SOME ALLOCATION IN MULTIVARIATE STRATIFIED SAMPLING UNDER COST FUNCTION USING MULTI-OBJECTIVE OPTIMIZATION

By

ATTA ULLAH

A THESIS SUBMITTED IN THE PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE M.PHIL DEGREE IN STATISTICS

DEPARTMENT OF STATISTICS QUAID-I-AZAM UNIVERSITY ISLAMABAD, PAKISTAN

March, 2013

CERTIFICATE

SOME ALLOCATION IN MULTIVARIATE STRATIFIED SAMPLING UNDER COST FUNCTION USING MULTI-OBJECTIVE OPTIMIZATION

By

ATTA ULLAH

A THESIS SUBMITTED IN THE PARTIAL FULFILLMENT OF THE REQUIREMTNS FOR THE DEGREE OF M.PHIL. IN

STATISTICS

We accept this thesis as conforming to the required standards

1.________________________
Prof. Dr. Javid Shabbir 6/3/2013
(Chairman)

2.________________________
Prof. Dr. Javid Shabbir
(Supervisor)

3.________________________
Dr. Masood Anwar
(External Examiner)

Date:________________

DEPARTMENT OF STATISTICS
QUAID-I-AZAM UNIVERSITY
ISLAMABAD, PAKISTAN.
2013

Dedicated to

Our Prophet

***HazratMuhammad** (PBUH)*

&

My parents

Without their efforts and support I am unable to complete this task of my life. Without their love & prayers I am nothing in this world.

Declaration

I "Mr.Atta Ullah" hereby solemnly declare that this thesis entitled "Some allocation in multivariate stratified sampling under cost function using multi-objective optimization", submitted by me for the partial fulfillment of Master of Philosophy in Statistics, is the original work and has not been submitted concomitantly or latterly to this or any other university for any other Degree.

(Signature) (Place, Date)

Acknowledgment

First and foremost thanks to Almighty ALLAH. All praises to Allah for the strengths and His blessing in completing this thesis, without His blessing it is impossible to complete this dissertation. I offer my humblest, sincerest and million Darood to Holy Prophet Hazrat MUHAMMAD (Peace Be Upon Him), who exhorts His followers to seek for Knowledge from cradle to grave. I would have never finished this thesis without the prayers, encouragement and help of many teachers, relatives, friends to whom I am very much indebted.

I have deeply appreciated the complementary perspective of my supervisor **Professor Dr. Javid Shabbir (Chairman)** and all he has contributed to the shape of this thesis. I thank him for his guidance, for great working relationship, and for the grasp of broad concept and attention to detail. His invaluable help of constructive comments and suggestions throughout the thesis works have contributed to the success of this research. **Professor Dr. Javid Shabbir** is a great supervisor, and he can always give out inspiring ideas and productive suggestions. Without his patience, encouragement and intuitive suggestion, I could not finish my dissertation so smoothly. What more, he taught me to have tremendous enthusiasm and confidence in research, which I know is priceless. I convey special acknowledgement to **Dr. Muhammad Yousaf Shad** for his understanding, advice, encouragement and constant support which have provided good and smooth basis for my M.Phil tenure. Words fail me to express my appreciation to **Dr. Muhammad Riaz** , **Dr. Zahid Asghar**, **Dr. Zawar Hussain**, **Dr. Freed Khan**, **Dr. Ijaz Hussain**, **Mr Abdul Haq**, and **Mr Manzoor Khan**, through their vast experience in this field of study, they have offered invaluable and constructed advice and guidance to make this study come to fulfillment. Their guidance has served me well and I could never have embarked and started all of this without their prior teachings in **Statistics** and thus opened up unknown areas to me. I learned a lot from them.

Abstract

The problem of allocation for multiple characteristics of interest in stratified sampling design is examined in this thesis. We use regression estimator for estimating the population mean of more than one characteristics in stratified sampling. We determine the sample size that minimize coefficients of variation of estimates of population means under deterministic and estimated, linear and nonlinear, cost functions. We propose a procedure to maximize the precision of more than one estimates of population means and minimize the variable cost of survey, jointly, under a given sample size. The additional condition is introduced on sample size to avoid over sampling and achieve minimum precision in each strata. The allocation problem in multivariate stratified sampling design is formulated in multi-objective integer nonlinear programming. The proposed multi-objective optimization methods are used to solve formulated allocation problem. A general method is proposed to solve allocation problem in multivariate stratified sampling. One data set is used to illustrate the procedure and compare the efficiency of allocation methods.

Contents

Chapter 1

Introduction and Literature Review

1.1 Introduction

By Deming (1950), "Sampling is not mere substitution of a partial coverage of a total coverage. Sampling is a science and art of controlling and measuring reliability useful statistical information through the theory of probability".

It is a procedure by which we select a representative sample from a given population. The main purpose of sampling is to make an efficient use of the budget and other resources of a study to obtain as precise estimate of population parameter as possible. One of the branches of statistics that is commonly used in all area of scientific inquiries is that of probability sampling. An effective sampling technique is one which produces a meaningful information of the important aspect of population. In probability sampling each units in the population has nonzero known probability of its being included in the sample. Stratified sampling is one among the designs of sampling surveys that is used for obtaining such knowledge.

In stratified sampling design, we divide the sampling frame of heterogeneous population into mutually exclusive and collectively exhaustive sub population or strata, each containing homogeneous elements with reference to study variables under consideration. That is, homogeneity within stratum based on characteristics under study. The stratified sampling enjoys with all of its benefits of convenience, flexibility, efficiency with respect to sampling variance as well as cost, has become necessity factor in all sampling surveys of practical importance.

In general, the following four important questions are raised in process of stratified sampling.

(a) How to form the strata ?

(b) How many number of strata to be formed ?

(c) How to select the samples from each stratum ?

(d) How to allocate or determine the sample size to various strata ?

Of course, these fundamental questions are addressed with objective to minimize sampling variance and cost of sample survey.

The procedure of selecting the best strata boundaries that form strata internally homogeneous, given some sample allocation, is known as optimum stratification. In order to form strata internally homogeneous, the strata boundaries points are selected in such a way the variances within strata should be as small as possible for all the characteristics under study and the distribution of study variable is known as optimum strata boundaries (OSB). One procedure for OSB consists of cutting the range of the distribution of study variable at suitable points. The problem of determining the OSB was discussed by Dalenius (1950), Dalenius and Gurney (1951), Mahalanobis (1952), Hansen et al (1953), Aoyama (1954), Ekman (1959), Hedlin (2000), Kozak (2004), Horgan (2006), Keskinturk and Er (2007) etc. When cost of conducting the survey varies considerably among the strata, the optimum stratification play significant role in controlling the sample size in each stratum that produce the precise estimate in the sense that the error variance of cost best linear unbiased predictor is minimum. Second procedure, called cumulative square root procedure, based on frequency distribution of highly study variable. This technique of stratification is known as cum$\sqrt{f}$ rule after Dalenius and Hodges (1959). The population is based on some auxiliary variable with a different measurement unit cost occurs in strata. Third type of technique to determine OSB for the problem discussed above is mathematical programming approach was discussed by Dalenius (1950), Ahsan and Khan (1977), Khan et al. (2009). Practical consideration like cost, administrative convenience, simplicity of methods are considered in stratification.

Decision on number of strata is quantitative part of stratification. One can make as many strata as sample size. But we should have at least two units per stratum to estimate the variance. Thus the maximum number of strata should be half of sample size, that is $\frac{n}{2}$. The efficiency will be increased as number of strata increases but fact that this increase goes on depressing with increasing number of strata. Break down the strata in situation where, estimation of variance is not possible due to increase in number of strata. More general and complex techniques of sampling can be used in each stratum individually and estimation of population parameters can be done respectively.

In fact in order to increase the precision of estimates, it is essential to select suitable allocation plan. In this procedure, strata sizes, within stratum variability and the cost of measuring a sample units within different strata are taking into account which can effect the selection of allocation plan. If the variability within strata is not known and difference in size of strata is small, we choose an equal sample size from each stratum. We allocate the sample to various strata in proportion to size of each strata where the size of strata is important. This method is simple to use with greater degree of precision of estimates. However, it does not consider an important aspect, namely the variability within strata, associated with stratified sampling.

The population with large variability have to be large sample size. We should select or allocate large sample size to strata which have large variability. An important criteria for improvement in precision of estimates of population means in stratified sampling is to allocate or determine the sample size that minimizes the variance of estimator for fixed total sample size. This considers both strata size and within strata variability. If measurement unit cost varies from stratum to stratum, Neyman (1934), Stuart (1954), Cochran (1977) and Sukhatme et al. (1984) used lagrange multiplier method to determine optimum sample size that minimizes a variance of an estimator for a fixed survey cost or minimize the cost for fixed a variance of an estimator.

If we observe more than one characteristics from each and every unit of population, than lagrange multiplier and other used methods for allocation of sample size for univariate are not useful because it consider a single characteristic among many characteristics. The allocation which is optimum for one characteristic can not be optimum for other characteristics. In such situation a compromise criterion is needed to work out a suitable allocation which is optimum for all characteristics in some sense. Such allocation is called a compromise allocation or mixed allocation in sampling literature.

The auxiliary information can be used to increase precision of an estimate of the parameter. Dalenius (1957), Ghosh (1958), Yates (1960), Aoyama (1963), Kokan and Khan (1967), Ahsan (1975-1976), Ahsan (1978), Ahsan and Khan (1977) , Ahsan and Khan (1982), Bethel (1985), Bethel (1989), Chromy (1987), Kreienbrock (1993), Jahan et el. (1994), Jahan et el. (2001), Jahan and Ahsan (1995), Khan et el. (1997), Khan et el. (2008), Khan et el. (2008), Singh (2003), Semiz (2004), Diaz and Cortez (2006), Diaz and Cortez (2008), Kozak (2006a), Kozak (2006b), Ansari el al. (2009), Pirzada and Maqbool (2003), Khowaja et al. (2011), Ghufran et al. (2011), Ali et al. (2011), Varshney and Ahsan (2011), Ghufran et al. (2012), Khan et al. (2012), Varshney et al. (2012) used lagrange multipliers and mathematical programming methods for solving allocation problems in stratified sampling design.

1.2 Literature review

Folks and Antle (1965) studied the optimum allocation in multivariate stratified as a nonlinear problem of matrix optimization of integers constrained by simple cost function or by fixed sample size under some assumptions. Folks and Antle (1965) proposed a procedure to solve the multi-objective optimization problem as a particular case of matrix optimization. Folks and Antle (1965) defined a particular vectorial function of the objective function of the matrix optimization problem and used value function approach and distance based method to solve non linear integer multi objective optimum allocation problem in multivariate stratified sampling.

Kokan and Khan (1967) used the non-integer nonlinear mathematical programming for obtaining an optimum allocation of sample size under simple cost function when several characteristics are under study. They used an analytical procedure to solve an allocation problem in stratified and two stage sampling designs. Cochran (1977) proposed the use of the average of individual optimum allocation obtained by applying lagrange multipliers method for various characteristics in stratified random sampling. Cochran (1977) minimized the variance of each characteristic under fixed simple cost function.

Jahan et el. (1994) considered the compromise sample allocation problem by optimizing the total relative increase in variances as compared to individual optimum allocation

Pirzada and Maqbool (2003) studied the problem of allocating the sample size to different strata that is minimized the variances of different characteristics subject to constraint of given budget and tolerance limits on certain variances. The problem was converted into non linear mathematical programming problem with various objective functions and single convex constraint. Pirzada and Maqbool (2003) handled the non linearity of convex constraint through cutting plane method and then solved the resulting linear programming problem by Tchebychev's approximation method.

Diaz and Cortez (2006) criticized approaches adopted by Cochran (1977), Sukhatme et al. (1984), Stuart (1954), Arthanari and Dodge (1981), Thompson (1997) for solving allocation problem for various characteristics under study in stratified sampling. Diaz and Cortez (2006) encountered some conditions that violate lagrange multipliers method used for sample allocation. Diaz and Cortez (2006) examined problem of minimizing the variances subject to cost constraint or a given sample size of optimum allocation in multivariate stratified sampling. They used different techniques: lexicographic method, e-constraint method and distance base method taking into account the prior knowledge of population which was classified as complete, partial or no information.

Khowaja et al. (2011) discussed a procedure to work out sample allocation in multivariate stratified sampling survey using compromise criteria. They minimized the sum of squared coefficients of variation for the estimate of population mean of all characteristics of interest under simple linear cost constraint and used lagrange multipliers method to solve formulated nonlinear mathematical programming problem.

Varshney and Ahsan (2011) considered a sample allocation problem in stratified sampling when more than one characteristics are measured from each and every unit of target population. Varshney and Ahsan (2011) considered the simple cost function. Varshney and Ahsan (2011) extended the idea of Ahsan et al. (2005) for several characteristics under study and a combination of compromise and mixed allocation using lagrange multiplier method for solving the nonlinear programming problem. Varshney and Ahsan (2011) compared the compromised mixed allocation with Cochran (1977) average allocation and Sukhatme et al. (1984) compromise allocation.

Ghufran et al. (2011) proposed a method to work out the compromised allocation in multivariate stratified sampling. Ghufran et al. (2011) used compromise criteria for minimizing the sum of sampling variances of simple estimators of the population means of several characteristics under probabilistic cost constraint. The optimum compromise allocation problem is formulated in stochastic nonlinear programming problem. They applied chance constraint programming procedure to transform the stochastic programming to deterministic non linear programming.

Khowaja et al. (2011) proposed a method to work out the compromise allocation in multivariate stratified sampling. Khowaja et al. (2011) considered the travel cost as well as measurement cost observing the sampled units. The cost function is quadratic in $\sqrt{n}$. They formulated the problem as all integer nonlinear programming problem minimizing the sampling variances of an estimators of the population means of various characteristics subject to quadratic cost constraint. They used value function approach, $\varepsilon-$constraint method and distance based method for solving multi-objective compromise allocation problem taking into account the prior knowledge about target population which can be categorized as complete, partial or no information.

Ali et al. (2011) used the probabilistic quadratic cost function in the sample allocation for the estimation of population means of several characteristics in stratified sampling. They formulated the allocation problem as stochastic optimization problem. Ali et al. (2011) used chance constrained programming technique and applied Chebyshev approximation method, goal programming and D_1 distance method to solve integer nonlinear deterministic multi-objective optimization problem

Ghufran et al. (2012) discussed an optimum allocation problem in multivariate stratified sample survey with an objective to minimize simultaneously the squared coefficients of variation of estimators of the population means of various characteristics under cost constraint. The cost function considered here consists of travel cost within stratum to reach the selected units. Ghufran et al. (2012) introduced an additional condition on sample size to safe over sampling and ensure the availability of estimates of stratum variances. They formulated an optimum allocation problem as multi-objective all integer nonlinear mathematical programming problem. They taking into account the complete, partial and no information about the population, proposed value function approach, $\varepsilon-$constraint method and distance based method to solve formulated all integer nonlinear multi-objective optimization problem

Varshney et al. (2012) proposed the compromise allocation based on minimization of individual coefficients of variation of ratio estimators of population means of various characteristics. Varshney et al. (2012) used cost function including measurement unit cost and nonlinear travel cost within stratum. They formulated the allocation problem as multi-objective nonlinear programming problem and used goal programming and weighted method to solve this problem. They used rounding rule to get integer solution.

Khan et al. (2012) discussed problem of sample allocation in multivariate stratified sampling. They stated the problem in two ways. first, optimization of variances of an estimates of population means under cost constraint, second optimization of cost of sample survey under fixed variances of an estimates of population means of each characteristics. They considered nonlinear probabilistic travel cost function. The cost function in nonlinear is $\sqrt{n}$. The chance constraint programming technique was used to get deterministic optimization model equivalent to stochastic optimization model. Khan et al. (2012) applied modified E-Model technique on both cases of sample allocation optimization problem to get optimum allocation in multivariate stratified sampling.

Chapter 2

Methodology

It is essential to select an allocation procedure in order to increase the precision of estimates of population parameters. In his chapter, we discuss various allocation procedure in stratified sampling design and multi-objective optimization methods for solving multi-objective mathematical programming problems.

2.1 Multivariate stratified sampling

Let a population consist of N units divided in to L mutually exclusive strata of size $N_h(h = 1, 2, ..., L)$ such that $\sum_{j=1}^{L} N_h = N$. The sample size n_h is drawn from each stratum independently. In univariate stratified sampling we measure only one characteristic $Y_i(i = 1, 2, ..., Nh)$. Let, we observe $Y_{pi}(i = 1, 2, ..., Nh)$, $p \geq 2$, characteristics from each units of the h^{th} stratum and we estimate the parameter of $p \geq 2$ characteristics of the population.

2.2 Proportionate allocation

In proportionate allocation, the number of units allocated to different strata is proportional to size of the strata in target population. That is, the sample size drawn from each stratum is proportional to the relative size of that stratum in target population. As such, it is self weighting sample allocation method. The sampling fraction $f_h = \frac{n_h}{N_h}$ is used in each stratum, giving each unit in population have an equal chance of selection.

2.3 Disproportionate allocation

Disproportionate allocation is a sample allocation procedure in which number of units selected from each stratum are not representative in the target population. The population units have not an equal chance to be selected in the sample and strata have different sampling fraction.

2.3.1 Disproportionate allocation for within strata analysis

The objective of study may require a researcher to conduct a detail analysis within strata. Since the proportional allocation may select a sample proportional to size of strata. Therefore, the researcher may not select an efficient sample from small strata for obtaining precise estimate within stratum . Disproportional allocation may be used to select over sample from small strata that would be efficient for within strata analysis to meet the objective of study.

2.3.2 Disproportionate allocation for between strata analysis

The objective of a study may require a researcher to make comparison among subgroups of population to each other. If this is the problem, sufficient equal number of units must be selected from each subgroups and an equal allocation or balance allocation of sample size is appropriate for such comparison.

2.3.3 Optimum allocation

Although proportional allocation may provide smaller mean square error than simple random sampling in estimation of population parameters. It may possible to use better allocation than proportional allocation. Optimum allocation is designed even greater over all accuracy than that gained by using proportional allocation. It allocates the sample size to different subgroups or strata considering two important aspects of doing research, cost and precision. The sampling fraction change with respect to cost and variability within different strata. Disproportionate allocation, especially optimum allocation, may be more appropriate for a study than proportional allocation. The allocation of sample to different subgroups differ in terms of cost and the variability of the variable of interest. Optimum allocation may be used concentrating on cost only, precision only or both cost and precision jointly. The optimum sample size to various strata can be determined by lagrange multipliers method and mathematical programming method.

2.4 Lagrange multipliers methods

The scope of lagrange multiplier method for constrained optimization has passed through radical transformation starting with an introduction of augmented lagrangian function and method of multipliers by Hestenes (1969). The method of lagrange multipliers gives a strategy for finding local maxima and minima of a function subject to equality constraints. The lagrange multiplier's method is used to solve single objective optimization problem. Let $f(Z)$ be a objective function of decision variable Z and $h_j(Z), (j = 1, 2, ..., m)$ are constraints on decision variable Z with equality upper bound b_j. The general algorithm of lagrange multipliers method for solving optimization problems with equality constraints is given as:

$$Min\ f(Z)$$

$$\text{Subject to} \qquad h_j(Z) = b, j = 1, 2, ..., m.$$

The lagrange function may be written as

$$L(Z, \lambda_j) = f(Z) + \sum_{j=1}^{m} \lambda_j \left[f(Z) - b \right]$$

where, λ is lagrange multipliers vector of order $m \times 1$ and b is constraint limit. We find the partial derivatives and equivalent to zero as:

$$\frac{\partial L(Z, \lambda_j)}{\partial Z} = 0$$

and

$$\frac{\partial L(Z, \lambda_j)}{\partial \lambda_j} = 0.$$

We solve above equations to find local minima which is global minima if following conditions hold.

$$\frac{\partial^2 L(Z, \lambda_j)}{\partial Z^2} < 0$$

and

$$\frac{\partial^2 L(Z, \lambda_j)}{\partial \lambda_j^2} < 0.$$

2.5 Mathematical programming

It concern with problems of allocating limit resources among contending activities in an optimal manner. The problems of allocation originate whenever one must pick out the level of certain activities or objectives which must contend for certain scare resources compulsory to achieve those objectives. Many type of situations on which mathematical programming can be applied is admit remarkable. However one common component in each of these situations is the essential for controlling resources to activities.

Multi-objective programming consult to techniques for solving the general class of optimization problems concerning with the interaction of many objectives subject to certain bounding conditions or constraints. In solving the multi objective problems such as supply, profit, production, costs, precision or other measure of characteristics obtained in the best possible or optimal manner subject to certain bounding constraints. Let $f_j(Z)(j = 1, 2, ..., p)$ are objective functions of decision variable Z, and $h_{1j}(Z)$ and $h_{2j}(Z)$ are constraints on decision variable Z with inequality upper bound b_{1j} and inequality lower bound b_{2j} respectively. Consider the multi objective optimization problem,

$$Minimize\ \ (f_1(Z), f_2(Z), ..., f_m(Z))$$

Subject to

$$\begin{aligned} h_{1j}(Z) &\leq b_{1j} \\ h_{2j}(Z) &\geq b_{2j} \\ Z^l \leq Z_j &\leq Z^u \\ Z_j &> 0 \\ j &= 1, 2, ..., m \end{aligned} \tag{2.1}$$

where, Z^l and Z^u are lower and upper limits of decision variable. Many methods are available for solving multi objective optimization problem. Some methods are being discussed here.

2.5.1 Goal programming method

Charnes et al. (1955); Charnes and Cooper (1961, 1977), Charnes and Cooper (1961), Ijiri (1965), Charnes et al. (1967) used the goal programming method for solving multi objective optimization problems. It is a technique used by decision makers in optimizing more than one conflicting objectives with each other under unavoidable conditions. In goal programming, all specified objectives are included in the model. The decision maker tries to minimize the sum of potential deviations from specified objectives. These deviations may be positive as above the objectives or negative as below the objective. Consider the following individual optimum problem,

$$Minimize\ f_j(Z)$$

Subject to

$$h_{1j}(Z) \leq b_{1j}$$

$$h_{2j}(Z) \geq b_{2j} \tag{2.2}$$

$$Z^l \leq Z_j \leq Z^u$$

$$Z_j > 0$$

$$j = 1, 2, ..., m$$

Let $f_j^*(Z)$ be the individual optimum values of the function $f_j(Z)$ obtained by solving above problem. These optimum values $f_j^*(Z)$ specify objectives. We try to achieve these objectives in multi objective optimization methods. Let $\hat{f}_j(z)$ is a value of objective function obtained by applying multi objective optimization method. It is obvious that $\hat{f}_j(Z) \geq f_j^*(Z)$ or $\hat{f}_j(Z) - f_j^*(Z) \geq 0$ is the increase in $f_j(Z)$. Suppose this increase is $d_j \geq 0$. To achieve these specified objectives, we must have

$$\hat{f}_j(Z) - f_j^*(Z) \leq d_j$$

or

$$\hat{f}_j(Z) - d_j \leq f_j^*(Z)$$

The sum of all deviations from specified objectives $f_j^*(Z)$ is $\sum_{j=1}^{m} d_j$. In goal programming method, we minimize this total deviations using above additional constraint. The multi objective optimization problem (2.1) can be written in Goal programming as:

$$Minimize \sum_{j=1}^{m} d_j$$

Subject to

$$\hat{f}_j(z) - d_j \leq f_j^*(Z)$$

$$h_{1j}(Z) \leq b_{1j}$$

$$h_{2j}(Z) \geq b_{2j}$$

$$Z^l \leq Z_j \leq Z^u$$

$$Z_j > 0$$

$$j = 1, 2, ..., m$$

where, $d_j (j = 1, 2, ..., m)$ are called deviation variables which are to be minimized.

2.5.2 Value function method

It is a function that indicates the preference of a decision maker among the objectives vector. The different decision makers have different value functions for same multi objective optimization vector. Mathematically, Steuer (1989) linked the decision maker's preference function to weight function. This method is used for solving multiple objective optimization problems when complete information about each objective is known so that the relative weights can be assigned to them. The multi objective minimization problem (2.1) under this weighted method may be expressed as:

$$Min \sum_{j=1}^{m} W_j f_j(Z)$$

Subject to

$$h_{1j}(Z) \leq b_{1j}$$

$$h_{2j}(Z) \geq b_{2j}$$

$$Z^l \leq Z_j \leq Z^u$$

$$\sum_{j=1}^{m} Wj = 1$$

$$Z_j > 0$$

$$j = 1, 2, ..., m$$

Here, $W_j (j = 1, 2, ..., m)$ are the weights which indicates the relative importance of each objectives.

2.5.3 $\varepsilon-$ Constraint method

Besides the value function or weighted approach, the $\varepsilon-$ constraint method is probably best known technique for solving multi objective optimization problems. This method was introduced by Haimes et al. (1971). This method is used in situations where, only partial information is available about objectives vector. The $\varepsilon-$ constraint method minimize only one objective and other objectives are transformed into constraints. The decision maker needs to identify most important objective to apply this method. The problem (2.1) can be written as:

$$Minimize\ f_k(Z)$$

Subject to

$$\hat{f}_j(z) \leq f_j^*(Z)$$

$$h_{1j}(Z) \leq b_{1j}$$

$$h_{2j}(Z) \geq b_{2j}$$

$$Z^l \leq Z_j \leq Z^u$$

$$Z_j > 0$$

$$j \neq k = 1, 2, ..., m$$

2.5.4 Hybrid method

Guddat et al. (1985) introduced a method called hybrid method that is mixture of weighted method and $\varepsilon-$ constraint method. This method for solving multi objective optimization problems has weighted sum objective function and constraints on all objectives. The problem (2.1) may be solved by hybrid method as:

$$Minimize\ \sum_{j=1}^{m} W_j f_j(Z)$$

$$\hat{f}_j(z) \leq f_j^*(Z)$$

$$h_{1j}(Z) \leq b_{1j}$$

$$h_{2j}(Z) \geq b_{2j}$$

$$Z^l \leq Z_j \leq Z^u$$

$$\sum_{j=1}^{m} Wj = 1$$

$$Z_j > 0$$

$$j = 1, 2, ..., m$$

Chapter 3

Optimization of Precision under Deterministic Cost Constraint

3.1 Introduction

An effective sampling procedure that selects a representative sample provides a meaningful information of the important characteristics of population. Stratified sampling is appropriate for obtaining such information from heterogeneous population. It is usual to make use of auxiliary information to improve precision of estimates. In order to increase precision of estimates, it is essential to select a suitable allocation procedure. Sample allocation to different strata has large effect on the precision of estimates. In general, allocation procedure aim to minimize the variance associated with estimating some overall population parameter subject to condition on sampling resources. The allocation problem under discussion is optimization problem. In multivariate stratified sampling, where more than on parameter is to be estimated, an allocation which is optimum for one study variable may not be optimum for other variables. In such situation some compromise criterion is need to allocate sample which is optimum for all the study variables in some sense.

Ghufran et al. (2012) formulated the allocation problem in multi-objective mathematical programming as:

$$Minimize \ \ ((CV_1)^2, (CV_2)^2, \ldots, (CV_p)^2)$$

Subject to

$$\sum_{h=1}^{L} C_h n_h + \sum_{h=1}^{L} t_h \sqrt{n_h} \leq C$$

$$2 \leq n_h \leq N_h$$

$$n_h \text{ are integers } (h = 1, 2, ..., L),$$

where, $CV_j (j = 1, 2, ..., p)$ are the coefficients of variation of a simple estimator of population mean of jth variable. Ghufran et al. (2012) used value function approach, ε-constraint method and distance based method to solve this problem. Varshney et al. (2012) formulated the allocation problem in mathematical programming given as:

$$Minimize \, (CV(\bar{y}_{1,st}), CV(\bar{y}_{2,st}), \ldots, CV(\bar{y}_{p,st}))$$

Subject to

$$\sum_{h=1}^{L} C_h n_h + \sum_{h=1}^{L} t_h \sqrt{n_h} \leq C$$

$$2 \leq n_h \leq N_h$$

$$n_h \text{ are integers} (h = 1, 2, ..., L).$$

$CV(\bar{y}_{j,st}) = \sqrt{\frac{MSE(\bar{y}_{j,st})}{\bar{Y}^2}}$ is coefficient of variation of ratio estimator of population mean of j^{th} characteristic in multivariate stratified sampling. They used goal programming method and weighted method to solve this problem to get compromise allocation.

In this chapter, we minimize the coefficients of variation of regression estimators of population means of $Y_j (j = 1, 2, ..., p)$ characteristics under different types of cost function in general form. For compromise allocation, we have used goal programming technique and proposed general method for solving integer multi-objective optimum allocation problems.

