OPTIMIZATION IN MULTIVARIATE SAMPLING

THESIS

SUBMITTED FOR THE AWARD OF THE DEGREE OF

Doctor of Philosophy

IN

STATISTICS

BY

MOHAMMAD VASEEM ISMAIL

UNDER THE SUPERVISION OF

PROFESSOR SANAULLAH KHAN

DEPARTMENT OF STATISTICS AND OPERATIONS RESEARCH
ALIGARH MUSLIM UNIVERSITY
ALIGARH (INDIA)

2002

Dedicated

to my

Beloved Parents

Phone No.: (O) 0571-2701251
(R) 0571-2403782
E-mail : prof_suk@rediffmail.com

Prof. Sanaullah Khan
M.Sc.(Alig.), Ph.D. (Paris)

DEPARTMENT OF STATISTICS & OPERATIONS RESEARCH
ALIGARH MUSLIM UNIVERSITY
ALIGARH-202002 (INDIA)

Dated 27/12/02

Certificate

I Certify that material contained in this thesis entitled "OPTIMIZATION IN MULTIVARIATE SAMPLING" submitted by MOHAMMAD VASEEM ISMAIL for the award of the degree of "DOCTOR OF PHILOSOPHY" in Statistics is original.

The work has been done under my supervision. In my opinion the work contained in this thesis is sufficient for consideration of the award of a Ph.D. degree in Statistics.

(PROF. SANAULLAH KHAN)
Supervisor

ACKNOWLEDGEMENT

Praise and thanks be to Allah, first and last, Lord and Cherisher of all the worlds who taught humankind everything they knew not.

I would like to express my deep thanks to my academic supervisor Professor Sanaullah Khan, Department of Statistics & Operations Research, Aligarh Muslim University, Aligarh, India. He suggested me to work on this field and throughout the work his meticulous and perceptive criticism and unfailing patient and enthusiastic encouragement have made the success of this attempt possible.

I am extremely grateful to Prof. S. Rehman, Chairman, Department of Statistics and Operations Research who very kindly provided me with necessary facilities to complete this work. I am also grateful to Prof. (Late) Zaheeruddin, Prof. A. H. Khan, Prof. M. Z. Khan, Dr S. N. Alam, Dr M. J. Ahsan for their encouragement which was a constant source of inspiration to me.

I am grateful to Professor T. P. Tripathi, Stat-Math Unit, Indian Statistical Institute, Calcutta for his valuable criticism, constructive discussion and important suggestions on my work. I also acknowledge his hearty co-operation in providing me with all his valuable reference materials on this area. To his generosity my gratitude is unbounded.

I am also thankful to all teaching and non-teaching members of the department for their kind help and co-operation.

I am grateful to the Aligarh Muslim University, Aligarh, India for awarding me University Research Fellowship for this course of study.

I would like to accord special thanks to Mr. Abdul Razzaq, my friend and colleague for his willing sacrifice of time and energy, prompt response to inquires and stimulating statistical discussions on many aspects of this thesis. He is by nature unassuming and generous.

I wish to express my gratitude to my friends and colleagues Dr. Qazi Shoeb, Miss Najmussaher and Miss Gulshan for encouragement and moral support.

A substantial amount of credit goes to Mr. Azhar, Mr. Kamal, Mr. Wasim, Mr. Babar, Mr. Kashif, Mr. Aftab and particularly little Fahal, Rida and Aiman for their forbearance, love and encouragement.

I hereby express my sincere gratitude to all my family members, my parents, my brothers and my sisters for their constant encouragement, assistance and invocation which enable me to reach this stage of my life and to complete this work.

Finally, my words of thanks are also due to Kr. Naseem Ahmad and Mr. Modabbir Azam of Global Services, Aligarh for typing this thesis.

Vaseem

(MOHAMMAD VASEEM ISMAIL)

PREFACE

This thesis entitled "**OPTIMIZATION IN MULTIVARIATE SAMPLING**" is submitted to the Aligarh Muslim University, Aligarh, India, to supplicate the degree of **Doctor of Philosophy** in **Statistics**. It consists of the research work carried out by me in the Department of Statistics and Operations Research, Aligarh Muslim University, Aligarh, India.

The fundamental problem of optimization is to arrive at the best possible decision in any given set of circumstances. However, sometimes what is best for one person is worst for another and more often we are not at all sure what is meant by best. The first step, therefore, in mathematical optimization is to choose some quality, typically a function of several variables, to be maximized or minimized, subject possibly to one or more constraints. The next step is to choose a mathematical method to solve the optimization problem; such methods are usually called optimization techniques or algorithms.

The problem of deriving statistical information on the population characteristics, based on sample data, can be formulated as an optimization problem in which we wish to minimize the cost of the survey, which is a function of the sample size, size of the sampling unit, the sampling scheme and the scope of the survey, subject to the restriction that the loss in precision arising out of making decisions on the basis of the survey results is within a certain prescribed limit. Or alternatively, we may minimize the loss in precision, subject to the restriction that the cost of the survey is within the given budget. Thus we are interested in finding the optimal sample size and the optimal sampling scheme which will enable us to obtain estimates of the population characteristics with prescribed properties.

In stratified sampling the population is first divided into groups called strata. These strata are mutually exclusive and exhaustive. Independent simple random samples are then drawn from these strata.

The procedure of stratified sampling is intended to give a better cross-section of the population than that of unstratified sampling. It follows that one would expect the precision of the estimates of the population characteristics to be higher in stratified than in unstratified sampling. Stratified sampling is also convenient in other ways like the selection of sampling units, the location and enumeration of the selected units, distribution and supervision of field-work. In general the whole administration of the survey is greatly simplified in stratified sampling.

An important problem in stratified sampling is the determination of sample sizes (allocation) for different strata. They may be chosen to minimize the sampling variance of the estimator for a fixed cost or to minimize the total cost of the survey for a desired precision. Such an allocation is called an optimum allocation.

The solution of the above problem for univariate case i.e. when a single characteristic is studied on each and every population unit, exits in sampling literature. However, the multivariate case is more complicated and few attempts have been made to attack the problem so far.

In multivariate sample surveys where more than one population characteristics are under study, the optimum allocation of the sample sizes to various strata becomes complicated due to the fact that an allocation that is optimal for one characteristic may be far from optimal for other characteristics.

In this thesis we have formulated some problems arising in multivariate sample survey designs as multiobjective convex programming problems. Attempt has also been made to develope procedures to solve these problems using Chebyshev goal programming approach.

The thesis consists of five chapters. Chapter-I provides an introduction to Multivariate Stratified Sampling, Optimization, Multiobjective Programming, Chebyshev and Fuzzy Goal Programming and also a brief history of the use of Auxiliary Information in Multivariate Sample Surveys.

In Chapter-II, we formulate the multiple character problems arising in the areas of Stratified Random Sampling, Two-stage Sampling, Double Sampling and Response Errors as multiobjective convex programming problems.

A solution procedure is developed in Chapter-III for the multiobjective convex programming problem by linearizing the convex objective functions at the respective optimal points obtained by minimizing the individual objective functions. The multiobjective linear problem is then solved by Chebyshev goal programming approach. A numerical example is also presented.

In Chapter-IV, we represent the allocation problem with multiple characters as a convex programming problem with several linear objective functions and a single convex constraint. The cutting plane technique is used for linearizing the single convex constraint and then the optimum allocation is obtained by using Chebyshev goal programming approach. A comparison has also been made with the fuzzy programming solution. A numerical example is solved to illustrate the procedure.

In Chapter-V, we discuss the simultaneous estimation of several finite population means under stratified sampling design, in the situations where mean vector of the auxiliary variables is known. An optimum estimator by using the criterion of preference coined by Tripathi and Chauby (2000) has been obtained.

CONTENTS

CHAPTER-I

INTRODUCTION

1.1 MULTIVARIATE STRATIFIED SAMPLING

Sampling theory deals with the problems associated with the selection of samples from a population according to certain probability mechanism. The purpose of survey is to obtain information about the population which is defined according to the aims and objects of the survey. Since the information on population is based on sample data, a stage is always reached in planning of a sample survey, at which a decision must be made about the size of the sample, size of the sampling unit, the sampling scheme, the scope of the survey, number of strata and stratum boundaries etc. These decisions have much significance, e.g., the decision regarding the size of sample to be selected is important because too large a sample implies a waste of resources and too small a sample diminishes the utility of the results obtained. The problem of deriving the maximum statistical information on a population characteristic has been formulated as an optimization problem by minimizing the cost of the survey subjected to the restriction that the loss of precision is within a certain prescribed limit or alternately by minimizing the loss in precision subject to the restriction that cost of the survey remains within the given budget.

Stratified sampling is the most popular among various sampling designs that are extensively used in sample surveys. The problem of determining the number of strata, the problem of cutting the stratum boundaries, the problem of optimum allocation of sample sizes to various strata are treated as optimization problems and solved by several authors.

In multivariate stratified sampling where more than one-population characteristics are to be measured on every selected unit, the above problems become more complicated because of the non availability of a single optimality criterion which is suitable for all the characteristics.

The problem of sample allocation in multivariate stratified sampling has drawn attention of researchers for long time starting apparently with Neyman (1934). It is felt that unless the strata variances for various character are distributed in the same way, the classical Neyman allocation based on the variances of a single character is of no use because an allocation which is optimum for one characteristic may not be acceptable for another. For this reason, there is no unique or even widely accepted solution to the problem of optimum allocation in multivariate stratified sampling. One way to resolve this problem is to search for a compromise allocation, which is in some sense optimum for all the characters.

Cochran (1963) suggested the use of the average of individual optimum allocations for various characters. Chatterjee (1967) worked out a compromise allocation by minimizing the sum of the proportional increases in the variances due to the use of non-optimum allocations. Both the above authors have assumed that the measurement cost with respect to the various characters in a particular stratum is constant.

