n_{1(prop)} = 76, n_{2(prop)} = 90, n_{3(prop)} = 68, n_{4(prop)} = 35, n_{5(prop)} = 37, n_{6(prop)} = 53,$
\[ n_{7{(prop)}} = 42 \text{ with } V_{prop}(\bar{y}_{1st}) = 0.80473883, V_{prop}(\bar{y}_{2st}) = 0.1189056398 \text{ with trace value 1.993795228.} \]

5.7.2 Cochran’s Average Compromise Allocation

Cochran (1977) gave the compromise criteria by averaging the individual optimum allocations \( n_{ih}^*; h = 1, 2, \ldots, L; \ l = 1, 2, \ldots, p \) that are solutions to the following NLPP for all the characteristics separately.

\[
\text{Minimize } V_l = \sum_{h=1}^{L} \frac{W_h^2 s_{ih}^2}{n_{ih}}
\]

subject to \( \sum_{h=1}^{L} \hat{c}_{ih} n_{ih} \leq C_{0l}; l = 1, 2, \ldots, p \) \hfill (5.48)

and \( 2 \leq n_{ih} \leq N_h; h = 1, 2, \ldots, L \)

where \( C_{0l} \) denote the amount assigned for measuring the \( l^{th} \) characteristics; \( l = 1, 2, \ldots, p \).

Let \( \bar{n}_l^* = (n_{1l}^*, n_{2l}^*, \ldots, n_{Ll}^*) \) denote the solution to the \( l^{th} \) NLPP in (5.48) with \( V_l^* \) as the value of the objective function.

Cochran’s average compromise allocation is then given by

\[
n_{h(a)} = \frac{1}{p} \sum_{l=1}^{p} n_{ih}^*; h = 1, 2, \ldots, L \hfill (5.49)
\]

where the suffix ‘a’ stands for ‘Average Compromise Allocation’.

Substituting the numerical values of the parameters the NLPP (5.48) and their optimal solutions \( n_{ih}^*; l = 1, 2 \) with the corresponding values of \( V_l^* \) using LINGO, are obtained as follows.

For \( l = 1 \)

With cost \( C_{01} = 2300 \) for the first characteristics
Minimize  
\[ V_1 = \left( \frac{0.979684146}{n_{11}} + \frac{1.700163584}{n_{12}} + \frac{0.7929293}{n_{13}} + \frac{4.932558351}{n_{14}} + \frac{4.274872038}{n_{15}} + \frac{3.927929807}{n_{16}} + \frac{18.09013978}{n_{17}} \right) \]  
subject to  
\[ (3n_{11} + 5n_{12} + 4n_{13} + 7n_{14} + 6n_{15} + 6n_{16} + 9n_{17}) \leq 2300 \]  
\[ 2 \leq n_{11} \leq 472 \]  
\[ 2 \leq n_{12} \leq 559 \]  
\[ 2 \leq n_{13} \leq 425 \]  
\[ 2 \leq n_{14} \leq 218 \]  
\[ 2 \leq n_{15} \leq 233 \]  
\[ 2 \leq n_{16} \leq 328 \]  
\[ 2 \leq n_{17} \leq 265 \]  

The optimum allocation \( n_1^* = (n_{11}^*, n_{12}^*, \ldots, n_{17}^*) \) is  
\[ n_{11}^* = 37.58944 \pm 37, \quad n_{12}^* = 38.35692 \pm 38, \quad n_{13}^* = 29.28672 \pm 29, \quad n_{14}^* = 55.21673 \pm 55, \]  
\[ n_{15}^* = 55.52257 \pm 55, \quad n_{16}^* = 53.22184 \pm 53, \quad n_{17}^* = 93.25741 \pm 93. \]  
The corresponding value of the variance ignoring fpc is \( V_1^* = 0.524518648 \).

