International Journal of Operations Research

International Journal of Operations Research Vol. 10, No. 4, 171−181 (2013)

Multi Choice for Precision in Multivariate Stratified Surveys: A Compromise Solution

Shamsher Khan* **and M.M. Khalid**

Department of Statistics and Operations Research, Aligarh Muslim University, India.

Received August 2013; Revised December 2013; Accepted December 2013

Abstract **—**The problem of optimum allocation for multivariate stratified surveys is considered here. The key idea behind the optimum allocation in multivariate surveys is to minimize the variances of the estimates for a given cost or to minimize the cost for specified variance tolerances of the estimates. But there is no any definite answer to the question that how much should be the precision of an estimate because it depends on the purpose of the survey. It is also noted that there is no any single criteria by which we can specify the aspiration level. In this manuscript, we have considered the situation in which, we are specifying more choice of precisions by using the idea of multi choice precisions for estimates with the condition that with every one of them we can tolerate. The problem of optimal allocation is considered as a multiobjective programming problem (MOPP). Compromise solutions are obtained by using multi choice goal programming approach (MCGP). A comparison of the proposed approach with goal programming and weighted goal programming approach is also made. To demonstrate the correctness of the proposed approach, a numerical illustration is also given solved by Lingo Software.

Keywords **—**multivariate stratified surveys, compromise allocation, multiobjective programming, goal programming, weighted goal programming, multi choice goal programming

1. INTRODUCTION

Following Hansen *et al.* (1953), "Precision or sampling error of a sample result, means that how closely we can reproduce from a sample the results which would be obtained if we should take a complete count under the same conditions". The purpose of sampling theory is the development of methods of sample selection and estimation that provide precise estimate at a lowest possible cost. To obtain precise estimates of the population parameters is an important goal in sampling surveys. The purpose of stratification is to reduce the heterogeneity of the population and to provide greater precision in sample estimates by utilizing the available information. In stratification the heterogeneous population is divided into homogeneous groups. These groups or subpopulations are called strata and each group is called stratum. The whole procedure of stratification and selecting independent random samples from each stratum is known as stratified sampling.

In stratified sampling the variance of the estimator depends on the size of sample allocated to various strata. Now the question arises that how the total sample should be allocated to various strata.

To allocate the size of samples, Neyman (1934) introduced the criteria "minimize the variance" subject to a fixed sample size and this allocation is known as "Neyman Allocation". In Yates and Zacopanay (1935), this criterion extended as "minimize the variance" subject to fixed cost or vice versa. The allocation of sizes of samples to various strata with these criterions is known as "optimum allocation". Stuart (1954) uses "Schwarz inequality for this purpose. Sample surveys, where we study more than one character/variable on a population unit, are known as multivariate surveys. If the stratum variances for different variates are distributed in the same way then "Neyman Allocation" gives the optimal allocation for all variates, but if stratum variances are not distributed in the same way then the optimal allocation for any variate may be quite unsuitable for another. Yates (1953) suggested an approach in which the variances of the estimates for different characteristics are equal to a certain specified level of precision to minimize the cost of the survey. Dalenius (1957), given a more reasonable criteria in the minimization of the total cost subject to the condition that the variances of the estimates for different characteristics do not exceed certain pre-assigned quantities (variance of the estimate for character under study).

The problem of optimum allocation of sample sizes to various strata treated as a mathematical programming problem firstly by Dalenius (1957). In multivariate surveys, each character cannot be estimate exactly so a certain margin of error in each estimate must be tolerated. With this concept Kokan (1963) consider an upper bound on the error of the estimate at some confidence level and proposes a solution using nonlinear programming problem. Kokan and Khan (1967) have given an analytical solution for the optimum allocation in multivariate surveys. Various authors such as Aoyama (1963), Folks and

*Corresponding author's email: shamsherstats@gmail.com

Antle (1965), Chatterjee (1967, 1968), Ahsan and Khan (1977, 1982), Bethel (1985), Chromy (1987), Jahan *et al.* (1994), Khan *et al.* (1997), Khan and Ahsan (2003) etc., used different approaches for sample allocation in multivariate surveys.

Following Bethel (1989), in some situations coefficient of variation (CV) can be used as a measure of sampling error of the estimate.

Chaddha *et al.* (1971) used dynamic programming technique to find the optimum allocation in univariate case. Omule (1985) used the same technique for the multivariate case. He minimized the total cost of the survey when the tolerance levels for the precision of the estimates of the various characteristics are predefined. Khan *et al.* (2003) also used the dynamic programming in multivariate case, when the population means of several characteristics are to be estimated.

In multivariate sampling, the problem of optimum allocation will be more complicated because optimal allocation for one character may be unsuitable for another. In such situations, we obtain compromise allocations which are optimal for all characters in some sense.

Khan *et al.* (2010) considered the problem of determining the integer optimum allocation as a multiobjective nonlinear programming problem and obtained compromise solutions by using goal programming technique, when the population means of various characteristics are of interest and auxiliary information is available for the separate and combined ratio and regression estimate.

Recently, Swain (2013) has given a note on optimum allocation in stratified random sampling in which a comparison is made between Chatterjee (1967) technique and goal programming technique for finding compromise allocation. Mathew *et al.* (2013) described the efficiency of Neyman allocation procedure over other allocation procedures in stratified random sampling.