Consider an estimator,

$$\bar{y}_{j,lrs} = \sum_{h=1}^{L} W_h \bar{y}_{j,lrh}.$$

where

$$\bar{y}_{j,lrh} = \bar{y}_{jh} + b_{jh}(\bar{X}_{jh} - \bar{x}_{jh}).$$

The mean square error (MSE) of $\bar{y}_{j,lrs}$ is given as:

$$MSE\left(\bar{y}_{j,lrs}\right) = \sum_{h=1}^{L} W_h^2 \left(\frac{1}{n_h} - \frac{1}{N_h}\right) \left[S_{Yjh}^2 - 2\beta_{jh} S_{YXjh} + \beta_{jh}^2 S_{Xjh}^2\right] \tag{3.1}$$

where

$\bar{y}_{jh}$ = sample mean of the j^{th} study characteristic in h^{th} stratum,

$\bar{x}_{jh}$ = sample mean of the j^{th} auxiliary characteristic in h^{th} stratum,

$\bar{X}_{jh}$ = population mean of the j^{th} auxiliary characteristic in h^{th} stratum,

S_{yjh}^2 = population variance of the j^{th} study characteristic in h^{th} stratum,

S_{xjh}^2 = population variance of the j^{th} auxiliary characteristic in h^{th} stratum,

S_{xyjh} = population covariance between the j^{th} study and the j^{th} auxiliary characteristic in the h^{th} stratum and $\beta_{jh} = \frac{S_{YXjh}}{S_{Xjh}^2}$ is population regression coefficient.

Rewrite (3.1)

$$MSE(\bar{y}_{j,lrs}) = \sum_{h=1}^{L} \frac{W_h^2 \acute{U}_{jh}}{n_h} - \sum_{h=1}^{L} \frac{W_h^2 \acute{U}_{jh}}{N_h} \tag{3.2}$$

where

$$\acute{U}_{jh} = S_{Yjh}^2 - 2\beta_{jh} S_{YXjh} + \beta_{jh}^2 S_{Xjh}^2.$$

We ignore the second term in R.H.S of (3.2) because it is independent of sample size n_h.

$$MSE\left(\bar{y}_{j,lrs}\right) = \sum_{h=1}^{L} \frac{W_h^2 \acute{U}_{jh}}{n_h}.$$

Since different characteristics are measured with different units, there is need to use estimate which is independent of measurement unit. Therefore, we use coefficient of variation instead of mean square error as:

$$C.V\left(\bar{y}_{j,lrs}\right) = \sqrt{\frac{MSE\left(\bar{y}_{j,lrs}\right)}{\bar{Y}_j^2}}$$

or

$$C.V\left(\bar{y}_{j,lrs}\right) = \sqrt{\sum_{h=1}^{L} \frac{W_h^2 \acute{U}_{jh}}{n_h \bar{Y}_j^2}} \tag{3.3}$$

or

$$C.V\ (\bar{y}_{j,lrs}) = \sqrt{\sum_{h=1}^{L} \frac{u'_{jh}}{n_h}} = Z_j, \tag{3.4}$$

where,

$$u'_{jh} = \frac{W_h^2 U'_{jh}}{\bar{Y}_j^2}.$$

Under the cost function, we determine the sample size n_h that minimize $Z_j (j = 1, 2, ..., p)$.

3.2 Allocation under simple cost constraint

Consider the cost function

$$\sum_{h=1}^{L} C_h n_h \leq C - C_0 = \acute{C}. \tag{3.5}$$

where, C_h is cost of observing a selected unit in the h^{th} stratum, C is total cost of survey, C_0 is fixed and $\acute{C}$ is variable cost of conducting the sample survey. We determine the sample size n_h under the cost function (3.5) that minimize coefficients of variation of the estimate of population mean for each characteristics $j (j = 1, 2, ..., p)$. This problem can be formulated in mathematical programming as:

$$Minimize\ \ (Z_1, Z_2, ..., Z_p)$$

Subject to

$$\sum_{h=1}^{L} C_h n_h \leq \acute{C} \tag{3.6}$$

$$2 \leq n_h \leq N_h$$

n_h are integers.

h=1, 2,...,L.

This problem is a multi-objective optimization problem.We use the following methods to solve this problem.

3.2.1 Allocation using individual optimum method

We find the allocation that optimize coefficient of variation of one characteristic of population among $Y_j (j = 1, 2, ..., p)$ characteristics and use that allocation for estimating other characteristics of the population. Let Z_j^* be the optimum value of objective function Z_j obtained by solving following integer nonlinear mathematical programming problem.

$$Minimize\ Z_j$$

Subject to

$$\sum_{h=1}^{L} C_h n_{jh} \leq \acute{C} \tag{3.7}$$

$$2 \leq n_{jh} \leq N_h$$

$$n_{jh} \text{ are integers.}$$

$$h = 1, 2, ..., L \text{ and } j = 1, 2, ..., p.$$

3.2.2 Allocation using goal programming

The multi-objective optimum allocation problem (3.6) may be solved with goal programming method (GP) as:

$$Minimize\ \textstyle\sum_{j=1}^{p} d_j$$

Subject to

$$Z_j - d_j \leq Z_j^* \tag{3.8}$$

$$\sum_{h=1}^{L} C_h n_{hc} \leq \acute{C}$$

$$2 \leq n_{hc} \leq N_h$$

$$n_{hc} \text{ are integers.}$$

$$h = 1, 2, ..., L \text{ and } j = 1, 2, ..., p,$$

where, $d_j (j = 1, 2, ..., p)$ are deviation variables.

3.2.3 Allocation using general method

The problem (3.6) having multiple objectives may be solved with general method as:

$$Minimize\ \sum_{k=1}^{p} W_k Z_k + \sum_{j=1}^{p} d_j$$

Subject to

$$Z_j - dj \leq Z_j^* \tag{3.9}$$

$$\sum_{h=1}^{L} C_h n_{hc} \leq \acute{C}$$

$$2 \leq n_{hc} \leq N_h$$

$$\sum_{k=1}^{p} W_k = 1$$

$$n_{hc} \text{ are integers.}$$

$$h = 1, 2, ..., L \text{ and } k = j = 1, 2, ..., p,$$

where, $W_k (k = 1, 2, ..., p)$ are weights which indicate relative importance of each characteristics and $d_j (j = 1, 2, ..., p)$ are deviation variables.

3.2.4 Numerical example

The data are taken from agricultural census in Iowa state 1997 and 2002 conducted by National Agricultural Statistics Service , USDA, Washington D. C as reported by Khan et al.(2010) (Source :http://wwww.agcensus.usda.gov).

Y_1 denote the quantity of corn harvested in 2002,

Y_2 denote the quantity of oats harvested in 2002,

X_1 denote the quantity of corn harvested in 1997,

X_2 denote the quantity of oats harvested in 1997.

where,X_1 and X_2 are auxiliary information on study variables Y_1 and Y_2 respectively.

$\bar{Y_1} = 474973.90$, $\bar{X_1} = 405654.19$, $\bar{Y_2} = 1576.25$, $\bar{X_2} = 2116.70$

h	N_h	W_h	S^2_{y1h}	S^2_{x1h}	S^2_{y2h}	S^2_{X2h}
1	8	0.0808	29267524195.5	21601503189.8	777174.1	1154134.2
2	34	0.3434	26079256582.8	19734615816.7	4987812.9	7056074.8
3	45	0.4545	42362842460.8	27129658750.0	1074510.6	2082871.3
4	12	0.1212	30728265336.9	17258237358.5	388378.5	732004.9

h	S_{x1y1h}	S_{x2x2h}	β_{1h}	β_{2h}	$\acute{u_{1h}}$	$\acute{u_{2h}}$
1	24360422802.3	902170.6	1.1249	0.7834	0.000066	0.000181
2	22003466630.3	5813439.5	1.1150	0.8239	0.000809	0.009411
3	33367597192.0	1285355.6	1.2300	0.6171	0.001212	0.023390
4	21033769867.3	456991.5	1.2188	0.4243	0.000332	0.000610

We assume that $\acute{C} = C - C_o = 500$ units and $C_1 = 11, C_2 = 13, C_3 = 9, C_4 = 10$. We determine the sample size n_h that minimize the coefficients of variation of estimate of population mean of Y_1 and Y_2. We formulate the problem as:

$$Minimize \begin{pmatrix} Z_1 = \sqrt{\frac{0.000066}{n_1} + \frac{0.000809}{n_2} + \frac{0.001212}{n_3} + \frac{0.000332}{n_4}} \\ Z_2 = \sqrt{\frac{0.000181}{n_1} + \frac{0.009411}{n_2} + \frac{0.023390}{n_3} + \frac{0.000610}{n_4}} \end{pmatrix}$$

Subject to

$$11n_1 + 13n_2 + 9n_3 + 10n_4 \leq 500$$

$$2 \leq n_1 \leq 8$$

$$2 \leq n_2 \leq 34$$

$$2 \leq n_3 \leq 45$$

$$2 \leq n_4 \leq 12$$

n_1, n_2, n_3 and n_4 are integers.

(a) Coefficients of variation using individual allocation

Optimum allocation for characteristic Y_1:

$$Minimize\ Z_1 = \sqrt{\tfrac{0.000066}{n_{11}} + \tfrac{0.000809}{n_{12}} + \tfrac{0.001212}{n_{13}} + \tfrac{0.000332}{n_{14}}}$$

Subject to

$$11n_{11} + 13n_{12} + 9n_{13} + 10n_{14} \leq 500$$

$$2 \leq n_{11} \leq 8$$

$$2 \leq n_{12} \leq 34$$

$$2 \leq n_{13} \leq 45$$

$$2 \leq n_{14} \leq 12$$

$$n_{11}, n_{12}, n_{13} \text{ and } n_{14} \text{ are integers.}$$

$$n_{11} = 3, n_{12} = 15, n_{13} = 18, n_{14} = 11.$$

Amount of cost used $\acute{C}$=500.

Optimum allocation for characteristic Y_2:

$$Minimize\ Z_2 = \sqrt{\tfrac{0.000181}{n_{21}} + \tfrac{0.009411}{n_{22}} + \tfrac{0.023390}{n_{23}} + \tfrac{0.000610}{n_{24}}}$$

Subject to

$$11n_{21} + 13n_{22} + 9n_{23} + 10n_{24} \leq 500$$

$$2 \leq n_{21} \leq 8$$

$$2 \leq n_{22} \leq 34$$

$$2 \leq n_{23} \leq 45$$

$$2 \leq n_{24} \leq 12$$

$$n_{21}, n_{22}, n_{23} \text{ and } n_{24} \text{ are integers.}$$

$$n_{21} = 2, n_{22} = 15, n_{23} = 25, n_{24} = 4.$$

Here Z_1^* and Z_2^* are coefficients of variation using individual allocation which are given in the Table 3.1. Amount of cost used $\acute{C}$=500.

Table 3.1: Coefficients of variation using individual allocation.

$C.V_j$	Y_1	Y_2
Z_1^*	0.01317	0.01478
Z_2^*	0.04520	0.04167
total	0.05837	0.05645

(b) Coefficients of variation using goal programming

For estimating the population means of characteristics Y_1 and Y_2, we determine sample size n_h applying goal programming. Let $\hat{Z}_1$ and $\hat{Z}_2$ are coefficients of variation using proposed allocation methods.

$$Minimize\ d_1 + d_2$$

Subject to

$$\sqrt{\frac{0.000066}{n_{1c}} + \frac{0.000809}{n_{2c}} + \frac{0.001212}{n_{3c}} + \frac{0.000332}{n_{4c}}} - d_1 \leq 0.01317$$

$$\sqrt{\frac{0.000181}{n_{1c}} + \frac{0.009411}{n_{2c}} + \frac{0.023390}{n_{3c}} + \frac{0.000610}{n_{4c}}} - d_2 \leq 0.04167$$

$$11n_{1c} + 13n_{2c} + 9n_{3c} + 10n_{4c} \leq 500$$

$$2 \leq n_{1c} \leq 8$$

$$2 \leq n_{2c} \leq 34$$

$$2 \leq n_{3c} \leq 45$$

$$2 \leq n_{4c} \leq 12$$

n_{1c}, n_{2c}, n_{3c} and n_{4c} are integers.

$$n_{1c} = 3, n_{2c} = 14, n_{3c} = 25, n_{4c} = 6.$$

We get $\hat{Z}_1 = 0.01355$ and $\hat{Z}_2 = 0.04177$ using above allocation. Amount of cost used $\acute{C}$=500.

(c) Coefficients of variation using general method.

We apply general method to determine sample size n_h for estimation of population mean of the characteristics Y_1 and Y_2.

$$Minimize \left(\begin{array}{c} W_1\sqrt{\frac{0.000066}{n_{1c}} + \frac{0.000809}{n_{2c}} + \frac{0.001212}{n_{3c}} + \frac{0.000332}{n_{4c}}} \\ +W_2\sqrt{\frac{0.000181}{n_{1c}} + \frac{0.009411}{n_{2c}} + \frac{0.023390}{n_{3c}} + \frac{0.000610}{n_{4c}}} \\ +d_1 + d_2 \end{array} \right)$$

Subject to

$$\sqrt{\frac{0.000066}{n_{1c}} + \frac{0.000809}{n_{2c}} + \frac{0.001212}{n_{3c}} + \frac{0.000332}{n_{4c}}} - d_1 \leq 0.01317$$

$$\sqrt{\frac{0.000181}{n_{1c}} + \frac{0.009411}{n_{2c}} + \frac{0.023390}{n_{3c}} + \frac{0.000610}{n_{4c}}} - d_2 \leq 0.04167$$

$$11n_{1c} + 13n_{2c} + 9n_{3c} + 10n_{4c} \leq 500$$

$$2 \leq n_{1c} \leq 8$$

$$2 \leq n_{2c} \leq 34$$

$$2 \leq n_{3c} \leq 45$$

$$2 \leq n_{4c} \leq 12$$

$$W_1 + W_2 = 1$$

$$n_{1c}, n_{2c}, n_{3c} \text{ and } n_{4c} \text{ are integers.}$$

Here $\hat{Z}_1$ and $\hat{Z}_2$ are coefficients of variation using general method given in following Table 3.2.

Table 3.2: Coefficients of variations using general method.

W_1	W_2	n_{1c}	n_{2c}	n_{3c}	n_{4c}	C	$\hat{Z}_1$	$\hat{Z}_2$	$\hat{Z}_1 + \hat{Z}_2$
0.1	0.9	2	13	27	6	494	0.01398	0.04222	0.05622
0.2	0.8	4	13	25	6	498	0.01351	0.04250	0.05601
0.3	0.7	3	16	23	5	498	0.01384	0.04228	0.05612
0.4	0.6	2	14	26	8	498	0.01388	0.04200	0.05588
0.6	0.4	4	15	23	5	496	0.01377	0.04256	0.05633
0.7	0.3	4	13	26	5	497	0.01385	0.04232	0.05617
0.8	0.2	2	13	27	6	494	0.01398	0.04222	0.05622
0.9	0.1	2	14	24	8	500	0.01352	0.04259	0.05611

3.2.5 Efficiency comparison

We compare the efficiency of goal programming and general method to the individual allocations for characteristic $j(j = 1, 2, ..., p)$. We use the following expression to obtain the percentage relative efficiency *(PRE)* as:

$$PRE = \frac{\sum_{j=1}^{p} Z_j^*}{\sum_{j=1}^{p} \hat{Z}_j} \times 100$$

The results based on goal programming and general method are given in Tables 3.3-3.4.

Table 3.3: *PRE* of *GP* method.

Y_1	Y_2
105.51	102.04

Table 3.4: *PRE* of general method.

W_1	W_2	Y_1	Y_2	W_1	W_2	Y_1	Y_2
0.1	0.9	103.824	100.400	0.6	0.4	103.622	100.213
0.2	0.8	104.214	100.780	0.7	0.3	103.917	100.490
0.3	0.7	104.009	100.588	0.8	0.2	103.824	100.400
0.4	0.6	104.456	101.020	0.9	0.1	104.028	100.600

3.2.6 Results

Table 3.1 shows that optimum allocation according to Y_1 is more efficient than Y_2. The Table 3.3-3.4 are indicate that GP method and general method give more precise estimates than both individual optimum allocation and the precision varies with respect to different weights.

3.3 Allocation under traveling cost function

We determine a sample size n_h under travel cost function $\sum_{h=1}^{L} t_h n_h^{\delta} \leq \acute{C}$ that minimize coefficients of variation Z_j defined by (3.4) for each characteristics $j(j = 1, 2, ..., p)$. Here t_h is travel cost between units within stratum and $\delta > 0$ represents the effect of travel to cost. This allocation problem becomes a multi-objective mathematical programming problem which can be formulated as:

$$Minimize\ \ (Z_1, Z_2, \ldots, Z_p)$$

Subject to

$$\sum_{h=1}^{L} t_h n_h^{\delta} \leq \acute{C} \tag{3.10}$$

$$2 \leq n_h \leq N_h$$

$$n_h \text{ are integers.}$$

$$h = 1, 2, ..., L \text{ and } j = 1, 2, ..., p.$$

We use following methods for solving multi-objective mathematical programming problem (3.10).

3.3.1 Allocation using individual optimum method

Let Z_j^* be the optimum value of Z_j obtained by solving following integer nonlinear mathematical programming problem.

$$Minimize\ Z_j$$

Subject to

$$\sum_{h=1}^{L} t_h n_{jh}^{\delta} \leq \acute{C} \tag{3.11}$$

$$2 \leq n_{jh} \leq N_h$$

$$n_{jh} \text{ are integers}$$

$$h = 1, 2, ..., L \text{ and } j = 1, 2, ..., p.$$

3.3.2 Allocation using goal programming method

The formulated multi-objective allocation problem (3.10) is solved by goal programming as:

$$Minimize\ \textstyle\sum_{j=1}^{p} d_j$$

Subject to

$$Z_j - d_j \leq Z_j^* \tag{3.12}$$

$$\sum_{h=1}^{L} t_h n_{hc}^{\delta} \leq \acute{C}$$

$$2 \leq n_{hc} \leq N_h$$

$$n_{hc} \text{ are integers}$$

$$h = 1, 2, ..., L \text{ and } j = 1, 2, ..., p,$$

where, $d_j (j = 1, 2, ..., p)$ are deviation variables.

3.3.3 Allocation using general method

The allocation problem (3.10) can be solved with general method as:

$$Minimize\ \textstyle\sum_{k=1}^{p} W_k Z_k + \sum_{j=1}^{p} d_j$$

Subject to

$$Z_j - d_j \leq Z_j^* \tag{3.13}$$

$$\sum_{h=1}^{L} t_h n_{hc}^{\delta} \leq \acute{C}$$

$$2 \leq n_{hc} \leq N_h$$

$$\sum_{k=1}^{p} W_k = 1$$

$$n_{hc} \text{ are integers.}$$

$$h = 1, 2, ..., L \text{ and} j = 1, 2, ..., p,$$

where, $W_k (k = 1, 2, ..., p)$ are relative weights of each characteristics and $d_j (j = 1, 2, ..., p)$ are deviation variables.

3.3.4 Numerical example

Data source[Khan et al.(2010)]. We assume that $t_1 = 10, t_2 = 8, t_3 = 9, t_4 = 11$

$$Minimize \begin{pmatrix} Z_1 = \sqrt{\frac{0.000066}{n_1} + \frac{0.000809}{n_2} + \frac{0.001212}{n_3} + \frac{0.000332}{n_4}} \\ Z_2 = \sqrt{\frac{0.000181}{n_1} + \frac{0.009411}{n_2} + \frac{0.023390}{n_3} + \frac{0.000610}{n_4}} \end{pmatrix}$$

Subject to

$$10n_1^{\delta} + 8n_2^{\delta} + 9n_3^{\delta} + 11n_4^{\delta} \leq \acute{C}$$

$$2 \leq n_1 \leq 8$$

$$2 \leq n_2 \leq 34$$

$$2 \leq n_3 \leq 45$$

$$2 \leq n_4 \leq 12$$

$$n_1, n_2, n_3 \text{ and } n_4 \text{ are integers.}$$

(a) Coefficients of variation using individual optimum allocation

Optimum allocation for characteristics Y_1:

$$Minimize\ Z_1 = \sqrt{\frac{0.000066}{n_{11}} + \frac{0.000809}{n_{12}} + \frac{0.001212}{n_{13}} + \frac{0.000332}{n_{14}}}$$

Subject to

$$10n_{11}^{\delta} + 8n_{12}^{\delta} + 9n_{13}^{\delta} + 11n_{14}^{\delta} \leq \acute{C}$$

$$2 \leq n_{11} \leq 8$$

$$2 \leq n_{12} \leq 34$$

$$2 \leq n_{13} \leq 45$$

$$2 \leq n_{14} \leq 12$$

$$n_{11}, n_{12}, n_{13} \text{ and } n_{14} \text{ are integers.}$$

Optimum allocation for characteristics Y_2:

$$Minimize\ Z_2 = \sqrt{\frac{0.000181}{n_{21}} + \frac{0.009411}{n_{22}} + \frac{0.023390}{n_{23}} + \frac{0.000610}{n_{24}}}$$

Subject to

$$10n_{21}^{\delta} + 8n_{22}^{\delta} + 9n_{23}^{\delta} + 11n_{24}^{\delta} \leq \acute{C}$$

$$2 \leq n_{21} \leq 8$$

$$2 \leq n_{22} \leq 34$$

$$2 \leq n_{23} \leq 45$$

$$2 \leq n_{24} \leq 12$$

$$n_{21}, n_{22}, n_{23} \text{ and } n_{24} \text{ are integers.}$$

Here Z_1^* and Z_2^* are coefficients of variation under individual allocation for different value of δ and$\acute{C}$ in Table 3.5.

Table 3.5: Coefficients of variation using individual allocation.

δ	$\acute{C}$	*Allocation*	n_1	n_2	n_3	n_4	*used* $\acute{C}$	Z_1^*	Z_2^*	$Z_1^*+Z_2^*$
0.5	120	Y_1	4	16	23	4	117.6	0.01424	0.04246	0.05676
		Y_2	2	22	34	2	119.70	0.01647	0.03887	0.05534
0.8	250	Y_1	3	20	15	8	248.57	0.01359	0.04655	0.06014
		Y_2	2	17	29	2	246.82	0.01498	0.04190	0.05888
1	340	Y_1	4	10	14	8	388.00	0.01501	0.05228	0.06729
		Y_2	2	13	20	4	338.00	0.01640	0.04682	0.06322
1.5	1000	Y_1	4	11	13	7	997.00	0.01518	0.05279	0.06797
		Y_2	2	12	16	3	993.99	0.01494	0.05040	0.06734
1.7	1670	Y_1	4	14	12	6	1662.27	0.01518	0.05261	0.06779
		Y_2	2	12	16	3	1658.22	0.01694	0.05040	0.06734
2	1800	Y_1	4	8	10	4	1748.00	0.01794	0.06094	0.07888
		Y_2	3	8	11	3	1790.00	0.01855	0.05972	0.07872

(b) Coefficients of variation using goal programming

Allocation considering the characteristics Y_1 and Y_2, we use goal programming method.

$$Minimize\ d_1 + d_2$$

Subject to

$$\sqrt{\frac{0.000066}{n_{1c}} + \frac{0.000809}{n_{2c}} + \frac{0.001212}{n_{3c}} + \frac{0.000332}{n_{4c}}} - d_1 \leq Z_1^*$$

$$\sqrt{\frac{0.000181}{n_{1c}} + \frac{0.009411}{n_{2c}} + \frac{0.023390}{n_{3c}} + \frac{0.000610}{n_{4c}}} - d_2 \leq Z_2^*$$

$$10n_{1c}^{\delta} + 8n_{2c}^{\delta} + 9n_{3c}^{\delta} + 11n_{4c}^{\delta} \leq \acute{C}$$

$$2 \leq n_{1c} \leq 8$$

$$2 \leq n_{2c} \leq 34$$

$$2 \leq n_{3c} \leq 45$$

$$2 \leq n_{4c} \leq 12$$

n_{1c}, n_{2c}, n_{3c} and n_{4c} are integers.

Here $\hat{Z}_1$ and $\hat{Z}_2$ are coefficients of variation for different values of $\acute{C}$ and δ using goal programming are given in Table 3.6.

Table 3.6: Coefficients of variation using GP method.

δ	$\acute{C}$	n_{1c}	n_{2c}	n_{3c}	n_{4c}	*used* $\acute{C}$	$\hat{Z}_1$	$\hat{Z}_2$	$\hat{Z}_1 + \hat{Z}_2$
0.5	120	2	18	23	3	118.83	0.01501	0.03906	0.05407
0.8	250	2	14	25	5	241.53	0.01434	0.04267	0.05701
1.0	340	2	14	18	4	338.00	0.01553	0.04706	0.06259
1.5	1000	2	13	14	5	997.69	0.01614	0.05085	0.06699
1.7	1670	2	14	14	4	1658.22	0.01575	0.05106	0.06681
2.0	1800	2	9	16	4	1764.00	0.01678	0.05245	0.06923

(c) Coefficients of variation using general method

We determine sample size n_h using general method.

$$Minimize \left(\begin{array}{c} W_1\sqrt{\frac{0.000066}{n_{1c}} + \frac{0.000809}{n_{2c}} + \frac{0.001212}{n_{3c}} + \frac{0.000332}{n_{4c}}} \\ +W_2\sqrt{\frac{0.000181}{n_{1c}} + \frac{0.009411}{n_{2c}} + \frac{0.023390}{n_{3c}} + \frac{0.000610}{n_{4c}}} \\ +d_1 + d_2 \end{array} \right)$$

Subject to

$$\sqrt{\frac{0.000066}{n_{1c}} + \frac{0.000809}{n_{2c}} + \frac{0.001212}{n_{3c}} + \frac{0.000332}{n_{4c}}} - d_1 \leq Z_1^*$$

$$\sqrt{\frac{0.000181}{n_{1c}} + \frac{0.009411}{n_{2c}} + \frac{0.023390}{n_{3c}} + \frac{0.000610}{n_{4c}}} - d_2 \leq Z_2^*$$

$$10n_{1c}^{\delta} + 8n_{2c}^{\delta} + 9n_{3c}^{\delta} + 11n_{4c}^{\delta} \leq \acute{C}$$

$$2 \leq n_{1c} \leq 8$$

$$2 \leq n_{2c} \leq 34$$

$$2 \leq n_{3c} \leq 45$$

$$2 \leq n_{4c} \leq 12$$

$$W_1 + W_2 = 1$$

n_{1c}, n_{2c}, n_{3c} and n_{4c} are integers. The coefficients of variation $\hat{Z}_1$ and $\hat{Z}_2$ for different values of $\acute{C}$ and δ under above allocation are given in Table 3.7.

Table 3.7: Coefficients of variation using general method.

δ	$\acute{C}$	W_1	W_2	n_{1c}	n_{2c}	n_{3c}	n_{4c}	*used* $\acute{C}$	$\hat{Z}_1$	$\hat{Z}_2$	$\hat{Z}_1 + \hat{Z}_2$
0.5	120	0.2	0.8	2	21	30	8	119.15	0.01239	0.03774	0.05103
		0.4	0.6	2	22	27	4	117.48	0.01406	0.03921	0.05327
		0.6	0.4	2	18	30	4	119.38	0.01419	0.03931	0.05350
0.8	250	0.2	0.8	3	15	26	4	249.20	0.01434	0.04171	0.05605
		0.4	0.6	3	15	26	4	249.20	0.01434	0.04171	0.05605
		0.6	0.4	3	17	22	5	24.82	0.01382	0.04242	0.05624
		0.8	0.2	3	17	22	5	24.82	0.01382	0.04242	0.05624
1	340	0.2	0.8	2	14	17	5	339.00	0.01512	0.04755	0.06267
		0.4	0.6	3	13	18	4	340.00	0.01532	0.04729	0.06261
		0.4	0.6	3	13	18	4	340.00	0.01532	0.04729	0.06261
		0.8	0.2	2	14	17	5	389.00	0.01512	0.04755	0.06267
1.5	1000	0.2	0.8	4	10	14	6	996.98	0.01600	0.05100	0.06700
		0.4	0.6	3	11	15	5	989.66	0.01558	0.05096	0.06654
		0.6	0.4	3	11	15	5	989.66	0.01558	0.05096	0.06654
		0.8	0.2	2	13	14	5	997.69	0.01575	0.05106	0.06681
1.7	1670	0.2	0.8	2	12	15	5	1647.47	0.01574	0.05056	0.06630
		0.4	0.6	2	12	15	5	1647.47	0.01574	0.05056	0.06630
		0.6	0.4	3	13	14	5	1659.93	0.01540	0.05076	0.06616
		0.8	0.2	3	10	16	5	1638.24	0.01565	0.05114	0.06679
2	1800	0.2	0.8	2	9	10	4	1764.00	0.01809	0.06023	0.07832
		0.4	0.6	3	8	11	3	1790.00	0.01855	0.05972	0.0787
		0.6	0.4	2	7	11	5	1796.00	0.01803	0.06092	0.07872
		0.8	0.2	2	7	11	5	1796.00	0.01803	0.06092	0.07872

3.3.5 Efficiency comparison

Tables 3.8-3.9 show the *PRE* of goal programming and general method, respectively, to individual allocation method used for the determination of sample size under nonlinear deterministic cost function.