For \( l = 2 \)

With cost \( C_{02} = 1700 \) for the second characteristics the results are  
\[ n_{21}^* = 39.08367 \pm 39, \quad n_{22}^* = 47.11587 \pm 47, \quad n_{23}^* = 34.74994 \pm 35, \quad n_{24}^* = 53.71892 \pm 54, \]  
\[ n_{25}^* = 61.43807 \pm 61, \quad n_{26}^* = 55.40147 \pm 55, \quad n_{27}^* = 89.96012 \pm 90. \]  
with \( V_2^* = 0.826655611 \).
Using formula (5.49) the rounded off Cochran’s Average Compromise Allocation \( n_{(a)} \) obtained by averaging \( n_1^* \) and \( n_2^* \) is given below.

\[
\begin{align*}
\hat{n}_1^* &= (n_{11}^*, n_{12}^*, n_{13}^*, n_{14}^*, n_{15}^*, n_{16}^*, n_{17}^*) \\
&= (37.58944, 38.35692, 29.28672, 55.21673, 55.52257, 53.22184, 93.24741) \\
\end{align*}
\]

and \( \hat{n}_2^* = (n_{21}^*, n_{22}^*, n_{23}^*, n_{24}^*, n_{25}^*, n_{26}^*, n_{27}^*) \)

\[
\begin{align*}
&= (39.08367, 47.11587, 34.74994, 53.71892, 61.43807, 55.40147, 89.96012) \\
\end{align*}
\]

\[
\begin{align*}
\hat{n}_{(a)} &= (n_{1(a)}^*, n_{2(a)}^*, n_{3(a)}^*, n_{4(a)}^*, n_{5(a)}^*, n_{6(a)}^*, n_{7(a)}^*) \\
\text{where } \hat{n}_{h(a)} &= \frac{n_{h1}^* + n_{h2}^*}{2}, \ h = 1, 2, ..., 7.
\end{align*}
\]

This gives the rounded off average allocation \( \hat{n}_{(a)} = (38, 43, 32, 54, 58, 54, 92) \) with variances (ignoring fpc) as

\[
V_a(\hat{y}_{1st}) = 0.524518648 \quad \text{and} \quad V_a(\hat{y}_{2nd}) = 0.826655611. \quad \text{Where the suffix ‘a’ corresponds to the ‘Average Compromise Allocation’.}
\]

The ‘Trace’ is equal to \( 0.524518648 + 0.826655611 = 1.35117426 \).

### 5.7.3 Chatterjee’s Compromise Allocation

Chatterjee (1967) obtained the compromise allocation by minimizing the sum of the relative increase \( E_i \) in the variances of the estimates \( \hat{y}_{ist} \) of the population means \( \hat{y}_i; i = 1, 2, ..., p \).

Chatterjee formulated the problem as

\[
\begin{align*}
\text{Minimize } & \sum_{i=1}^{p} E_i = \frac{1}{C_0} \sum_{i=1}^{p} \sum_{h=1}^{L} \tilde{c}_h (n_{ih}^* - n_h^*)^2 \\
\text{subject to } & \sum_{h=1}^{L} \tilde{c}_h n_h \leq C_0 \\
& 2 \leq n_h \leq N_h; h = 1, 2, ..., L \\
\end{align*}
\]

where \( n_{ih}^* \) is the usual optimum allocation for fixed budget \( C_0 \) for the \( i^{th} \) characteristic in \( h^{th} \) stratum.
On compromise mixed allocation...

Substituting the numerical values of the parameters the NLPP (5.53) becomes

Minimize \( E = \frac{1}{4000} \left[ \frac{6}{n_1} \left( 37 - n_1 \right)^2 \right. \left. + \frac{39 - n_1}{n_2} \right] \)
\[ + \frac{7}{n_3} \left( 39 - n_3 \right)^2 + \frac{35 - n_3}{n_4} \left( 55 - n_4 \right)^2 \right. \]
\[ + \frac{12}{n_5} \left( 61 - n_5 \right)^2 \left. + \frac{10}{n_6} \left( 53 - n_6 \right)^2 \right. \]
\[ + \left. \frac{15}{n_7} \left( 93 - n_7 \right)^2 \left( 90 - n_7 \right)^2 \right] \]

\[ = \frac{1}{4000} \left[ \frac{6}{n_1} \left[ 2890 + 2n_1^2 - 152n_1 \right] + \frac{8}{n_2} \left[ 3653 + 2n_2^2 - 170n_2 \right] \right. \]
\[ + \frac{7}{n_3} \left[ 2066 + 2n_3^2 - 128n_3 \right] + \frac{12}{n_4} \left[ 5941 + 2n_4^2 - 218n_4 \right] \]
\[ + \frac{11}{n_5} \left[ 6746 + 2n_5^2 - 232n_5 \right] + \frac{10}{n_6} \left[ 5834 + 2n_6^2 - 216n_6 \right] \]
\[ + \frac{15}{n_7} \left[ 16749 + 2n_7^2 - 366n_7 \right] \]