Almost all the previous authors used the common approach in optimum allocation that is minimize the variances of the estimates subject to a cost function or minimizes the cost subject to desired precisions of the estimates. But there is no any definite answer to the question that how much should be the precision of an estimate in a particular situation because it depends on the purpose of the survey. On the basis of above discussed approaches the desired precision can be specified as:

1. by the help of experts or from the values reported in the literature;
2. by specifying the upper bound on the error of the estimate with some confidence level;
3. by using the coefficient of variation as the error of estimates.

In other words we can say that, there is no any single criteria by which we can specify the precisions. Precision can be control by setting the outer bounds on the possible sampling errors, such that the probability of exceeding these bounds is very small. It is also noted that, uncertainty is inherent in statistical techniques which may be in the form of an error. If precision is concerned, no one can foretell exactly that how large amount of error will be present in an estimate in a particular situation. Consequently, the specification of the degree of precision wanted in the results is an important step. This step is the responsibility of the person who is going to use the data. It may present difficulties, since many administrators are unaccustomed to thinking in terms of the amount of error that can be tolerated in estimates, consistent with making good decisions (Cochran 1977).

To overcome these problems, it may be convenient to consider the situation of more than one choice of precisions for the estimates with the condition that, with every one of them we can tolerate. In other words we have a set of choices of acceptable precisions. So, the problem of optimal allocation will be the minimization of the variances of estimates for a given cost with multi choice of precisions. Thus, the problem of optimum allocation is considered as a multi objective programming problem (MOPP). In order to optimize all the objectives simultaneously, these multi choices of precisions for each characteristic are considered as goal by using the idea of multi choice aspiration levels (MCAL) in goal programming problem. To solve this problem, we are using the technique developed by Cheng (2007) and obtained compromise solutions.

Compromise solution means a solution that can be used to optimize all the objectives on a compromise basis. In multivariate surveys, by compromise allocation/solution we mean a common sample size used to estimate all the characteristics.

The purpose of this manuscript is to motivate the researchers for using the idea of multi choice desired precisions to specify the errors of the estimates, not to develop a procedure to specify them.

According to best of our knowledge, no work has been done with this approach in sample surveys till now. To specify the multi choice of desired precisions we are using the values reported in the literature. The remainder of the paper is organized as follows. Section 2 contains some useful notations. For the clearer description of the idea formulation of problem is given in section 3. In section 4 used solution technique and the techniques for comparison are described. A numerical illustration is also given to demonstrate the concept and solution technique in section 5 with solutions obtained by Lingo (2001) Software. Section 6 is the section of discussion in which we obtained the relative precision of the proposed approach as compare to other discussed approaches. Finally, conclusions are presented in section 7.

2. NOTATIONS

In stratified sampling the population having N units is divided into L subpopulations having $N_1, N_2, N_3, \ldots, N_L$ units respectively (symbols have their usual meaning from the standard book Cochran (1977), otherwise stated).

Let the suffix *h* denotes the stratum and *i* the unit within the stratum. Also let

N_h	total number of units in h^{th} stratum
n_h	number of units in sample from h^{th} stratum
y_{hi}	value obtained for the i^{th} unit in the h^{th} stratum
$W_h = \frac{N_h}{N}$	stratum weight
$f_h = \frac{n_h}{N_h}$	sampling fraction in the h^{th} stratum
$\bar{Y}_h = \frac{\sum_{i=1}^{N_h} y_{hi}}{N_h}$	true mean
$\bar{y}_h = \frac{\sum_{i=1}^{n_h} y_{hi}}{n_h}$	sample mean
$S_h^2 = \frac{\sum_{i=1}^{N_h} (y_{hi} - \bar{Y}_h)^2}{N_h - 1}$	true variance

For the population mean per unit, the estimate $\bar{y}_{st}$ is used and the variance of the estimate will be

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

If the sampling fractions $f_h = \frac{n_h}{N_h}$ are negligible in all strata, then

$$V(\bar{y}_{st}) = \frac{1}{N^2}\sum_{h=1}^{L} N_h^2 \frac{S_h^2}{n_h} \text{ , or}$$

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

In multivariate surveys we have more than one variate under study then the variance function for j^{th} characteristic will be

$$V(\bar{y}_{jst}) = \sum_{h=1}^{L} W_h^2 \frac{S_{jh}^2}{n_h}$$

Where S_{jh}^2 $(h = 1, 2, 3, \ldots, L)$ are the true population variances for j^{th} characteristic.

3. FORMULATION OF THE PROBLEM

Arthanari and Dodge (1981) have given the mathematical formulation of the problem of optimum allocation for minimizing the total cost, subject to a desired precision as follows

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

Subject to

$$V(\bar{y}_{st}) = \sum_{h=1}^{L} W_h^2 \frac{S_h^2}{n_h} \le v_0$$

and $1 \le n_h \le N_h$, n_h integer for $h = (1, 2, 3, \ldots, L)$; (1)

where v_0 is the specified variance tolerance of the estimator for population mean.