The first author to give the convex programming formulation to the allocation problem in multivariate stratified sampling was Kokan (1963). Kokan and Khan (1967) derived an analytical solution to this convex programming problem. They also showed how the sample allocation problem in other designs such as two-stage sampling, double sampling and response errors can be viewed as a convex programming problem. Chatterjee (1968) also considered the allocation problem for multivariate stratified surveys. An integer solution to this problem was given by Khan and Bari (1977).

Roy, B (1971) defined an unique objective function when a precise weight is known for each character in a survey. In the absence of such apriori knowledge of relative weights a problem cannot be exactly transformed to give a unique objective function and hence a best compromising solution.

The optimum allocation in multivariate stratified sampling using prior information about the population means within stratum can be obtained by assigning an L-variate normal prior distribution to the vector of within stratum population means, where L denotes number of strata. Ericsion (1965) stated the problem as to "minimize the posterior variance of the overall population mean subject to a total budgetary constraint". He also discussed the case when more than one population characteristics are to be estimated, under the assumption that the strata are sufficiently similar with respect to the various characteristics. Soland (1967) also treated the case of multivariate stratified sampling when there is prior information concerning the unknown stratum means of all the variates. He discussed the stratification problem proposed by Dalenius (1953) and formulated it as a non-linear programming problem and also formulated other multivariate stratified sampling problems that may be solved by non-linear programming.

Ahsan and Khan (1977) considered the multivariate allocation problem where the prior information about the unknown within stratum means of p characters is available in terms of a multivariate normal distribution with known parameters. Ahsan and Khan (1982) treated this problem by considering the posterior variances of the population means when the sampling is multipurpose.

Chaddha et.al. (1971) used dynamic programming technique to find the optimum allocation in univariate case. Omule (1985) used the same technique to obtain compromise allocation for multivariate case by minimizing the total cost of the survey when the sampling variances of the estimates of various characteristics are subjected to specified tolerances limits. Jahan et. al.(1994) applied the dynamic programming technique for obtaining the compromise allocation by minimizing the total relative increase in the variances as compared to the optimum allocation, when the costs for measuring the various characteristics are fixed in advance. Khan (1997) treated the multivariate problem as a multi-stage decision problem, in which the k-th stage of the solution provides the sample size for the k-th stratum.

Bethal (1989) expresses the optimal multi-character stratified sample allocation as a closed expression in terms of normalized lagrangian multipliers whereas Rahim (1994) proposed an alternative procedure based on distance function of the sampling errors of all the estimates. Various authors like Nandi and Aich (1995), Chernyak and Starytskyy (1998), Chernyak and Chornous (2000) either suggested new criteria or explored further the already existing criteria.

In chapter II of this thesis, we formulate the multiple character problems arising in the areas of Stratified Random Sampling, Two-Stage Sampling, Double Sampling and Response Errors as multiobjective convex programming problems with convex objective functions and linear constraints.

A solution procedure for the multiobjective convex programming problem formulated in chapter II is developed in chapter III by linearizing the convex objective functions at the respective optimal points when single objective is considered. The multiobjective problem is then solved by Chebyshev goal programming approach.

In chapter IV, we transform the allocation problem with multiple characters into a convex programming problem with several linear objective functions and a single convex constraint. The cutting plane technique is then used for linearizing the single convex constraint and then the optimum allocation is obtained again by using the Chebyshev goal programming approach.

1.2 OPTIMIZATION

The fundamental problem of optimization is to arrive at the best possible decision in any given set of circumstances. Of course, many situations arise where the best is unattainable for one reason or another; sometimes what is best for one person is worst for another; more often we are not at all sure what is meant by best. The first step, therefore, in mathematical optimization is to choose some quantity, typically a function of several variables, to be maximized or minimized, subject possibly to one or more constraints. The commonest type of constraints are equalities and inequalities which must be satisfied by the variables of the problem, but many other types of constraints are possible; for example a solution in integers may be required. The next step is to choose a mathematical method to solve the optimization problem; such methods are usually called optimization techniques or algorithms.

The theory and practice of optimization has developed rapidly since the advent of electronic computers in 1945. It came of age as a subject in the mathematical curriculum in the 1950's, when well established methods of the differential calculus and the calculus of variations were combined with the highly successful new techniques of mathematical programming which were being developed at that time.

The optimization problems that have been posed and solved in the recent years have tended to become more and more elaborate, not to say abstract. Perhaps the most outstanding example of the rapid development of optimization techniques occurred with the introduction of Dynamic programming by Bellman in 1957 and of the maximum principle by Pontryagin in 1958. The techniques were designed to solve the problems of the optimal control of dynamical systems.

The simply stated problem of maximizing or minimizing a given function of several variables attracted the attention of many mathematicians over the past fifty years or so for developing the solution techniques under mathematical programming.

1.3 MATHEMATICAL PROGRAMMING

A mathematical programming problem (MPP) can be stated as follows:

Maximize (or minimize) $Z = f(x_1, x_2, \ldots, x_n)$ (1.3.1)

Subject to the constraints

$$g_i(x_1, x_2, \ldots, x_n) \{\le, =, \ge\} 0 \; ; \; i = 1,2,\ldots,m \quad (1.3.2)$$

and $$x_j \ge 0; \; j = 1,2,\ldots,n \quad (1.3.3)$$

where in (1.3.2) one and only one sign among { $\le$, $=$, $\ge$ } holds true for each i. Usually, unless specified otherwise, in an MPP all the involved functions are assumed to be continuously differentiable.

The variables x_j, $j = 1,2,\ldots,n$ are called decision variables, the function $Z = f(x_1, x_2, \ldots, x_n)$ in (1.3.1) is called objective function, the conditions (1.3.2) are called the constraints and the additional restrictions in (1.3.3) are called non-negativity restrictions. Often

(1.3.3) is also included in (1.3.2) and the MPP takes a more simple expression as:

Maximize (or minimize) $f(\underline{x})$

Subject to $\quad g_i(\underline{x})\{\leq,=,\geq\}0;\ i=1,2,\ldots,m$

where $\underline{x}'=(x_1,x_2,\ldots,x_n)$ is the vector of decision variables.

To develop the theory of mathematical programming either of the maximization or minimization problems may be taken as standard form because of the simple reason that maximization of $f(\underline{x})$ is equivalent to minimization of $-f(\underline{x})$ and vise-versa. Furthermore all the constraints can be described with $\leq or = or \geq$ by simple operations of multiplying by -1 and /or addition or subtraction of some slack or surplus variables defined to have a ≥ 0 value and noting that an equation is equivalent to two inequalities, one with $\leq$ and the another with $\geq$ sign. Thus we may transform any given MPP in the following form:

$$\left.\begin{array}{ll} \text{Minimize} & f(\underline{x}), \\ \text{Subject to} & g_i(\underline{x})\geq \underline{0}, \\ \text{and} & \underline{x}\geq \underline{0}. \end{array}\right\} \quad (1.3.4)$$

Any $\underline{x}$ satisfying the constraints and non-negativity restriction to an MPP is called a feasible solution to the MPP. The set of all feasible solutions to an MPP is usually denoted by F. Thus the set F for MPP (1.3.4) is $F = \{\underline{x} \mid \underline{g}(\underline{x}) \geq \underline{0}, \underline{x} \geq \underline{0}\}$.

Any $\underline{x}^* \in F$ for which $f(\underline{x}^*) \leq f(\underline{x})$ for all $\underline{x} \in F$ is called an optimum solution for a minimization MPP.

The optimal value $\underline{x}^*$ of the decision variables is the function of various parameters appearing in MPP, such as: the availability of resources, costs or profits and technological coefficients (coefficients of decision variables in constraint functions). If some or all of the parameters of an MPP are stochastic variables rather than deterministic quantities then the MPP is called a Stochastic Programming Problem.

If all the functions involved in an MPP are linear functions of decision variables the MPP is called a Linear Programming Problem (LLP). On the other hand if some or all the functions are nonlinear, the MPP is called a Nonlinear Programming Problem (NLPP).

Depending upon the nature of the involved functions, restrictions on the decision variables and the objectives function(s), an MPP (Linear and /or Nonlinear) can further be placed in one or more of the several classes such as Integer Progamming Problem (IPP), Quadratic Programming Problem (QPP), Convex Programming Problem (CPP), Separable Programming Problem (SPP), Geometric Progamming Problem (GPP) and Multiobjective Progrmming Problem (MOPP).

1.4 MULTIOBJECTIVE OPTIMIZATION

The single-objective approach had been so ignored, and so widely accepted, that it may seem hard to believe that it has only seen widespread use since 1947. Further, it is easy to forget the fact that in 1947 the very notion of even a single-objective function was considered quite revolutionary. Specifically, until the development of LP, the typical mathematical model consisted of either a system of equations or a system of inequalities and, for the most part, one's attention was directed toward the determination of just a feasible solution (i.e., one that satisfied the system of constraints as opposed to one that both satisfied the constraint set and optimized a single measure of performance). As such, in 1947, the concept of the inclusion of an objective function was considered just as radical as some now view the inclusion of multiple-objective functions.

However, although the consideration of multiple objectives may seem a novel concept, virtually any nontrivial, real word problem invariably involves multiple objectives. For example, the success of an airplane is determined by such things as its cost (to be minimized), payload (to be maximized), speed (to be maximized), maximum range (to be maximized), weight (to be minimized), survivability (to be maximized) etc. And, in the design of an aircraft, we may actually hope to optimize each and every one of these parameters.

In the traditional LP model, each and every constraint is considered to be absolutely rigid. That is, if a solution does not satisfy each and every constraint it is termed infeasible. However, in real-world problems, the notion of strictly rigid constraints does not necessarily hold at least not for every constraint function. In real-world problems, we just may be able to tolerate a certain level of "violation" of a constraint. Such flexible constraints are termed "soft" constraints (or soft goals) and are frequently encountered when we deal with actual problems. Thus, a soft constraint is one that we would like to satisfy, but for which we would be able to accept some degree of "violation". On the other hand, a hard constraint (or hard goal) is one for which any degree of violation would be absolutely intolerable. However, from a traditional LP point of view, such notions as multiple objectives and soft constraints only serve to complicate the situation.