subject to \((6n_1 + 8n_2 + 7n_3 + 12n_4 + 11n_5 + 10n_6 + 15n_7) \leq 4000 \)
\[ 2 \leq n_{1i} \leq 472 \]
\[ 2 \leq n_{12} \leq 559 \]
\[ 2 \leq n_{13} \leq 425 \]
\[ 2 \leq n_{14} \leq 218 \]
\[ 2 \leq n_{15} \leq 233 \]
\[ 2 \leq n_{16} \leq 328 \]
\[ 2 \leq n_{17} \leq 265 \]

where the numerical values of \( n_i \) are as computed in section 5.7.2.
On compromise mixed allocation ...

Using LINGO, the chatterjee's compromise allocation (rounded off) are obtained as:

\[ n_1 = 38, n_2 = 43, n_3 = 32, n_4 = 54, n_5 = 58, n_6 = 54, n_7 = 91. \]

With variances

\[ V_C(\bar{x}_{1y}) = 0.0526679439, \quad V_C(\bar{x}_{2y}) = 0.829416763 \]

with a trace value of 1.356096202, where 'C' stands for Chatterjee.

5.7.4 Sukhatme's Compromise Allocation

Sukhatme (1984) obtained the compromise allocation by minimizing the sum of the variances for the \( p \) characteristics under linear cost constraints. The NLPP for this allocation is given as:

\[
\text{Minimize } V = \sum_{l=1}^{p} \sum_{h=1}^{L} \frac{W_h^2 \sigma_{lh}^2}{n_h}
\]

subject to \( \sum_{l=1}^{p} \sum_{h=1}^{L} \sigma_{lh} n_h \leq C_0 \)

and \( 2 \leq n_h \leq N_h, h = 1, 2, ..., L \) \( (5.57) \)

Using the values given in Table 5.1 the NLPP (5.57) may be given as

\[
\text{Minimize } V = \left( \frac{3.1613085}{n_1} + \frac{4.870634496}{n_2} + \frac{2.5175657}{n_3} + \frac{11.80154696}{n_4} \right.
\]

\[
+ \left. \frac{13.25977325}{n_5} + \frac{9.772743314}{n_6} + \frac{41.20650191}{n_7} \right]
\]

\( (5.58) \)

subject to \( (6n_1 + 8n_2 + 7n_3 + 12n_4 + 11n_5 + 10n_6 + 15n_7 ) \leq 4000 \)

\( (5.59) \)

\[ 2 \leq n_{11} \leq 472 \]

\[ 2 \leq n_{12} \leq 559 \]

\[ 2 \leq n_{13} \leq 425 \]

\[ 2 \leq n_{14} \leq 218 \]

\[ 2 \leq n_{15} \leq 233 \]

\[ 2 \leq n_{16} \leq 328 \]

\[ 2 \leq n_{17} \leq 265 \]

\( (5.60) \)

The rounded off solution to NLPP (5.58)-(5.60) is given as
On compromise mixed allocation...

\[ n_1 = 39, n_2 = 42, n_3 = 33, n_4 = 54, n_5 = 60, n_6 = 54, n_7 = 90 \] with the objective value as 1.351304.
The trace value is 1.351304 which is the value of the objective function of NLPP (5.58)-(5.60).

5.8 SUMMARY OF THE RESULTS

Table 5.3 gives the summary of the results obtained for the second numerical illustration. The 'Proportional', 'Cochran', 'Chatterjee', 'Sukhatme' and 'Proposed' allocations and the corresponding Trace values are summarised.

<table>
<thead>
<tr>
<th>Allocations</th>
<th>Sample Sizes under various Allocation</th>
<th>R.E. w.r.t. Proportional Allocation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( n_1 )</td>
<td>( n_2 )</td>
</tr>
<tr>
<td>(i) Proportional</td>
<td>76</td>
<td>90</td>
</tr>
<tr>
<td>(ii) Cochran</td>
<td>38</td>
<td>43</td>
</tr>
<tr>
<td>(iii) Chatterjee</td>
<td>38</td>
<td>43</td>
</tr>
<tr>
<td>(iv) Sukhatme</td>
<td>39</td>
<td>42</td>
</tr>
<tr>
<td>(v) Proposed</td>
<td>51</td>
<td>51</td>
</tr>
</tbody>
</table>

The last column of Table 5.3 provides the Relative Efficiencies (R.E.) of the four compromise allocations as compared to the proportional allocation. The Relative Efficiency (R.E.) of an allocation \( \eta' \) as compared to the proportional allocation \( \eta \) is defined as
\[
R.E. = \frac{Trace(\eta)}{Trace(\eta')}.
\]
(See Sukhatme (1984)). An examination of the R.E. recorded in the last column of Table 5.3 reveals that the proposed allocation is the most efficient among the considered allocations.