In multivariate surveys where we have more than one variate, say p, under study then the problem of optimum allocation will be

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

Subject to

$$V(\overline{y}_{jst}) = \sum_{h=1}^{L} W_h^2 \frac{S_{jh}^2}{n_h} \le v_j^0 \quad \forall\, j = 1,2,3,\ldots,p$$

and $2 \le n_h \le N_h$, n_h integer for $h = (1,2,3,\ldots,L)$; (2)

where v_j^0 is the specified variance tolerance of the estimator for population mean for j^{th} character. The restriction $2 \le n_h \le N_h$ is taken here to overcome the problem of oversampling. In the cost function $c_h = \sum_{h=1}^{L} c_{jh} n_h$ denotes the cost of measuring all the p characters on a sampled unit in the h^{th} stratum and c_{jh} is the per unit cost of measuring the j^{th} characteristic in h^{th} stratum. It is also noted that the overhead cost c_0 is not the part of optimization.

For a fixed budget C the above said problem as a Multiobjective Programming Problem (MOPP) can be formulated as

Minimize $[V(\overline{y}_{1st}), V(\overline{y}_{2st}), V(\overline{y}_{3st}), \ldots, V(\overline{y}_{pst})]$

Subject to

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

and $2 \le n_h \le N_h$, n_h integer for $h = (1,2,3,\ldots,L)$. (3)

Our aim is to optimize all the objectives simultaneously. If we specified the desired precisions for each objective i.e. fixed the target value/goal, then the problem of optimum allocation as a Goal Programming Problem (GPP) will be

(Goal j) $V(\overline{y}_{jst}) = \sum_{h=1}^{L} W_h^2 \frac{S_{jh}^2}{n_h} \le v_j^0 \quad \forall\, j = 1,2,3,\ldots,p$

Subject to

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

and $2 \le n_h \le N_h$, n_h integer for $h = (1,2,3,\ldots,L)$; (4)

where v_j^0 is the specified/targeted variance tolerance of the estimator for population mean for j^{th} character.

According to our assumption, we have considered the case of an either or choice of the desired precisions with the condition that, with every one of them we can tolerate.

Thus, the problem of optimal allocation as a goal programming problem with multi choice of aspiration levels (MCAL) will be

(Goal j) $V(\overline{y}_{jst}) = \sum_{h=1}^{L} W_h^2 \frac{S_{jh}^2}{n_h} \le \left\{ v_j^{(1)} \; or \; v_j^{(2)} \; or \; v_j^{(3)} \; or \ldots v_j^{(k_j)} \right\} \quad \forall\, j = 1,2,3,\ldots,p$

Subject to

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

and $2 \le n_h \le N_h$, n_h integer for $h = (1,2,3,\ldots,L)$. (5)

For each goal we have k_j number of choices i.e. aspiration levels.

4. SOLUTION APPROACHES

4.1 Multi Choice Goal Programming Approach

In this manuscript, we have considered the multi choice nature of desired precisions. It is a multiobjective programming problem with multi choices of goals. From the equation (5), we have the following problem

(Goal j) $V(\overline{y}_{jst}) = \sum_{h=1}^{L} W_h^2 \frac{S_{jh}^2}{n_h} \le \left\{ v_j^{(1)} \text{ or } v_j^{(2)} \text{ or } v_j^{(3)} \text{ or} \ldots v_j^{(k_j)} \right\} \; \forall\, j = 1,2,3,\ldots,p$;

Subject to

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

and $2 \le n_h \le N_h$, n_h integer for $h = (1,2,3,\ldots,L)$.

For each goal we have k_j number of choices (i.e., each goal mapping many aspiration levels).

Since we have multi choice of goals so this problem can't be solved by the exit techniques such as goal programming, weighted goal programming etc., in which decision maker specifies aspiration levels for the objective functions and any deviations from these aspiration levels are minimized Miettinen (1999). For solving such type of problems we need a technique by which this problem can be easily solved. Cheng (2007) introduced the concept of multi choice goal programming technique (MCGP). The problem of optimum allocation as a MCGP problem can be expressed as

$$\text{Minimize } \left| V(\overline{y}_{jst}) - v_j^{(1)} \text{ or } v_j^{(2)} \text{ or } v_j^{(3)} \text{ or} \ldots v_j^{(k_j)} \right|$$

Subject to

$$V(\overline{y}_{jst}) - x_j^+ + x_j^- = \left\{ v_j^{(1)} \text{ or } v_j^{(2)} \text{ or } v_j^{(3)} \text{ or} \ldots v_j^{(k_j)} \right\}$$

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

$$\text{and } 2 \le n_h \le N_h, \ n_h \text{ integer for } h = (1,2,3,\ldots,L); \tag{6}$$

where x_j^+ and x_j^- are the overachievement and underachievement variables respectively for j^{th} goal.

Consider the following cases:

Case (i) $k_j = 1$, this problem can be solved by any one techniques such as goal programming and weighted goal programming.