Table 3.8: *PRE* of goal programming method.

δ	$\acute{C}$	Y_1	Y_2	δ	$\acute{C}$	Y_1	Y_2
0.5	0120	104.864	100.349	1.5	1000	101.463	100.522
0.8	0250	112.507	103.280	1.7	1670	101.467	100.793
1.0	0340	107.509	101.007	2.0	1700	113.9939	113.708

Table 3.9: *PRE* of general method.

δ	$\acute{C}$	W_1	W_2	Y_1	Y_2	δ	$\acute{C}$	W_1	W_2	Y_1	Y_2
0.5	120	0.2	0.8	113.106	110.393	1.5	1000	0.2	0.8	101.448	100.507
		0.4	0.6	106.439	103.886			0.4	0.6	102.149	101.202
		0.6	0.4	105.980	103.439			0.6	0.4	102.149	101.202
		0.8	0.2	113.106	110.393			0.8	0.2	101.736	100.739
0.8	250	0.2	0.8	107.290	105.049	1.7	1670	0.2	0.8	102.247	101.569
		0.4	0.6	107.290	105.040			0.4	0.6	102.247	101.569
		0.6	0.4	106.930	104.694			0.6	0.4	102.264	101.784
		0.8	0.2	106.930	104.694			0.8	0.2	101.497	100.823
1.0	340	0.2	0.8	107.372	100.878	2.0	1800	0.2	0.8	100.715	100.511
		0.4	0.6	107.475	100.974			0.4	0.6	100.779	100.575
		0.6	0.4	107.475	100.974			0.6	0.4	100.203	100.000
		0.8	0.2	107.372	100.878			0.8	0.2	100.203	100.000

3.3.6 Results

The optimum allocation according to Y_1 gives more efficient results than Y_2 for different values of δ and $\acute{C}$ as shown in Table 3.5. The goal programming method and general method give efficient results as compare to individual allocation according to Y_2 and Y_2 as shown in Tables 3.8-3.9.

3.4 Allocation under nonlinear cost function

Consider following nonlinear cost function.

$$\sum_{h=1}^{L} C_h n_h + \sum_{h=1}^{L} t_h n_h^{\delta} \leq \acute{C}. \tag{3.14}$$

We minimize coefficients of variation $Z_j(j = 1, 2, ..., p)$ defined by (3.4) under cost function defined by (3.14).We formulate this multi-objective mathematical programming problem in the form of integer nonlinear mathematical programming problem as:

$$Minimize\ \ (Z_1, Z_2, Z_3, \ldots, Z_p)$$

Subject to

$$\sum_{h=1}^{L} C_h n_h + \sum_{h=1}^{L} t_h n_h^{\delta} \leq \acute{C} \tag{3.15}$$

$$2 \leq n_h \leq N_h$$

$$n_h \text{ are integers}$$

$$h = 1, 2, ..., L.$$

The multi-objective mathematical programming allocation problem (**??**) is solved by following methods

3.4.1 Allocation using individual optimum method

Let Z_j^* be the optimum value of Z_j obtained by solving following integer nonlinear mathematical programming problem.

$$Minimize\ Z_j$$

Subject to

$$\sum_{h=1}^{L} C_h n_h + \sum_{h=1}^{L} t_h n_{jh}^{\delta} \leq \acute{C} \tag{3.16}$$

$$2 \leq n_{jh} \leq N_h$$

$$n_{jh} \text{ are integers}$$

$$h = 1, 2, ..., L \text{ and } j = 1, 2, ..., p.$$

3.4.2 Allocation using goal programming method

The problem (3.15) can be solved with goal programming method as:

$$Minimize \ \sum_{j=1}^{p} d_j$$

Subject to

$$Z_j - d_j \leq Z_j^* \tag{3.17}$$

$$\sum_{h=1}^{L} C_h n_h + \sum_{h=1}^{L} t_h n_{hc}^{\delta} \leq \acute{C}$$

$$2 \leq n_{hc} \leq N_h$$

$$n_{hc} \text{ are integers.}$$

$$h = 1, 2, ..., L \text{ and } j = 1, 2, ..., p,$$

where, $d_j (j = 1, 2, ..., p)$ are deviation variables.

3.4.3 Allocation using general method

The problem (3.15) can be solved with general method as:

$$Minimize \ \sum_{k=1}^{p} W_k Z_k + \sum_{j=1}^{p} d_j$$

Subject to

$$Z_j \leq Z_j^* \tag{3.18}$$

$$\sum_{h=1}^{L} C_h n_h + \sum_{h=1}^{L} t_h n_{hc}^{\delta} \leq \acute{C}$$

$$2 \leq n_{hc} \leq N_h$$

$$\sum_{k=1}^{p} W_k = 1$$

$$n_{hc} \text{ are integers}$$

$$h = 1, 2, ..., L \text{ and} j = 1, 2, ..., p,$$

where, $W_k (k = 1, 2, ..., p)$ are relative weights of characteristics and $d_j (j = 1, 2, ..., p)$ are deviation variables.

3.4.4 Numerical example

Data source [Khan et al.(2010)].We assume that $C_1 = 12, C_2 = 8, C_3 = 6, C_4 = 10, t_1 = 6, t_2 = 4, t_3 = 3, t_4 = 5$.

$$Minimize \begin{pmatrix} Z_1 = \sqrt{\frac{0.000066}{n_1} + \frac{0.000809}{n_2} + \frac{0.001212}{n_3} + \frac{0.000332}{n_4}} \\ Z_2 = \sqrt{\frac{0.000181}{n_1} + \frac{0.009411}{n_2} + \frac{0.023390}{n_3} + \frac{0.000610}{n_4}} \end{pmatrix}$$

Subject to

$$12n_1 + 8n_2 + 6n_3 + 10n_4 + 6n_1^\delta + 4n_2^\delta + 3n_3^\delta + 5n_4^\delta \leq \acute{C}$$

$$2 \leq n_1 \leq 8$$

$$2 \leq n_2 \leq 34$$

$$2 \leq n_3 \leq 45$$

$$2 \leq n_4 \leq 12$$

n_1, n_2, n_3 and n_4 are integers.

(a) Coefficients of variation using individual optimum method.

Individual Optimum allocation for characteristic Y_1:

$$Minimize\ Z_1 = \sqrt{\frac{0.000066}{n_{11}} + \frac{0.000809}{n_{12}} + \frac{0.001212}{n_{13}} + \frac{0.000332}{n_{14}}}$$

Subject to

$$12n_{11} + 8n_{12} + 6n_{13} + 10n_{14} + 6n_{11}^\delta + 4n_{12}^\delta + 3n_{13}^\delta + 5n_{14}^\delta \leq \acute{C}$$

$$2 \leq n_{11} \leq 8$$

$$2 \leq n_{12} \leq 34$$

$$2 \leq n_{13} \leq 45$$

$$2 \leq n_{14} \leq 12$$

$$n_{11}, n_{12}, n_{13} \text{ and } n_{14} \text{ are integers.}$$

Individual Optimum allocation for characteristic Y_2:

$$Minimize\ Z_2 = \sqrt{\frac{0.000181}{n_{21}} + \frac{0.009411}{n_{22}} + \frac{0.023390}{n_{23}} + \frac{0.000610}{n_{24}}}$$

Subject to

$$12n_{21} + 8n_{22} + 6n_{23} + 10n_{24} + 6n_{21}^{\delta} + 4n_{22}^{\delta} + 3n_{23}^{\delta} + 5n_{24}^{\delta} \leq \acute{C}$$

$$2 \leq n_{21} \leq 8$$

$$2 \leq n_{22} \leq 34$$

$$2 \leq n_{23} \leq 45$$

$$2 \leq n_{24} \leq 12$$

$$n_{21}, n_{22}, n_{23} \text{ and } n_{24} \text{ are integers.}$$

Here Z_1^* and Z_2^* are coefficients of variation under individual allocation for different values of δ and$\acute{C}$ given in Table 3.10.

Table 3.10: Coefficients of variations using individual allocation.

δ	$\acute{C}$	*Allocation*	n_1	n_2	n_3	n_4	*used* $\acute{C}$	Z_1^*	Z_2^*	Z_1^*+Z_2^*
0.5	300	Y_1	2	9	18	5	298.00	0.01602	0.05057	0.06659
		Y_2	2	10	22	2	298.27	0.01830	0.04899	0.06729
0.7	400	Y_1	4	14	16	6	395.62	0.01433	0.04776	0.06209
		Y_2	2	15	25	3	399.71	0.01569	0.04309	0.05878
1	500	Y_1	4	11	21	7	498.00	0.01398	0.04584	0.05982
		Y_2	2	13	29	3	498.00	0.01601	0.04271	0.05872
1.5	850	Y_1	4	12	15	9	847.56	0.01420	0.04950	0.06376
		Y_2	2	14	21	3	833.18	0.01610	0.04560	0.06170
1.8	1300	Y_1	3	11	16	8	1295.18	0.01460	0.04954	0.06414
		Y_2	2	12	19	3	1272.43	0.01658	0.04805	0.06463
2	1500	Y_1	3	10	13	7	1470.00	0.01560	0.05374	0.06934
		Y_2	2	10	16	3	1467.00	0.01733	0.05193	0.06926

(b) Coefficient of variation using goal programming Method.

We use goal programming method for sample allocation to different strata taking into account two characteristics Y_1 and Y_2.

$$Minimize\ d_1 + d_2$$

Subject to

$$\sqrt{\frac{0.000066}{n_{1c}} + \frac{0.000809}{n_{2c}} + \frac{0.001212}{n_{3c}} + \frac{0.000332}{n_{4c}}} - d_1 \leq Z_1^*$$

$$\sqrt{\frac{0.000181}{n_{1c}} + \frac{0.009411}{n_{2c}} + \frac{0.023390}{n_{3c}} + \frac{0.000610}{n_{4c}}} - d_2 \leq Z_2^*$$

$$12n_{21} + 8n_{22} + 6n_{23} + 10n_{24} + 6n_{21}^{\delta} + 4n_{22}^{\delta} + 3n_{23}^{\delta} + 5n_{24}^{\delta} \leq \acute{C}$$

$$2 \leq n_{1c} \leq 8$$

$$2 \leq n_{2c} \leq 34$$

$$2 \leq n_{3c} \leq 45$$

$$2 \leq n_{4c} \leq 12$$

n_{1c}, n_{2c}, n_{3c} and n_{4c} are integers.

Here $\hat{Z}_1$ and $\hat{Z}_2$ are coefficients of variation using goal programming method for sample allocations given in Table 3.11.

Table 3.11: Coefficients of variation using *GP* method.

δ	$\acute{C}$	n_{1c}	n_{2c}	n_{3c}	n_{4c}	*used* $\acute{C}$	$\hat{Z}_1$	$\hat{Z}_2$	$\hat{Z}_1 + \hat{Z}_2$
0.5	300	2	11	19	3	299.48	0.01678	0.04879	0.06557
0.7	400	2	12	25	5	396.50	0.01467	0.043967	0.05863
1.0	500	2	17	22	4	498.00	0.01479	0.04313	0.05792
1.5	850	2	15	19	4	835.80	0.01529	0.04584	0.06113
1.8	1300	2	12	17	6	1271.06	0.01507	0.04850	0.06357
2.0	1500	2	10	15	5	1468.00	0.01616	0.05209	0.06825

(c) Coefficients of variation using general method

We determine sample size n_h taking into account two characteristics Y_1 and Y_2 using general method.

$$Minimize \left(\begin{array}{c} W_1\sqrt{\frac{0.000066}{n_{1c}} + \frac{0.000809}{n_{2c}} + \frac{0.001212}{n_{3c}} + \frac{0.000332}{n_{4c}}} \\ +W_2\sqrt{\frac{0.000181}{n_{1c}} + \frac{0.009411}{n_{2c}} + \frac{0.023390}{n_{3c}} + \frac{0.000610}{n_{4c}}} \\ +d_1 + d_2 \end{array} \right)$$

Subject to

$$\sqrt{\frac{0.000066}{n_{1c}} + \frac{0.000809}{n_{2c}} + \frac{0.001212}{n_{3c}} + \frac{0.000332}{n_{4c}}} - d_1 \leq Z_1^*$$

$$\sqrt{\frac{0.000181}{n_{1c}} + \frac{0.009411}{n_{2c}} + \frac{0.023390}{n_{3c}} + \frac{0.000610}{n_{4c}}} - d_2 \leq Z_2^*$$

$$12n_{21} + 8n_{22} + 6n_{23} + 10n_{24} + 6n_{21}^{\delta} + 4n_{22}^{\delta} + 3n_{23}^{\delta} + 5n_{24}^{\delta} \leq \acute{C}$$

$$2 \leq n_{1c} \leq 8$$

$$2 \leq n_{2c} \leq 34$$

$$2 \leq n_{3c} \leq 45$$

$$2 \leq n_{4c} \leq 12$$

$$W_1 + W_2 = 1$$

n_{1c}, n_{2c}, n_{3c} and n_{4c} are integers.

The coefficients of variation $\hat{Z}_1$ and $\hat{Z}_2$ obtained by general method are given in Table 3.12.

Table 3.12: Coefficients of variation using general method.

δ	$\acute{C}$	W_1	W_2	n_{1c}	n_{2c}	n_{3c}	n_{4c}	$used\ \acute{C}$	$\hat{Z}_1$	$\hat{Z}_2$	$\hat{Z}_1+\hat{Z}_2$
0.5	300	0.2	0.8	2	11	19	3	2999.48	0.01676	0.04878	0.06554
		0.35	0.65	2	11	19	3	299.48	0.016766	0.04878	0.06554
		0.7	0.3	2	10	18	4	295.86	0.01626	0.04984	0.06610
		0.85	0.15	2	11	17	4	298.12	0.01615	0.04974	0.06589
0.7	400	0.2	0.8	2	14	24	4	396.04	0.01498	0.04347	0.05845
		0.35	0.65	3	12	25	4	399.47	0.01486	0.04396	0.05822
		0.7	0.3	2	12	25	4	396.50	0.01523	0.04430	0.05953
		0.85	0.15	2	15	22	4	391.68	0.01500	0.04397	0.05897
1	500	0.2	0.8	2	12	27	5	498.00	0.01455	0.04316	0.05771
		0.35	0.65	2	17	22	4	498.00	0.01479	0.04313	0.05793
		0.7	0.3	2	15	23	5	498.00	0.01435	0.04309	0.05744
		0.85	0.15	2	13	24	6	498.00	0.01418	0.04348	0.05766
1.5	850	0.2	0.8	3	14	19	4	831.16	0.01505	0.04600	0.06105
		0.35	0.65	3	11	22	5	848.58	0.01473	0.04584	0.06057
		0.7	0.3	3	11	21	6	849.30	0.01444	0.04617	0.06061
		0.7	0.3	3	11	21	6	849.30	0.01444	0.04617	0.06061
1.7	1300	0.2	0.8	3	12	18	4	1279.66	0.01548	0.04792	0.06390
		0.35	0.65	5	12	16	5	1292.83	0.01493	0.04903	0.06396
		0.7	0.3	2	12	17	6	1271.05	0.01507	0.04850	0.06357
		0.85	0.15	2	12	17	6	1271.05	0.01507	0.04850	0.06357
2	1500	0.2	0.8	2	13	21	4	1499.90	0.01536	0.04562	0.06098
		0.35	0.65	4	12	20	5	1486.90	0.01452	0.04654	0.06106
		0.7	0.3	3	14	18	6	1492.88	0.01423	0.04619	0.06042
		0.85	0.15	3	13	19	6	1488.86	0.01426	0.04601	0.06027

3.4.5 Efficiency comparison

We compare the efficiency of goal programming method and general method to the individual allocations for characteristic Y_1 and Y_2. The PRE is given in Tables 3.13-3.14.

$$PRE = \frac{\sum_{j=1}^{p} Z_j^*}{\sum_{j=1}^{p} \hat{Z}_j} \times 100$$

Table 3.13: PRE of goal programming method.

δ	$\acute{C}$	Y_1	Y_2	δ	$\acute{C}$	Y_1	Y_2
0.5	0300	101.556	102.623	1.5	850	104.302	100.932
0.7	0400	105.901	100.256	1.8	1300	100.897	101.667
1.0	0500	103.280	101.381	2.0	1500	101.597	101.480

Table 3.14: PRE of general method.

δ	$\acute{C}$	W_1	W_2	Y_1	Y_2	δ	$\acute{C}$	W_1	W_2	Y_1	Y_2
0.5	120	0.2	0.8	101.602	102.670	1.5	850	0.2	0.8	104.439	101.065
		0.35	0.64	101.602	102.670			0.35	0.65	105.267	101.886
		0.7	0.3	100.740	100.800			0.7	0.3	105.197	101.798
		0.85	0.15	101.062	102.125			0.85	0.15	105.197	101.798
0.7	400	0.2	0.8	106.228	100.565	1.8	1300	0.2	0.8	100.376	101.142
		0.35	0.65	106.647	100.962			0.35	0.65	100.281	101.048
		0.7	0.3	104.300	098.740			0.7	0.3	100.897	101.667
		0.85	0.15	105.291	099.670			0.85	0.15	100.897	101.667
1.0	500	0.2	0.8	103.656	101.750	2.0	1500	0.2	0.8	113.708	113.615
		0.35	0.65	103.263	101.364			0.35	0.65	113.560	113.429
		0.7	0.3	104.143	102.228			0.7	0.3	114.763	114.631
		0.85	0.15	103.746	101.838			0.85	0.25	115.049	114.916

3.4.6 Results

Table 3.10 indicates that individual optimum allocation according to Y_1 gives larger coefficients of variation compare to Y_2. Table 3.13 exhibit that GP method is efficient allocation technique than individual allocation criteria.Tables 3.13-3.14 show that goal programming method and general method give more precise estimates as compare to individual optimum method according to Y_1 and Y_2 for different values of constants δ and $\acute{C}$.

3.5 Allocation under logarithmic traveling cost function

We determine sample size n_h under the logarithmic cost function $\sum_{h=1}^{L} t_h log(n_h^{\delta}) \leq \acute{C}$. We minimize coefficients of variation $Z_j (j = 1, 2, ..., p)$ defined by (3.4) under this logarithmic cost function.We formulate this sample determination problem as multi-objective optimization problem in the form of integer nonlinear mathematical programming as:

$$Minimize\ \ (Z_1, Z_2, \ldots, Z_p)$$

Subject to

$$\sum_{h=1}^{L} t_h log(n_h^{\delta}) \leq \acute{C} \tag{3.19}$$

$$2 \leq n_h \leq N_h$$

$$n_h \text{ are integers}$$

$$h = 1, 2, ..., L.$$

We use following methods for solving multi-objective allocation problem which is formulated as mathematical programming problem (3.19).

3.5.1 Allocation using individual optimum method

The individual optimum allocation method is used to determine the sample size n_h that minimized one coefficient of variation among $Z_j (j = 1, 2, ..., p)$ and other coefficients of variation are measured for that allocation.

$$Minimize\ Z_j$$

Subject to

$$\sum_{h=1}^{L} t_h log(n_h^{\delta}) \leq \acute{C} \tag{3.20}$$

$$2 \leq n_{jh} \leq N_h$$

$$n_{jh} \text{ are integers.}$$

$$h = 1, 2, ..., L \text{ and } j = 1, 2, ..., p.$$

Let Z_j^* be the optimum value of Z_j obtained by solving above allocation problem.

3.5.2 Allocation using goal Programming method

The multi-objective optimization problem (3.19) may be written in goal programming as:

$$Minimize\ \sum_{j=1}^{p} d_j$$

Subject to

$$Z_j - d_j \leq Z_j^* \tag{3.21}$$

$$\sum_{h=1}^{L} t_h log(n_h^{\delta}) \leq \acute{C}$$

$$2 \leq n_{hc} \leq N_h$$

$$n_{hc} \text{ are integers.}$$

$$h = 1, 2, ..., L \text{ and } j = 1, 2, ..., p,$$

where, $d_j (j = 1, 2, ..., p)$ are deviation variables.

3.5.3 Allocation using general method

The multi-objective allocation problem (3.19) may be solved with general method as:

$$Minimize\ \sum_{k=1}^{p} W_k Z_k + \sum_{j=1}^{p} d_j$$

Subject to

$$Z_j - d_j \leq Z_j^* \tag{3.22}$$

$$\sum_{h=1}^{L} t_h log(n_h^{\delta}) \leq \acute{C}$$

$$2 \leq n_{hc} \leq N_h$$

$$\sum_{k=1}^{p} W_k = 1$$

$$n_{hc} \text{ are integers.}$$

$$h = 1, 2, ..., L \text{ and } j = 1, 2, ..., p,$$

where, $W_k(k = 1, 2, ..., p)$ are relative weights which express importance of each characteristics and $d_j(j = 1, 2, ..., p)$ are deviation variables.

3.5.4 Numerical example

We assume that $t_1 = 10, t_2 = 8, t_3 = 7, t_4 = 9$.

$$Minimize \begin{pmatrix} Z_1 = \sqrt{\frac{0.000066}{n_1} + \frac{0.000809}{n_2} + \frac{0.001212}{n_3} + \frac{0.000332}{n_4}} \\ Z_2 = \sqrt{\frac{0.000181}{n_1} + \frac{0.009411}{n_2} + \frac{0.023390}{n_3} + \frac{0.000610}{n_4}} \end{pmatrix}$$

Subject to

$$10log(n_1^\delta) + 8log(n_2^\delta) + 7log(n_3^\delta) + 9log(n_4^\delta) \leq \acute{C}$$

$$2 \leq n_1 \leq 8$$

$$2 \leq n_2 \leq 34$$

$$2 \leq n_3 \leq 45$$

$$2 \leq n_4 \leq 12$$

$$n_1, n_2, n_3 \text{ and } n_4 \text{ are integers.}$$

(a) Coefficients of variation using individual optimum method

The individual allocation taking a characteristic Y_1:

$$Minimize\ Z_1 = \sqrt{\frac{0.000066}{n_{11}} + \frac{0.000809}{n_{12}} + \frac{0.001212}{n_{13}} + \frac{0.000332}{n_{14}}}$$

Subject to

$$10log(n_{11}^\delta) + 8log(n_{12}^\delta) + 7log(n_{13}^\delta) + 9log(n_{14}^\delta) \leq \acute{C}$$

$$2 \leq n_{11} \leq 8$$

$$2 \leq n_{12} \leq 34$$

$$2 \leq n_{13} \leq 45$$

$$2 \leq n_{14} \leq 12$$

n_{11}, n_{12}, n_{13} and n_{14} are integers.

The individual allocation using a characteristic Y_2:

$$Minimize\ Z_2 = \sqrt{\tfrac{0.000181}{n_{21}} + \tfrac{0.009411}{n_{22}} + \tfrac{0.023390}{n_{23}} + \tfrac{0.000610}{n_{24}}}$$

Subject to

$$10log(n_{21}^{\delta}) + 8log(n_{22}^{\delta}) + 7log(n_{23}^{\delta}) + 9log(n_{24}^{\delta}) \leq \acute{C}$$

$$2 \leq n_{22} \leq 34$$

$$2 \leq n_{23} \leq 45$$

$$2 \leq n_{24} \leq 12$$

n_{21}, n_{22}, n_{23} and n_{24} are integers.

Here Z_1^* and Z_2^* are coefficients of variation under using individual allocation for different value of δ and$\acute{C}$ given in Table 3.15.

Table 3.15: Coefficients of variation using individual allocation.

δ	$\acute{C}$	*Allocation*	n_1	n_2	n_3	n_4	*used* $\acute{C}$	Z_1^*	Z_2^*	$Z_1^*+Z_2^*$
0.5	30	Y_1	3	10	8	4	28.22	0.01837	0.06386	0.08223
		Y_2	3	10	8	4	28.22	0.01837	0.06386	0.08223
1.0	60	Y_1	3	10	8	4	56.44	0.01837	0.06386	0.08223
		Y_2	3	10	8	4	56.44	0.01837	0.06386	0.08223
1.5	85	Y_1	3	8	5	3	73.16	0.02182	0.07822	0.10004
		Y_2	2	19	15	2	71.98	0.02200	0.07462	0.09662
2	115	Y_1	3	8	6	4	105.28	0.02020	0.07271	0.09291
		Y_2	3	8	6	4	105.28	0.02020	0.07271	0.09291

(b) Coefficients of variation using goal programming method.

We use propose goal programming method for allocation of sample size to four strata considering two characteristics Y_1 and Y_2.

$$Minimize\ d_1 + d_2$$

Subject to

$$\sqrt{\frac{0.000066}{n_{1c}} + \frac{0.000809}{n_{2c}} + \frac{0.001212}{n_{3c}} + \frac{0.000332}{n_{4c}}} - d_1 \leq Z_1^*$$

$$\sqrt{\frac{0.000181}{n_{1c}} + \frac{0.009411}{n_{2c}} + \frac{0.023390}{n_{3c}} + \frac{0.000610}{n_{4c}}} - d_2 \leq Z_2^*$$

$$10log(n_{1c}^{\delta}) + 8log(n_{2c}^{\delta}) + 7log(n_{3c}^{\delta}) + 9log(n_{4c}^{\delta}) \leq \acute{C}$$

$$2 \leq n_{1c} \leq 8$$

$$2 \leq n_{2c} \leq 34$$

$$2 \leq n_{3c} \leq 45$$

$$2 \leq n_{4c} \leq 12$$

$$n_{1c}, n_{2c}, n_{3c} \text{ and } n_{4c} \text{ are integers.}$$

Here $\hat{Z}_1$ and $\hat{Z}_2$ are coefficients of variation obtained by solving problem which are given in Table 3.16.

Table 3.16: Coefficients of variation using *GP* method.

δ	$\acute{C}$	n_{1c}	n_{2c}	n_{3c}	n_{4c}	*used* $\acute{C}$	$\hat{Z}_1$	$\hat{Z}_2$	$\hat{Z}_1 + \hat{Z}_2$
0.5	30	2	10	15	2	25.27	0.01899	0.05381	0.07280
1.0	60	2	10	15	2	50.54	0.01899	0.05381	0.07280
1.5	85	2	34	5	2	78.97	0.02157	0.07315	0.09472
2.0	150	2	26	6	2	111.66	0.02079	0.06823	0.08902

(c) Coefficients of variation using general method

Two characteristics Y_1 and Y_2 are considered in allocation of sample size by general method.

$$Minimize\left(\begin{array}{c} W_1\sqrt{\frac{0.000066}{n_{1c}}+\frac{0.000809}{n_{2c}}+\frac{0.001212}{n_{3c}}+\frac{0.000332}{n_{4c}}} \\ +W_2\sqrt{\frac{0.000181}{n_{1c}}+\frac{0.009411}{n_{2c}}+\frac{0.023390}{n_{3c}}+\frac{0.000610}{n_{4c}}} \\ +d_1+d_2 \end{array}\right)$$

Subject to

$$\sqrt{\frac{0.000066}{n_{1c}}+\frac{0.000809}{n_{2c}}+\frac{0.001212}{n_{3c}}+\frac{0.000332}{n_{4c}}}-d_1\leq Z_1^*$$

$$\sqrt{\frac{0.000181}{n_{1c}}+\frac{0.009411}{n_{2c}}+\frac{0.023390}{n_{3c}}+\frac{0.000610}{n_{4c}}}-d_2\leq Z_2^*$$

$$10log(n_{1c}^{\delta})+8log(n_{2c}^{\delta})+7log(n_{3c}^{\delta})+9log(n_{4c}^{\delta})\leq \acute{C}$$

$$2\leq n_{1c}\leq 8$$

$$2\leq n_{2c}\leq 34$$

$$2\leq n_{3c}\leq 45$$

$$2\leq n_{4c}\leq 12$$

$$W_1+W_2=1$$

$$n_{1c}, n_{2c}, n_{3c} \text{ and } n_{4c} \text{ are integers.}$$

The coefficients of variation $\hat{Z}_1$ and $\hat{Z}_2$ are given in Table 3.17 obtained by solving above allocation problem.

Table 3.17: Coefficients of variation using general method.