There are several ways in which the multiobjective problem might be modeled.

(i) Conversion to a linear program via objective function Transformation (or deletion)

The traditionalist would most likely decide that, regardless of what management may have stated, a single objective model is going to be employed. Thus, one way to force the problem into the single objective format is to select one of the objectives, use it as the single objective, and then either ignore the other objectives or treat them as (rigid) constraints.

(ii) Conversion to a linear program via Utility Theory:

A Method of Aggregation

Theoretically (and only theoretically), it should be possible to combine any number of objectives into an equivalent, single objective if we can determine a common measure of effectiveness (i.e., a so-called "Proxy") by means of which each of the objectives may be expressed. The basis of such an approach is the aggregation of multiple objectives into a single and, it can be considered, equivalent function.

(iii). Conversion to a Goal Program (GP)

When one employs utility theory, the bulk of one's efforts is typically dedicated toward obtaining an adequate and rational representation of the decision maker's (theoretically) preference function. However, when one uses goal programming the effort shifts toward that of obtaining a better representation of the actual problem, through the development of the goal-programming model. Whichever approach is deemed "best" is strictly a function of one's personal perspective.

There are actually a number of types of goal programs, each espousing a somewhat different philosophy (i.e, with respect to how to measure the "goodness" of a solution to a problem involving multiple, conflicting goals). Three of the most popular (as well as the most practical) forms of GP are Archimedean GP (i.e., weighted GP), Non-Archimedean GP (i.e., lexicographic GP, or preemptive GP), and Chebyshev GP (or Minimax GP, or Fuzzy GP).

To form a goal-programming model, the very first thing that must be done is to convert all objectives into goals. When we convert the objectives into goals, we apply the following guidelines:

A maximizing (or a minimizing) objective is converted into a type II ($\geq$) (or type I ($\leq$)) inequality by means of the establishment and inclusion of a right-hand side, or aspiration level value. Indeed, we convert a goal into a constraint. Specifically, in a typical model, some of the goals will be hard (i.e, they absolutely must be attained) and some will be soft (i.e., some deviation is tolerable). Thus, we need a means to indicate the deviations from the right-hand sides of the constraints corresponding to each goal, whether hard or soft. To accomplish this, we shall add negative deviations and subtract positive deviations from the left-hand sides of each goal (and constraint).

Now, although the model is expressed in terms of goals (where some are hard and some are soft), we next need a function by means of which the achievement of the minimization of the unwanted goal deviations may be measured. This function, in fact, is termed the goal programming achievement function. Further, we need a philosophy upon which to develop such a function. Two approaches are found in the literature.

(a) Archimedean Goal Programming

In Archimedean, or weighted, GP, we shall form an achievement function consisting of precisely two terms. The first term represents the sum of all unwanted deviations for those goals that are hard (i.e., the rigid constraints). The second is composed of the weighted sum of all unwanted deviations for those goals that are soft. Thus, the achievement function for the general Archimedean GP model is given as

$$\text{Lexmin} \quad u = \left\{\left(\mu^{(1)}\eta^{(1)} + \omega^{(1)}\rho^{(1)}\right), \left(\mu^{(2)}\eta^{(2)} + \omega^{(2)}\rho^{(2)}\right)\right\}$$

where

Lexmin= lexicographic minimum (or achievement function)

u= achievement vector (or achievement function)

$\eta^{(k)}$= vector of negative deviations, at priority level k

$\rho^{(k)}$= vector of positive deviations, at priority level k

$\mu^{(k)}$= vector of weights for all negative deviations, at priority level k

$\omega^{(k)}$= vector of weights for all positive deviations, at priority level k.

(b) Non-Archimedean Goal Programming

In non-Archimedean GP (also called lexicographic or preemptive GP) as well, we form an achievement function. However, the number of terms in this achievement function will always be three or more. As before, the first term represents the sum of all unwanted deviations for those goals that are hard (i.e., rigid constraints). The second is composed of the weighted sum of all unwanted deviations for those goals at priority level two. The third is composed of the weighted sum of all unwanted deviations at priority level three, and so on. The general form of the achievement function for a non-Archimedean GP is given as

$$\text{Lexmin} \quad u = \left\{ \left(\mu^{(1)}\eta^{(1)} + \omega^{(1)}\rho^{(1)}\right), \ldots, \left(\mu^{(K)}\eta^{(K)} + \omega^{(K)}\rho^{(K)}\right) \right\}$$

Wherein the total number of priority levels is K (i.e., $k = 1,2,\ldots,K$).

A comprehensive presentation on goal programming and its extensions is given in Ignizio (1976), and a summary of different variations of goal programming is provided in Charnes and Cooper (1977). In addition, a wide survey of literature around goal programming up to the year 1983 is presented in Soyibo (1985).

The short comings and the solution of the goal programming were discussed by Khorramshahgol and Hooshiari (1991), Chakraborty and Sinha (1995), Neelam and Arora (1999), Chakraborty and Dubey (2001).

1.5 CHEBYSHEV GOAL PROGRAMMING

Charnes and Cooper (1961) introduced the idea of Goal Programming. Later Charnes and Cooper (1977) discussed the solution of multiple objective optimization problems through Goal Programming (GP). Ignizio (1983,1985) observed that the Chebyshev GP (or Minimax GP) and Fuzzy GP are (closely) related. Ignizio and Cavalier (1994) have illustrated the procedure of solving the multiobjective linear programming problem through an example by its formulation to Chebyshev Linear Goal Programming (LGP) and compared it by the Fuzzy LGP. They also discussed the Chebyshev multiplex model for solving multiobjective problem.

The Minimax or Chebyshev formulation implies the optimization of a utility function where the maximum deviation is minimized. The underlying philosophy of Chebyshev LGP is to find that solution that serves to minimize the single worst unwanted deviation from any (soft) goal. This particular notion also provides the basis of what is called Minimax GP and Fuzzy programming or Fuzzy GP.

As with any GP approach, the first step is to convert the problem into one containing nothing but goals. Next, we solve the problem as a conventional LP, using but one objective at a time. Once we have solved such a problem, we have determined the best possible value of the objective being considered as an aspiration level. An aspiration level is employed in order to convert an objective into goal. It represents a target level for the given objective- a level that is desired and/or acceptable. The use of aspiration levels to transform objectives (which are to be optimized) into goals (which are to be achieved) is known as the concept of "Satisficing". Satisficing, in turn, is a pragmatic approach based upon the manner in which most organizations, and most individuals, approach real-world decision making (Simon, 1957 and March and Simon, 1958). That is, rather than attempting to achieve solution optimality (which is actually only meaningful for static, deterministic, error free, single objective problems), we hope to find a solution that comes "as close as possible" to satisfying our goals.

Consider the multiobjective linear programming problem

$$\text{Minimize } Z_k, \quad k = 1,2,\ldots,p$$

$$\text{Subject to } \underline{A}\,\underline{x}(\leq, \geq, or =)\underline{b} \tag{1.5.1}$$

$$\underline{x} \geq \underline{0}.$$

The general form of the Chebyshev LGP model may be written as

Minimize δ

Subject to:

$$Z_k - \delta \leq L_k, \text{ for all } p \text{ objectives} \tag{1.5.2}$$

$$\underline{A}\underline{x}(\leq, \geq, or =)\underline{b}, \text{ for all m constraints}$$

$$\delta \geq 0, \underline{x} \geq \underline{0}$$

where

δ = dummy variable representing the worst deviation level

Z_k = a linear function representing the kth objective

L_k= minimum value that Z_k can take on while solving the various LPs in (1.5.1) individually for $Z_1,\ldots,Z_p$.

1.6 FUZZY PROGRAMMING

Zimmermann (1978, 1981) developed fuzzy mathematical programming to solve the problems with several functions. Narasimhan (1980) in one of his papers discussed goal programming in fuzzy environment. Sandipan Gupta and Chakraborty (1997), use the fuzzy programming approach to multiobjective linear programming problems. Several other authors such as Kassem and Ammer (1996), Mohan and Nguyen (1999), Han-Lin Li and Chian-Son Yu (2000), Aghezzaf and Ouaderhman (2000) and Aghezzaf (2001) etc have also discussed the fuzzy programming approach for solving multiobjective fuzzy programming problems.

Like Chebyshev goal programming, the basis of fuzzy programming approach is also to minimize the worst deviation from any (soft) goal. Using Zimmermann's (1978, 1985) approach to fuzzy programming, and assuming that all objectives are of the minimizing type, we may represent the general fuzzy linear programming model as:

Minimize δ (1.6.1)

Subject to:

$$\frac{Z_k - L_k}{d_k} \le \delta, \text{ for all } p \text{ objectives} \qquad (1.6.2)$$

$$\underline{A}\underline{x}(\leq, \geq, or =)\underline{b}, \text{ for all m constraints} \quad (1.6.3)$$

$$\delta \geq 0, \underline{x} \geq \underline{0} \quad (1.6.4)$$

where U_k=maximum value that Z_k can take on while solving the various LPs in (1.5.1) individually for $Z_1,....,Z_p$

L_k = minimum value that Z_k can take on while solving the various LPs in (1.5.1) individually for $Z_1,....,Z_p$.

$d_k = U_k - L_k$

and the left-hand side of (1.6.2) is termed the fuzzy membership function.

The purpose of the fuzzy goal programming approach is to find the solution that serves to minimize the largest fuzzy membership function [worst deviation level (δ)]. However the fuzzy programming model is identical to the Chebyshev programming model except for the weight given to δ.