Case (ii) $k_j = 2$, then the objectives will be

$$V(\overline{y}_{jst}) \le \left\{ v_j^{(1)} \text{ or } v_j^{(2)} \right\}$$

This is the case of an either-or choice. This problem cannot be solved by the above discussed approaches. In order to solve this problem, two binary variables should be added as described below.

$$\text{Minimize } \sum_{j=1}^{p} x_j$$

Subject to

$$V(\overline{y}_{jst}) - x_j \le v_j^{(1)} z_j^{(1)} + v_j^{(2)} \left(1 - z_j^{(1)}\right)$$

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

$$\text{and } 2 \le n_h \le N_h, \ n_h \text{ integer for } h = (1,2,3,\ldots,L); \tag{7}$$

where $x_j \ge 0$; $\forall j = 1,2,3,\ldots,p$; are the deviation variables such that

$\left| V(\overline{y}_{jst}) - v_j^{(1)} \text{ or } v_j^{(2)} \text{ or } v_j^{(3)} \text{ or} \ldots v_j^{(k_j)} \right|$ and $z_j^{(1)}$ $\forall j = 1,2,3,\ldots,p$; are the binary variables.

Our aim will be to minimize the sum of these deviation variables.

Since all the objectives are of less than inequality type, so we can leave the underachievement variable and minimized the sum of the overachievement variables only.

Case (iii) $k_j = 3$, then the objective will be

$$V(\overline{y}_{jst}) \le \left\{ v_j^{(1)} \text{ or } v_j^{(2)} \text{or } v_j^{(3)} \right\}$$

This is also the case of an either-or choice. This case cannot be solved by the above discussed approaches. In order to solve this problem, four binary variables should be added as described below.

$$\text{Minimize } \sum_{j=1}^{p} x_j$$

Subject to

$$V(\overline{y}_{jst}) - x_j \le v_j^{(1)} z_j^{(1)} z_j^{(2)} + v_j^{(2)} z_j^{(1)} \left(1 - z_j^{(2)}\right) + v_j^{(3)} \left(1 - z_j^{(1)}\right) z_j^{(2)}$$

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

$$\text{and } 2 \le n_h \le N_h, \ n_h \text{ integer for } h = (1,2,3,\ldots,L); \tag{8}$$

where $x_j \ge 0$ $\forall j = 1,2,3,\ldots,p$; are the deviation variables such that

$\left| V(\overline{y}_{jst}) - v_j^{(1)} \, or \, v_j^{(2)} \, or \, v_j^{(3)} \, or \ldots v_j^{(k_j)} \right|$ and $z_j^{(1)}, z_j^{(2)}$ $\forall j = 1,2,3,\ldots,p$; are the binary variables.

Our aim is to achieve all goals precisely but due to conflict nature of objectives we cannot optimize all the objectives simultaneously, so we get a compromise solution. Compromise means this solution will be optimal for all the objectives in some sense.

4.2 Goal Programming Approach

Consider the problem given in equation (4); we have single aspiration level so this problem can be solved by Goal Programming Technique. In goal programming technique the decision maker fixed his (her) aspiration levels for each of the objective and the deviations from these aspiration levels are minimized. In goal programming all the objectives considered simultaneously so we get a compromise solution. Compromise means this solution will be optimal for each objective in some sense. If the desired precision is considered as goal then the problem of optimum allocation can be solved by goal programming (GP) technique as follows

$$\text{Minimize } \sum_{j=1}^{p} x_j$$

Subject to

$$V(\overline{y}_{jst}) - x_j \le v_j^0$$

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

$$\text{and } 2 \le n_h \le N_h, \; n_h \text{ integer for } h = (1,2,3,\ldots,L); \tag{9}$$

where $x_j \ge 0 \;\; \forall\, j = 1,2,3,\ldots,p$; are the deviation variables such that $\left|V(\overline{y}_{jst}) - v_j^0\right|$.

4.3 Weighted Goal Programming Approach

In weighted goal programming approach we also considered all the objectives simultaneously and minimize the total weighted deviation from all the goals. These weights are not preemptive, but reflect the relative importance of each goal. Since we wish to optimize the most important objective precisely so we assign more weight to such objective.

The problem of optimum allocation can be solved by weighted goal programming technique as follows

$$\text{Minimize } \sum_{j=1}^{p} w_j x_j$$

Subject to

$$V(\overline{y}_{jst}) - x_j \le v_j^0$$

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

$$\text{and } 2 \le n_h \le N_h, \; n_h \text{ integer for } h = (1,2,3,\ldots,L); \tag{10}$$

where $x_j \ge 0 \;\; \forall\, j = 1,2,3,\ldots,p$; are the deviation variables such that $\left|V(\overline{y}_{jst}) - v_j^0\right|$, and $\sum_{j=1}^{p} w_j = 1$.

The techniques (4.2) & (4.3) are used only for the purpose of comparison in the section 6.

5. NUMERICAL ILLUSTRATION

To illustrate the procedure we use the data form Jessen (1942) on a farm survey in Iowa (also discussed in Cochran 1977, p. 119-121). The three items of most interest are the number of "cows milked per day", "the number of gallons of milk per day", and the total annual cash "receipts from dairy products". We have the information given in table 1.

To demonstrate the concept of developed approach we are using the choices of the variance for the estimated mean of each variate obtained by optimum, compromise and proportional allocation Cochran (1977) Section5.A, p. 119; for fixed sample size.