δ	$\acute{C}$	W_1	W_2	n_{1c}	n_{2c}	n_{3c}	n_{4c}	*used* $\acute{C}$	$\hat{Z}_1$	$\hat{Z}_2$	$\hat{Z}_1 + \hat{Z}_2$
0.5	30	0.2	0.8	2	14	39	2	29.96	0.01697	0.04083	0.05780
		0.4	0.6	2	16	32	2	29.80	0.01695	0.04141	0.05836
		0.6	0.4	2	12	45	2	29.84	0.01713	0.04123	0.05836
		0.8	0.2	2	15	21	3	29.90	0.01598	0.04511	0.06109
1.0	60	0.2	0.8	2	14	39	2	59.93	0.01697	0.04083	0.05780
		0.4	0.6	2	14	39	2	59.93	0.01697	0.04083	0.05780
		0.6	0.4	2	17	28	2	59.16	0.01703	0.04224	0.05927
		0.8	0.2	2	15	21	3	59.80	0.01598	0.04511	0.06109
1.5	85	0.2	0.8	2	16	11	2	78.20	0.01897	0.05577	0.07474
		0.4	0.6	2	16	11	2	78.20	0.01897	0.05577	0.07474
		0.6	0.4	2	15	12	2	78.34	0.01881	0.05452	0.07333
		0.8	0.2	2	14	13	2	78.35	0.01871	0.05354	0.07225
2	115	0.2	0.8	2	13	28	2	114.03	0.01745	0.04421	0.06166
		0.4	0.6	2	13	28	2	114.03	0.01745	0.04421	0.06166
		0.6	0.4	2	15	25	2	114.73	0.01736	0.04425	0.06161
		0.8	0.2	2	12	29	2	113.24	0.01756	0.04457	0.06213

3.5.5 Efficiency comparison

The percentage relative efficiency of goal programming method and general method as compare to individual optimum method for characteristic $Y_j (j = 1, 2)$ are given in Tables 3.18-3.19.

Table 3.18: *PRE* of goal programming method.

δ	$\acute{C}$	Y_1	Y_2	δ	$\acute{C}$	Y_1	Y_2
0.5	030	112.95	112.95	1.5	085	105.62	102.01
1.0	060	112.95	112.95	2.0	115	104.37	104.37

Table 3.19: PRE of general method.

δ	$\acute{C}$	W_1	W_2	Y_1	Y_2	δ	$\acute{C}$	W_1	W_2	Y_1	Y_2
0.5	30	0.2	0.8	142.27	142.27	1.5	85	0.2	0.8	133.85	129.27
		0.4	0.6	140.90	140.90			0.4	0.6	133.85	129.276
		0.6	0.4	140.90	140.90			0.6	0.4	136.42	131.76
		0.8	0.2	134.60	134.60			0.8	0.2	138.46	133.73
1.0	60	0.2	0.8	142.27	142.27	2.0	115	0.2	0.8	150.68	150.68
		0.4	0.6	142.27	142.27			0.35	0.65	150.68	150.68
		0.6	0.4	138.74	138.74			0.7	0.3	150.80	150.80
		0.8	0.2	134.60	134.60			0.85	0.25	149.54	149.54

3.5.6 Results

Individual optimum allocation according to Y_1 and Y_2 give equally precise estimates of population means as shown in Table 3.15. Goal programming method and general method of compromise allocation give more efficient estimates than individual optimum allocation according to Y_1 and Y_2 for different values of variable cost $\acute{C}$ and constant δ as given in Tables 3.18-3.19.

Chapter 4

Optimization of Precision under Probabilistic Cost Function

4.1 Introduction

In many surveys, the cost for obtaining information from selected units from different strata are not known exactly, rather than these costs are being estimated from sample. In this situation the cost function becomes probabilistic and chance constraint programming is used to solve formulated problem.

Ghufran et al (2011) and Ali et al (2011) formulated the allocation problem in mathematical programming given below.

$$Minimize\ \ [V(\bar{y}_{1,st}), V(\bar{y}_{2,st}), \ldots, V(\bar{y}_{p,st})]$$

Subject to

$$p\left(\sum_{h=1}^{L} C_h n_h + \sum_{h=1}^{L} t_h \sqrt{n_h} \leq C\right) \geq p_o$$

$$2 \leq n_h \leq N_h$$

$$h = 1, 2, ..., L.$$

where, $V(\bar{y}_{j,st})(j = 1, 2, ..., p)$ are variances of estimators of population means of Y_j characteristics in multivariate stratified sampling. Ali et al used chebyshev approximation method , goal programming method and D_1 distance method of optimization and Ghufran et all used weighted method to solve above multi-objective optimization problem. The integer solution was obtained to round the results.

In this chapter, We minimize coefficients of variation of regression estimators for the parameter of population means of $Y_j (j = 1, 2, ..., p)$ characteristics under different types of probabilistic cost function using goal programming and proposed general method to solve optimization problem. We also adopt charnes and copper (1959) procedure to convert integer stochastic optimization problem into deterministic optimization problem.

4.2 Allocation under simple probabilistic cost function

Consider the following cost function.

$$C_0 + \sum_{h=1}^{L} C_h n_h \leq C \tag{4.1}$$

We perceive that cost of observing the selected sample units within stratum cannot remain same but vary from unit to unit. We consider the measurement cost C_h as a variable and assume that it is normally distributed with mean μ_{ch} and variance σ^2_{ch}. Thus the linear cost function (4.1) becomes probabilistic which can be written as:

$$p\left(C_0 + \sum_{h=1}^{L} C_h n_h \leq C\right) \geq p_o \tag{4.2}$$

where, p_o is given probability. Under the probabilistic cost functions, we determine the sample size n_h that minimize $Z_j (j = 1, 2, 3, ..., p)$ defined by (3.4). We notice that sample size n_h has direct or positive relationship with variability S^2_h within stratum and inverse relationship with stratum measurement cost C_h. The objective of study may require a researcher to achieve at least some specified level of precision within each stratum. That is, some proportion of units from each stratum must be included in the sample for obtaining minimum level of precision in each stratum. The researcher need the sample allocation procedure to conduct a detail analysis within stratum. We can use mathematical programming techniques to handle this situation by introducing an additional condition on sample size n_h. We determine the sample size n_h that maximize the overall precision of the estimates of population means of characteristics $Y_j(j = 1, 2, ..., p)$ under the probabilistic cost function given in (4.2) and additional condition discussed as above. This multi-objective allocation problem can be formulated in chance constrained integer mathematical programming as:

$$Minimize \ (Z_1, Z_2, Z_3, \ldots, Z_p)$$

Subject to

$$p\left(C_0 + \sum_{h=1}^{L} c_h n_h \leq C\right) \geq p_o \tag{4.3}$$

$$l_h \leq n_h \leq u_h$$

$$l_h \geq 2$$

$$n_h \text{ are integers.} h = 1, 2, ..., L,$$

where, l_h is lower bound on sample size n_h that must be satisfied for within strata analysis and u_h is upper bound on sample size n_h. Now we transform problem (4.3) into deterministic mathematical programming problem as:

Let

$$g_1 = C_0 + \sum_{h=1}^{L} C_h n_h.$$

Applying expectation on both sides, we get

$$E\left(g_1\right) = \left(C_0 + \sum_{h=1}^{L} E(C_h) n_h\right)$$

or

$$E\left(g_1\right) = C_0 + \sum_{h=1}^{L} \mu_{ch} n_h \tag{4.4}$$

Now

$$Var\left(g_1\right) = Var\left(C_0 + \sum_{h=1}^{L} C_h n_h\right).$$

$$Var\left(g_1\right) = \sum_{h=1}^{L} \sigma_{ch}^2 n_h^2 \tag{4.5}$$

The function g_1 is normally distributed with mean $E(g_1)$ and variance $Var(g_1)$ are defined in (4.4) and (4.5) respectively.That is, $g_1 \sim N(E(g_1), V(g_1))$.The chance constraint (4.2) can be written as,

$$p(g \leq C) \geq p_o$$

$$p\left(\frac{g_1 - E(g_1)}{\sqrt{Var(g_1)}} \leq \frac{C - E(g_1)}{\sqrt{Var(g_1)}}\right) \geq p_o$$

where, $p\left(\frac{g_1 - E(g_1)}{\sqrt{Var(g_1)}}\right)$ is standard normal random variable. Thus the probability of realizing g_1 less than or equal to total cost C can be written as:

$$p(g_1 \leq C) = \phi\left(\frac{C - E(g_1)}{\sqrt{Var(g_1)}}\right) = \phi(z) \tag{4.6}$$

where, $\phi(z)$ represent the cumulative density function of normal random variable g_1 calculated at z. Let A_α represent standard normal variable at which $\phi(A_\alpha) = p_o$. Then the constraint (4.6) can be written as:

$$\phi\left(\frac{C - E(g_1)}{\sqrt{Var(g_1)}}\right) \geq \phi(A_\alpha).$$

This inequality will be satisfied only if

$$\frac{C - E(g_1)}{\sqrt{Var(g_1)}} \geq A_\alpha$$

or

$$E(g_1) + (A_\alpha)\sqrt{Var(g_1)} \leq C \tag{4.7}$$

Substituting (4.4) and (4.5) in (4.7), we get deterministic constraint equivalent to probabilistic constraint (4.2).

$$\sum_{h=1}^{L} \mu_{ch} n_h + (A_\alpha)\sqrt{\sum_{h=1}^{L} \sigma_{ch}^2 n_h^2} \leq C - C_o = \acute{C} \tag{4.8}$$

If parameters μ_{ch} and σ_{ch}^2 are unknown, we replace their estimates. Let $\hat{\mu_{ch}}$ and $\hat{\sigma_{ch}^2}$ are estimates of parameters μ_{ch} and σ_{ch}^2 respectively. The chance constraint mathematical programming problem (4.3) can be written in deterministic mathematical programming problem as:

$$Minimize\ \ (Z_1, Z_2, Z_3, \ldots, Z_p)$$

Subject to

$$\sum_{h=1}^{L} \mu_{ch} n_h + (A_\alpha) \sqrt{\sum_{h=1}^{L} \sigma_{ch}^2 n_h^2} \leq \acute{C} \tag{4.9}$$

$$l_h \leq n_h \leq u_h$$

$$l_h \geq 2$$

$$n_h \text{ are integers.}$$

$$h = 1, 2, ..., L \text{ and } j = 1, 2, ..., p.$$

We propose following methods for solving multi-objective integer mathematical programming problem (4.9)

4.2.1 Allocation using individual optimum method

Let Z_j^* be a optimum value of objective Z_j obtained by solving following nonlinear integer mathematical programming problem

$$Minimize\ Z_j$$

Subject to

$$\sum_{h=1}^{L} \mu_{ch} n_{jh} + (A_\alpha) \sqrt{\sum_{h=1}^{L} \sigma_{ch}^2 n_{jh}^2} \leq \acute{C} \tag{4.10}$$

$$l_h \leq n_j h \leq u_h$$

$$l_h \geq 2$$

$$n_h \text{ are integers.}$$

$$h = 1, 2, ..., L \text{ and } j = 1, 2, ..., p.$$

4.2.2 Allocation using goal programming

The problem (4.9) can be solved in goal programming (GP) as:

$$Minimize\ \sum_{j=1}^{p} d_j$$

Subject to

$$Z_j - d_j \leq Z_j^* \tag{4.11}$$

$$\sum_{h=1}^{L} \mu_{ch} n_j h + (A_\alpha) \sqrt{\sum_{h=1}^{L} \sigma_{ch}^2 n_j^2 h} \leq \acute{C}$$

$$l_h \leq n_{hc} \leq u_h$$

$$l_h \geq 2$$

$$n_{hc} \text{ are integers.}$$

$$h = 1, 2, ..., L \text{ and } j = 1, 2, ..., p,$$

where, $d_j (j = 1, 2, ..., p)$ are deviation variables.

4.2.3 Allocation using general method

The problem (4.9) may be solved with general method as:

$$Minimize\ \sum_{k=1}^{p} W_k Z_k + \sum_{j=1}^{p} d_j$$

Subject to

$$Z_j - d_j \leq Z_j^* \tag{4.12}$$

$$\sum_{h=1}^{L} \mu_{ch} n_j h + (A_\alpha) \sqrt{\sum_{h=1}^{L} \sigma_{ch}^2 n_j^2 h} \leq \acute{C}$$

$$l_h \leq n_{hc} \leq u_h$$

$$\sum_{k=1}^{p} W_k = 1$$

$$l_h \geq 2$$

$$n_{hc} \text{ are integers.}$$

$$h = 1, 2, ..., L \text{ and } j = 1, 2, ..., p,$$

where, $W_k(k = 1, 2, ..., p)$ are relative weights which indicates an importance of characteristics Y_j and $d_j(j = 1, 2, ..., p)$ are deviation variables.

4.2.4 Numerical example

Data Source [Khan et al.(2010)]. Let chance be required to be sanctified with 99.60 percent probability. The A_α is such that $\phi(A_\alpha) = 0.996$. A_α correspond to 99.60 percent confidence limit is 2.67. We assume that $\acute{C} = C - C_o = 350$ and

$$E(C_1) = 8, E(C_2) = 6, E(C_3) = 5, E(C_4) = 7, V(C_1) = 5, V(C_2) = 4.5, V(C_3) = 3.5, V(C_4) = 5$$

$l_h$is 25 percent of stratum size N_h and u_h is 75 percent of stratum size N_h

$$Minimize \begin{pmatrix} Z_1 = \sqrt{\frac{0.000066}{n_1} + \frac{0.000809}{n_2} + \frac{0.001212}{n_3} + \frac{0.000332}{n_4}} \\ Z_2 = \sqrt{\frac{0.000181}{n_1} + \frac{0.009411}{n_2} + \frac{0.023390}{n_3} + \frac{0.000610}{n_4}} \end{pmatrix}$$

Subject to

$$8n_1 + 6n_2 + 5n_3 + 7n_4 + 5n_1^2 + 4.5n_2^2 + 3.5n_3^2 + 5n_4^2 \leq 350$$

$$2 \leq n_1 \leq 6$$

$$9 \leq n_2 \leq 26$$

$$11 \leq n_3 \leq 34$$

$$3 \leq n_4 \leq 9$$

n_1, n_2, n_3 and n_4 are integers.

(a) Coefficients of variation using individual allocation

Optimum allocation for characteristic Y_1 :

$$Minimize\ Z_1 = \sqrt{\frac{0.000066}{n_{11}} + \frac{0.000809}{n_{12}} + \frac{0.001212}{n_{13}} + \frac{0.000332}{n_{14}}}$$

Subject to

$$8n_{11} + 6n_{12} + 5n_{13} + 7n_{14} + 5n_{11}^2 + 4.5n_{12}^2 + 3.5n_{13}^2 + 5n_{14}^2 \leq 350$$

$$2 \leq n_{11} \leq 6$$

$$9 \leq n_{12} \leq 26$$

$$11 \leq n_{13} \leq 34$$

$$3 \leq n_{14} \leq 9$$

n_{11}, n_{12}, n_{13} and n_{14} are integers.

Optimum allocation for characteristic Y_2 :

$$Minimize\ Z_2 = \sqrt{\frac{0.000181}{n_{21}} + \frac{0.009411}{n_{22}} + \frac{0.023390}{n_{23}} + \frac{0.000610}{n_{24}}}$$

Subject to

$$8n_{21} + 6n_{22} + 5n_{23} + 7n_{24} + 5n_{21}^2 + 4.5n_{22}^2 + 3.5n_{23}^2 + 5n_{24}^2 \leq 350$$

$$2 \leq n_{21} \leq 6$$

$$9 \leq n_{22} \leq 26$$

$$11 \leq n_{23} \leq 34$$

$$3 \leq n_{24} \leq 9$$

n_{21}, n_{22}, n_{23} and n_{24} are integers.

The coefficients of variation Z_1^* and Z_1^* using individual allocation are given in the Table 4.1.

Table 4.1: Coefficients of variation using individual allocation

$C.V_j$	Y_1	Y_2
Z_1^*	0.01458	0.01624
Z_2^*	0.04850	0.04617
total	0.06308	0.06241

(b) Coefficients of variation using goal programming

We determine sample size n_h taking into account the characteristics $Y_j(j = 1, 2)$ using goal programming.

$$Minimize\ d_1 + d_2$$

Subject to

$$\sqrt{\frac{0.000066}{n_{1c}} + \frac{0.000809}{n_{2c}} + \frac{0.001212}{n_{3c}} + \frac{0.000332}{n_{4c}}} - d_1 \leq 0.01458$$

$$\sqrt{\frac{0.000181}{n_{1c}} + \frac{0.009411}{n_{2c}} + \frac{0.023390}{n_{3c}} + \frac{0.000610}{n_{4c}}} - d_2 \leq 0.04617$$

$$8n_{1c} + 6n_{2c} + 5n_{3c} + 7n_{4c} + 5n_{1c}^2 + 4.5n_{2c}^2 + 3.5n_{3c}^2 + 5n_{4c}^2 \leq 350$$

$$2 \leq n_{1c} \leq 6$$

$$9 \leq n_{2c} \leq 26$$

$$11 \leq n_{3c} \leq 34$$

$$3 \leq n_{4c} \leq 9$$

n_{1c}, n_{2c}, n_{3c} and n_{4c} are integers.

$$n_{1c}^* = 3, n_{2c}^* = 13, n_{3c}^* = 19, n_{4c}^* = 4.$$

Here $\hat{Z}_1$ and $\hat{Z}_2$ are coefficients of variation obtained by using above allocation.

$$\hat{Z}_1 = 0.01551, \hat{Z}_2 = 0.04682, \hat{Z}_1 + \hat{Z}_2 = 0.06233.$$

Amount of cost used $\acute{C}$=348.77.

(c) Coefficients of variation using general method

For estimation of population mean of characteristics $Y_j (j = 1, 2)$, we determine sample size n_h using general method.

$$Minimize \left(\begin{array}{c} W_1\sqrt{\frac{0.000066}{n_{1c}} + \frac{0.000809}{n_{2c}} + \frac{0.001212}{n_{3c}} + \frac{0.000332}{n_{4c}}} \\ +W_2\sqrt{\frac{0.000181}{n_{1c}} + \frac{0.009411}{n_{2c}} + \frac{0.023390}{n_{3c}} + \frac{0.000610}{n_{4c}}} \\ +d_1 + d_2 \end{array} \right)$$

Subject to

$$\sqrt{\frac{0.000066}{n_{1c}} + \frac{0.000809}{n_{2c}} + \frac{0.001212}{n_{3c}} + \frac{0.000332}{n_{4c}}} - d_1 \leq 0.01458$$

$$\sqrt{\frac{0.000181}{n_{1c}} + \frac{0.009411}{n_{2c}} + \frac{0.023390}{n_{3c}} + \frac{0.000610}{n_{4c}}} - d_2 \leq 0.04617$$

$$8n_{1c} + 6n_{2c} + 5n_{3c} + 7n_{4c} + 5n_{1c}^2 + 4.5n_{2c}^2 + 3.5n_{3c}^2 + 5n_{4c}^2 \leq 350$$

$$2 \leq n_{1c} \leq 6$$

$$9 \leq n_{2c} \leq 26$$

$$11 \leq n_{3c} \leq 34$$

$$3 \leq n_{4c} \leq 9$$

$$W_1 + W_2 = 1$$

n_{1c}, n_{2c}, n_{3c} and n_{4c} are integers.

The coefficients of variation $\hat{Z}_1$ and $\hat{Z}_2$ are obtained by solving above problem given in Table 4.2

Table 4.2: Coefficients of variation using general method.

W_1	W_2	n_{1c}	n_{2c}	n_{3c}	n_{4c}	C	$\hat{Z}_1$	$\hat{Z}_2$	$\hat{Z}_1 + \hat{Z}_2$
0.2	0.8	2	14	19	4	349.52	0.01505	0.04633	0.06138
0.4	0.6	2	14	18	5	349.19	0.01498	0.04673	0.06171
0.6	0.4	2	13	18	6	348.19	0.01476	0.04707	0.06183
0.8	0.2	2	13	19	5	348.35	0.01501	0.04656	0.06157

4.2.5 Efficiency comparison

We compare the efficiency of goal programming method and general method to the individual optimum method for characteristic $Y_j (j = 1, 2)$. The *PRE* of goal programming method and general method to individual optimum method are given below in Tables 4.3-4.4.

Table 4.3: *PRE* of goal programming method.

Y_1	Y_2
101.12	100.05

Table 4.4: *PRE* of general method.

W_1	W_2	Y_1	Y_2	W_1	W_2	Y_1	Y_2
0.2	0.8	102.77	101.67	0.6	0.4	102.02	100.93
0.4	0.6	102.22	101.13	0.8	0.2	102.45	101.36

4.2.6 Results

Table 4.1 shows that optimum allocation with respect to characteristic Y_2 gives smaller coefficients of variation than Y_1. The *GP* method and general method give efficient estimates of population means as compare to individual optimum allocation according to Y_1 and Y_2 as given in Tables 4.3-4.4. The precision of estimates of general method differ from individual optimum allocation for different relative weights of variables Y_1 and Y_2.

4.3 Allocation under probabilistic traveling cost function

Consider the following cost function.

$$C_0 + \sum_{h=1}^{L} t_h n_h^{\delta} \leq C \tag{4.13}$$

We determine the sample size n_h under the cost function (4.13) that minimize the coefficients of variation of estimates of population means for characteristic $Y_j (j = 1, 2, ..., p)$.

Here, t_h is travel cost of measuring the selected units in the h^{th} stratum for all characteristics $Y_j(j = 1, 2, ..., p)$ and δ represents effect of travel to cost.C_o is fixed cost and C is total cost of survey. Practically, we observe that traveling cost t_h vary for taking information from unit to unit in the h^{th} stratum and consider as random. We assume that $t_h \sim N(\mu_{th}, \sigma_{th}^2)$. Thus traveling cost function (4.13) becomes probabilistic written as:

$$p\left(C_0 + \sum_{h=1}^{L} t_h n_h^{\delta} \leq C\right) \geq p_o \tag{4.14}$$

We determine the sample size n_h that minimize the coefficients of variation $Z_j(j = 1, 2, ..., p)$ defined by (3.4) under the constraint (4.14) and additional condition discussed in section 4.2. This multi-objective optimization problem can be formulated in chance constraint nonlinear mathematical programming problem as:

$$Minimize\ \ (Z_1, Z_2, \ldots, Z_p)$$

Subject to

$$p\left(C_0 + \sum_{h=1}^{L} t_h n_h^{\delta} \leq C\right) \geq p_o \tag{4.15}$$

$$l_h \leq n_h \leq u_h$$

$$l_h \geq 2$$

$$n_h \text{ are integers.}$$

$$h = 1, 2, ..., L \text{ and } j = 1, 2, ..., p.$$

Now we transform chance constraint nonlinear mathematical programming problem (4.15) into deterministic nonlinear mathematical programming problem as:

Let

$$g_2 = C_0 + \sum_{h=1}^{L} t_h n_h^{\delta}$$

Applying expectation on both sides,

$$E\left(g_2\right) = \left(C_0 + \sum_{h=1}^{L} E(t_h) n_h^{\delta}\right)$$

or

$$E\left(g_2\right) = C_0 + \sum_{h=1}^{L} \mu_{th} n_h^{\delta} \tag{4.16}$$

$$Var\left(g_2\right) = Var\left(C_0 + \sum_{h=1}^{L} t_h n_h^{\delta}\right)$$

$$Var\left(g_2\right) = \sum_{h=1}^{L} \sigma_{th}^2 n_h^{2\delta} \tag{4.17}$$

The function g_2 is normally distributed with mean $E(g_2)$ and variance $Var(g_2)$ are defined in (4.16) and (4.17) respectively. That is, $g_2 \sim N\left(E(g_2), Var(g_2)\right)$. The chance constraint (4.14) can be written as:

$$p\left(g_2 \leq C\right) \geq p_o$$

or

$$p\left(\frac{g_2 - E(g_2)}{\sqrt{Var(g_2)}} \leq \frac{C - E(g_2)}{\sqrt{Var(g_2)}}\right) \geq p_o$$

where, $p\left(\frac{g_2 - E(g_2)}{\sqrt{Var(g_2)}}\right)$ is standard normal random variable. Thus the probability of realizing g_2 less than or equal to total cost C can be written as:

$$p\left(g_2 \leq C\right) = \phi\left(\frac{C - E(g_2)}{\sqrt{Var(g_2)}}\right) = \phi(z) \tag{4.18}$$

where, $\phi\left(z\right)$ represents the cumulative density function of the standard normal distribution random variable calculated at z. Let A_α represent standard normal variable at which $\phi(A_\alpha) = p_o$. Then the constraint (4.18) can be written as:

$$\phi\left(\frac{C - E(g_2)}{\sqrt{Var(g_2)}}\right) \geq \phi\left(A_\alpha\right).$$

This inequality will be satisfied only if

$$\frac{C - E(g_2)}{\sqrt{Var(g_2)}} \geq A_\alpha.$$

or

$$E(g_2) + (A_\alpha)\sqrt{Var(g_2)} \leq C. \tag{4.19}$$

Substituting (4.16) and (4.17) in (4.19), we get deterministic constraint equivalent to probabilistic constraint (4.14).

$$\sum_{h=1}^{L} \mu_{th} n_h^{\delta} + (A_\alpha) \sqrt{\sum_{h=1}^{L} \sigma_{th}^2 n_h^{2\delta}} \leq C - C_o = \acute{C} \tag{4.20}$$

If parameters μ_{th} and σ_{th}^2 are unknown, we replace by their estimates. Let $\hat{\mu_{th}}$ and $\hat{\sigma_{th}^2}$ are estimates of parameters μ_{th} and σ_{th}^2 respectively. The chance constraint nonlinear mathematical programming problem (4.15) can be written in multi-objective deterministic optimization nonlinear mathematical programming problem as:

$$Minimize \ \ (Z_1, Z_2, \ldots, Z_p)$$

Subject to

$$\sum_{h=1}^{L} \mu_{th} n_h^{\delta} + (A_\alpha) \sqrt{\sum_{h=1}^{L} \sigma_{th}^2 n_h^{2\delta}} \leq \acute{C} \tag{4.21}$$

$$l_h \leq n_h \leq u_h$$

$$l_h \geq 2$$

$$n_h \text{ are integers}$$

$$h = 1, 2, ..., L.$$

We use the following optimization methods for solving deterministic multi-objective optimization problem (4.21).

4.3.1 Allocation using individual optimum method

Let Z_j^* be a optimum value of objective function obtained by solving following nonlinear integer mathematical programming problem.

$$Minimize \ Z_j$$

Subject to

$$\sum_{h=1}^{L} \mu_{th} n_{jh}^{\delta} + (A_\alpha) \sqrt{\sum_{h=1}^{L} \sigma_{th}^2 n_{jh}^{2\delta}} \leq \acute{C} \tag{4.22}$$

$$l_h \leq n_j h \leq u_h$$

$$l_h \geq 2$$

$$n_h \text{ are integers}$$

$$h = 1, 2, ..., L \text{ and } j = 1, 2, ..., p.$$

4.3.2 Allocation using goal programming method

The multi-objective allocation problem (4.21) can be solved with goal programming (GP) method as:

$$Minimize \ \sum_{j=1}^{p} d_j$$

Subject to

$$Z_j - d_j \leq Z_j^* \tag{4.23}$$

$$\sum_{h=1}^{L} \mu_{th} n_{hc}^{\delta} + (A_\alpha)\sqrt{\sum_{h=1}^{L} \sigma_{th}^2 n_{hc}^{2\delta}} \leq \acute{C}$$

$$l_h \leq n_{hc} \leq u_h$$

$$l_h \geq 2$$

$$n_{hc} \text{ are integers.}$$

$$h = 1, 2, ..., L \text{ and } j = 1, 2, ..., p,$$

where, $d_j (j = 1, 2, ..., p)$ are deviation variables.

4.3.3 Allocation using general method

We can determine the sample size n_h by solving multi-objective optimization problem (4.21) with general method as:

$$Minimize \ \sum_{k=1}^{p} W_k Z_k + \sum_{j=1}^{p} d_j$$

Subject to

$$Z_j - d_j \leq Z_j^* \tag{4.24}$$

$$\sum_{h=1}^{L} \mu_{th} n_{hc}^{\delta} + (A_\alpha)\sqrt{\sum_{h=1}^{L} \sigma_{th}^2 n_{hc}^{2\delta}} \leq \acute{C}$$

$$l_h \leq n_{hc} \leq u_h$$

$$\sum_{k=1}^{p} W_k = 1$$

$$l_h \geq 2$$

n_{hc} are integers.

$$h = 1, 2, ..., L \text{ and } j = 1, 2, ..., p,$$

where, $W_k(k = 1, 2, ..., p)$ are relative weights of characteristics Y_j and $d_j(j = 1, 2, ..., p)$ are deviation variables

4.3.4 Numerical example

Data Source [Khan et al.(2010)]. We assume that $(t_1) = 7, E(t_2) = 6,$

$$E(t_3) = 8, E(t_4) = 5, Var(t_1) = 4.5, Var(t_2) = 3.5, Var(t_3) = 2.5, Var(t_4) = 4.5.$$

The lower bound l_h on sample size n_h is 25 percent of stratum size and upper bound u_h on sample size n_h is 75 percent of stratum size.

$$Minimize \begin{pmatrix} Z_1 = \sqrt{\frac{0.000066}{n_1} + \frac{0.000809}{n_2} + \frac{0.001212}{n_3} + \frac{0.000332}{n_4}} \\ Z_2 = \sqrt{\frac{0.000181}{n_1} + \frac{0.009411}{n_2} + \frac{0.023390}{n_3} + \frac{0.000610}{n_4}} \end{pmatrix}$$

Subject to

$$7n_1^{\delta} + 6n_2^{\delta} + 8n_3^{\delta} + 5n_4^{\delta} + 2.67\sqrt{4.5n_1^{2\delta} + 3.5n_2^{2\delta} + 2.5n_3^{2\delta} + 5n_4^{2\delta}} \leq \acute{C}$$

$$2 \leq n_1 \leq 6$$

$$9 \leq n_2 \leq 26$$

$$11 \leq n_3 \leq 34$$

$$3 \leq n_4 \leq 9$$

n_1, n_2, n_3 and n_4 are integers.