To justify the above claim in the next section a simulation study is carried out.
5.9 A Simulation Study

For simulation study a stratified population with seven strata and two independent characteristics is considered. It is assumed that the two characteristics are distributed independently and normally within each stratum with strata means and strata standard deviations as given in Table 5.4 and Table 5.5 respectively.

Corresponding to the two characteristics, seven independent normal populations are generated, one for each stratum, through the software ‘R’ (Venables WN, Smith DM, the R Core Team (2012)).

**Table 5.4: Strata means ($\bar{Y}_{ih}$)**

<table>
<thead>
<tr>
<th>$h \rightarrow$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>$l \downarrow$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>65</td>
<td>60</td>
<td>62</td>
<td>39</td>
<td>38</td>
<td>45</td>
<td>62</td>
</tr>
<tr>
<td>2</td>
<td>55</td>
<td>50</td>
<td>57</td>
<td>50</td>
<td>55</td>
<td>53</td>
<td>52</td>
</tr>
</tbody>
</table>

**Table 5.5: Strata standard deviations ($S_{ih}$)**

<table>
<thead>
<tr>
<th>$h \rightarrow$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>$l \downarrow$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>25</td>
<td>22</td>
<td>15</td>
<td>40</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>8</td>
<td>7</td>
<td>30</td>
<td>32</td>
<td>18</td>
<td>45</td>
</tr>
</tbody>
</table>

It is assumed that the budget of the survey $C=4500$ units with an overhead cost $c_0=500$ units. The available amount for measurements $C_0=C-c_0=4500-500=4000$ units. This amount is bifurcated in proportion to $\sum_{h=1}^{7} c_{ih}$ and $\sum_{h=1}^{7} c_{2h}$ (approximately) as $C_{01}=2300$ and $C_{02}=1700$ for the first ($l=1$) and second ($l=2$) characteristics respectively.

By repeated sampling, from the generated populations four different samples of observations are randomly selected from each of the seven strata and two
characteristics and the values of the strata variances and the fourth moments about mean are computed and listed in Table 5.6. Table 5.7 gives the expected measurement cost and their estimated variances. The sample sizes $n_h, h=1,2,...,7$ are $n_1=472, n_2=559, n_3=425, n_4=218, n_5=233, n_6=328$ and $n_7=265$ with the total sample size $n = \sum_{h=1}^{7} n_h = 2500$. This gives the expected stratum weights as $\hat{w}_1=0.189, \hat{w}_2=0.224, \hat{w}_3=0.170, \hat{w}_4=0.087, \hat{w}_5=0.093, \hat{w}_6=0.131$ and $\hat{w}_7=0.106$ respectively.

| Samples | $h$ | $s_{ih}^2$ | $s_{2h}^2$ | $\hat{C}_{1h}^4$ | $\hat{C}_{2h}^4$ |
|---------|-----|---------|---------|<<<<<<|<<<<<<|
| 1       | 38.226 | 56.647 | 3999.737 | 10245.520 |
| 2       | 35.929 | 61.787 | 3981.273 | 10444.760 |
| 3       | 22.865 | 49.130 | 1637.434 | 7284.730 |
| First   | 4    | 676.791 | 959.956 | 1370394.000 | 2651515.000 |
| Sample  | 5    | 571.771 | 1030.628 | 912987.900 | 2740463.000 |
| 6       | 259.942 | 307.201 | 218262.200 | 303569.500 |
| 7       | 1533.817 | 2356.312 | 6829099.000 | 16474685.000 |
| Second  | 4    | 578.269 | 795.062 | 853578.000 | 2029287.000 |
| Sample  | 5    | 501.232 | 978.141 | 793075.100 | 2925479.000 |
| 6       | 286.625 | 344.055 | 217872.800 | 337723.000 |
| 7       | 1306.864 | 2049.082 | 4741409.000 | 12138822.000 |
Table 5.7: Expected cost with the estimates of their variances