In the multiobjective allocation problem, there are p non-linear objective functions which later turn into soft goals with a single linear constraint (hard goal). To apply Chebyshev/Fuzzy goal programming approach, all the hard and soft goals must be in linear form so that the worst deviation from the approximated linear goals is minimized. We thus approximate the non-linear soft goals by linear ones and use the linearized soft goals for minimizing the worst deviation in finding the Chebyshev/Fuzzy point. The aspiration levels being used in the Chebyshev/Fuzzy goal programming approach are taken as the optimal values of the respective non-linear programming problems instead of those of the linear programming problems.

1.7 AUXILIARY INFORMATION IN SAMPLE SURVEYS

This section presents the developments related to the utilization of auxiliary information in sample surveys for estimating the population means.

The works of Bowley (1926) and Neymen (1934,1938) can be referred to as the initial efforts to utilize the auxiliary information in sampling theory. The works of Watson (1937) and Cochran (1940,1942) initiated the use of auxiliary information in devising estimation procedures aimed at improvement of the precision of estimation. Hansen and Hurwitz (1943) were the first to suggest the use of auxiliary information to selecting the units with varying probabilities.

In most of the survey situations, the auxiliary information is always available in one form or the other or it can be made available by diverting for this purpose a part of survey resources at moderate cost. In whatever form the auxiliary information is available, one may always utilize it to devise sampling strategies which are better (if not uniformly then at least in a part of parametric space) than those in which no auxiliary information is used. The method of utilizing auxiliary information depends on the form in which it is available.

In sample surveys, the auxiliary information may be utilized in three basic ways [Tripathi (1970,1973,1976)]:

(i) The information on one or more auxiliary variables may be used at the planning or designing stage of the survey. For example, one may stratify the population according to the frequency distribution of an auxiliary variable.

(ii) The information on one or more auxiliary variables may be used at the sample selection stage of the survey i.e., in selecting units for sample with or without replacement and with varying probabilities proportional to some suitable measure of size

(iii) The information on one or more auxiliary variables may be used at the estimation stage e.g., through defining ratio, regression, difference and product estimators based on the auxiliary information.

The auxiliary information may also be used in mixed ways as well by combining any two or all of the above three basic ways.

The univariate ratio and regression estimators [Cochran (1940,1942)], difference estimator [Hansen et al. (1953)] and product estimator [Robson (1957), Murthy (1964)] for population mean of Y based on the knowledge of the population mean of an auxiliary character X are well known in sampling theory, and for their detailed study in the case of simple random sampling without replacement (SRSWOR) and in that of stratified sampling one may refer to the books by Cochran (1977), Sukhatme et al. (1984), Raj (1968), Murthy (1967), Kish (1965) and others.

The univariate ratio, regression, product and difference estimators [Murthy (1967)] for any general sampling design are defined respectively as

$$\overline{Y}_R = \frac{\hat{\overline{Y}}\,\overline{X}}{\hat{\overline{X}}},$$

$$\overline{Y}_{rg} = \hat{\overline{Y}} - \hat{\beta}\left(\hat{\overline{X}} - \overline{X}\right),$$

$$\overline{Y}_p = \frac{\hat{\overline{Y}}\,\hat{\overline{X}}}{\overline{X}},$$

$$\overline{Y}_d = \hat{\overline{Y}} - \lambda\left(\hat{\overline{X}} - \overline{X}\right)$$

where $\hat{\overline{Y}}$ and $\hat{\overline{X}}$ are the unbiased estimators of the population means $\overline{Y}$ and $\overline{X}$ of the estimation and auxiliary variables respectively, λ is a suitably chosen constant and $\hat{\beta}$ is the sample regression coefficient of $\hat{\overline{Y}}$ on $\hat{\overline{X}}$.

Das and Tripathi (1980) and Das (1988) gave the classes of estimators for $\overline{Y}$, for any sampling design, as

$$d_1 = \frac{\overline{Y} - t_1\left(\overline{X} - \overline{X}\right)}{\left[\overline{X} - t_2\left(\overline{X} - \overline{X}\right)\right]^{\alpha}}\left(\overline{X}\right)^{\alpha}$$

and $$d_2 = W\left[\hat{\overline{Y}} - t\left(\hat{\overline{X}} - \overline{X}\right)\right]$$

respectively, where t_1, t_2 and t are suitably chosen constants. The classes of estimators due to Srivastava (1971, 1980) for any general sampling design are given as

$$d_3 = \hat{\bar{Y}}\, h\left(\frac{\hat{\bar{X}}}{\bar{X}}\right)$$

and $$d_4 = g\left(\hat{\bar{Y}}, \frac{\hat{\bar{X}}}{\bar{X}}\right)$$

where h and g are suitably chosen functions.

In Sample Surveys, the use of multivariate auxiliary information in estimating mean $\bar{Y}$ of a study variable y has largely been made in the form of knowledge of population mean of a p-dimensional auxiliary vector.

Olkin (1958) and Raj (1965) extended the univariate ratio and difference estimators to the multivariate case for SRSWOR as

$$\hat{\bar{Y}}_{rm} = \sum_{i=1}^{p} \omega_i\, \alpha_i, \quad \alpha_i = \left(\frac{\bar{y}_n}{\bar{x}_{in}}\right)\bar{X}_i$$

and $$\hat{\bar{Y}}_{dm} = \sum_{i=1}^{p} \omega_i\, \alpha_i, \quad \alpha_i = \bar{y}_n - \lambda_i\left(\bar{x}_{in} - \bar{X}_i\right)$$

respectively, where ω_i's are weights such that $\sum_{i=1}^{p}\omega_i = 1$, $\bar{y}_n$ and $\bar{x}_{in}$ are the means of characters y and x_i based on a sample of n units. $\bar{X}_i$ is the population mean of x_i and λ_i a suitably chosen constant.

Khan and Tripathi (1967) defined the multivariate ratio estimator in double sampling as

$$\bar{y}_{rm} = \sum_{l=1}^{p} w_l \alpha_l , \quad \alpha_l = \left(\frac{\bar{y}_m}{\bar{x}_{lm}} \right) \bar{x}_{ln}$$

and multivariate regression estimator as

$$\bar{y}_{lrm} = \bar{y}_m + \hat{\beta}'_{1 \times p} \left(\underline{\bar{x}}_n - \underline{\bar{x}}_m \right)$$

where $\bar{x}_{ln}$ being mean of x_l based on s(1), $\bar{y}_m$ and $\bar{x}_{lm}$ being means of y and x_l based on s(2); $\underline{\bar{x}}_n = \left(\bar{x}_{1n}, \bar{x}_{2n}, \ldots, \bar{x}_{pn} \right)'$ and $\underline{\bar{x}}_m = \left(\bar{x}_{1m}, \bar{x}_{2m}, \ldots, \bar{x}_{pm} \right)'$.

Tripathi and khattree (1989) discussed the estimation of means of several principal variables under simple random sampling, in the situations where means of several auxiliary variables are known. Further, Tripathi (1989) extended the results to the case of two occasions. Tripathi and chaubey (1993) considered the problem of obtaining optimum probabilities of selection based on auxiliary variables, in PPS sampling for estimating the mean of several variables. Recently, Tripathi and chaubey (2000) discussed the estimation of finite population mean vector $\underline{y}$ of the principal variables, under the general sampling designs, in the situations where mean vector $\underline{x}$ of the auxiliary variable is known.

In chapter V of this thesis, we define the estimator of the finite population mean vector of several principal variables under *stratified sampling design*, in the situations where mean vector of the auxiliary variables is known. An optimum estimator by using the criterion of preference given by Tripathi and Chaubey (2000) has been obtained.

CHAPTER-II

SOME MULTIOBJECTIVE CONVEX PROGRAMMING PROBLEMS ARISING IN MULTIVARIATE SAMPLING

2.1 INTRODUCTION

In multivariate surveys there are more than one population characteristics to be estimated and usually these characteristics are of conflicting nature. The derivation of the optimal sample numbers among various strata or various stages thus requires some special treatment.

In this chapter, we formulate the problems of multivariate sampling arising in the areas of stratified random sampling, two-stage sampling, double sampling and response errors as multi-objective convex programming problems with convex objective functions and a single linear constraint with some upper and lower bounds.

2.2 MULTIVARIATE STRATIFIED SAMPLING

We consider a multivariate population partitioned into L strata. Suppose that p characteristics are measured on each unit of the population. We assume that the strata boundaries are fixed in advance Let n_i be the number of units drawn without replacement from i^{th} stratum ($i=1,2,...,L$). Let N_i be the size of i^{th} stratum. For j^{th} character, an unbiased estimate of the population mean $\overline{Y_j}(j=1,2,...,p)$, denoted by $\bar{y}_{jst}$, has its sampling variance

$$V(\bar{y}_{jst})=\sum_{i=1}^{L}\left(\frac{1}{n_i}-\frac{1}{N_i}\right)W_i^2 S_{ij}^2, \qquad j=1,2,..,p$$

where

$$W_i=\frac{N_i}{N}, \quad S_{ij}^2=\frac{1}{N_i-1}\sum_{h=1}^{N_i}\left(y_{ijh}-\overline{Y}_{ij}\right)^2 .$$

Substituting $a_{ij}=W_i^2 S_{ij}^2$, we get

$$V(\bar{y}_{jst}) = \sum_{i=1}^{L}\frac{a_{ij}}{n_i}-\sum_{i=1}^{L}\frac{a_{ij}}{N_i}, \qquad j=1,2,...,p. \qquad (2.2.1)$$

Let C_{ij} be the cost of enumerating the j^{th} character in the i^{th} stratum and let C be the upper limit on the total cost of the survey. Then assuming linear cost function one should have

$$\sum_{i=1}^{L}\sum_{j=1}^{p} C_{ij} n_i \le C,$$

$$or \sum_{i=1}^{L} C_i n_i \le C, \qquad (2.2.2)$$

where $C_i = \sum_{j=1}^{p} C_{ij}$, the cost of enumeration of all the p characters in the i^{th} stratum