(a) Coefficients of variation using individual allocation

Individual optimum allocation for characteristic Y_1 :

$$Minimize\ Z_1 = \sqrt{\frac{0.000066}{n_{11}} + \frac{0.000809}{n_{12}} + \frac{0.001212}{n_{13}} + \frac{0.000332}{n_{14}}}$$

Subject to

$$7n_{11}^{\delta} + 6n_{12}^{\delta} + 8n_{13}^{\delta} + 5n_{14}^{\delta} + 2.67\sqrt{4.5n_{11}^{2\delta} + 3.5n_{12}^{2\delta} + 2.5n_{13}^{2\delta} + 5n_{14}^{2\delta}} \leq \acute{C}$$

$$2 \leq n_{11} \leq 6$$

$$9 \leq n_{12} \leq 26$$

$$11 \leq n_{13} \leq 34$$

$$3 \leq n_{14} \leq 9$$

n_{11}, n_{12}, n_{13} and n_{14} are integers.

Individual optimum allocation for characteristic Y_2 :

$$Minimize\ Z_2 = \sqrt{\frac{0.000181}{n_{21}} + \frac{0.009411}{n_{22}} + \frac{0.023390}{n_{23}} + \frac{0.000610}{n_{24}}}$$

Subject to

$$7n_{21}^{\delta} + 6n_{22}^{\delta} + 8n_{23}^{\delta} + 5n_{24}^{\delta} + 2.67\sqrt{4.5n_{21}^{2\delta} + 3.5n_{22}^{2\delta} + 2.5n_{23}^{2\delta} + 5n_{24}^{2\delta}} \leq \acute{C}$$

$$2 \leq n_{21} \leq 6$$

$$9 \leq n_{22} \leq 26$$

$$11 \leq n_{23} \leq 34$$

$$3 \leq n_{24} \leq 9$$

n_{21}, n_{22}, n_{23} and n_{24} are integers.

Here Z_1^* and Z_2^* are coefficients of variation using individual allocation for different value of δ and $\acute{C}$ given in Table 4.5.

Table 4.5: Coefficients of variations using individual allocation.

δ	$\acute{C}$	*Allocation*	n_1	n_2	n_3	n_4	*used* $\acute{C}$	Z_1^*	Z_2^*	$Z_1^*+Z_2^*$
0.5	115	Y_1	3	15	15	9	113.38	0.01391	0.04811	0.06202
		Y_2	2	15	23	3	110.90	0.01528	0.04402	0.05930
1	315	Y_1	5	9	11	9	309.50	0.01582	0.05724	0.07306
		Y_2	2	12	15	3	310.55	0.01708	0.05136	0.06844
1.5	1730	Y_1	6	17	17	9	1709.54	0.01291	0.04503	0.05794
		Y_2	5	14	21	5	1709.80	0.01397	0.04409	0.05806
2	2700	Y_1	4	10	11	7	2696.88	0.01596	0.05657	0.07253
		Y_2	3	10	12	4	2688.65	0.01694	0.05570	0.07264

(b) Coefficients of variation using goal programming method

We use goal programming method for determination of sample size considering the characteristics $Y_j (j = 1, 2)$.

$$Minimize\ d_1 + d_2$$

Subject to

$$\sqrt{\frac{0.000066}{n_{1c}} + \frac{0.000809}{n_{2c}} + \frac{0.001212}{n_{3c}} + \frac{0.000332}{n_{4c}}} - d_1 \leq Z_1^*$$

$$\sqrt{\frac{0.000181}{n_{1c}} + \frac{0.009411}{n_{2c}} + \frac{0.023390}{n_{3c}} + \frac{0.000610}{n_{4c}}} - d_2 \leq Z_2^*$$

$$7n_{1c}^{\delta} + 6n_{2c}^{\delta} + 8n_{3c}^{\delta} + 5n_{4c}^{\delta} + 2.67\sqrt{4.5n_{1c}^{2\delta} + 3.5n_{2c}^{2\delta} + 2.5n_{3c}^{2\delta} + 5n_{4c}^{2\delta}} \leq \acute{C}$$

$$2 \leq n_{1c} \leq 6$$

$$9 \leq n_{2c} \leq 26$$

$$11 \leq n_{3c} \leq 34$$

$$3 \leq n_{4c} \leq 9$$

$$n_{1c}, n_{2c}, n_{3c} \text{ and } n_{4c} \text{ are integers.}$$

Here $\hat{Z}_1$ and $\hat{Z}_2$ are coefficients of variation given in Table 4.6 for different values of $\acute{C}$ and δ using goal programming methods for allocation.

Table 4.6: Coefficients of variation using *GP* method.

δ	$\acute{C}$	n_{1c}	n_{2c}	n_{3c}	n_{4c}	used $\acute{C}$	$\hat{Z}_1$	$\hat{Z}_2$	$\hat{Z}_1 + \hat{Z}_2$
0.5	115	2	12	26	5	114.07	0.01461	0.04355	0.05816
1.0	315	2	9	16	6	314.69	0.01594	0.05196	0.06790
1.5	1730	2	15	20	9	1719.34	0.01358	0.04422	0.05780
2.0	2700	4	9	12	6	2693.58	0.01621	0.05605	0.07226

(c) Coefficients of variation using general method

For estimation of population mean of characteristics $Y_j (j = 1, 2)$, we determine sample sizes n_{1c}, n_{2c}, n_{3c} and n_{4c} using general method.

$$Minimize \begin{pmatrix} W_1\sqrt{\frac{0.000066}{n_{1c}} + \frac{0.000809}{n_{2c}} + \frac{0.001212}{n_{3c}} + \frac{0.000332}{n_{4c}}} \\ +W_2\sqrt{\frac{0.000181}{n_{1c}} + \frac{0.009411}{n_{2c}} + \frac{0.023390}{n_{3c}} + \frac{0.000610}{n_{4c}}} \\ +d_1 + d_2 \end{pmatrix}$$

Subject to

$$\sqrt{\frac{0.000066}{n_{1c}} + \frac{0.000809}{n_{2c}} + \frac{0.001212}{n_{3c}} + \frac{0.000332}{n_{4c}}} - d_1 \leq Z_1^*$$

$$\sqrt{\frac{0.000181}{n_{1c}} + \frac{0.009411}{n_{2c}} + \frac{0.023390}{n_{3c}} + \frac{0.000610}{n_{4c}}} - d_2 \leq Z_2^*$$

$$7n_{1c}^{\delta} + 6n_{2c}^{\delta} + 8n_{3c}^{\delta} + 5n_{4c}^{\delta} + 2.67\sqrt{4.5n_{1c}^{2\delta} + 3.5n_{2c}^{2\delta} + 2.5n_{3c}^{2\delta} + 5n_{4c}^{2\delta}} \leq \acute{C}$$

$$2 \leq n_{1c} \leq 6$$

$$9 \leq n_{2c} \leq 26$$

$$11 \leq n_{3c} \leq 34$$

$$3 \leq n_{4c} \leq 9$$

n_{1c}, n_{2c}, n_{3c} and n_{4c} are integers.

Table 4.7 shows the coefficients of variation obtained by solving above problem of allocation.

Table 4.7: Coefficients of variation using general method.

δ	$\acute{C}$	W_1	W_2	n_{1c}	n_{2c}	n_{3c}	n_{4c}	*used* $\acute{C}$	$\hat{Z}_1$	$\hat{Z}_2$	$\hat{Z}_1 + \hat{Z}_2$
0.5	113	0.2	0.8	2	13	25	5	114.25	0.01450	0.04327	0.05777
		0.4	0.6	2	16	22	5	114.47	0.01432	0.04317	0.05749
		0.6	0.4	2	12	25	6	114.58	0.01429	0.04373	0.05802
		0.8	0.2	2	12	25	6	114.58	0.01429	0.04373	0.05802
1	315	0.2	0.8	3	12	14	4	313.81	0.01609	0.05165	0.06774
		0.4	0.6	2	11	15	5	314.21	0.01593	0.05126	0.06719
		0.6	0.4	3	12	14	4	313.81	0.01609	0.05165	0.06774
		0.8	0.2	3	12	14	4	313.81	0.01609	0.05165	0.06774
1.5	1730	0.2	0.8	4	17	19	6	1716.39	0.01354	0.04395	0.05749
		0.4	0.6	4	15	20	8	1728.49	0.01314	0.04380	0.05694
		0.6	0.4	4	14	21	7	1729.33	0.01340	0.04380	0.05780
		0.8	0.2	4	15	20	8	1728.49	0.01314	0.04380	0.05694
2	2700	0.2	0.8	4	9	12	6	2693.58	0.01621	0.05605	0.07226
		0.4	0.6	4	9	12	6	2693.58	0.01621	0.05605	0.07226
		0.6	0.4	4	9	12	6	2693.58	0.01621	0.05605	0.07226
		0.8	0.2	4	9	12	6	2693.58	0.01621	0.05605	0.07226

4.3.5 Efficiency comparison

The efficiency comparison of goal programming method and general method to the individual optimum method for characteristic $Y_j (j = 1, 2)$ are given in Tables 4.8-4.9 respectively.

Table 4.8: *PRE* of gaol programming method.

δ	$\acute{C}$	Y_1	Y_2	δ	$\acute{C}$	Y_1	Y_2
0.5	0113	106.64	101.96	1.5	1000	100.24	100.45
1.0	0340	107.60	100.80	2.0	1700	100.38	100.54

Table 4.9: PRE of general method.

δ	$\acute{C}$	W_1	W_2	Y_1	Y_2	δ	$\acute{C}$	W_1	W_2	Y_1	Y_2
0.5	120	0.2	0.8	107.36	102.65	1.5	1000	0.2	0.8	100.78	100.99
		0.4	0.6	107.88	103.15			0.4	0.6	101.76	101.97
		0.6	0.4	106.89	102.61			0.6	0.4	101.29	101.50
		0.8	0.2	107.36	102.65			0.8	0.2	101.76	101.97
1.0	340	0.2	0.8	107.85	101.03	2.0	1800	0.2	0.8	100.38	100.54
		0.4	0.6	108.74	101.86			0.4	0.6	100.38	100.54
		0.6	0.4	107.85	101.03			0.6	0.4	100.38	100.54
		0.8	0.2	107.85	101.03			0.8	0.2	100.38	100.54

4.3.6 Results

Table 4.5 shows the results of coefficients of variation of the estimates of mean of study variables Y_1 and Y_2 under general travel cost function. Optimum allocation according to Y_2 gives efficient results as compare to Y_1. The compromise allocation provide efficient results under goal programming method and general method as compete to allocation using individual optimum method for different values of constants δ and $\acute{C}$ as shown in Tables 4.8-4.9 respectively.

4.4 Allocation under probabilistic nonlinear cost function

Consider the following cost function.

$$C_0 + \sum_{h=1}^{L} C_h n_h + \sum_{h=1}^{L} t_h n_h^{\delta} \leq C \tag{4.25}$$

Practically, we observe that per unit measurement cost C_h and traveling cost t_h are not fixed but vary from unit to unit in each stratum and consider as random. We assume that $C_h \sim N\left(\mu_{ch}, \sigma_{ch}^2\right)$ and $t_h \sim N\left(\mu_{th}, \sigma_{th}^2\right)$. Thus nonlinear cost function (4.25) become probabilistic and given as:

$$p\left(C_0+\sum_{h=1}^{L}C_hn_h+\sum_{h=1}^{L}t_hn_h^{\delta}\leq C\right)\geq p_o \tag{4.26}$$

We determine the sample size n_h that minimize $Z_j(j=1,2,...,p)$ under probabilistic cost function (4.26) and condition on sample size n_h discussed in section 4.1. This multi-objective optimization problem can be formulated in chance constraint integer nonlinear mathematical programming problem as:

$$Minimize\ \ (Z_1, Z_2, \ldots, Z_p)$$

Subject to

$$p\left(C_0+\sum_{h=1}^{L}C_hn_h+\sum_{h=1}^{L}t_hn_h^{\delta}\leq C\right)\geq p_o \tag{4.27}$$

$$l_h\leq n_h\leq u_h$$

$$l_h\geq 2$$

$$n_h \text{ are integers.}$$

$$h=1,2,...,L \text{ and} j=1,2,...,p.$$

We transform chance constraint mathematical programming problem (4.27) into deterministic mathematical programming problem as:

Let

$$g_3=C_0+\sum_{h=1}^{L}C_hn_h+\sum_{h=1}^{L}t_hn_h^{\delta}.$$

Applying expectation on both sides, we get

$$E\left(g_3\right)=C_0+\sum_{h=1}^{L}E(C_h)n_h+\sum_{h=1}^{L}E(t_h)n_h^{\delta}.$$

$$E\left(g_3\right)=C_0+\sum_{h=1}^{L}\mu_{ch}n_h+\sum_{h=1}^{L}\mu_{th}n_h^{\delta} \tag{4.28}$$

$$Var\left(g_3\right)=Var\left(C_0+\sum_{h=1}^{L}C_hn_h+\sum_{h=1}^{L}t_hn_h^{\delta}\right).$$

$$Var(g_3)=\sum_{h=1}^{L}\sigma_{ch}^2n_h^2+\sum_{h=1}^{L}\sigma_{th}^2n_h^{2\delta} \tag{4.29}$$

The function g_3 is normally distributed with mean $E(g_3)$ and variance $Var(g_3)$ are defined in (4.28) and (4.29)respectively.That is, $g_3 \sim N\left(E(g_3), Var(g_3)\right)$. The chance constraint is,

$$p\left(g_3 \leq C\right) \geq p_o$$

$$p\left(\frac{g_3 - E(g_3)}{\sqrt{Var(g_3)}} \leq \frac{C - E(g_3)}{\sqrt{Var(g_3)}}\right) \geq p_o$$

where, $p\left(\frac{g_3 - E(g_3)}{\sqrt{Var(g_3)}}\right)$ is standard normal random variable.Thus the probability of realizing g_3 less than or equal to total cost C can be written as:

$$p\left(g_3 \leq C\right) = \phi\left(\frac{C - E(g_3)}{\sqrt{Var(g_3)}}\right) = \phi(z) \tag{4.30}$$

where, $\phi\left(z\right)$ represent the cumulative density function of the standard normal random variable calculated at z. Let A_α represent standard normal variable at which $\phi\left(A_\alpha\right) = p_o$. Then the constraint (4.26) can be written as

$$\phi\left(\frac{C - E(g_3)}{\sqrt{Var(g_3)}}\right) \geq \phi\left(A_\alpha\right)$$

This inequality will be satisfied only if

$$\frac{C - E(g_3)}{\sqrt{Var(g_3)}} \geq A_\alpha$$

or

$$E\left(g_3\right) + \left(A_\alpha\right)\sqrt{Var\left(g_3\right)} \leq C \tag{4.31}$$

Substituting (4.28) and (4.29) in (4.31), we get deterministic constraint equivalent to probabilistic constraint (4.26).

$$\sum_{h=1}^{L} \mu_{ch} n_h + \sum_{h=1}^{L} \mu_{th} n_h^{\delta} + \left(A_\alpha\right)\sqrt{\sum_{h=1}^{L} \sigma_{ch}^2 n_h^2 + \sum_{h=1}^{L} \sigma_{th}^2 n_h^{2\delta}} \leq C - C_o = \acute{C} \tag{4.32}$$

If parameters μ_{ch} , μ_{th}, σ^2_{ch} and σ^2_{th} are unknown, we replace by their estimates. Let $\hat{\mu_{ch}}$, $\hat{\mu_{th}}$, $\hat{\sigma^2_{ch}}$ and $\hat{\sigma^2_{th}}$ are estimates. The chance constraint multi-objective nonlinear integer mathematical programming problem (4.27) can be written in deterministic multi-objective nonlinear mathematical programming problem as:

$$Minimize\ \ (Z_1, Z_2, \ldots, Z_p)$$

Subject to

$$\sum_{h=1}^{L} \mu_{ch} n_h + \sum_{h=1}^{L} \mu_{th} n_h^{\delta} + (A_\alpha)\sqrt{\sum_{h=1}^{L} \sigma^2_{ch} n_h^2 + \sum_{h=1}^{L} \sigma^2_{th} n_h^{2\delta}} \leq \acute{C} \tag{4.33}$$

$$l_h \leq n_h \leq u_h$$

$$l_h \geq 2$$

$$n_h \text{ are integers}$$

$$h = 1, 2, ..., L.$$

This multi-objective allocation problem is solved using multi-objective optimization methods.

4.4.1 Allocation using individual optimum method

Let Z_j^* be a optimum value of Z_j obtained by solving following nonlinear integer mathematical programming problem.

$$Minimize\ Z_j$$

Subject to

$$\sum_{h=1}^{L} \mu_{ch} n_h + \sum_{h=1}^{L} \mu_{th} n_h^{\delta} + (A_\alpha)\sqrt{\sum_{h=1}^{L} \sigma^2_{ch} n_h^2 + \sum_{h=1}^{L} \sigma^2_{th} n_h^{2\delta}} \leq \acute{C} \tag{4.34}$$

$$l_h \leq n_j h \leq u_h$$

$$l_h \geq 2$$

$$n_h \text{ are integers}$$

$$h = 1, 2, ..., L \text{ and } j = 1, 2, ..., p.$$

4.4.2 Allocation using goal programming method

The multi-objective allocation problem (4.31) may be solved with goal programming (GP) as:

$$Minimize \ \sum_{j=1}^{p} d_j$$

Subject to

$$Z_j - d_j \leq Z_j^* \tag{4.35}$$

$$\sum_{h=1}^{L} \mu_{ch} n_h + \sum_{h=1}^{L} \mu_{th} n_h^{\delta} + (A_\alpha)\sqrt{\sum_{h=1}^{L} \sigma_{ch}^2 n_h^2 + \sum_{h=1}^{L} \sigma_{th}^2 n_h^{2\delta}} \leq \acute{C}$$

$$l_h \leq n_{hc} \leq u_h$$

$$l_h \geq 2$$

$$n_{hc} \text{ are integers.}$$

$$h = 1, 2, ..., L \text{ and } j = 1, 2, ..., p,$$

where, $d_j (j = 1, 2, ..., p)$ are deviation variables.

4.4.3 Allocation using general method

The multi-objective optimization problem (4.33) can be solved with general method as:

$$Minimize \ \sum_{k=1}^{p} W_k Z_k + \sum_{j=1}^{p} d_j$$

Subject to

$$Z_j - d_j \leq Z_j^* \tag{4.36}$$

$$\sum_{h=1}^{L} \mu_{ch} n_h + \sum_{h=1}^{L} \mu_{th} n_h^{\delta} + (A_\alpha)\sqrt{\sum_{h=1}^{L} \sigma_{ch}^2 n_h^2 + \sum_{h=1}^{L} \sigma_{th}^2 n_h^{2\delta}} \leq \acute{C}$$

$$l_h \leq n_{hc} \leq u_h$$

$$\sum_{k=1}^{p} W_k = 1$$

$$l_h \geq 2$$

$$n_{hc} \text{ are integers}$$
$$h = 1, 2, ..., L \text{ and } k = j = 1, 2, ..., p,$$

where, $W_k(k = 1, 2, ..., p)$ are relative weights which indicates the importance of characteristics Y_j and $d_j(j = 1, 2, ..., p)$ are deviation variables.

4.4.4 Numerical example

Data Source [Khan et al.(2010)]. We assume that

$$E(C_1) = 6, E(C_2) = 3.5, E(C_3) = 3, E(C_4) = 5.5$$

$$V(C_1) = 2.5, V(C_2) = 1.5, V(C_3) = 2, V(C_4) = 1.75$$

$$E(t_1) = 4.5, E(t_2) = 2.5, E(t_3) = 3, E(t_4) = 4$$

$$V(t_1) = 2.5, V(t_2) = 1.25, V(t_3) = 1.5, V(t_4) = 2.30$$

$$Minimize \begin{pmatrix} Z_1 = \sqrt{\frac{0.000066}{n_1} + \frac{0.000809}{n_2} + \frac{0.001212}{n_3} + \frac{0.000332}{n_4}} \\ Z_2 = \sqrt{\frac{0.000181}{n_1} + \frac{0.009411}{n_2} + \frac{0.023390}{n_3} + \frac{0.000610}{n_4}} \end{pmatrix}$$

Subject to

$$6n_1 + 3.5n_2 + 3n_3 + 5.5n_4 + 4.5n_1^\delta + 2.5n_2^\delta + 3n_3^\delta + 4n_4^\delta$$
$$+2.67\sqrt{2.5n_1^2 + 1.5n_2^2 + 2n_3^2 + 1.75n_4^2 + 2.5n_1^{2\delta} + 1.25n_2^{2\delta} + 1.5n_3^{2\delta} + 2.3n_4^{2\delta}} \leq \acute{C}$$

$$2 \leq n_1 \leq 6$$

$$9 \leq n_2 \leq 26$$

$$11 \leq n_3 \leq 34$$

$$3 \leq n_4 \leq 9$$

$$n_1, n_2, n_3 \text{ and } n_4 \text{ are integers.}$$

(a) Coefficients of variation using individual optimum method

Individual optimum allocation taking into account the characteristic Y_1 :

$$Minimize\ Z_1 = \sqrt{\frac{0.000066}{n_{11}} + \frac{0.000809}{n_{12}} + \frac{0.001212}{n_{13}} + \frac{0.000332}{n_{14}}}$$

Subject to

$$6n_{11} + 3.5n_{12} + 3n_{13} + 5.5n_{14} + 4.5n_{11}^{\delta} + 2.5n_{12}^{\delta} + 3n_{13}^{\delta} + 4n_{14}^{\delta}$$

$$+2.67\sqrt{2.5n_{11}^2 + 1.5n_{12}^2 + 2n_{13}^2 + 1.75n_{14}^2 + 2.5n_{11}^{2\delta} + 1.25n_{12}^{2\delta} + 1.5n_{13}^{2\delta} + 2.3n_{14}^{2\delta}} \leq \acute{C}$$

$$2 \leq n_{11} \leq 6$$

$$9 \leq n_{12} \leq 26$$

$$11 \leq n_{13} \leq 34$$

$$3 \leq n_{14} \leq 9$$

n_{11}, n_{12}, n_{13} and n_{14} are integers.

Individual optimum allocation taking into account the characteristic Y_2 :

$$Minimize\ Z_2 = \sqrt{\frac{0.000181}{n_{21}} + \frac{0.009411}{n_{22}} + \frac{0.023390}{n_{23}} + \frac{0.000610}{n_{24}}}$$

Subject to

$$6n_{21} + 3.5n_{22} + 3n_{23} + 5.5n_{24} + 4.5n_{21}^{\delta} + 2.5n_{22}^{\delta} + 3n_{23}^{\delta} + 4n_{24}^{\delta}$$

$$+2.67\sqrt{2.5n_{21}^2 + 1.5n_{22}^2 + 2n_{23}^2 + 1.75n_{24}^2 + 2.5n_{21}^{2\delta} + 1.25n_{22}^{2\delta} + 1.5n_{23}^{2\delta} + 2.3n_{24}^{2\delta}} \leq \acute{C}$$

$$2 \leq n_{21} \leq 6$$

$$9 \leq n_{22} \leq 26$$

$$11 \leq n_{23} \leq 34$$

$$3 \leq n_{24} \leq 9$$

n_{21}, n_{22}, n_{23} and n_{24} are integers.

Here Z_1^* and Z_2^* are coefficients of variation using individual optimum method for different value of δ and $\acute{C}$ given in Table 4.10.

Table 4.10: Coefficients of variations using individual allocation.

δ	$\acute{C}$	*Allocation*	n_1	n_2	n_3	n_4	*used* $\acute{C}$	Z_1^*	Z_2^*	$Z_1^*+Z_2^*$
0.5	300	Y_1	3	17	15	8	295.59	0.01385	0.04743	0.06128
		Y_2	2	14	25	3	298.16	0.01580	0.04361	0.05941
1	450	Y_1	4	14	18	9	441.30	0.01336	0.04565	0.05901
		Y_2	2	17	25	3	448.62	0.01548	0.04223	0.05771
1.5	675	Y_1	4	10	12	8	672.08	0.01549	0.05487	0.07036
		Y_2	2	12	15	3	665.97	0.01708	0.05136	0.06864
2	1900	Y_1	6	10	12	8	1887.92	0.01531	0.05474	0.07005
		Y_2	2	12	14	3	1890.19	0.01725	0.05242	0.06969

(b) Coefficients of variation using goal programming

We use propose goal programming method to find the compromise allocation for the characteristics $Y_j(j = 1, 2)$.

$$Minimize\ d_1 + d_2$$

Subject to

$$\sqrt{\frac{0.000066}{n_{1c}} + \frac{0.000809}{n_{2c}} + \frac{0.001212}{n_{3c}} + \frac{0.000332}{n_{4c}}} - d_1 \leq Z_1^*$$

$$\sqrt{\frac{0.000181}{n_{1c}} + \frac{0.009411}{n_{2c}} + \frac{0.023390}{n_{3c}} + \frac{0.000610}{n_{4c}}} - d_2 \leq Z_2^*$$

$$6n_{1c} + 3.5n_{2c} + 3n_{3c} + 5.5n_{4c} + 4.5n_{1c}^{\delta} + 2.5n_{2c}^{\delta} + 3n_{3c}^{\delta} + 4n_{4c}^{\delta}$$

$$+2.67\sqrt{2.5n_{1c}^2 + 1.5n_{2c}^2 + 2n_{3c}^2 + 1.75n_{4c}^2 + 2.5n_{1c}^{2\delta} + 1.25n_{2c}^{2\delta} + 1.5n_{3c}^{2\delta} + 2.3n_{4c}^{2\delta}} \leq \acute{C}$$

$$2 \leq n_{1c} \leq 6$$

$$9 \leq n_{2c} \leq 26$$

$$11 \leq n_{3c} \leq 34$$

$$3 \leq n_{4c} \leq 9$$

n_{1c}, n_{2c}, n_{3c} and n_{4c} are integers.

$\hat{Z}_1$ and $\hat{Z}_2$ are coefficients of variations for different values of $\acute{C}$ and δ given in Table 4.11.

Table 4.11: Coefficients of variation using *GP* method.

δ	$\acute{C}$	n_{1c}	n_{2c}	n_{3c}	n_{4c}	*used* $\acute{C}$	$\hat{Z}_1$	$\hat{Z}_2$	$\hat{Z}_1 + \hat{Z}_2$
0.5	300	2	14	25	3	299.07	0.01448	0.04361	0.05809
1.0	450	3	14	23	6	446.15	0.01370	0.04302	0.05672
1.5	675	2	12	14	5	669.29	0.01691	0.05164	0.06755
2.0	1900	4	12	13	5	1885.29	0.01560	0.05244	0.06804

(c) coefficients of variation using general method

The compromise allocation for $Y_j (j = 1, 2)$, we use general method.

$$Minimize \begin{pmatrix} W_1\sqrt{\frac{0.000066}{n_{1c}} + \frac{0.000809}{n_{2c}} + \frac{0.001212}{n_{3c}} + \frac{0.000332}{n_{4c}}} \\ +W_2\sqrt{\frac{0.000181}{n_{1c}} + \frac{0.009411}{n_{2c}} + \frac{0.023390}{n_{3c}} + \frac{0.000610}{n_{4c}}} \\ +d_1 + d_2 \end{pmatrix}$$

Subject to

$$\sqrt{\frac{0.000066}{n_{1c}} + \frac{0.000809}{n_{2c}} + \frac{0.001212}{n_{3c}} + \frac{0.000332}{n_{4c}}} - d_1 \leq Z_1^*$$

$$\sqrt{\frac{0.000181}{n_{1c}} + \frac{0.009411}{n_{2c}} + \frac{0.023390}{n_{3c}} + \frac{0.000610}{n_{4c}}} - d_2 \leq Z_2^*$$

$$6n_{1c} + 3.5n_{2c} + 3n_{3c} + 5.5n_{4c} + 4.5n_{1c}^{\delta} + 2.5n_{2c}^{\delta} + 3n_{3c}^{\delta} + 4n_{4c}^{\delta}$$

$$+2.67\sqrt{2.5n_{1c}^2 + 1.5n_{2c}^2 + 2n_{3c}^2 + 1.75n_{4c}^2 + 2.5n_{1c}^{2\delta} + 1.25n_{2c}^{2\delta} + 1.5n_{3c}^{2\delta} + 2.3n_{4c}^{2\delta}} \leq \acute{C}$$

$$2 \leq n_{1c} \leq 6$$

$$9 \leq n_{2c} \leq 26$$

$$11 \leq n_{3c} \leq 34$$

$$3 \leq n_{4c} \leq 9$$

n_{1c}, n_{2c}, n_{3c} and n_{4c} are integers.

