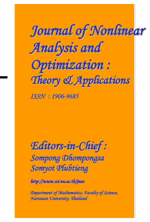

Journal of **N**onlinear **A**nalysis and **O**ptimization
Vol. 5, No. 1, (2014), 67-80
ISSN : 1906-9685
http://www.math.sci.nu.ac.th

A COMPROMISE ALLOCATION IN MULTIVARIATE STRATIFIED SAMPLING IN PRESENCE OF NON-RESPONSE

YASHPAL SINGH RAGHAV, IRFAN ALI*, ABDUL BARI

Department of Statistics & Operations Research, Aligarh Muslim University Aligarh, 202002, India

ABSTRACT. In this paper, we have considered the problem of determining the optimum allocation of sample sizes and sampling fractions for all the characteristics, which minimizes the coefficients of variance among the non-respondent to various characteristics under study. The problem is formulated as a Multi objective Programming Problem (MNLPP) and Multi objective Goal Programming Problem. A solution procedure is developed by using Lagrange Multiplier's Technique (LMT) and Goal Programming technique. A numerical example is presented to illustrate the computational details.

KEYWORDS : Compromise allocation; Compromise criterion; Non-response; Lagrange Multiplier's Technique; Multivariate Stratified Sampling; Non-linear programming problem; Goal programming.
AMS Subject Classification:

1. INTRODUCTION

While conducting a mailed survey, the problem of non-response generally occurs. In such cases, Hansen and Hurwitz [18] have suggested the method of sub-sampling from non-respondents to provide an estimator for population mean. They selected a preliminary sample and mailed the questionnaires to all the selected units. Non-respondents are identified and a second attempt was made by interviewing a sub-sample of non-respondents. They constructed the estimate of the population mean by combining the data from the two attempts and derived the expression for the sampling variance of the estimate. The optimum sampling fraction among the non-respondents is also obtained. El-Badry [11] has extended Hansen and Hurwitz's technique based on the experience that an appreciable in response rates to mail questionnaires can be secured by sending waves of questionnaires to the non-respondents group. Foradari [13] has generalized El-Badry's approach and has also studied the uses of Hansen and Hurwitz's technique under

* *Corresponding author.*
Email address : irfii.ali@gmail.com(Irfan Ali).
Article history : Received ...... Accepted .....

different designs. Srinath [37] described the selection of subsamples by making several attempts in non-respondents group.

Stratified sampling design is most widely sampling design for estimating the population parameters of a heterogeneous population. It deals with the properties of the estimates constructed from a stratified random sample and with the optimum choice of the sizes of the samples to be selected from various strata either to maximize the precision of the constructed estimate for a fixed cost or to minimize the cost of the survey for a fixed precision of the estimate. The sample sizes allocated according to either of the above criteria is called an "Optimum allocation". Khare [26] discussed 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. Khan et al. [25], has worked on optimum allocation and optimum size of subsamples to various strata in multivariate stratified sampling in presence of non-response by Lagrange Multiplier's Techniques.

For a population the coefficient of variation (C.V.) is represented by the ratio of population standard deviation to the population mean. It is useful statistic for comparing the degree of variation from one data series to another, even if the means are drastically different from each other. It also eliminates the effect of the scale of the observations.

The "optimum allocation" in stratified random sampling and its solution is well known in sampling literature for univariate population (see Cochran, [6]; Shukhatme et al., [38]. In multivariate populations where more than one characteristic are to be studied on every selected unit of the population. The problem of finding an optimum allocation becomes more complex due to complicating behavior of characteristics. Various another's such as Dalenius [7, 8], Ghosh [14], Yates [42], Aoyama [3], Gren [16, 17], Folks and Antle [12], Hertley [19], Kokan and Khan [28], Chromy [5], Wywial [41], Bethel [4], Kreienbrock [30], Jahan et al. [22], Khan et al. [23, 24], Ahsan et al. [1], Díaz-García and Ulloa [9, 10], Ansari et al. [2] etc. used different compromise criteria to work out a compromise allocation that is optimum for all characteristics in some sense.

When some auxiliary information is available, it can be used to increase the precision of estimates. Ige and Tripathi [21], Rao [35], Tripathi and Bahl [39] and some other discussed the use of auxiliary information in stratified sampling using double sampling technique.

The problem of optimum allocation, where the strata weights are unknown and non-response also occures have been studied by some authors. Okafor [34], solved the above problem for stratified population in univariate case using a double sampling strategy (DDS). The same problem was also formulated by Najmussehar and Bari [33] using dynamic programming technique to obtain a solution. A comparative study has also been done by Varshney et al. [40] by developing a goal programming to solve the problem.

In sampling literature, the allocation based on compromise criteria is termed as the "compromise allocation". Holmberg [20], discussed three compromise criteria to work out a compromise allocation.

- **(i):** Minimizing the sum of sampling variances of the estimators of the population parameters of various characteristics.
- **(ii):** Minimizing the sum of squared coefficients of variation of the estimator over the characteristics.
- **(iii):** Minimizing the sum of efficiency losses for not using the individual optimum allocation.

Kozak [29], used the concept of minimizing some function of squared coefficient of variations as an objective function. Latest Ghufran et al. [15] also discussed this concept.

In the present manuscript, we have developed a method to work out the compromise allocation in multivariate stratified sampling in presence of non-response using the criterion "Minimizing the sum of squared coefficients of variations of the estimators over the characteristics". It is assumed that the estimation of population means is of interest. The problem is formulated as a Multi-objective Non-Linear Programming Problem (MNLPP) and Multi-objective Non Linear Goal Programming Problem (MNLGPP). The Lagrange Multipliers Technique and goal programming are used to obtain a solution of formulated problem.