Further one should have

$$1 \le n_i \le N_i, \quad i = 1,2,...,L. \qquad (2.2.3)$$

We determine the optimum values of n_i, by minimizing (in some sense) all the p variances (2 2.1) for a fixed budget (2.2.2) i.e we have to

$$\begin{aligned} &Minimize\ V_j = \sum_{i=1}^{L}\frac{a_{ij}}{n_i} - \sum_{i=1}^{L}\frac{a_{ij}}{N_i}, && j = 1,2,...,p \\ &Subject\ to \quad \sum_{i=1}^{L} C_i n_i \le C && (2.2.4) \\ &and \quad 1 \le n_i \le N_i, && i=1,2, \ ,L \end{aligned}$$

Since N_i's are given, it is enough to minimize

$$V_j = \sum_{i=1}^{L} \frac{a_{ij}}{n_i}, \qquad j = 1,2,\ldots,p$$

Using X_i for n_i, the problem (2.2.4) can be written as the following multiobjective non-linear programming problem:

$$\left.\begin{array}{llll} \textit{Minimize} & V_j = \sum_{i=1}^{L} \frac{a_{ij}}{X_i}, & j = 1,2,\ldots,p & (a) \\ \textit{Subject to} & \sum_{i=1}^{L} C_i X_i \le C & & (b) \\ \textit{and} & 1 \le X_i \le N_i, & i=1,2,\ldots,L & (c) \end{array}\right\} \quad (2.2.5)$$

The objective functions in (2.2.5) are convex [see Kokan and Khan (1967)], the single constraint is linear and the bounds are also linear. The problem (2.2.5) is, therefore a multiobjective convex programming problem

If some tolerance limits, say v_j, are given on variances of the p characters then the allocation problem reduces to the single objective convex programming problem

$$\text{Minimize} \sum_{i=1}^{L} C_i X_i$$

$$\text{Subject to} \sum_{i=1}^{L} \frac{a_{ij}}{X_i} \le v_j, j = 1,2, .., p \qquad (2.2.6)$$

$$1 \le X_i \le N_i, \quad i = 1,2,...,L.$$

2.3 TWO-STAGE SAMPLING

Let us consider a population which consists of N Primary Stage Units (PSU's) and the i^{th} PSU consists of M_i Secondary Stage Units (SSU's). A sample of n PSU's is to be selected and from the i^{th} selected PSU, a sample of m_i SSU's is to be selected

Let us denote

y_{irj} = value obtained for the r^{th} SSU in the i^{th} PSU for j^{th} character

M_i = number of SSU's in the i^{th} PSU, ($i = 1,2,...,N$).

$M_0 = \sum_{i=1}^{N} M_i$ = total number of SSU's in the population.

$\overline{M} = \frac{M_0}{N}$ = average number of SSU's.

$$m_0 = \sum_{i=1}^{n} m_i = \text{total number of SSU's in the sample.}$$

$$\bar{Y}_{ij} = \sum_{r=1}^{M_i} \frac{y_{irj}}{M_i} = \text{the } i^{th} \text{ PSU population mean for } j^{th} \text{ character.}$$

$$\bar{Y}_{Nj} = \sum_{i=1}^{N} \frac{\bar{Y}_{ij}}{N} = \text{the overall population mean of PSU means for } j^{th}$$

character.

$$\bar{Y}_j = \frac{\sum_{i=1}^{N} M_i \bar{Y}_{ij}}{M_0} = \sum_{i=1}^{N} W_i \bar{Y}_{ij} = \text{population mean per SSU for } j^{th}$$

character.

$$\bar{y}_{ij} = \sum_{r=1}^{m_i} \frac{y_{irj}}{m_i} = \text{sample mean per SSU for } j^{th} \text{ character.}$$

$$\bar{y}_j = \frac{\sum_{i=1}^{n} M_i \bar{y}_{ij}}{n\bar{M}} = \text{sample mean per SSU in the } i^{th} \text{ PSU for } j^{th}$$

character.

Define

$$S_{bj}^2 = \frac{\sum_{i=1}^{N}\left(u_i\bar{Y}_{ij} - Y_j\right)^2}{N-1} = \text{population variance between PSU's means for } j^{th} \text{ character.}$$

$$S_{wj}^2 = \frac{\sum_{r=1}^{M_i}\left(y_{irj} - \bar{Y}_{ij}\right)^2}{(M_i - 1)} = \text{population variance within PSU's for } j^{th} \text{ character.}$$

where

$$u_i = \frac{M_i}{\bar{M}}.$$

For j^{th} character ($j = 1,2,...,p$), the unbiased estimate of the population mean $\bar{Y}_j$ is $\bar{y}_j$ which has the sampling variance as

$$V(\bar{y}_j) = \left(\frac{1}{n} - \frac{1}{N}\right)S_{bj}^2 + \sum_{i=1}^{N}\frac{M_i^2}{nN\bar{M}^2}\left(\frac{1}{m_i} - \frac{1}{M_i}\right)S_{wj}^2$$

$$= \frac{1}{n}S_{bj}^2 + \sum_{i=1}^{N}\frac{M_i^2}{nN\bar{M}^2}\frac{S_{wj}^2}{m_i} + \text{constant terms}$$

$$= \frac{a_{0j}}{n} + \sum_{i=1}^{N} \frac{a_{ij}}{nm_i} + \text{constant terms} \qquad (2.3.1)$$

where

$$a_{0j} = S_{bj}^2, \; a_{ij} = \frac{M_i^2}{N\overline{M}^2} S_{wj}^2 .$$

Let C be the upper limit on total cost of the survey Assuming the cost of the survey to be linear, we should have

$$nC_0 + \frac{nC_1}{N} \sum_{i=1}^{N} m_i \leq C \qquad (2.3.2)$$

where C_0 is the average cost of selection per PSU and C_1 is the average cost of sampling per SSU. In practice, C_o is likely to be larger than C_1

Now the problem is to determine the optimum values of n and m_i so as to minimize the variances (2.3.1) of the various characters for a fixed budget C. Ignoring the constant terms in (2.3.1), and using X_o for n & X_i for nm_i, we get the following multiobjective convex programming problem

$$\text{Minimize} \quad V_j = \sum_{i=0}^{N} \frac{a_{ij}}{X_i}, \qquad j = 1,2,\ldots,p$$

$$\text{subject to} \quad \sum_{i=0}^{N} C_i X_i \le C \qquad (2.3.3)$$

$$\text{and} \quad X_o \le N,\ X_i \le NM_i,\ \ i = 1,2,\ldots,N$$

where $C_i = \frac{C_1}{N}$ *for* $i = 1,2,\ldots,N.$

Case of Equal Primary-Stage Units

The equal Primary-Stage Units problem can be considered as a particular case of the unequal Primary-Stage Units problem where $M_i = M$ for $i = 1,2,\ldots,N$.

Let $X_1 = n$ and $X_2 = nm$ then the problem in case of equal primary-stage units reduces to the following multiobjective convex programming problem in only two variables:

$$\text{Minimize} \quad V_j = \sum_{i=1}^{2} \frac{a_{ij}}{X_i}, \qquad j = 1,2,\ldots,p$$

$$\text{Subject to} \quad \sum_{i=1}^{2} C_i X_i \le C \qquad (2.3.4)$$

$$\text{and} \quad X_1 \le N,\ X_2 \le NM$$

2.4 DOUBLE SAMPLING

Consider the problem of double sampling for stratification in which the population is to be stratified into L strata. The first sample of size n' is selected by simple random sampling without replacement to estimate the strata weights. A second sample of n units with n_i units from the i^{th} stratum is selected in which p characters $y_1, y_2, \ldots, y_p$ are observed. In allocating the sample size n to different strata, we use Neyman allocation where $n = \sum_{i=1}^{L} n_i$.

Let $W_i = \frac{N_i}{N}$ be the proportion of population units falling in the i^{th} stratum and $w_i = \frac{n_i'}{n'}$ be the proportion of first sample units falling in the i^{th} stratum. W_i being unknown is estimated by w_i.

Let $\bar{y}_{ij}$ be the sample mean of the j^{th} character in the i^{th} stratum, $i = 1,2,\ldots,L$; $j = 1,2,\ldots,p$ and $\bar{Y}_{ij}$ be the population mean of the j^{th} character in the i^{th} stratum. For j^{th} character $(j = 1,2,\ldots,p)$, an unbiased estimate of the population mean $\bar{Y}_j$, is $\bar{y}_j = \sum_{i=1}^{L} W_i \bar{y}_{ij}$, which, for large populations, has the sampling variance

$$V\left(\bar{y}_j\right)=\sum_{i=1}^{L}\left[W_i^2+\frac{W_i\left(1-W_i\right)}{n'}\right]\frac{S_{ij}^2}{n_i}+\sum_{i=1}^{L}\frac{W_i\left(\bar{Y}_{ij}-\bar{Y}\right)^2}{n'},$$

where

$$S_{ij}^2=\sum_{i=1}^{L}\frac{\left(y_{ijr}-\bar{Y}_{ij}\right)^2}{\left(N_i-1\right)},\quad i=1,2,\ldots,L;\, j=1,2,\ldots,p.$$

For the proportional allocation $n_i = nW_i$, the variance of $\bar{y}_j$ is approximately given by

$$V_j=\frac{1}{n}\sum_{i=1}^{L}W_iS_{ij}^2+\frac{1}{n'}\sum_{i=1}^{L}W_i\left(\bar{Y}_{ij}-\bar{Y}\right)^2,\quad j=1,2,\ldots,p. \tag{2.4.1}$$

An approximate expression of minimum variance under Neyman allocation for j^{th} character is

$$V_j=\frac{v_{1j}}{n'}+\frac{v_{2j}}{n},$$

where $\quad v_{1j}=\sum_{i=1}^{L}W_i\left(\bar{Y}_{ij}-\bar{Y}\right)^2 \text{ and } v_{2j}=\sum_{i=1}^{L}W_iS_{ij}^2,\quad i=1,2.$

Let C be the upper limit on total cost of the survey. Assuming the cost of the survey to be linear, we should have

$$C_1 n' + C_2 n \le C \tag{2.4.2}$$

where C_1 is the cost per unit of measuring the auxiliary variate and C_2 is the cost per unit of measuring all the study variates. C_1 is generally smaller than C_2.