Let $k_1 = 2, k_2 = 3, k_3 = 3$ i.e. for $j = 1$ we have two choices; for $j = 2$ we have three choices and for $j = 3$ we also have three choices.

In other words

$$v_1^{(2)} = \{0.0127 \text{ or } 0.0128\}$$

$$v_2^{(3)} = \{0.0800 \text{ or } 0.0802 \text{ or } 0.0837\}$$

$$v_3^{(3)} = \{76.9 \text{ or } 77.6 \text{ or } 80.9\}$$

It is also assumed that the cost c_h, of measuring all the p characters of the sampled unit in the h^{th} stratum are $c_1 = \$3, c_2 = \$4, c_3 = \$5, c_4 = \$6$ and $c_5 = \$7$ and the total budget of the survey is $C = \$6000$.

Table 1. Data of Farm Survey

Stratum h	N_h	W_h	S_{1h} Cows Milked	S_{2h} Gallons of Milk	S_{3h} Receipts of Dairy Products
1	39574	0.197	4.6	11.2	332
2	38412	0.191	3.4	9.8	357
3	44017	0.219	3.3	7.0	246
4	36935	0.184	2.8	6.5	173
5	41832	0.208	3.7	9.8	279

With these assumptions, the problem of optimal allocation as a goal programming problem with choice for precisions, for a fixed budget will be

(Goal 1) $$\frac{0.82119}{n_1} + \frac{0.42172}{n_2} + \frac{0.52229}{n_3} + \frac{0.26543}{n_4} + \frac{0.59228}{n_5} \le \left(0.0127\, or\, 0.0128\right);$$

(Goal 2) $$\frac{5.31256}{n_1} + \frac{3.50363}{n_2} + \frac{2.35008}{n_3} + \frac{1.43041}{n_4} + \frac{4.15507}{n_5} \le \left(0.0800 \text{ or } 0.0802 \text{ or } 0.0837\right);$$

(Goal 3) $$\frac{4277.68321}{n_1} + \frac{4649.46696}{n_2} + \frac{2902.40787}{n_3} + \frac{1013.27622}{n_4} + \frac{3367.71302}{n_5} \le \left(76.9 \text{ or } 77.6 \text{ or } 80.9\right);$$

Subject to

$$3n_1 + 4n_2 + 5n_3 + 6n_4 + 7n_5 \le 6000$$

$$2 \le n_1 \le 39574; 2 \le n_2 \le 38412; 2 \le n_3 \le 44017; 2 \le n_4 \le 36935; 2 \le n_5 \le 41832$$

and n_h must be integer for $h = 1,2,3,\dots,L$.

By using the multi choice goal programming technique discussed in subsection (4.1), this problem can be solved as

Minimize $x_1 + x_2 + x_3$

Subject to

$$\frac{0.82119}{n_1} + \frac{0.42172}{n_2} + \frac{0.52229}{n_3} + \frac{0.26543}{n_4} + \frac{0.59228}{n_5} - x_1 \le 0.0127 z_1^{(1)} + 0.0128\left(1 - z_1^{(1)}\right)$$

$$\frac{5.31256}{n_1} + \frac{3.50363}{n_2} + \frac{2.35008}{n_3} + \frac{1.43041}{n_4} + \frac{4.15507}{n_5} - x_2 \le 0.0800 z_2^{(1)} z_2^{(2)} + 0.0802 z_2^{(1)}\left(1 - z_2^{(2)}\right) + 0.0837\left(1 - z_2^{(1)}\right) z_2^{(2)};$$

$$\frac{4277.68321}{n_1} + \frac{4649.46696}{n_2} + \frac{2902.40787}{n_3} + \frac{1013.27622}{n_4} + \frac{3367.71302}{n_5} - x_3 \le 76.9 z_3^{(1)} z_3^{(2)} + 77.6 z_3^{(1)}\left(1 - z_3^{(2)}\right) + 80.9\left(1 - z_3^{(1)}\right) z_3^{(2)};$$

$$3n_1 + 4n_2 + 5n_3 + 6n_4 + 7n_5 \le 6000;$$

$$2 \le n_1 \le 39574; 2 \le n_2 \le 38412; 2 \le n_3 \le 44017; 2 \le n_4 \le 36935; 2 \le n_5 \le 41832$$

and n_h must be integer for $h = 1,2,3,\dots,L$.

$x_j \geq 0 \; \forall \; j = 1,2,3$; are the deviations as defined in subsection 4.1, and $z_1^{(1)}, z_2^{(1)}, z_2^{(2)}, z_3^{(1)}$ and $z_3^{(2)}$ are the binary variables.

After solving this problem by LINGO (2001), we obtain the compromise solutions as $\left(n_1, n_2, n_3, n_4, n_5, z_1^{(1)}, z_2^{(1)}, z_2^{(2)}, z_3^{(1)}, z_3^{(2)}\right) = \left(437, 516, 231, 98, 127, 0, 1, 1, 1, 1\right)$ and the variances for the estimates are $\left(V_1, V_2, V_3\right) = \left(0.0124, 0.0772, 68.8624\right)$.