2. SAMPLING STRATEGIES AND ESTIMATION PROCEDURE

Let us consider a population consisting of N units divided into k strata. Let N_i, $\bar{Y}_{ji}$, S^2_{ji} and $p_i = \frac{N_i}{N}(i = 1, 2, \ldots, k, j = 1, 2, \ldots, p)$ denote respectively, the stratum size, stratum mean, stratum variance and stratum weight of j^{th} character in i^{th} stratum, where p is the number of variables under study. Assume that every stratum is divided into two disjoint groups of respondents and non-respondents, with N_{i1} and $N_{i2} = N - N_{i1}$ as the sizes of respondents and non-respondents in the i^{th} stratum respectively. We decide to select a sample of size n from the entire population in such a way that n_i units are selected from the N_i units in the i^{th} stratum with $\sum_{i=1}^{k} n_i = n$. Let $p \geq 2$, characteristics be defined on each population unit and the estimation of the p unknown population means $\bar{Y}_j;\ j = 1, 2, \ldots, p$ are of interest.
The population mean $\bar{Y}_j$ of the j^{th} characteristic is given by:

$$\bar{Y}_j = \frac{1}{N} \sum_{i=1}^{k} \sum_{h=1}^{N_i} y_{jih};\ j = 1, 2, \ldots, p \tag{2.1}$$

$$= \sum_{i=1}^{k} p_i \bar{Y}_{ji} \tag{2.2}$$

where y_{jih} is the value of the j^{th} variable (characteristic) of the h^{th} element in the i^{th} stratum, $j = 1, 2, \ldots, p;\ i = 1, 2, \ldots, k;\ h = 1, 2, \ldots, N_i$
$\bar{Y}_{ji} = \frac{1}{N_i} \sum_{h=1}^{N_i} y_{jih}$ is the i^{th} stratum mean for the j^{th} characteristic
Let n_{i1} be the number of units of the sample in the i^{th} stratum that provide the data sought and n_{i2} be the number of units of non-respondents. By extensive efforts, the data are later obtained from a random sample of u_{i2} out of n_{i2} unit such that

$$n_{i2} = K_i u_{i2} (K_i > 1) \tag{2.3}$$

where $\frac{1}{K_i}$, denotes the sampling fraction among non-respondents in the i^{th} stratum.
Now we have

$$E\left(\frac{n_{i1}}{N_{i1}}\right) = E\left(\frac{n_{i2}}{N_{i2}}\right) = K_i E\left(\frac{u_{i2}}{N_{i2}}\right) \tag{2.4}$$

Let $\bar{y}^*_{ji}$ be the unbiased estimator of population mean $\bar{Y}_{ji}$ in i^{th} stratum for j^{th} characteristics which is given by

$$\bar{y}^*_{ji} = \frac{n_{i1}\bar{y}_{jn_{i1}} + n_{i2}\bar{y}_{ju_{i2}}}{n_i} \tag{2.5}$$

where $\bar{y}_{jn_{i1}}$ and $\bar{y}_{ju_{i2}}$ are the means based on n_{i1} units of response group and u_{i2} units of sub-sample of non-response group respectively for the j^{th} characteristics in i^{th} stratum.
Using Hansen-Hurwitz technique, an unbiased estimator of population mean $\bar{Y}_j$ for the j^{th} characteristics is given by

$$\bar{y}^*_{jst} = \sum_{i=1}^{k} p_i \bar{y}^*_{ji} \forall j = 1, 2, \ldots, p. \tag{2.6}$$

$$i.e.,\ E(\bar{y}^*_{jst}) = \bar{Y}_j \forall j = 1, 2, \ldots, p. \tag{2.7}$$

Ignoring the terms independent of n_i and K_i of the variance of the estimator is given by the expression

$$V(\bar{y}^*_{jst}) = \sum_{i=1}^{k} \frac{p_i^2 S_{ji}^2}{n_i} + \sum_{i=1}^{k} \left(\frac{k_i - 1}{n_i}\right) W_{i2} p_i^2 S_{ji2}^2 \forall j = 1, 2, \ldots, p. \tag{2.8}$$

where S_{ji}^2 are the population mean-square errors of entire group for the j^{th} characteristic in the i^{th} stratum. And S_{ji2}^2 are mean-square errors of non-response group for the j^{th} characteristic in the i^{th} stratum. The $W_{i2} = \frac{N_{i2}}{N_i}$ is the Non-Response rate in the i^{th} stratum.It is assumed that information on all the units of sub sample selected from the non-response group is available.
The values of n_i and K_i are to be chosen so as to give the maximum precision for fixed cost. Let C_{i0} be the cost per unit of selecting n_i units, $C_{i1} = \sum_{j=1}^{p} C_{ji1}$ be the cost per unit in enumerating n_{i1} units and $C_{i2} = \sum_{j=1}^{p} C_{ji2}$ be the cost per unit in enumerating u_{i2} units from non-respondent group. Then the total cost for the i^{th} stratum is given by

$$C_i = C_{i0} n_i + C_{i1} n_{i1} + C_{i2} u_{i2}, \forall i = 1, 2, \ldots, k. \tag{2.9}$$

Since the values of n_{i1} and n_{i2} are not known until the first attempt is made, the expected cost is used in planning the sample. The expected value of n_{i1} and u_{i2} are $W_{i1} n_i$ and $\frac{W_{i2} n_i}{K_i}$ respectively.
The average cost for the i^{th} stratum is

$$E(C_i) = \left(C_{i0} + C_{i1} W_{i1} + C_{i2} \frac{W_{i2}}{K_i}\right) n_i \tag{2.10}$$

where $W_{i1} = \frac{N_{i1}}{N_i}$ is the Response rate in the i^{th} stratum.
Thus the total cost over all the strata is given by

$$C_0 = \sum_{i=1}^{k} E(C_i) = \sum_{i=1}^{k} \left[C_{i0} + C_{i1} W_{i1} + C_{i2} \frac{W_{i2}}{K_i}\right] n_i \tag{2.11}$$