Here it is required to find the values of n' and n so that the total cost does not exceed the given budget and at the same time, the variances for various characters are minimized.

The problem then again reduces to the following multiobjective convex programming problem in two variables:

$$\begin{aligned}
&\text{Minimize} \quad V_j = \sum_{i=1}^{2} \frac{v_{ij}}{X_i}, \qquad j = 1,2,..,p \\
&\text{Subject to} \quad \sum_{i=1}^{2} C_i X_i \le C \\
&\text{and} \quad 1 \le X_i \le N, \qquad i = 1,2
\end{aligned} \tag{2.4.3}$$

where $n' = X_1$ and $n = X_2$.

If the upper tolerance limits $v_j, (j = 1,2,...,p)$ are given on the variances of the various characters and it is required to minimize the cost of the survey, then we get the following single objective problem

$$\text{Minimize} \sum_{i=1}^{2} C_i X_i$$

$$\text{Subject to} \sum_{i=1}^{2} \frac{v_{ij}}{X_i} \le v_j, j = 1,2,\ldots,p \tag{2.4.4}$$

$$1 \le X_i \le N, \quad i = 1,2.$$

2.5 RESPONSE ERRORS

Let an individual be selected at random from the population of N individuals and an interviewer be picked up at random out of M interviewers and assigned to the selected individual. Denote by y_{abc} the response value obtained for c^{th} sample individual by b^{th} sample interviewer in the a^{th} (population) group. The expected value of y_{abc} will be $\bar{Y}$. The sample mean is

$$\bar{y} = \sum_{a=1}^{L} \sum_{b}^{k_a} \sum_{c}^{na} y_{abc} \cdot$$

In many surveys, interviewers are available to interview only certain classes of the population and only in certain geographical areas. We shall, therefore, conceive of our interviewers as divided into L groups with M_a interviewers in the a^{th} group who are available to interview a particular N_a individuals and no others.

When all the interviewers are available to interview all individuals, we have $L = 1; M_a = M; N_a = N$.

Now n of the N individuals in the population are selected at random and m_a interviewers are selected at random from the a^{th} interviewer group to interview those sample individuals who are available for interview by this interviewer group. Let $m = \sum_a^L m_a$ be the total number of interviewers selected. Hensen & Hurwitz (1951) derive the total variance of individual responses around the mean of all individual responses in the population as

$$V(\bar{y}) = \frac{\left(\sigma_y^2 - \sigma_{yI}\right)}{n} + \frac{\sigma_{yI}}{m}.$$

Suppose a population of M interviewers is available to enumerate a population of N individuals on each of which p characters are defined. For j^{th} character $(j = 1,2,\ldots,p)$, the total variance of the sample mean $\bar{y}_j$ is given by

$$V_j = \frac{\left(\sigma_{yj}^2 - \sigma_{yjI}\right)}{n} + \frac{\sigma_{yjI}}{m} \qquad (2\ 5\ 1)$$

where σ_{yjI} is the covariance between responses obtained from different individuals by the same interviewer for j^{th} character (this covariance being taken within interviewer groups, since independent selections of interviewers are made from each interviewer group) and σ_{yj}^2 are the variances of over all responses for all individuals to all interviewers for the j^{th} character

With the ordinary survey which has a fixed total budget, increasing the number of interviewers will increase cost and will require a reduction of expenditures at some other point, e g , reducing the expenditure per interviewer or per individual or reducing the number of individuals included in the sample

Let C be the total budget available for field work on the survey Assuming the cost of the survey to be linear, we should have

$$C_1 n + C_2 m \leq C \qquad (2\ 5\ 2)$$

where C_1 is the cost per individual in the sample and C_2 is the cost per interviewer

The problem is to determine the values of n and m which can be found by minimizing the variances (2.5.1) for a fixed cost (2.5.2).

The problem of finding the optimal number of interviewers who should be assigned the job and the optimal number of individuals to be selected is finally formulated as

$$\begin{aligned} &\text{Minimize} \quad V_j = \frac{v_{1j}}{n} + \frac{v_{2j}}{m}, \qquad j = 1,2,\dots,p \\ &\text{Subject to} \quad C_1 n + C_2 m \le C \\ &\text{and} \quad n \le N,\ m \le M \end{aligned} \tag{2.5.3}$$

where we have used

$$v_{1j} = \left(\sigma_{yj}^2 - \sigma_{yjI}\right),\ v_{2j} = \sigma_{yjI}.$$

Using X_1 for n and X_2 for m, the problem (2.5.3) reduces again to the following form of multiobjective convex programming problem:

$$\begin{aligned} &\text{Minimize} \quad V_j = \sum_{i=1}^{2} \frac{v_{ij}}{X_i}, \qquad j = 1,2,\dots,p \\ &\text{Subject to} \quad \sum_{i=1}^{2} C_i X_i \le C \\ &\text{and} \quad X_1 \le N,\ X_2 \le M. \end{aligned} \tag{2.5.4}$$

In case we are interested in minimizing the cost of the survey while the tolerance limits are given on the variances for the various characters, the problem takes the form similar to (2.4.4).

CHAPTER–III

CHEBYSHEV SOLUTION TO A MULTIVARIATE STRATIFIED SAMPLING PROBLEM

3.1 INTRODUCTION

Usually in sample surveys more than one population characteristics of conflicting nature are estimated. When stratified sampling is to be used, an allocation among various strata that is optimum for one character is generally not so for the others. A suitable overall optimality criterion is required for dealing with such situations.

Various authors either suggested new criteria or explored further the already existing criteria such as Neyman (1934), Peter and Bucher (1940), Geary (1949), Dalenius (1957), Ghosh (1958), Yates (1960), Aoyama (1963), Chatterjee (1968). Kokan and Khan (1967), Huddleston, et al (1970), Arvanitis and Afonja (1971), Chromy (1987), Bethel (1985, 1989) etc., discuss the use of convex programming in relation to the multivariate optimal allocation problem. Each approach has its advantages and disadvantages. The weighted average method is computationally simple, intuitively appealing and can be solved under a fixed cost assumption, but the choice of the weights is arbitrary and the optimality properties are not clear. The convex programming approach gives the optimal solution to the defined problem where the upper limits are given on the variances and the cost is to be minimized. But if the variances are to be minimized a further search is usually required for an optimal solution which falls within the budgetary constraint.

In this chapter, we consider the problem of minimizing the variances for the various characters with fixed (given) budget. Each convex objective function is first linearized at its minimal point where it meets the linear cost constraint. The resulting multiobjective linear programming problem is then solved by Chebyshev goal programming.

3.2 MULTIVARIATE ALLOCATION PROBLEM

The multivariate allocation problem formulated in section 2.2, (2.2.5) is

$$\text{Minimize} \quad V_j = \sum_{i=1}^{L} \frac{a_{ij}}{X_i}, \qquad j = 1,2,\ldots,p$$

$$\text{Subject to} \quad \sum_{i=1}^{L} C_i X_i \leq C \qquad (3.2.1)$$

$$\text{and} \qquad 1 \leq X_i \leq N_i, \qquad i = 1,2,\ldots,L.$$

Each objective function in (3.2.1) is convex and the single constraint as well as the upper and lower bounds are linear. The problem (3.2.1) for $j = k$ is, therefore, a convex programming problem which can be solved by using any method of convex programming. Each of the p problems for $k = 1,2,...,p$ may have a different solution. A unique solution, suitable for all the p problems is obtained here by using the criterion of Chebyshev goal programming. In order to be able to apply the Chebyshev goal programming approach we approximate the convex objective functions in (3.2.1) by linear ones and then solve the resulting LPPs. The criterion behind the Chebyshev goal programming is to find a solution that minimizes the single worst unwanted deviation from any (soft) goal. In other words, it is a minimax goal programming approach.

3.3 TRANSFORMATION INTO A MULTIOBJECTIVE LINEAR PROGRAMMING PROBLEM

In the multiobjective allocation problem (3.2.1) there are p non-linear objective functions which later turn into soft goals with a single linear constraint (hard goal). To apply Chebyshev goal programming approach, all the hard and soft goals must be in linear form so that the worst deviation from the approximated linear goals in minimized. We thus approximate the non-linear soft goals by linear ones.

It may be noted that an analytic solution of the problem (3.2.1) for single character, say, $j = k$ is given (see Kokan and Khan (1967)) as

$$x'_{ik} = C\sqrt{a_{ik}C_i} \Big/ C_i\left\{\sum_{i=1}^{L}\sqrt{a_{ik}C_i}\right\}, \quad i = 1,2,\ldots,L \tag{3.3.1}$$

provided that $1 \le x'_{ik} \le N_i$, $i = 1,2,\ldots,L$.

In case the lower and/ or upper bounds are violated for some i (which is a very extreme case and rarely occurs in practice), some extra efforts are needed as explained in the above reference. However, since at this stage we need only approximate points, we may fix such x'_{ik} at the corresponding bounds.

Our strategy will be to approximate the convex objective surface V_k by the tangent hyperplane at the point (3.3.1).