Goal Programming Approach

If we have a single precision rather than the multi choice for precisions then this problem can be solved by goal programming technique discussed in subsection 4.2. We are considering three cases for our purpose. In each case we are specifying only one aspiration level selected form the multi choices specified for each character in multi choice goal programming approach.

Case (i)

Minimize $x_1 + x_2 + x_3$

Subject to

$$\frac{0.82119}{n_1} + \frac{0.42172}{n_2} + \frac{0.52229}{n_3} + \frac{0.26543}{n_4} + \frac{0.59228}{n_5} - x_1 \leq 0.0127\,;$$
$$\frac{5.31256}{n_1} + \frac{3.50363}{n_2} + \frac{2.35008}{n_3} + \frac{1.43041}{n_4} + \frac{4.15507}{n_5} - x_2 \leq 0.0800\,;$$
$$\frac{4277.68321}{n_1} + \frac{4649.46696}{n_2} + \frac{2902.40787}{n_3} + \frac{1013.27622}{n_4} + \frac{3367.71302}{n_5} - x_3 \leq 76.9\,;$$
$$3n_1 + 4n_2 + 5n_3 + 6n_4 + 7n_5 \leq 6000\,;$$
$$2 \leq n_1 \leq 39574\,; 2 \leq n_2 \leq 38412\,; 2 \leq n_3 \leq 44017\,; 2 \leq n_4 \leq 36935\,; 2 \leq n_5 \leq 41832\,.$$

n_h must be integer for $h = 1,2,3,\ldots,L$ and $x_j \geq 0 \; \forall \; j = 1,2,3$ as defined in section 4.2.

After solving this problem by LINGO (2001), we obtain the compromise solutions as $\left(n_1, n_2, n_3, n_4, n_5\right) = \left(691, 204, 327, 103, 122\right)$ and the variances for the estimates are $\left(V_1, V_2, V_3\right) = \left(0.0123, 0.0800, 75.2998\right)$.

Case (ii)

Minimize $x_1 + x_2 + x_3$

Subject to

$$\frac{0.82119}{n_1} + \frac{0.42172}{n_2} + \frac{0.52229}{n_3} + \frac{0.26543}{n_4} + \frac{0.59228}{n_5} - x_1 \leq 0.0128\,;$$
$$\frac{5.31256}{n_1} + \frac{3.50363}{n_2} + \frac{2.35008}{n_3} + \frac{1.43041}{n_4} + \frac{4.15507}{n_5} - x_2 \leq 0.0802\,;$$
$$\frac{4277.68321}{n_1} + \frac{4649.46696}{n_2} + \frac{2902.40787}{n_3} + \frac{1013.27622}{n_4} + \frac{3367.71302}{n_5} - x_3 \leq 77.6\,;$$
$$3n_1 + 4n_2 + 5n_3 + 6n_4 + 7n_5 \leq 6000\,;$$
$$2 \leq n_1 \leq 39574\,; 2 \leq n_2 \leq 38412\,; 2 \leq n_3 \leq 44017\,; 2 \leq n_4 \leq 36935\,; 2 \leq n_5 \leq 41832\,.$$

n_h must be integer for $h = 1,2,3,\ldots,L$ and $x_j \geq 0 \; \forall \; j = 1,2,3$ as defined in section 4.2.

After solving this problem by LINGO (2001), we obtain the compromise solutions as $\left(n_1, n_2, n_3, n_4, n_5\right) = \left(648, 281, 308, 104, 109\right)$ and the variances for the estimates are $\left(V_1, V_2, V_3\right) = \left(0.0124, 0.0802, 73.2104\right)$.

Case (iii)

Minimize $x_1 + x_2 + x_3$

Subject to

$$\frac{0.82119}{n_1} + \frac{0.42172}{n_2} + \frac{0.52229}{n_3} + \frac{0.26543}{n_4} + \frac{0.59228}{n_5} - x_1 \le 0.0128\,;$$

$$\frac{5.31256}{n_1} + \frac{3.50363}{n_2} + \frac{2.35008}{n_3} + \frac{1.43041}{n_4} + \frac{4.15507}{n_5} - x_2 \le 0.0837\,;$$

$$\frac{4277.68321}{n_1} + \frac{4649.46696}{n_2} + \frac{2902.40787}{n_3} + \frac{1013.27622}{n_4} + \frac{3367.71302}{n_5} - x_3 \le 80.9\,;$$

$$3\,n_1 + 4\,n_2 + 5\,n_3 + 6\,n_4 + 7\,n_5 \le 6000\,;$$

$$2 \le n_1 \le 39574\,; 2 \le n_2 \le 38412\,; 2 \le n_3 \le 44017\,; 2 \le n_4 \le 36935\,; 2 \le n_5 \le 41832\,.$$

n_h must be integer for $h = 1,2,3,\ldots,L$ and $x_j \ge 0\ \forall\ j = 1,2,3$ as defined in subsection 4.2.

After solving this problem by LINGO (2001), we obtain the compromise solutions as $\left(n_1, n_2, n_3, n_4, n_5\right) = \left(712, 219, 324, 102, 108\right)$ and the variances for the estimates are $\left(V_1, V_2, V_3\right) = \left(0.0128, 0.0832, 77.3131\right)$.