3. STATEMENT OF THE PROBLEM

Obviously, the best compromise allocation will be that which minimizes all the p variances given by (2.8) simultaneously. Thus, for finding the optimum compromise allocation we need to solve the following Multi-objective Non-linear Programming

Problem (MNLPP):

$$\left.\begin{aligned} &Minimize \begin{pmatrix} V(\bar{y}^*_{1st}) \\ \vdots \\ V(\bar{y}^*_{pst}) \end{pmatrix} \\ &Subject\ to\ \sum_{i=1}^{k}\left(C_{i0}+C_{i1}W_{i1}+C_{i2}\frac{W_{i2}}{K_i}\right)n_i \le C_0 \\ &\qquad K_i>1,\ 2\le n_i\le N_i \\ &\qquad and\ n_i\ are\ integers; i=1,2,\dots,k \end{aligned}\right\} \tag{3.1}$$

Kish [27], Rios et al. [36], Miettinen [32], Khan et al. [24] studied the problem of multi-objective optimization in detailed. Khan et al. [24] and García and Ulloa [10] expressed the problem (3.1) under the value function technique as

$$\left.\begin{aligned} &Minimize\ \vartheta(V(\bar{y}^*_{jst})) \\ &Subject\ to\ \sum_{i=1}^{k}\left(C_{i0}+C_{i1}W_{i1}+C_{i2}\frac{W_{i2}}{K_i}\right)n_i \le C_0 \\ &\qquad K_i>1,\ 2\le n_i\le N_i \\ &\qquad and\ n_i\ are\ integers; i=1,2,\dots,k \end{aligned}\right\} \tag{3.2}$$

where $\vartheta(.)$ is a scalar function that summarizes the importance of each of the variance of the p characteristics. Usually, ϑ is taken as weighted sum of the p variances that is

$$\vartheta = \sum_{j=1}^{p} w_j V(\bar{y}^*_{jst}) \tag{3.3}$$

where $\sum_{j=1}^{p} w_j = 1,\ w_j \ge 0,\ j=1,2,\dots,p$
w_j are the weights according to the importance of each characteristics.
Many authors object on using the weighted sum of variances because variances are not unit free thus cannot be added. In this paper, instead of variances we use the squared coefficient of variation that is unit free and positive. Thus, our problem under the value function technique may be expressed as:

$$\left.\begin{aligned} &Minimize\ Z = \sum_{j=1}^{p} w_j (CV)_j^2 \\ &Subject\ to\ \sum_{i=1}^{k}\left(C_{i0}+C_{i1}W_{i1}+C_{i2}\frac{W_{i2}}{K_i}\right)n_i \le C_0 \\ &\qquad K_i>1,\ 2\le n_i\le N_i \\ &\qquad and\ n_i\ are\ integers; i=1,2,\dots,k \end{aligned}\right\} \tag{3.4}$$

where $(CV)_j = CV(\bar{y}^*_{j,st}) = \frac{SD(\bar{y}^*_{j,st})}{\bar{Y}_j};\ j=1,2,\dots,p$

$$(CV)_j^2 = \frac{V(\bar{y}^*_{j,st})}{\bar{Y}_j^2};\ j=1,2,\dots,p \tag{3.5}$$

$\bar{Y}_j$, $\bar{y}^*_{jst}$ and $V(\bar{y}^*_{jst})$ are as defined in (2.2), (2.6) and (2.8), respectively. AINLPP (3.4) may now be restated as

$$\left.\begin{array}{l} Minimize\ Z = \sum_{j=1}^{p} \bar{Y}_j^{-2}\{\sum_{i=1}^{k}\left(\frac{N_i-n_i}{N_i n_i}\right)p_i^2 S_{ji}^2 + \sum_{i=1}^{k}\left(\frac{k_i-1}{n_i}\right)W_{i2}p_i^2 S_{ji2}^2\} \\ Subject\ to\ \sum_{i=1}^{k}\left(C_{i0}+C_{i1}W_{i1}+C_{i2}\frac{W_{i2}}{K_i}\right)n_i \le C_0 \\ K_i > 1,\ 2 \le n_i \le N_i \\ and\ n_i\ are\ integers; i = 1,2,\ldots,k \end{array}\right\} \tag{3.6}$$

ignoring the terms independent of n_i and K_i the equation (3.6) can be rewritten as:

$$\left.\begin{array}{l} Minimize\ Z = \sum_{i=1}^{k}\frac{p_i^2 a_i^2}{n_i} + \sum_{i=1}^{k}\left(\frac{k_i-1}{n_i}\right)W_{i2}p_i^2 b_i^2 \\ Subject\ to\ \sum_{i=1}^{k}\left(C_{i0}+C_{i1}W_{i1}+C_{i2}\frac{W_{i2}}{K_i}\right)n_i \le C_0 \\ K_i > 1,\ 2 \le n_i \le N_i \\ and\ n_i\ are\ integers; i = 1,2,\ldots,k \end{array}\right\} \tag{3.7}$$

where

$$a_i^2 = \sum_{j=1}^{p}\bar{Y}_j^{-2}S_{ji}^2,\ b_i^2 = \sum_{j=1}^{p}\bar{Y}_j^{-2}S_{ji2}^2 \tag{3.8}$$

and all the p weights are assumed to be equal, that is $w_1 = w_2 = \cdots = w_p = \frac{1}{p}$ without loss of generality the common factor $\frac{1}{p}$ may be dropped from the objective function.