<table>
<thead>
<tr>
<th>$h$</th>
<th>$\hat{c}_{1h}$</th>
<th>$\hat{c}_{2h}$</th>
<th>$\hat{c}<em>h = \hat{c}</em>{1h} + \hat{c}_{2h}$</th>
<th>$\sigma^2_{c_h}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>3</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>3</td>
<td>8</td>
<td>12</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>3</td>
<td>7</td>
<td>15</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>5</td>
<td>12</td>
<td>20</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>5</td>
<td>11</td>
<td>22</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>4</td>
<td>10</td>
<td>27</td>
</tr>
<tr>
<td>7</td>
<td>9</td>
<td>6</td>
<td>15</td>
<td>38</td>
</tr>
</tbody>
</table>
5.9.1 Analysis of the Simulation Study

In the following the detailed analysis of the first simulated sample has been carried out. The remaining three samples may also be analyzed similarly.

For the sake of simplicity it is assumed that the weights \( k_1 = k_2 = 0.5 \), \( k_1' = k_2' = 0.5 \) and \( a_1 = a_2 = 0.5 \).

The NLPP (5.32)-(5.34) with data values from Table 5.6 and 5.7 for the first sample will be

\[
\text{Minimize } \phi(\alpha) = \\
0.5 \left[ 0.5 \left( \frac{10.37261195}{(\alpha_1 - 1)} + \frac{12.38853804}{(0.087 \alpha_2 - 1)} + \frac{13.85914895}{(0.093 \alpha_2 - 1)} + \frac{9.732741023}{(0.698 \alpha_3 - 1)} + \frac{43.70948944}{(1.207 \alpha_3 - 1)} \right) \\
+ 0.5 \left( \frac{40.67506595}{(\alpha_1 (\alpha_1 - 1)^2)} + \frac{151.3794522}{0.087 \alpha_2 (0.087 \alpha_2 - 1)^2} + \frac{0.093 \alpha_2 (0.093 \alpha_2 - 1)^2}{169.384016} \right) \\
+ \frac{105.9874005}{0.698 \alpha_3 (0.698 \alpha_3 - 1)^2} + \frac{1944.086146}{1.207 \alpha_3 (1.207 \alpha_3 - 1)^2} \right]^{1/2}
\]

subject to

\[
0.5[21 \alpha_1 + 2.067 \alpha_2 + 25.085 \alpha_3 ] \\
+ 0.5 \sqrt{(37 \alpha_1^2 + 0.341658 \alpha_2^2 + 68.51477 \alpha_3^2 )} \leq 4000
\]

\[
2 \leq \alpha_1 \leq 425
\]

\[
23 \leq \alpha_2 \leq 2505
\]

\[
3 \leq \alpha_3 \leq 220
\]

Using optimization software LINGO the solution to the NLPP (5.61)-(5.63) is obtained as:

\( \alpha_1 = 50.32964 \), \( \alpha_2 = 620.6144 \) and \( \alpha_3 = 72.05850 \).

With these values of \( \alpha_j; j = 1, 2, 3 \), the compromise mixed allocation given by (5.3) is

For \( j = 1 \) \( n_{1(m)} = n_{2(m)} = n_{3(m)} = \alpha_1 = 50.32964 \equiv 50 \)
For \( j = 2 \) \( n_{4(m)} = \alpha_2 \beta_4 = 620.6144 \times 0.087 = 53.9934 \equiv 54 \)
\[ n_{5(m)} = \alpha_2 \beta_5 = 620.6144 \times 0.093 = 57.7171 \equiv 58 \]
For \( j = 3 \) \( n_{6(m)} = \alpha_3 \beta_6 = 72.05850 \times 0.698 = 50.2968 \equiv 50 \)
\[ n_{7(m)} = \alpha_3 \beta_7 = 72.05850 \times 1.207 = 86.9746 \equiv 87. \]

The total sample size \( n = \sum_{h=1}^{7} n_{h(m)} = 399. \)

The estimated variances \( \nu(\overline{y}_{1st}) \) (ignoring fpc) under the compromise mixed allocation for the two characteristics are

\( \nu(\overline{y}_{1st}) = 0.544016165 \) and \( \nu(\overline{y}_{2st}) = 0.828867038 \) with a trace value of 1.372883199.