Weighted Goal Programming Approach

Assuming that the third objective is most important as compared to others and we wish to optimize the third objective more precisely so we assign more weight to such objective. In weighted goal programming our aim will be to minimize the total weighted deviation from all the goals considered in the problem. Thus the problem of optimum allocation can be solved by weighted goal programming technique as follows

Minimize $0.30x_1 + 0.30x_2 + 0.40x_3$

Subject to

$$\frac{0.82119}{n_1} + \frac{0.42172}{n_2} + \frac{0.52229}{n_3} + \frac{0.26543}{n_4} + \frac{0.59228}{n_5} - x_1 \le 0.0128\,;$$

$$\frac{5.31256}{n_1} + \frac{3.50363}{n_2} + \frac{2.35008}{n_3} + \frac{1.43041}{n_4} + \frac{4.15507}{n_5} - x_2 \le 0.0802\,;$$

$$\frac{4277.68321}{n_1} + \frac{4649.46696}{n_2} + \frac{2902.40787}{n_3} + \frac{1013.27622}{n_4} + \frac{3367.71302}{n_5} - x_3 \le 80.9\,;$$

$$3\,n_1 + 4\,n_2 + 5\,n_3 + 6\,n_4 + 7\,n_5 \le 6000\,;$$

$$2 \le n_1 \le 39574\,; 2 \le n_2 \le 38412\,; 2 \le n_3 \le 44017\,; 2 \le n_4 \le 36935\,; 2 \le n_5 \le 41832\,.$$

n_h must be integer for $h = 1,2,3,\ldots,L$; and $x_j \ge 0\ \forall\ j = 1,2,3$ as defined in section 4.3.

It is also noted that $w_1 + w_2 + w_3 = 1$.

After solving this problem by LINGO (2001), we obtain the compromise solutions as $\left(n_1, n_2, n_3, n_4, n_5\right) = \left(692, 215, 323, 102, 119\right)$ and the variances for the estimates are $\left(V_1, V_2, V_3\right) = \left(0.0123, 0.0802, 75.0270\right)$.

6. DISCUSSION

In this section, all the results obtained in section 5 are summarized in the table 2. A comparative study of the proposed compromise allocation approach has been made with wellknown goal programming approach and weighted goal programming approach by using the criterion "Minimizing Trace" (Sukhatme *et al.*, 1967).

We have considered the case (ii) of goal programming approach for the purpose of comparison. Column 6 represents the trace of variance-covariance matrix. In column 7, we obtain the relative precision of the proposed approach to the discussed approaches, which is the ratio, $V_{other}(\overline{y}_{st}) / V_{proposed}(\overline{y}_{st})$, expressed as a percentage, Cochran (1977; p. 102).

It is interesting to note that the results obtained by using the multi choice of precision are more precise as compared to single precision.

Table 2. Comparison of the proposed approach with other approaches

(1) S.No.	(2) Compromise Allocation	(3) $V(\overline{y}_{1st})$	(4) $V(\overline{y}_{2st})$	(5) $V(\overline{y}_{3st})$	(6) Trace $V_1+V_2+V_3$	(7) Relative Precision	(8) Cost of Survey
1	Goal Programming	0.0124	0.0802	73.2104	73.3030	106	5995
2	Weighted Goal Programming	0.0123	0.0802	75.0270	75.1195	109	5996
3	Proposed Approach	0.0124	0.0772	68.8624	68.9520	100	5986

The results given in table 2; provide an answer to the problem of working out with multi choice of precision as compared to single precision since the trace of variances for the estimates by the proposed approach is 6% more precise as compared to GP and 9% as compared to WGP. In addition, regarding the cost, it can be seen that the cost incurred in the survey is least by the proposed approach as compared to the other approaches. Thus, we can say that the proposed approach provides better results as compared to other discussed approaches.

7. CONCLUSION

The problem of optimum allocation for multivariate stratified surveys considered here with multi choice of precision and solved by Multi Choice Goal Programming (MCGP) technique. A comparison of the proposed approach with the Goal Programming approach and Weighted Goal Programming Approach is also discussed. For comparison, we consider the trace of variance-covariance matrix and the relative precisions of the proposed approach with the other approaches are calculated. It is interesting to note that the trace of variances is least by the proposed approach as compared to the other discussed approaches. On the other hand, cost incurred in the surveys is least by using the proposed approach. We have considered maximum three choices of the precisions for our goals with a linear cost function in this article. This work can be extended for more than three choices of precisions with nonlinear cost function and can be considered as the future work direction of the proposed approach.