This is obtained as

$$V_k \approx V_{k(x'_{ik})} + \nabla V'_{k(x'_{ik})}(X_i - x'_{ik}), \quad i = 1,2,\ldots,L$$

where $\nabla V'_{k(x'_{ik})}$ is the vector of partial derivatives,

$$\nabla V'_k(x'_{ik}) = \left[-\frac{a_{1k}}{(x'_{1k})^2}, -\frac{a_{2k}}{(x'_{2k})^2}, \ldots, -\frac{a_{Lk}}{(x'_{Lk})^2}\right].$$

Then

$$\nabla V_k'(x_{ik}')(X_i - x_{ik}') = \sum_{i=1}^{L} \frac{a_{ik}}{x_{ik}'} - \sum_{i=1}^{L} \frac{a_{ik} X_i}{(x_{ik}')^2}.$$

This gives

$$V_k \approx 2\sum_{i=1}^{L} \frac{a_{ik}}{x_{ik}'} - \sum_{i=1}^{L} \frac{a_{ik} X_i}{(x_{ik}')^2} = v_k \text{ (say)}.$$

Then the multiobjective convex programming problem (3.2.1) reduces to the following approximate multiobjective linear programming problem:

$$\text{Minimize} \quad v_j = 2\sum_{i=1}^{L} \frac{a_{ij}}{x_{ij}'} - \sum_{i=1}^{L} \frac{a_{ij} X_i}{(x_{ij}')^2}, \quad j = 1,2,\ldots,p$$

$$\text{Subject to} \quad \sum_{i=1}^{L} C_i X_i \le C \tag{3.3.2}$$

$$\text{and} \qquad 1 \le X_i \le N_i, \qquad i = 1,2,\ldots,L.$$

3.4 SOLUTION USING CHEBYSHEV GOAL PROGRAMMING

It can be noted that for individual objective functions the solutions of the respective problems in (3.2.1) and those in (3.3.2) coincide for $j = 1,2,...,p$ and are given by (3.3.1).

To solve the multiobjective LPP (3.3.2), we use the Chebyshev goal programming approach in which the p objective functions are put in the form of constraints, termed as soft goals, with upper bounds called aspiration levels. Aspiration level L_k is nothing but the minimum value of V_k obtained by solving the convex programming problem (3.2.1) individually for the k^{th} objective function. The explicit solutions for these p problems can again be obtained by using (3.3.1).

The Chebyshev goal programming model for solving (3.3.2) is given (as explained in (1.5.2)) as

Minimize δ

Subject to $$\sum_{i=1}^{L} C_i X_i \leq C \tag{3.4.1}$$

$$2\sum_{i=1}^{L} \frac{a_{ij}}{x_{ij}} - \sum_{i=1}^{L} \frac{a_{ij} X_i}{\left(x'_{ij}\right)^2} - \delta \leq L_j, \; j = 1,2,...,p$$

and $1 \leq X_i \leq N_i, \; i = 1,2,...,L$

where δ (dummy variable) represents the worst deviation level.

Our practical experience shows that the solution X^*_{ch} by transforming the multiobjective convex programming problem to the multiobjective linear programming problem and using the Chebyshev approach for it's solution, provides us a satisfactory point in the sense that the values of the various objective functions at this point remain very close to the optimal values obtained by individually solving the convex programming problems (3.2.1) for various $j = 1,2,....,p$.

This observation is evident also from the numerical example given below.

3.5 NUMERICAL EXAMPLE

Consider a population, divided into two strata with three characters under study for which the values of N_i, W_i, S_{i1}, S_{i2} and S_{i3} are given in the following table:

TABLE-3.1

Stratum i	N_i	W_i	S_{i1}	S_{i2}	S_{i3}	C_{i1}	C_{i2}	C_{i3}
1	180	0.40	1.5	2.25	0.75	0.6	0.9	1.5
2	270	0.60	3.0	4.75	5.25	0.8	1.2	2.0

The variance coefficients matrix is obtained by $a_{ij} = W_i^2 S_{ij}^2$ as

$$\left(a_{ij}\right) = \begin{pmatrix} 0.36 & 0.81 & 0.09 \\ 3.24 & 8.12 & 9.92 \end{pmatrix}.$$

Let us fix the budget at 100 units.

The above problem is transformed to the multiobjective convex programming problem as

$$\text{Minimize } V_1 = \frac{0.36}{X_1} + \frac{3.24}{X_2},\; V_2 = \frac{0.81}{X_1} + \frac{8.12}{X_2} \textit{ and } V_3 = \frac{0.09}{X_1} + \frac{9.92}{X_2}$$

$$\text{Subject to } 3X_1 + 4X_2 \le 100 \tag{3.5.1}$$

and
$$1 \le X_1 \le 180$$
$$1 \le X_2 \le 270$$

First we find out the solutions of the problems of minimizing V_1, V_2 and V_3 individually, subject to the only linear constraint $3X_1 + 4X_2 \le 100$ by using (3.3.1).

For V_1 the solution is

$$x'_{11} = 100\sqrt{0.36 \times 3} / 3\left\{\sqrt{0.36 \times 3} + \sqrt{3.24 \times 4}\right\}$$

$$= 7.47.$$

$$x'_{21} = 100\sqrt{3.24 \times 4} / 4\left\{\sqrt{0.36 \times 3} + \sqrt{3.24 \times 4}\right\}$$

$=19.40$.

Similarly the solutions of V_2 and V_3 are given by (7.16, 19.63) and (2.54, 23.10) respectively.

Now, Linearized form of the objective function V_1 at the point (7.47, 19.40) is obtained as

$$v_1 \simeq -0.0065X_1 - 0.0086X_2 + 0.4304$$

Similarly the linearized forms of the objective functions V_2 and V_3 at the respective points are obtained as

$$v_2 \simeq -0.0158X_1 - 0.0211X_2 + 1.0540$$

$$v_3 \simeq -0.0140X_1 - 0.0186X_2 + 0.9300$$

The values of L_1, L_2, and L_3 (aspiration levels) at the points (7.47, 19.40), (7.16, 19.63) and (2.54, 23.10) are obtained as 0.2152, 0.5270 and 0.4650 respectively.

Now, the approximated multiobjective linear programming problem to the multiobjective convex programming problem (3.5.1) is

$$\text{Minimize } v_1 = -0.0065X_1 - 0.0086X_2 + 0.4304,$$

$$v_2 = -0.0158X_1 - 0.0211X_2 + 1.0540$$

$$\textit{and } v_3 = -0.0140X_1 - 0.018X_2 + 0.9300 \qquad (3.5.2)$$

Subject to : $3X_1+4X_2 \leq 100$

and $1 \leq X_1 \leq 180, \; 1 \leq X_2 \leq 270.$

The Chebyshev model of the problem (3.5.2), becomes as to

Minimize δ

Subject to:

$$\begin{aligned} &-0.0065X_1 - 0.0086X_2 - \delta \leq -0.2152 \\ &-0.0158X_1 - 0.0211X_2 - \delta \leq -0.5270 \\ &-0.0140X_1 - 0.0186X_2 - \delta \leq -04650 \\ &3X_1 + 4X_2 \leq 100 \end{aligned} \qquad (3.5.3)$$

and
$$\begin{aligned} &1 \leq X_1 \leq 180 \\ &1 \leq X_2 \leq 270 \\ &\delta \geq 0. \end{aligned}$$

The Chebyshev point by solving the LPP (3.5.3) is $X_{ch}^* = (12.15, 15.89)$ with $\delta = 0$. The values of sample sizes n_1 and n_2, rounded to the nearest integers, are 12 and 16 respectively.

The solution print out of the problem through MATLAB is:

$X =$

12.1442

15.8825

0

$Lambda =$

0

0

0

0

0

0

$How =$

ok

$Z =$

0

This solution is being summarized in table-3.2.

The percent increases in the three variances for the Chebyshev point as compared to the respective individual variance minimization points are 104.78%, 110.23% and 136.04%.

Table-3.2

Values of V_j at the individual optimal points and at the Chebyshev point

	Optimization w.r.t. V_1	Optimization w.r.t. V_2	Optimization w.r.t. V_3	Chebyshev point
Rounded n_1 & n_2	(7,19)	(7,20)	(3,23)	(12,16)
Value of V_1	0.2219	0.2134	0.2609	0.2325
Value of V_2	0.5432	0.5218	0.6232	0.5752
Value of V_3	0.5351	0.5090	0.4614	0.6277

3.6 SOLUTION OF A TWO DIMENSIONAL MULTIVARIATE PROBLEM WHEN THE COST IS MINIMIZED

Let us consider the problem (2.2.6). Due to its special character (only two dimension), we give in the following an easy method of solution by using the Analytical approach of Kokan & Khan (1967).

The problem is to

$$\begin{aligned} &\text{Minimize } C = \sum_{i=1}^{2} C_i X_i \\ &\text{Subject to } \sum_{i=1}^{2} \frac{a_{ij}}{X_i} \le v_j, \quad j = 1,2,\dots,p \\ &1 \le X_i \le N_i . \end{aligned} \tag{3.6.1}$$

Using the transformation $X_i = \frac{1}{x_i}$, this reduces to

$$\text{Minimize } \sum_{i=1}^{2} \frac{C_i}{x_i}$$

$$\text{Subject to } \sum_{i=1}^{2} a_{ij} x_i \le v_j, \; j = 1,2,\ldots,p \qquad (3.6.2)$$

$$\frac{1}{N_i} \le x_i \le 1.$$

First we identify the linear constraints k_1 and k_2 such that

$$\left.\begin{aligned} \min_j \frac{v_j}{a_{1j}} &= \frac{v_{k1}}{a_{1k1}} \\ \min_j \frac{v_j}{a_{2j}} &= \frac{v_{k2}}{a_{2k2}}. \end{aligned}\right\} \qquad (3.6.3)$$

Let us denote the minimum of C subject to the constraint(j) by $\underline{x}^{(j)}$. An explicit expression for $\underline{x}^{(j)} = \left(x_1^{(j)}, x_2^{(j)}\right)$ is given by

$$x_i^{(j)} = v_j \sqrt{a_{ij} C_i} \Big/ a_{ij} \left\{ \sum_{i=1}^{2} \sqrt{a_{ij} C_i} \right\}, i = 1,2. \qquad (3.6.4)$$

We illustrate the method by an (hypothetical) example represented in the following figure in which we have taken four constraints. The level curves of the objective functions touching the various constraints are also traced.