4. LAGRANGE MULTIPLIER'S TECHNIQUE

Taking care of the integer restrictions by rounding off the continuous solution the NLPP (3.7) (AINLPP (3.7) without integer restrictions) may be solved by Lagrange Multipliers Technique. For applying Lagrange Multipliers Technique (LMT), the restrictions $2 \le n_i \le N_i;\ i = 1,2,\ldots,k$ are ignored. We take cost constraint as an equation. To determine the optimum values of n_i and K_i for the cost function (2.11), we consider the function

$$\phi = \vartheta(V(\bar{y}^*_{jst})) + \mu(C_0) \tag{4.1}$$

where μ is the Lagrange Multiplier.

$$\phi(n_i,K_i,\mu) = \sum_{i=1}^{k}\left(\frac{1}{n_i}-\frac{1}{N_i}\right)p_i^2 a_i^2 + \sum_{i=1}^{k}\frac{(k_i-1)W_{i2}p_i^2 b_i^2}{n_i} + \mu\left[\sum_{i=1}^{k}\left(C_{i0}+C_{i1}W_{i1}+C_{i2}\frac{W_{i2}}{K_i}\right)n_i - C_0\right] \tag{4.2}$$

Now differentiating $\phi(n_i,K_i,\mu)$ w.r.t. K_i and equating to zero, noting that $\frac{\partial^2\phi}{\partial K_i^2} > 0$, we have

$$\frac{\partial\phi}{\partial K_i} = \frac{W_{i2}^2 p_i^2 b_i^2}{n_i} - \frac{\mu n_i C_{i2} W_{i2}}{K_i^2} = 0 \tag{4.3}$$

$$n_i = \frac{k_i p_i b_i}{\sqrt{\mu C_{i2}}} \tag{4.4}$$

again differentiating $\phi(n_i, K_i, \mu)$ w.r.t. n_i and equating to zero, noting that $\frac{\partial^2 \phi}{\partial n_i^2} > 0$, we have

$$\frac{\partial \phi}{\partial n_i} = -\frac{p_i^2 a_i^2}{n_i^2} - \frac{(k_i - 1) W_{i2} p_i^2 b_i^2}{n_i^2} + \mu \left(C_{i0} + C_{i1} W_{i1} + C_{i2} \frac{W_{i2}}{K_i} \right) = 0 \tag{4.5}$$

Now eliminating μ from (4.5) by putting its value from (4.4), we have

$$K_i = \sqrt{\frac{C_{i2}(a_i^2 + W_{i2} b_i^2)}{b_i^2 (C_{i0} + C_{i1} W_{i1})}} \tag{4.6}$$

which shows that K_i increases with the increase in C_{i2} and subsequently the number of units to be repeated from non-response group of the i^{th} stratum decreases. The total cost C_0 is fixed. So we have

$$C_0 = \sum_{i=1}^{k} \left(C_{i0} + C_{i1} W_{i1} + C_{i2} \frac{W_{i2}}{K_i} \right) n_i \tag{4.7}$$

Now putting the value of n_i from (4.4) in (4.7) and using the values of K_i from (4.6), we have

$$\frac{1}{\sqrt{\mu}} = \frac{C_0}{\sum_{i=1}^{k} p_i \left[\{ \sqrt{(C_{i0} + C_{i1} W_{i1}) * (a_i^2 - W_{i2} b_i^2)} \} + W_{i2} b_i \sqrt{C_{i2}} \right]} \tag{4.8}$$

Hence the value of n_i is given by

$$n_i = \frac{p_i C_0 \sqrt{(a_i^2 + W_{i2} b_i^2)/(C_{i0} + C_{i1} W_{i1})}}{\sum_{i=1}^{k} p_i \left[\{ \sqrt{(C_{i0} + C_{i1} W_{i1}) * (a_i^2 - W_{i2} b_i^2)} \} + W_{i2} b_i \sqrt{C_{i2}} \right]} \tag{4.9}$$

If the compromise allocation (4.9) satisfies the restriction $2 \le n_i \le N_i;\ i = 1, 2, \ldots, k$ then it will provide a continuous solution of NLPP (3.7). The continuous solution, rounded off to the nearest integer values of n_i, will then give an integer solution. After rounding off one has to be careful in rechecking that the rounded off values satisfy the cost constraints $\sum_{i=1}^{k} \left(C_{i0} + C_{i1} W_{i1} + C_{i2} \frac{W_{i2}}{K_i} \right) n_i \le C_0$. If the cost constraints are violated by the rounded off solution we may use some integer non linear programming technique. Ignoring the finite population correction (fpc) n_i given by (4.9) satisfies the restrictions restriction $2 \le n_i \le N_i;\ i = 1, 2, \ldots, k$.
However, if any $n_i > N_i$ it can be taken as equal to N_i and for the remaining *(k-1)* strata n_i are recalculated. Similarly, if $n_i < 2$ it can be put equal to 2 and the remaining n_i are recalculated.
As an alternative, the constraints $2 \le n_i \le N_i;\ i = 1, 2, \ldots, k$ may also be included and the Integer Non- Linear Programming Problem (3.7) may also be solved by other integer NLPP technique. Softwares are also available to solve AINLPP. One such software is LINGO. LINGO is a user's friendly package for constrained optimization developed by LINDO Systems Inc. A user's guide LINGO [31] is also available. For more information one can visit the site http://www.lindo.com.