Here $\hat{Z}_1$ and $\hat{Z}_2$ are coefficients of variation obtained for different values of δ and $\acute{C}$ solving above problem given in Table 4.12.

Table 4.12: Coefficients of variation using general method.

δ	$\acute{C}$	W_1	W_2	n_{1c}	n_{2c}	n_{3c}	n_{4c}	$used\ \acute{C}$	$\hat{Z}_1$	$\hat{Z}_2$	$\hat{Z}_1 + \hat{Z}_2$
0.5	300	0.2	0.8	2	15	22	5	297.94	0.01443	0.04362	0.05805
		0.4	0.6	2	15	21	6	298.61	0.01414	0.04397	0.05811
		0.6	0.4	3	15	22	4	298.81	0.01500	0.04362	0.05862
		0.8	0.2	3	16	19	6	299.49	0.01385	0.04451	0.05836
1	450	0.2	0.8	3	14	24	5	445.77	0.01402	0.04277	0.05679
		0.4	0.6	2	17	22	6	449.56	0.01382	0.04253	0.05635
		0.6	0.4	3	18	20	6	449.51	0.01352	0.04306	0.05658
		0.8	0.2	2	17	22	6	449.56	0.01382	0.04253	0.05635
1.5	675	0.2	0.8	3	12	14	4	666.21	0.01609	0.05165	0.06774
		0.4	0.6	2	12	14	5	669.29	0.01591	0.05164	0.06755
		0.6	0.4	3	12	14	4	666.21	0.01609	0.05165	0.06774
		0.8	0.2	4	10	14	5	1859.04	0.01582	0.05271	0.06853
2	1900	0.2	0.8	4	12	13	5	1885.29	0.01560	0.05244	0.06844
		0.4	0.6	2	11	14	5	1874.33	0.01610	0.05233	0.06843
		0.6	0.4	2	12	13	6	1873.47	0.01577	0.05268	0.06845
		0.8	0.2	3	11	13	7	1873.66	0.01536	0.05293	0.06827

4.4.5 Efficiency comparison

The percentage relative efficiency PRE of goal programming method and general method to individual optimum method are given in Tables 4.13-4.14 respectively.

Table 4.13: PRE of goal programming method.

δ	$\acute{C}$	Y_1	Y_2	δ	$\acute{C}$	Y_1	Y_2
0.5	0300	105.49	102.27	1.5	0675	104.16	101.32
1.0	0450	104.04	101.75	2.0	1900	102.95	102.43

Table 4.14: PRE of general method.

δ	$\acute{C}$	W_1	W_2	Y_1	Y_2	δ	$\acute{C}$	W_1	W_2	Y_1	Y_2
0.5	300	0.2	0.8	105.56	102.34	1.5	675	0.2	0.8	104.41	101.03
		0.4	0.6	105.46	102.24			0.4	0.6	104.71	101.32
		0.6	0.4	104.54.89	101.35			0.6	0.4	104.41	101.03
		0.8	0.2	105.00	101.80			0.8	0.2	104.41	101.03
1.0	450	0.2	0.8	103.91	101.62	2.0	1900	0.2	0.8	102.22	101.69
		0.4	0.6	104.72	102.41			0.4	0.6	102.37	101.84
		0.6	0.4	104.29	100.00			0.6	0.4	102.34	101.81
		0.8	0.2	104.72	102.41			0.8	0.2	102.58	102.05

4.4.6 Results

The efficiency of individual optimum method, GP method and general method for sample allocation to estimate mean of the variables Y_1 and Y_2 are displayed by the Tables(4.10-4.14) respectively. The individual optimum allocation subject to Y_2 give more efficient estimates than Y_1. The efficiency of compromise allocation using GP method and general method is greater than individual allocation method according to Y_1 and Y_2 as given in Tables 4.13-4.14

4.5 Allocation probabilistic logarithmic cost function

Consider the following cost function.

$$p\left(C_0 + \sum_{h=1}^{L} t_h n_h^{\delta} \leq C\right) \geq p_o \tag{4.37}$$

We determine the sample size n_h that minimize the coefficients of variation $Z_j (j = 1, 2, ..., p)$ under the probabilistic cost function (4.37) and additional constraint discussed in section 4.2. This multi-objective optimization problem can be formulated in chance constraint integer nonlinear mathematical problem as:

$$Minimize \ (Z_1, Z_2, \ldots, Z_p)$$

Subject to

$$p\left(C_0+\sum_{h=1}^{L}t_h\log n_h^{\delta}\leq C\right)\geq p_o \tag{4.38}$$

$$l_h\leq n_h\leq u_h$$

$$l_h\geq 2$$

$$n_h \text{ are integers.}$$

$$h=1,2,...,L \text{ and } j=1,2,...,p,$$

Now we transform multi-objective chance constraint mathematical programming problem (4.38) into multi-objective deterministic integer nonlinear mathematical programming problem as:

Let

$$g_4=C_0+\sum_{h=1}^{L}t_h\log n_h^{\delta}.$$

Applying expectation on both sides, we get

$$E\left(g_4\right)=C_0+\sum_{h=1}^{L}E(t_h)\log n_h^{\delta}.$$

$$E\left(g_4\right)=C_0+\sum_{h=1}^{L}\mu_{th}\log n_h^{\delta} \tag{4.39}$$

$$Var\left(g_4\right)=Var\left(C_0+\sum_{h=1}^{L}t_h\log n_h^{\delta}\right).$$

$$Var(g_4)=\sum_{h=1}^{L}\sigma_{th}^2\left(\log n_h^{\delta}\right)^2 \tag{4.40}$$

The function g_4 is normally distributed with mean $E(g_4)$ and variance $Var(g_4)$ are defined in (4.39) and (4.40) respectively.That is, $g_4\sim N\left(E(g_4),Var(g_4)\right)$. The chance constraint (4.38) can be written as,

$$p\left(g_4\leq C\right)\geq p_o$$

$$p\left(\frac{g_4-E(g_4)}{\sqrt{Var(g_4)}}\leq\frac{C-E(g_4)}{\sqrt{Var(g_4)}}\right)\geq p_o$$

where, $p\left(\frac{g_4-E(g_4)}{\sqrt{Var(g_4)}}\right)$ is standard normal random variable. Thus the probability of realizing g_4 less than or equal to total cost C can be written as:

$$p\left(g_4 \leq C\right) = \phi\left(\frac{C - E(g_4)}{\sqrt{Var(g_4)}}\right) = \phi(z) \tag{4.41}$$

where, $\phi(z)$ represents the cumulative density function of the normal random variable calculated for z. Let A_α represents standard normal variable at which $\phi(A_\alpha) = p_o$. Then the constraint (4.37) can be written as:

$$\phi\left(\frac{C - E(g_4)}{\sqrt{Var(g_4)}}\right) \geq \phi\left(A_\alpha\right).$$

This inequality will be satisfied only if

$$\frac{C - E(g_4)}{\sqrt{Var(g_4)}} \geq A_\alpha.$$

or

$$E(g_4) + (A_\alpha)\sqrt{Var(g_4)} \leq C \tag{4.42}$$

Substituting (4.39) and (4.40) in (4.42), we get deterministic constraint equivalent to probabilistic constraint (4.37).

$$\sum_{h=1}^{L} \mu_{th} \log n_h^\delta + (A_\alpha)\sqrt{\sum_{h=1}^{L} \sigma_{th}^2 \left(\log n_h^\delta\right)^2} \leq C - C_o = \acute{C} \tag{4.43}$$

If parameters μ_{th} and σ^2_{th} are unknown, we replace them by their estimate.Let $\hat{\mu_{th}}$ and $\hat{\sigma^2_{th}}$ are estimates of parameters μ_{th} and σ^2_{th} respectively. The multi-objective chance constraint integer nonlinear mathematical programming problem (4.38) can be written in multi-objective deterministic integer nonlinear mathematical programming problem as:

$$Minimize \;\; (Z_1, Z_2, Z_3 \ldots, Z_p)$$

Subject to

$$\sum_{h=1}^{L} \mu_{th} \log\left(n_h^\delta\right) + (A_\alpha)\sqrt{\sum_{h=1}^{L} \sigma_{th}^2 \left(\log n_h^\delta\right)^2} \leq \acute{C} \tag{4.44}$$

$$l_h \leq n_h \leq u_h$$

$$l_h \geq 2$$

$$n_h \text{ are integers}$$

$$h = 1, 2, ..., L.$$

The multi-objective optimization problem (4.44) is solved by following optimization methods.

4.5.1 Allocation using individual optimum method

Let Z_j^* be a optimum value of Z_j obtained by solving the following nonlinear integer mathematical programming problem.

$$Minimize\ Z_j$$

Subject to

$$\sum_{h=1}^{L} \mu_{th} \log\left(n_{jh}^{\delta}\right) + (A_\alpha)\sqrt{\sum_{h=1}^{L} \sigma_{th}^2 \left(\log n_{jh}^{\delta}\right)^2} \leq \acute{C} \tag{4.45}$$

$$l_h \leq n_j h \leq u_h$$

$$l_h \geq 2$$

$$n_h \text{ are integers}$$

$$h = 1, 2, ..., L \text{ and } j = 1, 2, ..., p.$$

4.5.2 Allocation using goal programming method

The multi-objective optimization problem (4.44) can be solved with goal programming (GP) method as:

$$Minimize\ \sum_{j=1}^{p} d_j$$

Subject to

$$Z_j - d_j \leq Z_j^* \tag{4.46}$$

$$\sum_{h=1}^{L} \mu_{th} \log\left(n_{hc}^{\delta}\right) + (A_\alpha)\sqrt{\sum_{h=1}^{L} \sigma_{th}^2 \left(\log n_{hc}^{\delta}\right)^2} \leq \acute{C}$$

$$l_h \leq n_{hc} \leq u_h$$

$$l_h \geq 2$$

$$n_{hc} \text{ are integers}$$

$$h = 1, 2, ..., L \text{ and } j = 1, 2, ..., p,$$

where, $d_j (j = 1, 2, ..., p)$ are deviation variables.

4.5.3 Allocation using general method

The multi-objective allocation problem (4.44) may be solved with general method as:

$$Minimize \ \sum_{k=1}^{p} W_k Z_k + \sum_{j=1}^{p} d_j$$

Subject to

$$Z_j - d_j \leq Z_j^* \tag{4.47}$$

$$\sum_{h=1}^{L} \mu_{th} \log\left(n_{hc}^{\delta}\right) + (A_\alpha)\sqrt{\sum_{h=1}^{L} \sigma_{th}^2 \left(\log n_{hc}^{\delta}\right)^2} \leq \acute{C}$$

$$l_h \leq n_{hc} \leq u_h$$

$$\sum_{k=1}^{p} W_k = 1$$

$$l_h \geq 2$$

$$n_{hc} \text{ are integers}$$

$$h = 1, 2, ..., L \text{ and } j = k = 1, 2, ..., p,$$

where, $W_k (k = 1, 2, ..., p)$ are relative weights indicate the importance for each characteristics and $d_j (j = 1, 2, ..., p)$ are deviation variables.

4.5.4 Numerical example

Data Source [Khan et al.(2010)]. We assume that

$$E(t_1) = 10, E(t_2) = 8, E(t_3) = 9, E(t_4) = 13, V(t_1) = 6, V(t_2) = 4.5, V(t_3) = 5.5, V(t_4) = 5$$

$$Minimize \begin{pmatrix} Z_1 = \sqrt{\frac{0.000066}{n_1} + \frac{0.000809}{n_2} + \frac{0.001212}{n_3} + \frac{0.000332}{n_4}} \\ Z_2 = \sqrt{\frac{0.000181}{n_1} + \frac{0.009411}{n_2} + \frac{0.023390}{n_3} + \frac{0.000610}{n_4}} \end{pmatrix}$$

Subject to

$$10 \log n_1^\delta + 8 \log n_2^\delta + 9 \log n_3^\delta + 13 \log n_4^\delta$$

$$+2.67\sqrt{6\left(\log n_1^\delta\right)^2 + 4.5\left(\log n_2^\delta\right)^2 + 5.5\left(\log n_3^\delta\right)^2 + 5\left(\log n_4^\delta\right)^2} \leq \acute{C}$$

$$2 \leq n_1 \leq 6$$

$$9 \leq n_2 \leq 26$$

$$11 \leq n_3 \leq 34$$

$$3 \leq n_4 \leq 9$$

n_1, n_2, n_3 and n_4 are integers.

(a) Coefficients of variation using individual allocation

Individual optimum allocation taking into consideration characteristic Y_1 :

$$Minimize\ Z_1 = \sqrt{\frac{0.000066}{n_{11}} + \frac{0.000809}{n_{12}} + \frac{0.001212}{n_{13}} + \frac{0.000332}{n_{14}}}$$

Subject to

$$10 \log n_{11}^\delta + 8 \log n_{12}^\delta + 9 \log n_{13}^\delta + 13 \log n_{14}^\delta$$

$$+2.67\sqrt{6\left(\log n_{11}^\delta\right)^2 + 4.5\left(\log n_{12}^\delta\right)^2 + 5.5\left(\log n_{13}^\delta\right)^2 + 5\left(\log n_{11}^\delta\right)^2} \leq \acute{C}$$

$$2 \leq n_{11} \leq 6$$

$$9 \leq n_{12} \leq 26$$

$$11 \leq n_{13} \leq 34$$

$$3 \leq n_{14} \leq 9$$

n_{11}, n_{12}, n_{13} and n_{14} are integers.

Individual optimum allocation taking characteristic Y_2 :

$$Minimize\ Z_2 = \sqrt{\frac{0.000181}{n_{21}} + \frac{0.009411}{n_{22}} + \frac{0.023390}{n_{23}} + \frac{0.000610}{n_{24}}}$$

Subject to

$$10\log n_{21}^{\delta} + 8\log n_{22}^{\delta} + 9\log n_{23}^{\delta} + 13\log n_{24}^{\delta}$$

$$+2.67\sqrt{6\left(\log n_{21}^{\delta}\right)^2 + 4.5\left(\log n_{22}^{\delta}\right)^2 + 5.5\left(\log n_{23}^{\delta}\right)^2 + 5\left(\log n_{24}^{\delta}\right)^2} \le \acute{C}$$

$$2 \le n_{21} \le 6$$

$$9 \le n_{22} \le 26$$

$$11 \le n_{23} \le 34$$

$$3 \le n_{24} \le 9$$

n_{21}, n_{22}, n_{23} and n_{24} are integers.

Here Z_1^* and Z_2^* are coefficients of variation using individual optimum allocation for different value of δ and $\acute{C}$ given in the Table 4.15.

Table 4.15: Coefficients of variations using individual allocation.

δ	$\acute{C}$	*Allocation*	n_1	n_2	n_3	n_4	*used* $\acute{C}$	Z_1^*	Z_2^*	$Z_1^*+Z_2^*$
0.5	45	Y_1	2	9	15	4	44.949	0.01693	0.05337	0.07030
		Y_2	2	11	15	3	43.956	0.01726	0.05205	0.06931
	47	Y_1	2	9	16	4	45.420	0.01678	0.05245	0.06923
		Y_2	2	11	16	3	44.390	0.01712	0.05110	0.06822
	49	Y_1	2	9	16	5	47.121	0.01628	0.05215	0.06843
		Y_2	2	11	17	3	44.809	0.01699	0.05205	0.06724
	51	Y_1	2	15	14	5	49.097	0.01549	0.05011	0.06560
		Y_2	2	21	20	3	49.615	0.01558	0.04372	0.05930

(b) Coefficients of variation using goal programming method

The goal programming method is used for determination of sample size considering the characteristics $Y_j (j = 1, 2)$.

$$Minimize\ d_1 + d_2$$

Subject to

$$\sqrt{\frac{0.000066}{n_{1c}} + \frac{0.000809}{n_{2c}} + \frac{0.001212}{n_{3c}} + \frac{0.000332}{n_{4c}}} - d_1 \leq Z_1^*$$

$$\sqrt{\frac{0.000181}{n_{1c}} + \frac{0.009411}{n_{2c}} + \frac{0.023390}{n_{3c}} + \frac{0.000610}{n_{4c}}} - d_2 \leq Z_2^*$$

$$10 \log n_{1c}^{\delta} + 8 \log n_{2c}^{\delta} + 9 \log n_{3c}^{\delta} + 13 \log n_{4c}^{\delta}$$

$$+2.67\sqrt{6\left(\log n_{1c}^{\delta}\right)^2 + 4.5\left(\log n_{2c}^{\delta}\right)^2 + 5.5\left(\log n_{3c}^{\delta}\right)^2 + 5\left(\log n_{4c}^{\delta}\right)^2} \leq \acute{C}$$

$$2 \leq n_{1c} \leq 6$$

$$9 \leq n_{2c} \leq 26$$

$$11 \leq n_{3c} \leq 34$$

$$3 \leq n_{4c} \leq 9$$

n_{1c}, n_{2c}, n_{3c} and n_{4c} are integers.

Here $\hat{Z}_1$ and $\hat{Z}_2$ are coefficients of variation for different values of δ and $\acute{C}$ obtained by solving above problem given in Table 4.16.

Table 4.16: Coefficients of variation using *GP* method.

δ	$\acute{C}$	n_{1c}	n_{2c}	n_{3c}	n_{4c}	*used* $\acute{C}$	$\hat{Z}_1$	$\hat{Z}_2$	$\hat{Z}_1 + \hat{Z}_2$
0.5	45	2	9	17	3	43.690	0.01746	0.05211	0.06957
	47	2	23	12	3	46.790	0.01670	0.05150	0.06820
	49	2	23	13	3	47.310	0.01649	0.05002	0.06651
	51	2	26	19	3	50.528	0.01545	0.04344	0.05889

(c) Coefficients of variation using goal general method

The compromise allocation for the characteristics $Y_j(j = 1, 2)$ is obtained by general method.

$$Minimize \begin{pmatrix} W_1\sqrt{\frac{0.000066}{n_{1c}} + \frac{0.000809}{n_{2c}} + \frac{0.001212}{n_{3c}} + \frac{0.000332}{n_{4c}}} \\ +W_2\sqrt{\frac{0.000181}{n_{1c}} + \frac{0.009411}{n_{2c}} + \frac{0.023390}{n_{3c}} + \frac{0.000610}{n_{4c}}} \\ +d_1 + d_2 \end{pmatrix}$$

Subject to

$$\sqrt{\frac{0.000066}{n_{1c}} + \frac{0.000809}{n_{2c}} + \frac{0.001212}{n_{3c}} + \frac{0.000332}{n_{4c}}} - d_1 \leq Z_1^*$$

$$\sqrt{\frac{0.000181}{n_{1c}} + \frac{0.009411}{n_{2c}} + \frac{0.023390}{n_{3c}} + \frac{0.000610}{n_{4c}}} - d_2 \leq Z_2^*$$

$$10 \log n_{1c}^{\delta} + 8 \log n_{2c}^{\delta} + 9 \log n_{3c}^{\delta} + 13 \log n_{4c}^{\delta}$$

$$+2.67\sqrt{6\left(\log n_{1c}^{\delta}\right)^2 + 4.5\left(\log n_{2c}^{\delta}\right)^2 + 5.5\left(\log n_{3c}^{\delta}\right)^2 + 5\left(\log n_{4c}^{\delta}\right)^2} \leq \acute{C}$$

$$2 \leq n_{1c} \leq 6$$

$$9 \leq n_{2c} \leq 26$$

$$11 \leq n_{3c} \leq 34$$

$$3 \leq n_{4c} \leq 9$$

$$n_{1c}, n_{2c}, n_{3c} \text{ and } n_{4c} \text{ are integers.}$$

Here Z_1^* and Z_2^* are coefficients of variation using general method for different value of δ and $\acute{C}$ given in Table 4.17.

Table 4.17: Coefficients of variation using general method.

δ	$\acute{C}$	W_1	W_2	n_{1c}	n_{2c}	n_{3c}	n_{4c}	*used* $\acute{C}$	$\hat{Z}_1$	$\hat{Z}_2$	$\hat{Z}_1 + \hat{Z}_2$
0.5	45	0.2	0.8	2	9	19	3	44.461	0.01724	0.05070	0.06794
		0.4	0.6	2	11	17	3	44.809	0.01699	0.05025	0.06724
		0.6	0.4	2	11	17	3	44.809	0.01699	0.05025	0.06724
		0.8	0.2	2	11	17	3	44.809	0.01699	0.05025	0.06724
	47	0.2	0.8	2	9	27	3	46.920	0.01669	0.04697	0.06366
		0.4	0.6	2	10	24	3	46.67	0.01659	0.04701	0.06360
		0.6	0.4	2	9	27	3	46.920	0.01669	0.04697	0.06366
		0.8	0.2	2	10	24	3	46.67	0.01659	0.04701	0.06360
	49	0.8	0.2	2	16	20	3	48.043	0.01596	0.04529	0.06125
		0.4	0.6	2	15	21	3	48.007	0.01598	0.04511	0.06109
		0.6	0.4	2	12	21	3	46.748	0.01640	0.04682	0.06322
		0.8	0.2	2	12	19	4	48.187	0.01572	0.04752	0.06324
	51	0.2	0.8	2	17	29	3	50.928	0.01527	0.04067	0.05594
		0.4	0.6	2	11	14	5	1874.33	0.01610	0.05233	0.06843
		0.6	0.4	2	14	24	4	50.643	0.01498	0.04347	0.05845
		0.8	0.2	2	16	22	4	50.928	0.01489	0.04352	0.05850

4.5.5 Efficiency comparison

The efficiency comparison of goal programming method and general method to individual optimum method is given in the Table 4.18-4.19 respectively.

Table 4.18: *PRE* of goal programming method.

δ	$\acute{C}$	Y_1	Y_2	δ	$\acute{C}$	Y_1	Y_2
0.5	45	101.05	099.63	0.5	49	102.89	101.10
	47	101.51	100.03		51	111.39	100.70

Table 4.19: PRE of general method.

δ	$\acute{C}$	W_1	W_2	Y_1	Y_2	δ	$\acute{C}$	W_1	W_2	Y_1	Y_2
0.5	45	0.2	0.8	103.47	102.02	0.5	49	0.8	0.2	111.72	109.80
		0.4	0.6	104.55	103.08			0.4	0.6	111.02	110.67
		0.6	0.4	104.55	103.08			0.6	0.4	108.24	106.35
		0.8	0.2	104.55	103.08			0.8	0.2	108.20	106.32
	47	0.2	0.8	108.75	107.16		51	0.2	0.8	117.27	106.01
		0.4	0.6	108.85	107.26			0.4	0.6	117.27	106.01
		0.6	0.4	108.75	107.16			0.6	0.4	112.23	101.45
		0.8	0.2	108.85	107.26			0.8	0.2	112.14	101.37

4.5.6 Results

Table 4.15 indicates the results of individual optimum methods under estimated travel cost function for constant $\delta = 0.5$ and different values of $\acute{C}$. The optimum allocation according to characteristic Y_2 provides smaller total coefficients of variation as compared to characteristic Y_2 to estimate population means of $Y_j (j = 1, 2)$. Tables 4.18-4.19 show that compromise allocation using GP method and general method produce more precise estimates than individual optimum allocation according to variables Y_1 and Y_2 respectively.

Chapter 5

Optimization of Cost and Precision

5.1 Introduction

In order to increase the precision of estimates, it is essential to select a suitable allocation procedure that fulfil the given conditions. Many allocation plans are available by which we allocate fixed sample size that increase the precision of estimates of population parameters or determine the sample size that minimize the cost of survey under given bound on variance of estimator or minimize the variance of overall estimate of population characteristic under given cost of survey. Khan et al (2012) used E-Model technique to solve following optimum allocation problem.

$$Minimize \;\; [V(\bar{y}_{1,st}), V(\bar{y}_{2,st}), \ldots, V(\bar{y}_{p,st})]$$

Subject to

$$p\left(\sum_{h=1}^{L} t_h\sqrt{n_h} + C_o \leq C\right) \geq p_o$$

$$2 \leq n_h \leq N_h$$

$$h = 1, 2, ..., L, 0 < p_o < 1$$

The formulated optimum cost problem is given as:

$$Minimize\ p\left(\sum_{h=1}^{L} t_h\sqrt{n_h} + C_o \leq C\right) \geq p_o$$

Subject to

$$V(\bar{y}_{j,st}) \leq V_j^*$$

$$2 \leq n_h \leq N_h$$

$$h = 1, 2, ..., L, j = 1, 2, ..., p.$$

where, $V(\bar{y}_{j,st})(j = 1, 2, ..., p)$ are variances of estimators of population means of characteristics in multivariate stratified sampling.

In this chapter, we propose an allocation procedure to find the compromise allocation among conflicting objectives, costs of survey and precision of estimates. We distribute a fixed sample size to various strata that minimize, jointly, the cost function and coefficients of variation of regression estimators of population means of $Y_j(j = 1, 2, ..., p)$ characteristics in multivariate stratified sampling. The coefficients of variation for the estimate of population mean of each characteristics $Y_j(j = 1, 2, ..., p)$ define in (3.4) is,

$$C.V\ (\bar{y}_{j,lrs}) = \sqrt{\frac{u'_{jh}}{n_{jh}}} = Z_j.$$

We allocate given sample size n to each stratum that optimize precision for estimate of population mean of each characteristics $Y_j(j = 1, 2, ..., p)$ and total variable cost (Z_{p+1}). where, $(Z_{p+1}) = C - C_o$

5.2 Traveling cost and precision

Consider the following traveling cost function

$$Z_{p+1} = \sum_{h=1}^{L} t_h n_h^{\delta} \tag{5.1}$$

where, t_h is per unit traveling cost within h^{th} stratum and constant $\delta > 0$ represents the effect of travel to cost. We have prior knowledge about t_h. We allocate given sample size n to different strata such that total variable cost Z_{p+1} and coefficients of variation of estimate of population means of each characteristics $Y_j(j = 1, 2, ..., p)$ are minimized. This multi-objective optimization problem can be formulated in nonlinear integer mathematical programming problem as:

$$Minimize \ \left(Z_1, Z_2, \ldots, Z_{p+1)}\right)$$

Subject to

$$\sum_{h=1}^{L} n_h = n \tag{5.2}$$

$$2 \leq n_h \leq N_h$$

$$h = 1, 2, ..., L \text{ and } j = 1, 2, ..., p + 1.$$

The above problem has $(p+1)$ objectives. We use following methods to solve this multiple objective nonlinear integer mathematical programming problem.

5.2.1 Allocation using individual optimum method

$$Minimize \ Z_j$$

Subject to

$$\sum_{h=1}^{L} n_h = n \tag{5.3}$$

$$2 \leq n_h \leq N_h$$

$$h = 1, 2, ..., L \text{ and } j = 1, 2, ..., p + 1.$$

Here $Z_j^*(j = 1, 2, ..., p, p + 1)$ is optimum value of $Z_j, (j = 1, 2, ..., p + 1)$ obtained by solving above integer mathematical programming problem.

5.2.2 Allocation using goal programming method

The multi-objective optimization problem 5.2 may be solved with goal programming as:

$$Minimize\ \sum_{j=1}^{p+1} d_j$$

Subject to

$$Z_j - d_j \leq Z_j^* \tag{5.4}$$

$$\sum_{h=1}^{L} n_h = n$$

$$2 \leq n_h \leq N_h$$

$$h = 1, 2, ..., L \text{ and } j = 1, 2, ..., p+1,$$

where, $d_j (j = 1, 2, ..., p+1)$ are deviation variables.

5.2.3 Allocation using weighted method

We solve the multi-objective allocation problem 5.2 with weighted method as:

$$Minimize\ \sum_{j=1}^{p+1} W_j Z_j$$

Subject to

$$\sum_{h=1}^{L} n_h = n \tag{5.5}$$

$$2 \leq n_h \leq N_h$$

$$\sum_{j=1}^{p+1} W_j = 1$$

$$h = 1, 2, ..., L \text{ and } j = 1, 2, 3, ..., p+1,$$

where, $W_j (j = 1, 2, ..., p+1)$ are relative weights of characteristics.

5.2.4 Allocation using general method

We propose the general method to solve multi-objective allocation problem 5.2

$$Minimize\ \sum_{k=1}^{p+l} W_k Z_k + \sum_{j=1}^{p+1} d_j$$

Subject to

$$Z_j - d_j \leq Z_j^* \tag{5.6}$$
$$\sum_{h=1}^{L} n_h = n$$
$$2 \leq n_h \leq N_h$$
$$\sum_{k=1}^{p+1} W_k = 1$$
$$h = 1, 2, ..., L,\ j \neq k = 1, 2, ..., p+1,$$

where, $W_k(k = 1, 2, ..., p+1)$ are relative weights of characteristics and $d_j(j = 1, 2, ..., p+1)$ are deviation variables.