A similar analysis for the remaining three samples has also been performed. The details are omitted for want of space. The results of all the four samples are summarized in Table 5.8.

### Table 5.8: Summary of results of the simulation study for the proposed allocation

<table>
<thead>
<tr>
<th>Sample</th>
<th>Proposed Allocation</th>
<th>Trace</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( n_1 )</td>
<td>( n_2 )</td>
</tr>
<tr>
<td>First Sample</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>Second Sample</td>
<td>53</td>
<td>53</td>
</tr>
<tr>
<td>Third Sample</td>
<td>51</td>
<td>51</td>
</tr>
<tr>
<td>Fourth Sample</td>
<td>50</td>
<td>50</td>
</tr>
</tbody>
</table>

### 5.10 CONCLUSION

To draw the conclusion from the above study in this section a comparison has been made between the 'Proportional', 'Cochran', 'Chatterjee', 'Sukhatme' and 'Proposed' allocations obtained via a simulation study. For want of space the details of the computations regarding 'Proportional', 'Cochran' and 'Proposed' allocations are
omitted except for the first sample. However, these calculations are exactly similar as given in section 5.6.2, for the illustrated numerical example. Summary of the results obtained from the simulation study are given in Table 5.9.

Table 5.9: Summary of the results of the simulation study

<table>
<thead>
<tr>
<th>Allocations</th>
<th>Sample Sizes</th>
<th>Trace</th>
<th>R.E. w.r.t. Proportional Allocation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n_1$</td>
<td>$n_2$</td>
<td>$n_3$</td>
</tr>
<tr>
<td>First Sample</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(i) Proportional</td>
<td>76</td>
<td>89</td>
<td>68</td>
</tr>
<tr>
<td>(ii) Cochran</td>
<td>40</td>
<td>42</td>
<td>29</td>
</tr>
<tr>
<td>(iii) Chatterjee</td>
<td>40</td>
<td>42</td>
<td>28</td>
</tr>
<tr>
<td>(iv) Sukhatme</td>
<td>40</td>
<td>42</td>
<td>29</td>
</tr>
<tr>
<td>(v) Proposed</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>Second Sample</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(i) Proportional</td>
<td>77</td>
<td>91</td>
<td>69</td>
</tr>
<tr>
<td>(ii) Cochran</td>
<td>41</td>
<td>45</td>
<td>30</td>
</tr>
<tr>
<td>(iii) Chatterjee</td>
<td>41</td>
<td>45</td>
<td>29</td>
</tr>
<tr>
<td>(iv) Sukhatme</td>
<td>42</td>
<td>44</td>
<td>30</td>
</tr>
<tr>
<td>(v) Proposed</td>
<td>53</td>
<td>53</td>
<td>53</td>
</tr>
<tr>
<td>Third Sample</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(i) Proportional</td>
<td>76</td>
<td>90</td>
<td>68</td>
</tr>
<tr>
<td>(ii) Cochran</td>
<td>40</td>
<td>42</td>
<td>30</td>
</tr>
<tr>
<td>(iii) Chatterjee</td>
<td>40</td>
<td>42</td>
<td>30</td>
</tr>
<tr>
<td>(iv) Sukhatme</td>
<td>41</td>
<td>42</td>
<td>30</td>
</tr>
<tr>
<td>(v) Proposed</td>
<td>51</td>
<td>51</td>
<td>52</td>
</tr>
</tbody>
</table>
Fourth Sample

(i) Proportional  75  89  67  35  37  52  42  2.084101963  1.000000000
(ii) Cochran  40  42  28  53  60  51  95  1.365279004  1.526502610
(iii) Chatterjee  40  42  27  53  59  50  95  1.375073235  1.515629793
(iv) Sukhatme  41  42  28  53  61  50  94  1.366151000  1.525528264
(v) Proposed  50  50  50  53  57  48  89  1.355874391  1.537090734

The last column of Table 5.9 provides a within sample comparison for the four simulated samples on the basis of the trace values. The relative efficiencies (R.E.) as compared to the proportional allocation are given in the last column of Table 5.9. An observation of the R.E.s reveals that in all the four simulated samples the proposed allocation is the most efficient. Thus we conclude that the authors’ proposed compromise mixed allocation is the most efficient among the compared compromise allocations.