5. THE GOAL PROGRAMMING TECHNIQUE

The problem (3.4) may be stated separately for all the p characteristics as

$$\left.\begin{array}{l} Minimize\ Z = [(CV)_1^2, (CV)_2^2, \ldots, (CV)_p^2] \\ Subject\ to\ \sum_{i=1}^{k}\left(C_{i0} + C_{i1}W_{i1} + C_{i2}\frac{W_{i2}}{K_i}\right) n_i \le C_0 \\ K_i > 1,\ 2 \le n_i \le N_i \\ and\ n_i\ are\ integers; i = 1, 2, \ldots, k \end{array}\right\} \quad (5.1)$$

Using (3.5), the equation (5.1) can be written as

$$\left.\begin{array}{l} Minimize\ Z = \left[\frac{V(\bar{y}_{1st}^*)}{\bar{Y}_1^2}, \frac{V(\bar{y}_{2st}^*)}{\bar{Y}_2^2}, \ldots, \frac{V(\bar{y}_{pst}^*)}{\bar{Y}_p^2}\right] \\ Subject\ to\ \sum_{i=1}^{k}\left(C_{i0} + C_{i1}W_{i1} + C_{i2}\frac{W_{i2}}{K_i}\right) n_i \le C_0 \\ K_i > 1,\ 2 \le n_i \le N_i \\ and\ n_i\ are\ integers; i = 1, 2, \ldots, k \end{array}\right\} \quad (5.2)$$

where $V(\bar{y}_{1st}^*),\ j = 1, 2, \ldots, p$ is defined in (2.8).
The equation (5.2) can be written as

$$\left.\begin{array}{l} Minimize\ Z = [Z_1, Z_2, \ldots, Z_p] \\ Subject\ to\ \sum_{i=1}^{k}\left(C_{i0} + C_{i1}W_{i1} + C_{i2}\frac{W_{i2}}{K_i}\right) n_i \le C_0 \\ K_i > 1,\ 2 \le n_i \le N_i \\ and\ n_i\ are\ integers; i = 1, 2, \ldots, k \end{array}\right\} \quad (5.3)$$

where $Z_j = \frac{V(\bar{y}_{jst}^*)}{\bar{Y}_j^2} = \bar{Y}_j^{-2}\{\sum_{i=1}^{k}\frac{p_i^2 S_{ji}^2}{n_i} + \sum_{i=1}^{k}\left(\frac{k_i-1}{n_i}\right) W_{i2}p_i^2 S_{ji2}^2\},\ j = 1, 2, \ldots, p.$
or
$Z_j = \sum_{i=1}^{k}\frac{p_i^2 S_{ji}^2 \bar{Y}_j^{-2}}{n_i} + \sum_{i=1}^{k}\left(\frac{k_i-1}{n_i}\right) W_{i2}p_i^2 S_{ji2}^2 \bar{Y}_j^{-2},\ j = 1, 2, \ldots, p.$ Let Z_j^* be the optimum value of Z_j obtained by solving the following Nonlinear Programming Problem(NLPP)

$$\left.\begin{array}{l} Minimize\ Z_j,\ j = 1, 2, \ldots, p \\ Subject\ to\ \sum_{i=1}^{k}\left(C_{i0} + C_{i1}W_{i1} + C_{i2}\frac{W_{i2}}{K_i}\right) n_i \le C_0 \\ K_i > 1,\ 2 \le n_i \le N_i \\ and\ n_i\ are\ integers; i = 1, 2, \ldots, k \end{array}\right\} \quad (5.4)$$

Further let

$$\tilde{Z}_j = \tilde{Z}_j(n_1, n_2, \ldots, n_i, \ldots, n_k) = \sum_{i=1}^{k}\frac{p_i^2 S_{ji}^2 \bar{Y}_j^{-2}}{n_i} + \sum_{i=1}^{k}\left(\frac{k_i - 1}{n_i}\right) W_{i2}p_i^2 S_{ji2}^2 \bar{Y}_j^{-2} \quad (5.5)$$

Denote the variance under the compromise allocation, where n_i and $K_i;\ i = 1,2,\ldots,k$ are to be worked out.
Obviously $\tilde{Z}_j \geq Z_j^*$ and $\tilde{Z}_j - Z_j^* \geq 0;\ j = 1,2,\ldots,p$ will give the increase in the variance due to not using the individual optimum allocation for j^{th} characteristic.
Now consider the following goal:
"Find n_i and K_i such that the increase in the value of the variance for each characteristic due to the use of compromise allocation, n_i instead of individual optimum allocation, should not greater than $x_j(j = 1,2,,p)$". Where $x_j \geq 0(j = 1,2,,p)$ are the unknown goal variables.
To achieve these goals n_i must satisfy

$$\tilde{Z}_j - Z_j^* \leq x_j;\ j = 1,2,\ldots,p \tag{5.6}$$

or

$$\tilde{Z}_j - x_j \leq Z_j^*;\ j = 1,2,\ldots,p$$

or

$$\sum_{i=1}^{k} \frac{p_i^2 S_{ji}^2 \bar{Y}_j^{-2}}{n_i} + \sum_{i=1}^{k} \left(\frac{k_i - 1}{n_i}\right) W_{i2} p_i^2 S_{ji2}^2 \bar{Y}_j^{-2} - x_j \leq Z_j^*;\ j = 1,2,\ldots,p \tag{5.7}$$

The value $\sum_{j=1}^{p} x_j$ will give us the total increase in variances by not using the individual optimum allocations.
This suggests the following Goal Programming Problem (GPP) to solve.

$$\left.\begin{array}{c} Minimize\ \sum_{j=1}^{p} x_j \\ \\ Subject\ to \\ \\ \sum_{i=1}^{k} \frac{p_i^2 S_{ji}^2 \bar{Y}_j^{-2}}{n_i} + \sum_{i=1}^{k} \left(\frac{k_i-1}{n_i}\right) W_{i2} p_i^2 S_{ji2}^2 \bar{Y}_j^{-2} - x_j \leq Z_j^*;\ j = 1,2,\ldots,p \\ \\ \sum_{i=1}^{k} \left(C_{i0} + C_{i1} W_{i1} + C_{i2} \frac{W_{i2}}{K_i}\right) n_i \leq C_0 \\ \\ K_i > 1,\ 2 \leq n_i \leq N_i \\ \\ and\ n_i\ are\ integers; i = 1,2,\ldots,k \end{array}\right\} \tag{5.8}$$