5.2.5 Numerical example

Data source [Khan et al.(2010)]. We assume that

$$t_1 = 10, t_2 = 9, t_3 = 7, t_4 = 8$$

$$Minimize \begin{pmatrix} Z_1 = \sqrt{\frac{0.000066}{n_1} + \frac{0.000809}{n_2} + \frac{0.001212}{n_3} + \frac{0.000332}{n_4}} \\ Z_2 = \sqrt{\frac{0.000181}{n_1} + \frac{0.009411}{n_2} + \frac{0.023390}{n_3} + \frac{0.000610}{n_4}} \\ Z_3 = 10n_{1c}^{\delta} + 9n_{2c}^{\delta} + 7n_{3c}^{\delta} + 8n_{4c}^{\delta} \end{pmatrix}$$

Subject to

$$\sum_{h=1}^{L} n_{hc} = n$$
$$2 \leq n_1 \leq 6$$
$$9 \leq n_2 \leq 26$$
$$11 \leq n_3 \leq 34$$

$$3 \leq n_4 \leq 9$$

$$n_1, n_2, n_3 \text{ and } n_4 \text{ are integers.}$$

(a) Coefficients of variation and cost using individual optimum method

Optimum allocation for characteristic Y_1 :

$$Minimize\ Z_1 = \sqrt{\frac{0.000066}{n_{11}} + \frac{0.000809}{n_{12}} + \frac{0.001212}{n_{13}} + \frac{0.000332}{n_{14}}}$$

Subject to

$$\sum_{h=1}^{L} n_{1h} = n$$

$$2 \leq n_{11} \leq 8$$

$$2 \leq n_{12} \leq 34$$

$$2 \leq n_{13} \leq 45$$

$$2 \leq n_{14} \leq 12$$

$$n_{11}, n_{12}, n_{13} \text{ and } n_{14} \text{ are integers.}$$

Optimum allocation for characteristic Y_2 :

$$Minimize\ Z_2 = \sqrt{\frac{0.000181}{n_{21}} + \frac{0.009411}{n_{22}} + \frac{0.023390}{n_{23}} + \frac{0.000610}{n_{24}}}$$

Subject to

$$\sum_{h=1}^{L} n_{2h} = n$$

$$2 \leq n_{21} \leq 8$$

$$2 \leq n_{22} \leq 34$$

$$2 \leq n_{23} \leq 45$$

$$2 \leq n_{24} \leq 12$$

$$n_{21}, n_{22}, n_{23} \text{ and } n_{24} \text{ are integers.}$$

Optimum allocation for total traveling cost C :

$$Minimize\ Z_3 = 10n_{1c}^{\delta} + 9n_{2c}^{\delta} + 7n_{3c}^{\delta} + 8n_{4c}^{\delta}$$

Subject to

$$\sum_{h=1}^{L} n_{hc} = n$$

$$2 \leq n_{1c} \leq 8$$

$$2 \leq n_{2c} \leq 34$$

$$2 \leq n_{3c} \leq 45$$

$$2 \leq n_{4c} \leq 12$$

$$n_{c1}, n_{2c}, n_{3c} \text{ and } n_{4c} \text{ are integers.}$$

Here Z_1^*, Z_2^* are coefficients of variation and Z_3^* is total traveling cost for different value of δ and n using individual optimum method is given in Table 5.1.

Table 5.1: Cost and Coefficients of variation using individual optimum method.

δ	n	*Allocation*	n_1	n_2	n_3	n_4	Z_1^*	Z_2^*	Z_3^*	$Z_1^* + Z_2^*$
0.5	50	Y_1	4	16	20	10	0.01268	0.04317	112.60	0.05586
		Y_2	2	16	28	4	0.01449	0.04082	103.18	0.05531
		C	2	2	44	2	0.02512	0.07505	84.62	0.10017
1.0	45	Y_1	5	14	14	12	0.01360	0.04929	370.00	0.06289
		Y_2	2	15	24	4	0.01485	0.04295	355.00	0.05780
		C	2	2	39	2	0.02519	0.07550	327.00	0.10069
1.5	40	Y_1	4	12	14	10	0.01427	0.05060	1073.79	0.06488
		Y_2	2	12	23	3	0.01624	0.04577	1216.10	0.06201
		C	7	8	14	11	0.01508	0.05411	1047.39	0.06919
2.0	35	Y_1	3	11	14	7	0.01515	0.05170	2943.00	0.06685
		Y_2	2	11	19	3	0.01676	0.04879	3728.00	0.06555
		C	7	8	11	9	0.01605	0.05828	2561.00	0.07433

(b) Coefficients of variation and cost using goal programming method

We use goal programming method for allocation of sample size n to four strata considering three characteristics Z_1, Z_2 and $C = Z_3$.

$$Minimize\ d_1 + d_2 + d_3$$

Subject to

$$\sqrt{\frac{0.000066}{n_1} + \frac{0.000809}{n_2} + \frac{0.001212}{n_3} + \frac{0.000332}{n_4}} - d_1 \le Z_1^*$$

$$\sqrt{\frac{0.000181}{n_1} + \frac{0.009411}{n_2} + \frac{0.023390}{n_3} + \frac{0.000610}{n_4}} - d_2 \le Z_2^*$$

$$10n_1^\delta + 9n_2^\delta + 7n_3^\delta + 8n_4^\delta - d_3 \le C^*$$

$$\sum_{h=1}^{L} n_h = n$$

$$2 \le n_1 \le 8$$

$$2 \le n_2 \le 34$$

$$2 \le n_3 \le 45$$

$$2 \le n_4 \le 12$$

$$n_1, n_2, n_3 \text{ and } n_4 \text{ are integers.}$$

Here $\hat{Z}_1$ and $\hat{Z}_2$ are coefficients of variation and $\hat{Z}_3$ is total traveling cost for different values of n and δ solving above allocation problem given in Table 5.2.

Table 5.2: Cost and Coefficients of variation using *GP* method.

δ	n	n_1	n_2	n_3	n_4	$\hat{Z}_1$	$\hat{Z}_2$	$\hat{Z}_3$	$\hat{Z}_1 + \hat{Z}_2$
0.5	50	2	34	12	2	.01800	0.05120	102.18	0.06919
1.0	45	2	2	39	2	0.02519	0.07550	327.00	0.10069
1.5	40	7	8	14	11	0.01508	0.05411	1047.39	0.06919
2.0	35	7	8	11	9	0.01605	0.05828	2561.00	0.07433

(c) Coefficients of variation and cost using weighted method

The compromise allocation for $\hat{Z}_1$, $\hat{Z}_2$ and $\hat{Z}_3$ is obtained by weighted method.

$$Minimize \left(\begin{array}{c} W_1\sqrt{\frac{0.000066}{n_{1c}} + \frac{0.000809}{n_{2c}} + \frac{0.001212}{n_{3c}} + \frac{0.000332}{n_{4c}}} \\ +W_2\sqrt{\frac{0.000181}{n_{1c}} + \frac{0.009411}{n_{2c}} + \frac{0.023390}{n_{3c}} + \frac{0.000610}{n_{4c}}} \\ +W_3\left(10n_1^{\delta} + 9n_2^{\delta} + 7n_3^{\delta} + 8n_4^{\delta}\right) \end{array} \right)$$

Subject to

$$\sum_{h=1}^{L} n_h = n$$

$$2 \leq n_1 \leq 8$$

$$2 \leq n_2 \leq 34$$

$$2 \leq n_3 \leq 45$$

$$2 \leq n_4 \leq 12$$

$$W_1 + W_2 + W_3 = 1$$

n_1, n_2, n_3 and n_4 are integers.

(W_1, W_2, W_3)= (0.2, 0.3, 0.5), (0.2, 0.5, 0.3), (0.5, 0.2, 0.3).

Here $\hat{Z}_1, \hat{Z}_2$ and $\hat{Z}_3$ are minimum values for different values of $alpha$ and n given in Table 5.3 obtained by solving above problem using different relative weights given above.

Table 5.3: Cost and Coefficients of variation using weighted method.

δ	n	n_1	n_2	n_3	n_4	$\hat{Z}_1$	$\hat{Z}_2$	$\hat{Z}_3$	$\hat{Z}_1 + \hat{Z}_2$
0.5	50	2	2	44	2	0.02512	0.07505	84.62	0.10017
1.0	45	2	2	39	2	0.02519	0.07550	327.00	0.10069
1.5	40	7	8	14	11	0.01508	0.05411	1047.39	0.06919
2.0	35	7	8	11	9	0.01605	0.05828	2561.00	0.07433

(d) Coefficients of variation and cost using general method

We use general method to allocate given sample size n taking into account Z_1, Z_2 and Z_3. We consider three cases here.

Case-1

We minimize travel cost Z_3 to compromise optimum value of Z_2 and Z_1.

$$Minimize\ Z_3 = 10n_1^{\delta} + 9n_2^{\delta} + 7n_3^{\delta} + 8n_4^{\delta}$$

Subject to

$$\sqrt{\frac{0.000066}{n_1} + \frac{0.000809}{n_2} + \frac{0.001212}{n_3} + \frac{0.000332}{n_4}} - d_1 \leq Z_1^*$$

$$\sqrt{\frac{0.000181}{n_1} + \frac{0.009411}{n_2} + \frac{0.023390}{n_3} + \frac{0.000610}{n_4}} - d_2 \leq Z_2^*$$

$$\sum_{h=1}^{L} n_h = n$$

$$2 \leq n_1 \leq 8$$

$$2 \leq n_2 \leq 34$$

$$2 \leq n_3 \leq 45$$

$$2 \leq n_4 \leq 12$$

$$n_1, n_2, n_3 \text{ and } n_4 \text{ are integers.}$$

Here $\hat{Z}_1$ and $\hat{Z}_2$ are coefficients of variation and $\hat{Z}_3$ is total travel cost given in Table 5.4 obtained by solving above problem for different values of δ and n.

Table 5.4: Cost and Coefficients of variation using general method.

δ	n	n_1	n_2	n_3	n_4	$\hat{Z}_1$	$\hat{Z}_2$	$\hat{Z}_3$	$\hat{Z}_1 + \hat{Z}_2$
0.5	50	2	18	23	7	0.01334	0.04144	107.06	0.05478
1.0	45	2	14	25	5	0.01441	0.04312	354.00	0.5753
1.5	40	3	12	20	5	0.01471	0.04622	1141.63	0.06093
2.0	35	3	11	17	4	0.01580	0.04944	3330.00	0.06524

Case-2

We minimize travel cost Z_3 to compromise optimum value of Z_2 and Z_1. We use propose general method to solve multi-objective allocation problem.

$$Minimize \begin{pmatrix} W_1\sqrt{\frac{0.000066}{n_1}+\frac{0.000809}{n_2}+\frac{0.001212}{n_3}+\frac{0.000332}{n_4}} \\ +W_2\left(10n_1^{\delta}+9n_2^{\delta}+7n_3^{\delta}+8n_4^{\delta}\right) \end{pmatrix}$$

Subject to

$$\sqrt{\frac{0.000181}{n_1}+\frac{0.009411}{n_2}+\frac{0.023390}{n_3}+\frac{0.000610}{n_4}} \leq Z_2^*$$

$$\sum_{h=1}^{L} n_h = n$$

$$2 \leq n_1 \leq 8$$

$$2 \leq n_2 \leq 34$$

$$2 \leq n_3 \leq 45$$

$$2 \leq n_4 \leq 12$$

$$W_1 + W2 = 1$$

n_1, n_2, n_3 and n_4 are integers.

Here $\hat{Z}_1$ and $\hat{Z}_2$ are coefficients of variation and $\hat{Z}_3$ is total travel cost given in Table 5.5 obtained by solving above problem for different values of δ and n.

Table 5.5: Minimum $Z_3$and Z_1 under optimum Z_2

δ	n	W_1	W_2	n_1	n_2	n_3	n_4	$\hat{Z}_1$	Z_2^*	$\hat{Z}_3$	$\hat{Z}_1 + Z_2^*$
0.5	50	0.2	0.8	2	15	30	3	0.01543	0.04082	101.20	0.05625
		0.4	0.6	2	15	30	3	0.01543	0.04082	101.20	0.05625
		0.6	0.4	2	15	30	3	0.01543	0.04082	101.20	0.05625
		0.8	0.2	2	15	30	3	0.01543	0.04082	101.20	0.05625
1	45	0.2	0.8	2	13	26	4	0.01499	0.04295	351.00	0.05794
		0.4	0.6	2	13	26	4	0.01499	0.04295	351.00	0.05794
		0.4	0.6	2	12	18	3	0.01595	0.04295	348.00	0.05890
		0.8	0.2	2	12	18	3	0.01595	0.04295	348.00	0.05890
1.5	40	0.2	0.8	2	12	20	6	0.01471	0.04577	1146.09	0.06048
		0.4	0.6	2	12	20	6	0.01471	0.04577	1146.09	0.06048
		0.6	0.4	4	12	19	5	0.01519	0.04577	1077.5	0.06096
		0.8	0.2	2	12	20	6	0.01471	0.04577	1146.09	0.06048
2	35	0.2	0.8	3	12	17	3	0.01647	0.04879	3481.00	0.06526
		0.4	0.6	2	12	17	4	0.01596	0.04879	3487.00	0.06475
		0.6	0.4	2	12	17	4	0.01596	0.04879	3487.00	0.06475
		0.8	0.2	2	13	16	4	0.01594	0.04879	3481.00	0.06472

Case-3

Minimization of Z_3 and Z_2 under optimum value of Z_1.

$$Minimize \begin{pmatrix} W_1\sqrt{\frac{0.000181}{n_1}+\frac{0.009411}{n_2}+\frac{0.023390}{n_3}+\frac{0.000610}{n_4}} \\ +W_2\left(10n_1^\delta+9n_2^\delta+7n_3^\delta+8n_4^\delta\right) \end{pmatrix}$$

Subject to

$$\sqrt{\frac{0.000066}{n_1}+\frac{0.000809}{n_2}+\frac{0.001212}{n_3}+\frac{0.000332}{n_4}} \le Z_2^*$$

$$\sum_{h=1}^{L} n_h = n$$

$$2 \le n_1 \le 8$$

$$2 \le n_2 \le 34$$

$$2 \leq n_3 \leq 45$$

$$2 \leq n_4 \leq 12$$

$$W_1 + W2 = 1$$

n_1, n_2, n_3 and n_4 are integers.

Here $\hat{Z}_2$ and $\hat{Z}_2$ are coefficients of variation and $\hat{Z}_3$ is total travel cost given in Table 5.6 obtained by solving above problem for different values of δ and n.

Table 5.6: Minimum Z_3 and Z_2 under optimum Z_1

δ	n	W_1	W_2	n_1	n_2	n_3	n_4	Z_1^*	$\hat{Z}_2$	$\hat{Z}_3$	$Z_1^* + \hat{Z}_2$
0.5	50	0.2	0.8	2	14	26	8	0.01268	0.04170	106.14	0.05438
		0.4	0.6	3	11	29	7	0.01268	0.04254	104.04	0.05522
		0.6	0.4	3	11	29	7	0.01268	0.04254	104.04	0.05522
		0.8	0.2	2	14	27	7	0.01268	0.04143	105.36	0.05411
1	45	0.2	0.8	2	9	24	10	0.01360	0.03809	529.00	0.05169
		0.4	0.6	2	9	24	10	0.01360	0.03809	529.00	0.05169
		0.4	0.6	2	9	24	10	0.01360	0.03809	529.00	0.05169
		0.8	0.2	2	9	24	10	0.01360	0.03809	529.00	0.05169
1.5	40	0.2	0.8	7	9	14	10	0.01427	0.05297	1047.87	0.06722
		0.4	0.6	7	9	13	11	0.01427	0.05410	1148.18	0.06837
		0.6	0.4	7	9	13	11	0.01427	0.05410	1148.18	0.06837
		0.8	0.2	7	9	13	11	0.01427	0.05410	1148.18	0.06837
2	35	0.2	0.8	4	10	13	8	0.01515	0.05350	2755.00	0.06865
		0.4	0.6	4	10	13	8	0.01515	0.05350	2755.00	0.06865
		0.6	0.4	4	10	13	8	0.01515	0.05350	2755.00	0.06865
		0.8	0.2	4	10	13	8	0.01515	0.05350	2755.00	0.06845

5.2.6 Results

Table 5.1 presents coefficients of variation and expected total travel cost for different values of δ and given sample size n. Individual allocation method minimize only one objective either travel cost Z_3 or coefficient of variation Z_1 or Z_2. GP method and weights are used to allocate the given sample size that compromise among cost and precision of estimates as shown in Table 5.2 for $\delta = 0.5$ and $n = 50$ in first row and Table 5.3 for $\delta = 2$ and $n = 35$ in 4^{th} row. For other values of constant δ and given sample size n, both methods give same results as we obtain to minimize the travel cost under given sample size n.Tables 5.4-5.6 indicate the results of Case(1-3) respectively, of propose general method. Table 5.4 gives the compromise allocation to minimize the travel cost to compromise optimum precision of estimates of population means of variables $Y_j (j = 1, 2)$. Mixed allocation given in Table 5.5 minimize coefficient of variation of estimate of population mean of Y_2 and travel cost to maintain minimum efficiency of estimator of mean of variable Y_1 for various values of δ, n, W_1 and W_2. The cost and precision of estimate of population mean of Y_2 are optimized to compromise on precision of estimate of population mean of Y_1 shown in Table 5.6.

5.3 Optimization of nonlinear cost and precision

Consider the following traveling cost function.

$$Z_{p+1} = \sum_{h=1}^{L} C_h n_h + \sum_{h=1}^{L} t_h n_h^{\delta} \tag{5.7}$$

where, t_h is per unit traveling cost within stratum and constant $\delta > 0$ represent the effect of travel to cost, C_h is per unit measurement cost in h_{th} stratum. We have prior knowledge about t_h and C_h. We allocate a given sample size n to each stratum such that total variable cost of survey and total coefficient of variation for estimate of population means of characteristics $Y_j (j = 1, 2, ..., p)$ are minimized. This problem become a multi-objective mathematical programming problem with $(p + 1)$ objectives. This problem can be formulated in nonlinear integer mathematical programming problem as:

$$Minimize \ (Z_1, Z_2, \ldots, Z_{p+1})$$

Subject to

$$\sum_{h=1}^{L} n_h = n \tag{5.8}$$

$$2 \leq n_h \leq N_h$$

$$h = 1, 2, ..., L \text{ and } j = 1, 2, ..., p + 1.$$

The above problem has $(p + 1)$ objectives. We use following methods to solve this multi-objective optimization programming problem.

5.3.1 Allocation using individual optimum method

Lat, Z_j^* be a optimum value of Z_j obtained by solving following integer mathematical programming problem.

$$Minimize \ Z_j$$

Subject to

$$\sum_{h=1}^{L} n_h = n \tag{5.9}$$

$$2 \leq n_h \leq N_h$$

$$h = 1, 2, ..., L \text{ and } j = 1, 2, ..., p + 1.$$

5.3.2 Allocation using goal programming method

The multi-objective optimization problem 5.8 can be solved by goal programming as:

$$Minimize \ \textstyle\sum_{j=1}^{p+1} d_j$$

Subject to

$$Z_j - d_j \leq Z_j^* \tag{5.10}$$

$$\sum_{h=1}^{L} n_h = n$$

$$2 \leq n_h \leq N_h$$

$$h = 1, 2, ..., L \text{ and } j = 1, 2, ..., p + 1,$$

where, $d_j(j = 1, 2, ..., p + 1)$ are deviation variables.

5.3.3 Allocation using weighted method

We solve the problem 5.8 with weighted method as:

$$Minimize \ \sum_{j=1}^{p+l} W_j Z_j$$

Subject to

$$\sum_{h=1}^{L} n_h = n \tag{5.11}$$
$$2 \leq n_h \leq N_h$$
$$\sum_{j=1}^{p+1} W_j = 1$$
$$h = 1, 2, ..., L \text{ and } j = 1, 2, 3, ..., p + 1,$$

where, $W_j(j = 1, 2, ..., p + 1)$ are relative weights which indicates the importance of different objectives.

5.3.4 Allocation using general method

We solve multi-objective optimization problem 5.8 with general as:

$$Minimize \ \sum_{k=1}^{p+l} W_j Z_j + \sum_{j=1}^{p+1} d_j$$

Subject to

$$Z_j - d_j \leq Z_j^* \tag{5.12}$$
$$\sum_{h=1}^{L} n_h = n$$
$$2 \leq n_h \leq N_h$$
$$\sum_{k=1}^{p+1} W_k = 1$$
$$h = 1, 2, ..., L, \ j \neq k = 1, 2, 3, ..., p + 1,$$

where, $W_k(k = 1, 2, ..., p + 1)$ are relative weights which indicates the importance of each objective and $d_j(j = 1, 2, ..., p + 1)$ are deviation variables.

5.3.5 Numerical example

Data source [Khan et al.(2010)].We assume that

$$C_1 = 15, C_2 = 9, C_3 = 5, C_4 = 7, t_1 = 5, t_2 = 4, t_3 = 2, t_4 = 3.$$

$$Minimize \begin{pmatrix} Z_1 = \sqrt{\frac{0.000066}{n_1} + \frac{0.000809}{n_2} + \frac{0.001212}{n_3} + \frac{0.000332}{n_4}} \\ Z_2 = \sqrt{\frac{0.000181}{n_1} + \frac{0.009411}{n_2} + \frac{0.023390}{n_3} + \frac{0.000610}{n_4}} \\ Z_3 = 15n_1 + 7n_2 + 5n3 + 9n_4 + 5n_1^{\delta} + 4n_2^{\delta} + 2n_3^{\delta} + 3n_4^{\delta} \end{pmatrix}$$

Subject to

$$\sum_{h=1}^{L} n_h = n$$

$$2 \le n_1 \le 8$$

$$2 \le n_2 \le 34$$

$$2 \le n_3 \le 45$$

$$2 \le n_4 \le 12$$

n_1, n_2, n_3 and n_4 are integers.

(a) Coefficients of variation and cost using individual optimum allocation

Individual optimum allocation for characteristic Z_1:

$$Minimize\ Z_1 = \sqrt{\frac{0.000066}{n_{11}} + \frac{0.000809}{n_{12}} + \frac{0.001212}{n_{13}} + \frac{0.000332}{n_{14}}}$$

Subject to

$$\sum_{h=1}^{L} n_{1h} = n$$

$$2 \le n_{11} \le 8$$

$$2 \le n_{12} \le 34$$

$$2 \le n_{13} \le 45$$

$$2 \le n_{14} \le 12$$

n_{11}, n_{12}, n_{13} and n_{14} are integers.

Individual optimum allocation for characteristic Z_2:

$$Minimize\ Z_2 = \sqrt{\frac{0.000181}{n_{21}} + \frac{0.009411}{n_{22}} + \frac{0.023390}{n_{23}} + \frac{0.000610}{n_{24}}}$$

Subject to

$$\sum_{h=1}^{L} n_{2h} = n$$

$$2 \leq n_{21} \leq 8$$

$$2 \leq n_{22} \leq 34$$

$$2 \leq n_{23} \leq 45$$

$$2 \leq n_{24} \leq 12$$

n_{21}, n_{22}, n_{23} and n_{24} are integers.

Individual optimum allocation for characteristic Z_3 :

$$Minimize\ Z_3 = C = 15n_{1c} + 7n_{2c} + 5n3c + 9n_{4c} + 5n_{1c}^{\delta} + 4n_{2c}^{\delta} + 2n_{3c}^{\delta} + 3n_{4c}^{\delta}$$

Subject to

$$\sum_{h=1}^{L} n_{hc} = n$$

$$2 \leq n_{1c} \leq 8$$

$$2 \leq n_{2c} \leq 34$$

$$2 \leq n_{3c} \leq 45$$

$$2 \leq n_{4c} \leq 12$$

n_{c1}, n_{2c}, n_{3c} and n_{4c} are integers.

Here Z_1^* and Z_2^* are coefficients of variation and Z_3^* is total variable cost using individual optimum method for different value of δ and n given in Table 5.7.

Table 5.7: Cost and Coefficients of variation using individual allocation.

δ	n	*Allocation*	n_1	n_2	n_3	n_4	Z_1^*	Z_2^*	Z_3^*	$Z_1^* + Z_2^*$
0.5	44	Y_1	3	14	15	12	0.01372	0.04840	367.77	0.06212
		Y_2	2	14	25	3	0.01581	0.04360	317.24	0.05941
		C	2	2	38	2	0.02521	0.07561	281.30	0.10082
1.0	38	Y_1	4	13	14	7	0.01459	0.05027	405.00	0.06486
		Y_2	2	12	20	4	0.01562	0.04687	360.00	0.06249
		C	2	2	32	2	0.02533	0.07637	310.00	0.10170
1.5	32	Y_1	3	10	12	7	0.01585	0.05512	529.18	0.07096
		Y_2	2	10	17	3	0.01720	0.05110	508.41	0.06830
		C	2	4	21	5	0.01896	0.06066	480.16	0.07962
2.0	26	Y_1	2	10	9	5	0.01775	0.06126	847.00	0.07907
		Y_2	2	9	13	2	0.01955	0.05692	870.00	0.07647
		C	3	5	11	7	0.01849	0.06447	732.00	0.08296

(b) Coefficients of variation and cost using goal programming method

For minimization of Z_1, Z_2 and Z_3, we use goal programming method.

$$Minimize\ d_1 + d_2 + d_3$$

Subject to

$$\sqrt{\frac{0.000066}{n_1} + \frac{0.000809}{n_2} + \frac{0.001212}{n_3} + \frac{0.000332}{n_4}} - d_1 \leq Z_1^*$$

$$\sqrt{\frac{0.000181}{n_1} + \frac{0.009411}{n_2} + \frac{0.023390}{n_3} + \frac{0.000610}{n_4}} - d_2 \leq Z_2^*$$

$$C = 15n_1 + 7n_2 + 5n3 + 9n_4 + 5n_1^\delta + 4n_2^\delta + 2n_3^\delta + 3n_4^\delta - d_3 \leq C^*$$

$$\sum_{h=1}^{L} n_h = n$$

$$2 \leq n_1 \leq 8$$

$$2 \leq n_2 \leq 34$$

$$2 \leq n_3 \leq 45$$

$$2 \leq n_4 \leq 12$$

$$n_1, n_2, n_3 \text{ and } n_4 \text{ are integers.}$$

Here $\hat{Z}_1$ and $\hat{Z}_2$ are coefficients of variation and $\hat{Z}_3$ is total variable cost for different values of n and δ using goal programming method of given in Table 5.8.

Table 5.8: Cost and Coefficients of variation using GP method.

δ	n	n_1	n_2	n_3	n_4	$\hat{Z}_1$	$\hat{Z}_2$	$\hat{Z}_3$	$\hat{Z}_1+\hat{Z}_2$
0.5	50	2	2	38	2	0.02521	0.07561	281.30	0.10082
1.0	45	2	2	32	2	0.02533	0.07637	310.00	0.10170
1.5	40	2	4	21	5	0.01896	0.06066	480.16	0.07962
2.0	35	3	5	11	7	0.01849	0.06447	732.00	0.08296

(c) Coefficients of variation and cost using weighted method

For minimization of Z_1, Z_2 and Z_3, we use weighted method.

$$Minimize \left(\begin{array}{c} W_1\sqrt{\frac{0.000066}{n_1}+\frac{0.000809}{n_2}+\frac{0.001212}{n_3}+\frac{0.000332}{n_4}} \\ +W_2\sqrt{\frac{0.000181}{n_1}+\frac{0.009411}{n_2+\frac{0.023390}{n_3}+\frac{0.000610}{n_4}}} \\ +W_3\left(15n_1+7n_2+5n3+9n_4+5n_1^{\delta}+4n_2^{\delta}+2n_3^{\delta}+3n_4^{\delta}\right) \end{array} \right)$$

Subject to

$$\sum_{h=1}^{L} n_h = n$$

$$2 \le n_1 \le 8$$

$$2 \le n_2 \le 34$$

$$2 \le n_3 \le 45$$

$$2 \le n_4 \le 12$$

$$W_1 + W_2 + W_3 = 1$$

$$n_1, n_2, n_3 \text{ and } n_4 \text{ are integers.}$$

$(W_1, W_2, W_3 =)$ (0.2, 0.3, 0.5), (0.2, 0.5, 0.3), (0.5, 0.2, 0.3).

$\hat{Z}_1$ and $\hat{Z}_2$ are coefficients of variation and $\hat{Z}_3$ is total variable cost for different values of n and δ using weighted method given in Table 5.9.

Table 5.9: Cost and Coefficients of variation using weighted method.