REFERENCES

1. Ahsan, M.J. and Khan, S.U. (1977). Optimum Allocation in Multivariate Stratified Random Sampling Using Prior Information. *Journal of Indian Statistical Association*, 15: 57–67.
2. Ahsan, M.J. and Khan, S.U. (1982). Optimum Allocation in Multivariate Stratified Random Sampling with Overhead Cost. *Metrika*, 29: 71–78.
3. Aoyama, H. (1963). Stratified Random Sampling with Optimum Allocation for Multivariate Populations. *Annals of the Institute of Statistical Mathematics*, 14: 251–258.
4. Arthanari, T.S. and Dodge, Y. (1981). *Mathematical Programming in Statistics*. New York: John Wiley.
5. Bethel, J. (1985). An Optimum Allocation Algorithm for Multivariate Surveys. Proceedings of the American Statistical Association, Section on Survey Methods Research, pp. 209–212.
6. Bethel, J. (1989). Sample Allocation in Multivariate Surveys. *Survey Methodology*, 15: 47–57.
7. Chaddha, R.L., Hordgrave, W.W., Hudson, D.J., Segal, M. and Suraballe, J.W. (1971). Allocation of total sample when only the strata means are of interest. *Technometrics* 13: 817–831.
8. Chang, C.-T. (2007). Multi choice goal programming. *Omega – The International Journal of Management Science*, 35: 389–396.
9. Chatterjee, S. (1967). A Note on Optimum Allocation. *Skandinavisk Aktuarietidskrift*, 50: 40–44.
10. Chatterjee, S. (1968). Multivariate Stratified Surveys. *Journal of the American Statistical Association*, 63: 530–534.
11. Chromy, J. (1987). Design Optimization with Multiple Objectives. Proceedings of the American Statistical Association, Section on Survey Methods Research, pp. 194–199.
12. Cochran, W.G. (1977). *Sampling Techniques*. New York: John Wiley and Sons, Inc.
13. Dalenius, T. (1957). *Sampling in Sweden: Contributions to the Methods and Theories of Sample Survey Practice*. Stockholm: Almqvistoch Wiksell.

14. Folks, J.L. and Antle, C.E. (1965). Optimum Allocation of Sampling Units to the Strata when there are R Responses of sampling Units of Interest. *Journal of the American Statistical Association*, 60: 225–233.
15. Hansen, M. H., Hurwitz, W.N. and Madow, W.G. (1953). *Sample Survey Methods and Theory, Volume I*. New York: John Wiley and Sons, Inc.
16. Jahan, N., Khan, M.G.M., and Ahsan, M.J. (1994). A Generalized Compromise Allocation. *Journal of Indian Statistical Association*, 32: 95–101.
17. Jessen, R.J. (1942). Statistical Investigation of a Sample Survey for Obtaining Farm Facts. *Iowa State College of Agriculture and Mechanic Arts, Agricultural Experiment Station, Research Bulletin*, 304.
18. Khan, M.G.M., Maiti T. and Ahsan M.J. (2010). An Optimal Multivariate Stratified Sampling Design Using Auxiliary Information: An Integer Solution Using Goal Programming Approach. *Journal of Official Statistics*, 26: 695–708.
19. Khan, M.G.M. and Ahsan, M.J. (2003). A Note on Optimum Allocation in Multivariate Stratified Sampling. *The South Pacific Journal of Natural Science*, 21: 91–95.
20. Khan, M.G.M., Ahsan, M.J. and Jahan, N. (1997). Compromise Allocation in Multivariate Stratified Sampling: An Integer Solution. *Naval Research Logistics*, 44: 69–79.
21. Khan, M.G.M., Khan, E.A., and Ahsan, M.J. (2003). An Optimal Multivariate Stratified Sampling Design Using Dynamic Programming. *Australian & New Zealand Journal of Statistics*, 45: 107–113.
22. Kokan, A.R. (1963). Optimum allocation in multivariate surveys. *Journal of the Royal Statistical Society, Series A*, 126: 557-565.
23. Kokan, A.R. and Khan, S.U. (1967). Optimum Allocation in Multivariate Surveys: An Analytical Solution. *Journal of the Royal Statistical Society, Series B*, 29: 115–125.
24. LINGO (2001). *User's Guide*. Lindo Systems Inc., Chicago, IL, USA.
25. Mathew, O.O., Sola, A.F., Oladiran, B.H. and Amos, A.A. (2013). Efficiency of Neyman Allocation procedure over other allocation procedures in stratified random sampling. *American Journal of Theoretical and Applied Statistics*, 2(5):122–127.
26. Miettinen, K. (1999). *Nonlinear Multiobjective Optimization*. Kluwer Academic Publishers.
27. Neyman, J. (1934). On the two different aspects of the representative method: the method of stratified sampling and the method of purposive selection. *Journal of the Royal Statistical Society*, 97: 558–606.
28. Omule, S.A.Y. (1985). Optimum design in multivariate stratified sampling. *Biometrical Journal*, 27: 907–912.
29. Stuart, A. (1954). A simple presentation of optimum sampling results. *Journal of the Royal Statistical Society, Series B*, 16: 239–241.
30. Sukhatme, P.V., Sukhatme, B.V., Sukhatme, S. and Asok, C. (1984). *Sampling Theory of Surveys with Applications*. Indian Society of Agricultural Statistics, New Delhi, India, and Iowa State University Press, Ames, IA.
31. Swain, A.K.P.C. (2013). A note on optimum allocation in stratified random sampling. *Revista Investigacion Operacional*, 34(2): 161–165.
32. Yates, F and Zacopanay, I. (1935). The estimation of the efficiency of the sampling, with special reference to sampling for yield in cereal experiments. *Journal of Agricultural Science*, 25: 545–577.
33. Yates, F. (1953). *Sampling Methods for Census and Surveys*. London: Charles Griffin and Co. Ltd.