6. NUMERICAL ILLUSTRATION

The following numerical example is presented to illustrate the practical use and the computational details of working out the optimum sampling fractions among non-respondents $1/K_i$ (where K_i are defined in 4.6) and proposed allocation defined by (4.9). Consider a population of size N=3850 divided into four strata. Let the population means of the two characteristics defined on each unit of the population are assumed to be known as $\bar{Y}_1 = 24.73$ and $\bar{Y}_2 = 31.53$. It is also assumed that the relative values of variances of the non-respondent and respondents, that is, $S_{ji2}^2/S_{ji}^2 = 0.25$ for $j = 1,2$ and $i = 1,2,\ldots,4$. Further, let the total amount available for the survey is $C_0 = 5000$ units. Table 5.1 shows the relevant information.
Substituting the values from Table 5.1 in (3.8), we get

$$a_1^2 = 16.05 \qquad a_2^2 = 17.89$$

TABLE 1. Data for four strata and two characteristics

i	N_i	p_i	S_{1i}^2	S_{2i}^2	W_{i1}	W_{i2}	C_{i0}	C_{i1}	C_{i2}
1	1214	0.32	4817.72	8121.15	0.70	0.30	1	2	3
2	822	0.21	6251.26	7613.52	0.80	0.20	1	3	4
3	1028	0.27	3066.16	1456.40	0.75	0.25	1	4	5
4	786	0.20	6207.25	6977.72	0.72	0.28	1	5	6

$$a_3^2 = 6.48 \qquad a_4^2 = 17.18$$

$$b_1^2 = 4.02 \qquad b_2^2 = 4.47$$

$$b_3^2 = 1.62 \qquad b_4^2 = 4.29$$

We redefine $A_i = a_i^2 - W_{i2}b_i^2$ and $B_i = C_{i0} + C_{i1}W_{i1},\ i = 1, 2, \ldots, 4$

$$A_1 = 14.85 \qquad A_2 = 16.99$$

$$A_3 = 6.08 \qquad A_4 = 15.97$$

$$B_1 = 2.4 \qquad B_2 = 3.4$$

$$B_3 = 4.0 \qquad B_2 = 4.6$$

From column (11) &(12) of the Table 5.2, Sampling fractions among the non-respondents are $\frac{1}{K_1} = 0.46, \frac{1}{K_2} = 0.47, \frac{1}{K_3} = 0.46$ and $\frac{1}{K_4} = 0.45$ and the rounded off compromise allocations using the proposed method are obtained as: $n_1 = 528, n_2 = 311, n_3 = 221$ and $n_4 = 247$ with the value of the objective function as $Z = .0113$.

Solution by using Goal Programming:

Using the equation (5.4) the non linear programming problem for the first characteristic is

$$\left.\begin{array}{c} Minimize\ Z_1 = \frac{.8066}{n_1} + \frac{.4507}{n_2} + \frac{.3655}{n_3} + \frac{.4059}{n_4} \\ + \frac{(K_1-1)}{n_1} * 0.06 + \frac{(K_2-1)}{n_2} * 0.02 + \frac{(K_3-1)}{n_3} * 0.02 + \frac{(K_4-1)}{n_4} * 0.02 \\ Subject\ to \\ (2.4 + \frac{0.9}{K_1})n_1 + (3.4 + \frac{0.8}{K_2})n_2 + (4 + \frac{1.25}{K_3})n_3 + (4.6 + \frac{1.68}{K_4})n_4 \leq 5000 \\ 2 \leq n_1 \leq 1214;\ 2 \leq n_2 \leq 822 \\ 2 \leq n_3 \leq 1028;\ 2 \leq n_4 \leq 786 \\ K_i > 1,\ and\ n_i\ are\ integers; i = 1, 2, \ldots, 4 \end{array}\right\}$$

The optimal solution provided by LINGO is
$n_{1,1}^* = 482, n_{1,2}^* = 307, n_{1,3}^* = 253, n_{1,4}^* = 247, K_1 = 2.15, K_2 = 2.12, K_3 = 2.17, K_4 = 2.20$ with the value of the objective function as $Z_1 = 0.0067$. Where $n_{1,i}^*$ denote the optimum allocation for the first characteristic.

Similarly using (5.4) the non linear programming problem for the second characteristic is

$$\left.\begin{array}{c} Minimize\ Z_2 = \frac{.8365}{n_1} + \frac{.3377}{n_2} + \frac{.1068}{n_3} + \frac{.2807}{n_4} \\ + \frac{(K_1-1)}{n_1} * 0.06 + \frac{(K_2-1)}{n_2} * 0.02 + \frac{(K_3-1)}{n_3} * 0.07 + \frac{(K_4-1)}{n_4} * 0.02 \\ Subject\ to \\ (2.4 + \frac{0.9}{K_1})n_1 + (3.4 + \frac{0.8}{K_2})n_2 + (4 + \frac{1.25}{K_3})n_3 + (4.6 + \frac{1.68}{K_4})n_4 \leq 5000 \\ 2 \leq n_1 \leq 1214;\ 2 \leq n_2 \leq 822 \\ 2 \leq n_3 \leq 1028;\ 2 \leq n_4 \leq 786 \\ K_i > 1,\ and\ n_i\ are\ integers; i = 1, 2, \ldots, 4 \end{array}\right\}$$