δ	n	n_1	n_2	n_3	n_4	$\hat{Z}_1$	$\hat{Z}_2$	$\hat{C}$	$\hat{Z}_1 + \hat{Z}_2$
0.5	50	2	2	38	2	0.02521	0.07561	281.30	0.10082
1.0	45	2	2	32	2	0.02533	0.07637	310.00	0.10170
1.5	40	2	4	21	5	0.01896	0.06066	480.16	0.07962
2.0	35	3	5	11	7	0.01849	0.06447	732.00	0.08296

(d) Coefficients of variation and cost using general method

We use propose general method for minimization of Z_1, Z_2 and Z_3 for different values of n and δ.

Case-1

We minimize travel cost Z_3 to compromise optimum value of Z_2 and Z_1.

$$Minimize\ Z_3 = 15n_1 + 7n_2 + 5n3 + 9n_4 + 5n_1^{\delta} + 4n_2^{\delta} + 2n_3^{\delta} + 3n_4^{\delta}$$

Subject to

$$\sqrt{\frac{0.000066}{n_1} + \frac{0.000809}{n_2} + \frac{0.001212}{n_3} + \frac{0.000332}{n_4}} - d_1 \leq Z_1^*$$

$$\sqrt{\frac{0.000181}{n_1} + \frac{0.009411}{n_2} + \frac{0.023390}{n_3} + \frac{0.000610}{n_4}} - d_2 \leq Z_2^*$$

$$\sum_{h=1}^{L} n_h = n$$

$$2 \leq n_1 \leq 8$$

$$2 \leq n_2 \leq 34$$

$$2 \leq n_3 \leq 45$$

$$2 \leq n_4 \leq 12$$

n_1, n_2, n_3 and n_4 are integers.

$\hat{Z}_1$ and $\hat{Z}_2$ are coefficients of variation and $\hat{Z}_3$ is total variable cost for different values of n and δ using ε-constraint method given in Table 5.10.

Table 5.10: Cost and coefficients of variation using ε-constraint method.

δ	n	n_1	n_2	n_3	n_4	$\hat{Z}_1$	$\hat{Z}_2$	$\hat{Z}_3$	$\hat{Z}_1 + \hat{Z}_2$
0.5	44	2	14	22	6	0.01418	0.04390	330.77	0.05808
1.0	38	3	11	20	4	0.01546	0.04730	369.00	0.6276
1.5	32	2	9	17	4	0.01665	0.05161	500.33	0.06825
2.0	26	2	8	13	3	0.01838	0.05717	819.00	0.07556

Case-2

Minimization of Z_3 and Z_1 under optimum value of Z_2.

$$Minimize \begin{pmatrix} W_1\sqrt{\frac{0.000066}{n_1} + \frac{0.000809}{n_2} + \frac{0.001212}{n_3} + \frac{0.000332}{n_4}} \\ +W_2\left(15n_1 + 7n_2 + 5n3 + 9n_4 + 5n_1^\delta + 4n_2^\delta + 2n_3^\delta + 3n_4^\delta\right) \end{pmatrix}$$

Subject to

$$\sqrt{\frac{0.000181}{n_1} + \frac{0.009411}{n_2} + \frac{0.023390}{n_3} + \frac{0.000610}{n_4}} \leq Z_2^*$$

$$\sum_{h=1}^{L} n_h = n$$

$$2 \leq n_1 \leq 8$$

$$2 \leq n_2 \leq 34$$

$$2 \leq n_3 \leq 45$$

$$2 \leq n_4 \leq 12$$

$$W_1 + W2 = 1$$

$$n_1, n_2, n_3 \text{ and } n_4 \text{ are integers.}$$

Here $\hat{Z}_1$ and $\hat{Z}_2$ are coefficients of variation and $\hat{Z}_3$ is total variable cost for different values of n and δ using general method given in Table 5.11.

Table 5.11: Z_3 and Z_1 under optimum Z_2

δ	n	W_1	W_2	n_1	n_2	n_3	n_4	$\hat{Z}_1$	Z_2^*	$\hat{Z}_3$	$\hat{Z}_1 + Z_2^*$
0.5	44	0.2	0.8	2	13	27	2	0.01750	0.04456	310.13	0.06206
		0.4	0.6	2	13	27	2	0.01750	0.04456	310.13	0.06206
		0.6	0.4	2	13	27	2	0.01750	0.04456	310.13	0.06206
		0.8	0.2	2	13	27	2	0.01750	0.04456	310.13	0.06206
1	38	0.2	0.8	2	12	22	2	0.01793	0.04735	350.00	0.06528
		0.4	0.6	2	12	22	2	0.01793	0.04735	350.00	0.06528
		0.4	0.6	2	11	23	2	0.01803	0.04762	346.00	0.06565
		0.8	0.2	2	11	23	2	0.01803	0.04762	346.00	0.06565
1.5	32	0.2	0.8	2	9	18	3	0.01735	0.05137	500.47	0.06872
		0.4	0.6	2	9	18	3	0.01735	0.05137	500.47	0.06872
		0.6	0.4	2	9	18	3	0.01735	0.05137	500.47	0.06872
		0.8	0.2	2	9	18	3	0.01735	0.05137	500.47	0.06872
2	26	0.2	0.8	2	8	13	3	0.01838	0.05718	819.00	0.07556
		0.4	0.6	2	8	13	3	0.01838	0.05718	819.00	0.07556
		0.6	0.4	2	8	13	3	0.01838	0.05718	819.00	0.07556
		0.8	0.2	2	8	13	3	0.01838	0.05718	819.00	0.07556

Case-3

We minimize Z_3 and $Z_2$to compromise optimum value of Z_1.

$$Minimize \begin{pmatrix} W_1\sqrt{\frac{0.000181}{n_1} + \frac{0.009411}{n_2} + \frac{0.023390}{n_3} + \frac{0.000610}{n_4}} \\ +W_2\left(15n_1 + 7n_2 + 5n3 + 9n_4 + 5n_1^\delta + 4n_2^\delta + 2n_3^\delta + 3n_4^\delta\right) \end{pmatrix}$$

Subject to

$$\sqrt{\frac{0.000066}{n_1} + \frac{0.000809}{n_2} + \frac{0.001212}{n_3} + \frac{0.000332}{n_4}} \leq Z_1^*$$

$$\sum_{h=1}^{L} n_h = n$$

$$2 \leq n_1 \leq 8$$

$$2 \leq n_2 \leq 34$$

$$2 \leq n_3 \leq 45$$

$$2 \leq n_4 \leq 12$$

$$W_1 + W2 = 1$$

n_1, n_2, n_3 and n_4 are integers.

$\hat{Z}_1$ and $\hat{Z}_2$ are coefficients of variation and $\hat{Z}_3$ is total variable cost for different values of n and δ using general method given in Table 5.12.

Table 5.12: Z_3 and Z_2 under optimum Z_1

δ	n	W_1	W_2	n_1	n_2	n_3	n_4	Z_1^*	$\hat{Z}_2$	$\hat{Z}_3$	$Z_1^* + \hat{Z}_2$
0.5	50	0.2	0.8	2	10	25	7	0.01448	0.04532	326.66	0.05980
		0.4	0.6	2	10	25	7	0.01448	0.04532	326.66	0.05980
		0.6	0.4	2	10	25	7	0.01448	0.04532	326.66	0.05980
		0.8	0.2	2	10	25	7	0.01448	0.04532	326.66	0.05980
1	45	0.2	0.8	2	10	20	6	0.01516	0.04798	362.00	0.06315
		0.4	0.6	2	10	20	6	0.01516	0.04798	362.00	0.06315
		0.4	0.6	2	10	20	6	0.01516	0.04798	362.00	0.06315
		0.8	0.2	2	10	20	6	0.01516	0.04798	362.00	0.06315
1.5	40	0.2	0.8	2	7	18	5	0.01680	0.05344	488.50	0.07024
		0.4	0.6	2	7	18	5	0.01680	0.05344	488.50	0.07024
		0.6	0.4	2	7	18	5	0.01680	0.05344	488.50	0.07024
		0.8	0.2	2	7	18	5	0.01680	0.05344	488.50	0.07024
2	35	0.2	0.8	3	7	10	6	0.01772	0.06201	747.00	0.07973
		0.4	0.6	3	5	11	7	0.01847	0.06447	732.00	0.08294
		0.6	0.4	3	5	11	7	0.01847	0.06447	732.00	0.08294
		0.8	0.2	3	5	11	7	0.01847	0.06447	732.00	0.08294

5.3.6 Results

Table 5.7 presents the allocation of a given sample size n that give individual optimum value of one objective among nonlinear cost and coefficients of variation of estimates of population means of $Y_j (j = 1, 2)$ using individual optimum method. First, second and third row of the table presents the optimum allocation of a sample size $n = 44$ and value of constant $\delta = 0.5$ according to study variables Y_1 and Y_2 and nonlinear cost C. GP method and weighted method give similar results to individual optimum allocation according to nonlinear cost C irrespective of different values of weights W_1, W_2 and W_3 given in Table 5.8-5.9 accordingly. The optimum values of nonlinear cost C, coefficients of variation of the estimates of population means of variables Y_1 and Y_2 using propose general method are given in Table 5.10-5.12. For different values of given sample size n and constant δ, Table 5.10 presents the results of proposed allocation that minimize nonlinear cost C to cooperate optimum precision of estimates of population means of Y_1 and Y_2. Table 5.11 displays the results of compromise allocation that minimize nonlinear cost C and coefficient of variation of estimate of population mean of Y_1 to compromise optimum precision of estimate of mean of Y_2. Similarly, Table 5.12 gives the optimum value of coefficient of variation of estimate of population mean of variable Y_2 and nonlinear cost C to bargain the coefficient of variation of estimate of population mean of variable Y_1. It is shown that compromise allocation using proposed general method is efficient than individual optimum method, goal programming method and weighted method to some extent.

Chapter 6

Conclusion and Recommendation

The study was conducted to solve sample allocation problem in multivariate stratified sampling design. The problem was formulated in integer multi-objective mathematical programming. The individual optimum method (IOM), goal programming method (GPM) and general method (GM) were used to solve allocation problem in following three situations.

(a) Maximization of precision under fixed given cost of sample survey.

(b) Minimization of coefficients of variation under estimated cost of sample survey.

(c) Optimization of precision and cost of survey under given sample size.

The above three situations ware discussed in chapters 3-5.

In section 3.2, the compromise allocation was given to maximize the precision of estimates of population means using regression estimator satisfying total budget available measurement of units. Tables 3.1-3.2 showed C.V using IOM and GM respectively and total C.V using GPM is 0.05572. The compromise allocation using GPM produced 5.51% and 2.04% efficient estimates and GM gave 4.46% and 1.02% better result for $(W_1, W_2) = (0.4, 0.6)$ as compared to IOM according to characteristics Y_1 and Y_2 as showed in Tables 3.3-3.4 respectively. The allocation problem under given fixed budget available for travel cost among units was solved in section 3.3. The formulated mathematical programming problem was given in (3.9) and its solution using IOM, GPM and GM was given in (3.10-3.12). The C.V under these methods were displayed in Tables 3.5-3.7 respectively for different values of constant δ and $\acute{C}$. We observed that IOM according to C yield 14% and 13.70%

less efficient results as compared to GPM for $\delta = 2.0$, $\acute{C} = 1700$ and 13.11% and 10.40% less efficient estimate as compared to GM for $\delta = 0.5$, $\acute{C} = 120$ and $(W_1, W_2) = (0.2, 0.8)$ respectively as showed in Tables 3.8-3.9. Section 3.4 illustrated the solution of allocation problem to minimize the coefficients of variation satisfying total budget available for gaining information from units and travel among units. The formulated allocation problem (3. 13) was solved by IOM (3.14),GPM (3.15) and GM (3.16) and C.V using these methods was given in Tables 3.10-3.12 accordingly. The GPM produced 3.28% and 1.40% precise estimates for $\delta = 1.0$, $\acute{C} = 500$ and GM gave 14.76% and 14.63% efficient results for $\delta = 2.0$, $\acute{C} = 1500$ and $(W_1, W_2) = (0.7, 0.3)$ as compared to IOM according to Y_1 and Y_2 as showed in Tables 3.13-3.14. In section 3.5, coefficients of variation were minimized to determine sample size under logarithmic travel cost function. The formulated allocation problem was given in (3.17) and its solution was given in (3.18-3.19) using IOM, GPM and GM. Tables 3.18-3.19 displayed the efficiency of GPM and GM to IOM for different values of constants δ and $\acute{C}$. The GPM provided 12.5% precise estimate as compared to IOM according to Y_1 and Y_2 for $\delta = 0.5$ and $\acute{C} = 30$. The GM delivered 50.68% efficient estimates of population means for $\delta = 2.0$, $\acute{C} = 115$ than IOM according to both characteristics Y_1 and Y_2.

Chapter 4 provided the solution of allocation problems under four types of probabilistic or estimated cost functions. In section 4.2, we found compromise allocation to minimize coefficients of variation under specified budget for estimated per unit measurement cost of sample survey. The allocation problem was formulated in stochastic mathematical programming (4.5) and solved using IOM (4.13), GPM (4.14) and GM (4.15). Tables 4.1-4.2 demonstrated coefficients of variation using mentioned allocation's methods. Table 4.3 showed that GPM provided 1.12% and 0.5% efficient results than IOM subject to Y_1 and Y_2. GM delivered 2.77% and 1.67% better estimates as compared to IOM Y_1 and Y_2 for $(W_1, W_2) = (0.2, 0.8)$. The allocation problem under given budget for travel cost among units estimated from sample cost was discussed in section 4.3. The problem was formulated in stochastic programming (4.17) and its equivalent deterministic programming problem was given in (4.24) and solution using proposed IOM, GPM and GM was given in (4.25-4.27). GPM produced 6.64% and 1.96% better estimates relative to IOM according to characteristics Y_1 and Y_2 using $\delta = 0.5$ and $\acute{C} = 115$ as given in Table 4.8. The GM provided 7.88% and 3.15% precise estimates of population mean as compared to IOM according to Y_1 and Y_2 for $\delta = 0.5$, $\acute{C} = 115$ and $(W_1, W_2) = (0.4, 0.6)$ respectively as showed in Table 4.9. The sample allocation that maximize the efficiency of regression estimator of mean under specific budget based on measurement unit cost and traveling cost estimated from sample costs discussed in section 4.4. The problem was formulated in chance constrained mathematical programming (4.29) and IOM (4.37), GPM (4.38) and GM (4.39) used to solve converted deterministic constraint programming problem (4.36). From Tables 4.13-4.14 it is clear that GPM and GM provided more precise estimates as compared to IOM according to characteristics Y_1 and Y_2 for different values of δ and $\acute{C}$. Section 4.5 produced efficient compromise allocation to minimize coefficients of variation under estimated logarithmic travel cost function. The problem was formulated in stochastic programming (4.41) and IOM (4.49), GPM (4.50) and GM (4.50) was used to solve this problem. GP produced 11.39% and 0.70% efficient results as compared to IOM subject to Y_1 and Y_2 for $\delta = 0.5$ and $\acute{C} = 45$ as given in Table 4.18. GM gave 11.02% and 10.67% precise estimates of population means as compared to IOM according to Y_1 and Y_2 for $\delta = 0.5$, $\acute{C} = 49$ and $(W_1, W_2) = (0.4, 0.6)$ as showed in Table 4.19.

In chapter 5, We discussed allocation of a given sample size to various strata taking coefficients of variation and cost of sample survey as variables. Section 5.2 discussed the allocation that maximize precision of regression estimator of population means of several study variables and minimize total traveling cost of sample survey under a given sample size. The allocation problem was formulated as integer mathematical programming problem (5.2) and IOM (5.3) GPM (5.4), weighted method (5. 5) and GM (5.6) were used to solve this allocation problem. Tables 5.1-5.7 displayed the coefficients of variation and total travel cost obtained using proposed optimization methods. IOM gave total travel cost $C = 1073.79$ units and total coefficients of variation $Z_1 + Z_2 = 0.06488$ according to Y_1, total travel cost $C = 1216$ units and total coefficients of variation $Z_1 + Z_2 = 0.06201$ according to Y_2 and total travel cost $C = 1047$ units and total coefficients of variation $Z_1 + Z_2 = 0.06919$ according to cost function C for $\delta = 1.5$ and $\acute{C} = 40$ as showed in Table 5.1. The GPM and weighted method gave same results as IOM according to cost function C as displayed in the Tables 5.2-5.3 respectively. The GM gave total travel cost $C = 1144.63$ units and total coefficients of variation $Z_1 + Z_2 = 0.06093$ according to Case-1 as showed in Table 5.4, total travel cost $C = 1146.09$ units and total coefficients of variation $Z_1 + Z_2 = 0.06048$ according to Case-2 as presented in Table 5.5 and total travel cost $C = 1148.18$ units and total coefficients of variation $Z_1 + Z_2 = 0.06837$ according to Case-3 as demonstrated in Table 5.6. The problem of compromise allocation to minimize the coefficients of variation and nonlinear cost function which consists measurement unit cost and travel cost under given sample size discussed in section 5.2. Table 5.7-5.12 showed the coefficients of variation and total nonlinear cost obtained using IOM, GPM, weighted method and GM with three cases. The proposed GM provide better solution to compromise between cost of sample survey and precision of estimates of population means of several characteristics.

We conclude that proposed multi-objective optimization methods provide efficient compromise solution of allocation problems in stratified sampling scheme when more than one characteristics of interest was studied under different deterministic and probabilistic cost functions. The proposed general method provide compromise allocation that maximize precision of estimate of population mean and minimize variable cost of survey in multivariate stratified sampling design. We achieved different level of precision in each strata by using additional condition ($l_h \leq n_h \leq u_h$). The multi-objective optimization technique can be used to solve allocation problems, especially conflicting in nature, for other sampling designs.

Bibliography

Ahsan, M. J., Najmussehar and Khan, M. G. M. (2005): Mixed allocation in stratified sampling. *Aligarh Journal of Statistics,* 25, 98-97.

Ahsan, M. J. and Khan, S. U. (1977): Optimum allocation in multivariate stratified random sampling using prior information. *Journal of the Indian Statistical Association*, 15, 57-67.

Ahsan, M. J. and Khan, S. U. (1982): Optimum allocation in multivariate stratified random sampling with overhead cost. *Metrika,* 29, 71-78.

Ahsan, M. J. (1975-1976): A procedure for the problem of optimum allocation inmultivariate stratified random sampling stratified random sampling. *The Aligarh Bulletin of Mathematics,* 5-6, 37-42.

Ahsan, M. J. (1978): Allocation problem in multivariate stratified random sampling. *Journal of the Indian Statistical Assocociation,* 16, 1-5.

Ali, I., Raghav, Y. S. and Bari, A. (2011): Compromise allocation in multivariate stratified surveys with stochastic quadratic cost function. *Journal of Statistical Computation and Simulation,* 1, 1-15.

Ansari, A. H., Najmussehar and Ahsan, M. J. (2009): On multiple response stratified random sampling design. *Journal of Statistical Society Kolkata, India,* 1(1), 45-54.

Aoyama, H. (1954): A study of stratified random sampling. *Annal of the Institute of Statistical Mathematics*, 6, 1-36.

Aoyama, H. (1963): Stratified random sampling with optimum allocation for multivariate population. *Annal of the Institute of Statistical Mathematics,* 14, 251-258.

Arthanari, T. S. and and Dodge, Y. (1981): *Mathematical Programming in Statistics. A Wiley-Interscience Publication, John Wiley and Sons, Inc.*

Bethel, J. (1985): An optimum allocation algorithm for multivariate surveys. Proceding of the Survvey Research Methods Section. *Amercian Statistical Association,* 209-212.

Bethel, J. (1989): Sample allocation in multivariate surveys. *Survey Methodology*, 15, 47-57.

Buhler, W. and Deutler, T. (1975): Optimal stratification and grouping by dynamic programming. *Metrika,* 22, 161-175.

Chankong, V. and Haimes, Y. Y. (1983): *Multi-objective Decision Making Theory and Methodology. Elsevier Science Publishing,New York.*

Charnes, A., Cooper, W. W. and Ferguson, R. O. (1955): Optimal estimation of executive compensation by linear programming. *Management Science,* 1, 138-151

Charnes, A. and Cooper, W. W. (1961): *Management Models and Industrial Applications of Linear Programming. John Wiley and Sons, New York.*

Charnes, A. and Cooper, W. W. (1977): Goal programming and multiple objective optimization; part 1. *European Journal of Operation Research*, 1, 39-54.

Charnes, A., Clower, R. W. and Kortanek, K. O. (1967): Efective control through coherent decentralization with preemptive goals. *Econometrica*, 35, 294-320.

Chatterjee, S. (1967): A note on optimum allocation. *Scandinavian Actuarial Journal*, 50, 40-44.

Chatterjee, S. (1968): Multivariate stratified surveys. *Journal of the Amercian Statistical Association*, 63, 530-534.

Chromy, J. R. (1987): *Design optimization with multiple objectives. Proceeding of the Survey Research Methods Section, Amercian Statistical Association, 194-199.*

Cochran, W. G. (1977): *Sampling Techniques, Jhon Wiley and Sons: New York.*

Dalenius, T. and Gurney, M. (1951): The problem of optimum stratification. *Scandinavian Actuarial Journal*, 34, 133-148.

Dalenius, T. (1950): The problem of optimum stratification-II. *Scandinavian Actuarial Journal*, 33, 203-213.

Dalenius, T. (1957): *Sampling in Sweden: Contributions to themethods and theories of sample survey practice. Almqvist and Wiksell, Stockholm.*

Dalenius, T. and Hodges, J. L. (1959): Minimum variance stratification. *Journal of the Amercian Statistical Association,* 54, 88-101.

Diaz-Garcia, J. A. and Cortez, L. U. (2006): Optimum allocation in multivariate stratified sampling: multi-objective programming. *Comunicacion Tecnica No. I (PE/CIMAT)*

Diaz-Garcia, J. A. and Cortez, L. U. (2008): Multi-objective optimisation for optimum allocation in multivariate stratified sampling. *Survey Methodology*, 34(2), 215-222.

Edwards, W. (1950, 1986): *Some Theory of Sampling. John Wiley and Sons, New York.*

Ekman, G. (1959): Approximate expression for conditional mean and variance over small intervals of a continuous distribution. *Annal of Mathematical Statistics*, 30(4),1131-1134.

Folks, J. L. and Antle, C. E. (1965): Optimum allocation of sampling units to strata when there are R responses of interest. *Journal of the Amercian Statistical Association*, 60, 225-233.

Ghosh, S. P. (1958): A note on stratified random sampling with multiple characters. *Calcutta Statistical Association Bulletin*, 8, 81-89.

Ghufran, S., Khowaja, S. and Ahsan, M. J. (2011): multi-objective optimum allocation problem with probablistic non linear cost constraint. *International Journal of Engeneering, Science and Tecnnology,* 3(6), 135-145.

Ghufran, S., Khowaja, S. and Ahsan, M. J. (2012): Optimum multivariate stratified sampling designs with travel cost: A multi-objective integer nonlinear programming approach. *Communication in Statistics-Simulation and Computation,* 41, 598-610.

Guddat, J., Guerre, V. F., Tammer, K. and Wendler, K. (1985): *Multi-objective and stochastic optimization based on parametric optimiztion, Akademie-Verlag, Burlin.*

Haimes, Y. Y., Lasdon, L. S. and Wismer, D. A. (1971): On a bicriterion formulation of the problems of integrated system identification and system optimization. *IEEE Transactions on System, Man and Cybernetics. SMC*-1, 296-297.

Hansen, M. H., Hurwitz, W. H. and Madow, W. G. (1953): *Sampling Survey Methods and Theory, Vol. 1. John Wiley and Sons: New York.*

Hedlin, D. (2000): A procedure for stratification by an extended Ekman rule. *Journal of Official Statistics*, 16, 15-29.

Hestenes, M. R. (1969): Multiplier and gradient methods. *Journal of Optimization Theory and Application,* 4. 303-320.

Horgan, J. M. (2006): Stratification of skewed populations. *International Statistical Review*, 74(1), 67-76.

Ijiri, Y. (1965): *Management Goals and Accounting for Control, Amsterdam: North-Holland.*

Jahan, N. and Ahsan, M. J. (1995): Optimum allocation using separable programming. *Dhaka University Journal of Science*, 43(1), 157-164.

Jahan, N., Khan, M. G. M. and Ahsan, M. J. (1994): A generalized compromise allocation. *Journal of the Indian Statistical Association*, 32, 95-101.

Jahan, N., Khan, M. G. M. and Ahsan, M. J. (2001) : Optimum compromise allocation using dynamic programming. *Dhaka University Jounal of Science,* 49(2), 197-202.

Keskinturk, T. and Er, S. (2007): A genetic algorithm approach to determine stratum boundaries and sample sizes of each stratum in stratified sampling. *Computational Statitical and Data Analysis*, 52, 53-67.

Khan, M. G. M., Ahmad, N. and Khan, S. (2009): Determining the optimum stratum boundries using mathematical programming. *Journal of Mathematical Modeling and Algorithms,* 8, 409-423.

Khan, M. F., Ali, I., Raghav, Y. S. and Bari, A. (2012): Allocation in multivariate stratified surveys with nonlinear random cost function. *Amercian Journel of Operation Research,* 2, 100-105.

Khan, M. G. M., Ahsan, M. J. and Jahan, N. (1997): Compromise allocation in multivariate stratified sampling: an integer solution. *Navel Research Logistics,* 44, 69-79.

Khan, M. G. M., Khan, E. A. and Ahsan, M. J. (2003): An optimal multivariate stratified sampling design using dynamic programming. *Australian and New Zealand Journal of Statistics*, 45(1), 107-113.

Khan, M. G. M., Khan, E. A. and Ahsan, M. J. (2008): Optimum allocation inmultivariate stratified sampling in presence of non-response. *Journal of the Indian Society of Agricultural Statistics*, 62(1), 42-48.

Khan, E. A., Khan, M. G. M. and Ahsan, M. J. (2002): Optimum stratification: a mathematical programming approach. *Calcutta Statistical Association Bulletin, 52(Special volume)*, 323-333.

Khowaja, S., Ghufran, S. and Ahsan, M. J. (2011): Estimation of population means in multivariate stratified random sampling. *Communication in Statistics-Simulation and Computation*, 40, 710-718.

Khowaja, S., Ghufran, S. and Ahsan, M. J. (2011): Multi-objective optimization for optimum allocation in multivariate stratified sampling with quadratic cost. *Journal of Statistical Computation and Simulation,* 1, 1-10.

Kokan, A. R. and Khan, S. U. (1967): Optimum allocation in multivariate surveys: an analytical solution. *Journal of Royal Statistical Society B*, 29(1): 115-125.

Kozak, M. (2004): Optimal stratification using random search method in agricultural surveys. *Statistics in Transition,* 6(5), 797-806.

Kozak, M. (2006a): On sample allocation in multivariate surveys. *Communication in Statitics- Simulation and Computation,* 35, 901-910.

Kozak, M. (2006b): Multivariate sample allocation: application of random search method. *Statistics in Transition,* 7(4), 889-900.

Kreienbrock, L. (1993): Generalized measures of dispersion to solve the allocation problem in multivariate stratified random sampling. *Communications in Statistics- Theory and Methods*, 22(1), 219-239.

Mahalanobis, P. C. (1952): Some aspect of the design of sample surveys. *Sankhya*, 12, 1-17

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 Royal Statistical Society*, 97(4), 558-625.

Pirzada, S. and Maqbool, S. (2003): Optimum allocation in multivriate sampling through chebyshev approximation. *Bulletin of the Malaysian Sciences Society*, 26, 221-230.

Rahim, M. A. (1995): *Use of distance function to optimize sample allocation in multivariate surveys: a new perspective. Proceeding Survey Research Methods Section, Amercian Journal of Statistical Assocociation,* 365-369.

Semiz, M. (2004): Determination of compromise integer strata sample sizes using goal programming. *Hacttepe Journal of Mathematics and Statistics*, 33: 91-96

Singh, S. (2003): *Advanced Sampling Theory with Applications. Kluwer Academic Publisher.*

Steuer, R. E. (1989): *Multiple Criteria Optimization: Theory, Computation, and Application. Krieger Publishing.*

Stuart, A. (1954): A simple presentation of optimum sampling results. *Journal of Royal Statistical Society B*, 16(2), 239-241.

Sukhatme, P. V., Sukhatme, B. V., Sukhatme, S. and Asok, C. (1984): *Sampling Theory of Surveys with Application. Iowa State University Press.*

Thompson, M. E. (1997): *Theory of Sample Surveys. Chapman and Hall.*

Varshney, R. and Ahsan, M. J. (2011): Compromise mixed allocation in multivariate stratified sampling. *Journal of the Indian Society of Agricultural Statistics,* 65(3), 291-296.

Varshney, R., Najmussehar. and Ahsan, M. J. (2012): Estimation of more than one parameters in stratified sampling with fixed budget. *Mathematical Methhod of Operation Research*, 75(2), 185-197.

Yates F. (1960): *Sampling Methods for Censuses and Surveys. Charles Griffin and Co Ltd., London.*