The solution of NLPP provided by LINGO is given as
$n^*_{2,1} = 594, n^*_{2,2} = 320, n^*_{2,3} = 149, n^*_{2,4} = 249, K_1 = 2.16, K_2 = 2.11, K_3 = 1.0, K_4 = 2.21$ with the value of the objective function as $Z_2 = 0.0046$. Where $n^*_{2,i}$ denote the optimum allocation for the second characteristic.
After finding the optimum values of $Z^*_j; j = 1, 2$ the GPP (5.8) takes the form

$$\left.\begin{array}{c} Minimize\ x_1 + x_2 \\ Subject\ to \\ \frac{.8066}{n_1} + \frac{.4507}{n_2} + \frac{.3655}{n_3} + \frac{.4059}{n_4} \\ + \frac{(K_1-1)}{n_1} * 0.06 + \frac{(K_2-1)}{n_2} * 0.02 + \frac{(K_3-1)}{n_3} * 0.02 + \frac{(K_4-1)}{n_4} * 0.02 - x_1 \leq 0.0067 \\ \frac{.8365}{n_1} + \frac{.3377}{n_2} + \frac{.1068}{n_3} + \frac{.2807}{n_4} \\ + \frac{(K_1-1)}{n_1} * 0.06 + \frac{(K_2-1)}{n_2} * 0.02 + \frac{(K_3-1)}{n_3} * 0.07 + \frac{(K_4-1)}{n_4} * 0.02 - x_2 \leq 0.0046 \\ (2.4 + \frac{0.9}{K_1})n_1 + (3.4 + \frac{0.8}{K_2})n_2 + (4 + \frac{1.25}{K_3})n_3 + (4.6 + \frac{1.68}{K_4})n_4 \leq 5000 \\ 2 \leq n_1 \leq 1214;\ 2 \leq n_2 \leq 822 \\ 2 \leq n_3 \leq 1028;\ 2 \leq n_4 \leq 786 \\ x_j \geq 0;\ j = 1, 2;\ K_i > 1,\ and\ n_i\ are\ integers; i = 1, 2, \ldots, 4 \end{array}\right\}$$

Using the LINGO, the optimum compromise allocation is found to be
$n^*_1 = 524, n^*_2 = 309, n^*_3 = 204, n^*_4 = 246, K_1 = 2.15, K_2 = 2.11, K_3 = 1.16, K_4 = 2.21$. Sampling fractions among the non-respondents are $\frac{1}{K_1} = 0.46, \frac{1}{K_2} = 0.47, \frac{1}{K_3} = 0.86$ and $\frac{1}{K_4} = 0.45$. The values of the objective function as $Z = 0.0115$.

7. CONCLUSION

In this paper, we have formulated the problem of non-response in multivariate stratified sample surveys as a problem of multi-objective mathematical programming problem (MOMPP). The formulated MOMPP has been solved by using two different optimization techniques namely goal programming and Lagrange Multiplier's Technique (LMT). The compromise allocations obtained by LMT gives the minimum coefficients of variation as compared to the goal programming technique

ACKNOWLEDGEMENT

The authors are grateful to the Editor-in-Chief and the referees for their valuable comments and suggestions that helped in improving the manuscript to its present form.

REFERENCES

1. ANK M. J. Ahsan, Najmussehar, M. G. M. Khan, Mixed allocation in stratified sampling. Aligarh Journal of Statistics, 25, (2005). 87-97.
2. CS A. H. Ansari, Najmussehar, M. J. Ahsan, On multiple response stratified random sampling design. International Journal of Statistical Sciences, Kolkata, India, 1(1), (2009). 1-11.
3. GC H. Aoyama, Stratified random sampling with optimum allocation for multivariate populations. Ann. Inst. Statist. Math., 14, (1963). 251-258.
4. GK J. Bethel, Sample allocation in multivariate surveys. Survey Methodology, 15, (1989). 47-57.
5. L J. R. Chromy, Design optimization with multiple objectives. Proceedings of the Survey Research Methods, (1987). 194-199.
6. S W. G. Cochran, Sampling Techniques.2nd ed., John Wiley, New York. (1977).
7. SZ T. Dalenius, The multivariate sampling problem. Skandinavisk Actuarietidskrift, 36, (1953). 92 - 102.
8. Sun T. Dalenius, Sampling in Sweden. Contributions to the Methods and Theories of Sample Survey Practice. Stockholm: Almqvist and Wicksell. (1957).
9. J. A. Díaz-García, C. L. Ulloa, Optimum allocation in multivariate stratified sampling: Multi-objective programming. Comunicación Técnica No. I-06-07/28-03-206 (PE/CIMAT), Guanajuato, México. (2006).
10. J. A. Diaz-Garcia, C. L. Ulloa, Multi-objective optimization for optimum allocation in multivariate stratified sampling. Survey Methodology, 34(2), (2008). 215-222.
11. M. A. El-Badry, A sampling procedure for mailed questionnaires.J. Amer. Stat. Associ.51, (1956). 209-227.
12. TX J. L. Folks, C. E. Antle, Optimum allocation of sampling units to the strata when there are R responses of interest. J. Amer. Statist. Assoc., 60, (1965). 225-233.
13. G.T. Foradari, Some non-response sampling theory for two stage designs. North Caralina State College.Mimeographed series, No. 297, Raleigh. (1961).
14. S. P. Ghosh, A note on stratified random sampling with multiple characters. Calcutta Statist. Assoc. Bull. 8, (1958). 81-89.
15. S. Ghufran, S. Khowaja, and J. Ahsan, Optimum multivariate stratified sampling design with travel cost: A multiobjective integer nonlinear programming approach. Communication in Statistics Simulation and Computation, 41, (2012). 598-610.
16. J. Gren, Some methods of sample allocation in multivariate stratified sampling, Przeglad Statystyczny, 11, (1964). 361-369 (in Polish).
17. J. Gren, Some application of non-linear programming in sampling methods, Przeglad Statystyczny, 13, (1966). 203-217 (in Polish).
18. M. H. Hansen, and W. N. Hurwitz, The problem of non-response in sample surveys. J. Amer. Stat. Assoc. 41, (1946). 517-529.
19. H. O. Hartley, Multiple purpose optimum allocation in stratified sampling. Proc. Amer. Statist. Assoc. Social Statist., Sec. (1965). 258-261.
20. A. Holmberg, A multi parameter perspective on the choice of sampling design in surveys. Statistics in Transition, 5, (2002). 969-994.
21. A. F. Ige, T. P.Tripathi, On double sampling for stratification and use of auxiliary information. J. Indian Soc. Agricultural Statist., 39(2), (1987). 191-201.
22. N. Jahan, M. G. M. Khan, M. J. Ahsan, A generalized compromise allocation. Journal of the Indian Statistical Association, 32, (1994). 95-101.
23. M. G. M. Khan, M. J. Ahsan, N. Jahan, Compromise allocation in multivariate stratified sampling: An integer solution. Naval Research Logistics, 44, (1997). 69-79.
24. M. G. M. Khan, E. A. Khan and M. J. Ahsan, An optimal stratified sampling design using dynamic programming. Australian & New Zealand J. Statist., 45 (1), (2003). 107- 113.

25. M. G. M. Khan, E. A. Khan and M. J. Ahsan, Optimum allocation in multivariate stratified sampling in presence of non-response. J.Ind. Soc. Agril. Statist. 62(1), (2008). 42-48.
26. B. B. Khare, Allocation in stratified sampling in presence of non- response.Metron, XLV (1-2), (1987). 213-221.
27. L. Kish, Multipurpose sample designs. Survey Methodology, 14, (1988). 19-32.
28. A. R. Kokan, S. U. Khan, Optimum allocation in multivariate surveys: An analytical solution. Journal of Royal Statistical Society, Ser. B, 29, (1967). 115-125.
29. M. Kozak, On sample allocation in multivariate surveys. Communication in Statistics Simulation and Computation, 35, (2006). 901-910.
30. L. Kreienbrock, Generalized measures of dispersion to solve the allocation problem in multivariate stratified random sampling. Communication in Statistics-Theory and Methods, 22, (1993). 219-239.
31. LINGO -User's Guide: LINGO -User's Guide. Published by LINDO SYSTEM INC.,1415, North Dayton Street. Chicago,Illinois, 60622, USA (2001).
32. K. M. Miettinen, Non Linear Multiobjective Optimization. Boston: Kluwer Academic Publishers. (1999).
33. Najmussehar, A. Bari, Double sampling for stratification with sub-sampling the Non respondents: a dynamic programming approach, Aligarh J. Statist., 22, (2002). 27-41.
34. F. C. Okafor, On double sampling for stratification with sub-sampling the non repondents. Aligarh J. Statist., 14, (1994). 13-23.
35. J. N. K. Rao, On double sampling for stratification and analytical surveys. Biometrika, 60, (1973). 125-133.
36. S. Rios, S. Rios Insua, M. J. Rios Insua, Procesos de decision Multicriterio. Madrid: EUDEMA (in Spanish). (1989).
37. K. P. Srinath, Multiphase sampling in non-response problems.J. Amer. Stat. Assoc. 66, (1971). 583-586.
38. P. V. Sukhatme, B. V. Sukhatme, Sukhatme, C. Asok, Sampling theory of Surveys with Applications. Iowa State University Press, Iowa, U. S. A. and Indian Society of Agricultural Statistics, New Delhi, India. (1984).
39. T. P. Tripathi, S. Bahl, Estimation of mean using double sampling for stratification and multivariate auxiliary information. Comn. Statist. Theory Methods, 20(8), (1991). 2589-2602.
40. R. Varshney, Najmussehar and M. J. Ahsan, An optimum multivariate stratified double sampling design in non-reponse. Optimization letters, Verlag, 6, (2012). 993-1008.
41. J. Wywial, Minimizing the spectral radius of means vector from sample variance covariance matrix sample allocation between strata. Prace Naukowe Akademii Ekonomicznej we Wroclawiu, 404, (1988). 223-235 (in Polish).
42. F. Yates, Sampling methods for Census and Surveys: Charles Griffin & Co. Ltd., London. (1960).

TABLE 2. Calculation for proposed allocation n_i and optimum sampling fractions among Non-respondents $1/K_i$n(where K_i is defined in **??**)

(1) i	(2) A_i	(3) B_i	(4) $A_i * B_i$	(5) $\sqrt{(4)}$	(6) A_i/B_i	(7) $\sqrt{(6)}$	(8) $\sqrt{C_{i2}/b_i^2}$	(9) $p_i C_0 \sqrt{(6)}$	(10) $p_i\{\sqrt{(5)} + W_{i2} b_i \sqrt{C_{i2}}\}$	(11) $K_i = (7) * (8)$	(12) $n_i = (9)/\sum(10)$
1	14.85	2.4	35.64	5.97	6.19	2.49	0.86	3980.06	2.24	2.15	527.92
2	16.99	3.4	57.78	7.60	4.99	2.23	0.94	2347.40	1.77	2.11	311.36
3	6.07	4.0	24.30	4.93	1.52	1.23	1.76	1663.68	1.52	2.16	220.67
4	15.97	4.6	73.48	8.57	3.47	1.86	1.18	1863.44	1.99	2.20	247.17