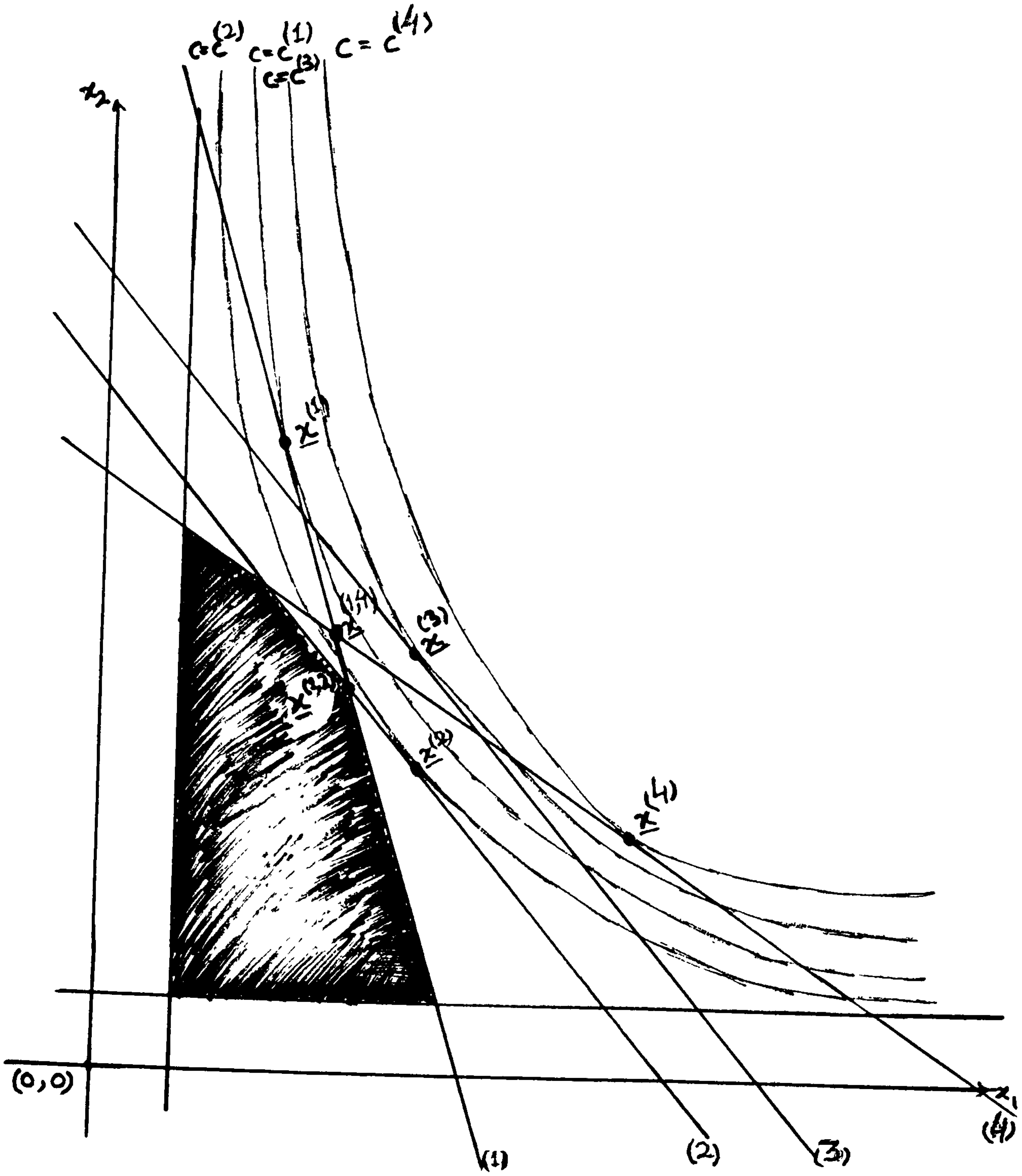

Figure: Graph for the hypothetical example.

The minimum intercept on x_1 is cut by the constraint (1) and the minimum intercept on x_2 is cut by the constraint (4).

Now $\underline{x}^{(4)}$ violates the constraint (1) and $\underline{x}^{(1)}$ violates the constraint (4). A dangling solution, will then be the point of intersection of the lines (1) & (4), viz $\underline{x}^{(1,4)}$.

This new point, however violates the constraint (2). So we test $\underline{x}^{(2)}$ which violates the constraint (1). Since $\underline{x}^{(1)}$ also violates the constraint (2), the intersection of the lines (1) & (2) is tested, which satisfies all the constraints and thus gives the optimal solution.

Let us consider the numerical example of 3.5 in which we are given the upper bounds on the three variances respectively as 0.30,0.60 and 0.50.

Then the problem to be solved is

$$\begin{aligned} &\textit{Minimize} \quad \frac{3}{x_1}+\frac{4}{x_2} \\ &\textit{Subject to} \quad 0.36x_1+3.24x_2 \le 0.30 \\ &\qquad 0.81x_1+8.12x_2 \le 0.60 \\ &\qquad 0.09x_1+9.92x_2 \le 0.50 \\ &\qquad 0.0056 \le x_1 \le 1 \\ &\qquad 0.0037 \le x_2 \le 1 \end{aligned} \tag{3.6.5}$$

We identify the linear constraints (2) & (3) by using (3.6.3).

By using (3.6.4), we obtain $\underline{x}^{(2)}$ and $\underline{x}^{(3)}$ as (0.1591, 0.0580) & (0.4233, 0.0466).

Now, $\underline{x}^{(3)}$ violates the constraint (2) and $\underline{x}^{(2)}$ violates the constraint (3). Then the solution $\underline{x}^{(2,3)}$ is obtained as the point of intersection of the lines (2) & (3) ie $\underline{x}^{(2,3)} = (0.2585, 0.0481)$. This point also satisfies the constraint (1). Hence it is the optimal solution to the given problem.

The values of sample sizes n_1 and n_2 are found respectively as 3.87 & 20.79 which rounded to the nearest integers are 4 & 21. The value of the objective function at the optimal point is 96. The same numerical example has been solved in section 3.5 where we fixed the cost at 100 and minimized the variances. The optimal solution given in table-3.2 may be compared with this solution.

CHAPTER-IV

USE OF CUTTING PLANE TECHNIQUE FOR SOLVING THE MULTIVARIATE SAMPLING PROBLEMS

4.1 INTRODUCTION

In this chapter, we again consider the sampling problems of chapter II where p convex objective functions are to be minimized subject to the linear cost constraint. The problem is first transformed to a multiobjective nonlinear programming problem with several linear objective functions and a single convex constraint. The non-linearity of the single non-linear constraint is handled through linearizing it by the cutting plane technique. The resulting LPP is then solved by Chebyshev goal programming approach. A comparison of Chebyshev solution with the fuzzy programming solution has also been made.

4.2 MULTIVARIATE SAMPLING PROBLEMS

The multivariate sampling problems formulated in chapter-II have the form

$$\text{Minimize} \quad V_j = \sum_{i=1}^{L} \frac{a_{ij}}{X_i}, \qquad j = 1,2,\dots, p$$

$$\text{Subject to} \quad \sum_{i=1}^{L} C_i X_i \le C$$

$$\text{and} \quad 1 \le X_i \le N_i, \qquad i = 1,2,\dots, L.$$

Using $\frac{1}{x_i}$ for X_i the problem gets transformed to

$$\text{Minimize} \quad Z_j = \sum_{i=1}^{L} a_{ij} x_i \;, \; j = 1,2,\dots,p \tag{4.2.1}$$

$$\text{Subject to. } g(X) = \sum_{i=1}^{L} \frac{C_i}{x_i} - C \le 0 \tag{4.2.2}$$

$$\frac{1}{N_i} \le x_i \le 1 \;, \qquad i = 1,2,\dots,L. \tag{4.2.3}$$

In order to be able to find a Chebyshev point, we will linearize the only convex constraint (4.2.2) by using the cutting plane technique of J.E. Kelly (1960).

4.3 OBTAINING AN EQUIVALENT PROBLEM BY LINEARIZING THE CONVEX CONSTRAINT

Let $X^{k(0)} = (x_1^{k(0)},\dots,x_L^{k(0)})$ be the solution of LPP, which minimizes (4.2.1) for $j = k$ subject to the bounds (4.2.2).

Then we compute

$$g(X^{k(0)}) = \sum_{i=1}^{L} \frac{C_i}{x_i^{k(0)}} - C .$$

Define $\in_1 \& \in_2$ to be two small positive tolerance limits for convergence.

If $\left|g(X^{k(0)})\right| \leq \in_1$, this means that (4.2.2) is satisfied to the tolerance limit and thus $X^{k(0)}$ solves the convex programming problem (4.2.1)-(4.2.3) for $j = k$.

If $\left|g(X^{k(0)})\right| > \in_1$, we linearize the convex constraint $g(X) \leq 0$ about the point $X^{k(0)}$ as :

$$G(X) \approx g(X^{k(0)}) + \nabla g(X^{k(0)})'(X - X^{k(0)}) \leq 0,$$

where $g\left(X^{k(0)}\right)$ is the value of $g(X)$ at the point $X^{k(0)}$ and

$$\nabla g(X^{k(0)})' = \left[\frac{\delta}{\delta x_1}\left\{\sum_{i=1}^{L} \frac{C_i}{x_i} - C\right\}, \frac{\delta}{\delta x_2}\left\{\sum_{i=1}^{L} \frac{C_i}{x_i} - C\right\}, \ldots\ldots, \frac{\delta}{\delta x_L}\left\{\sum_{i=1}^{L} \frac{C_i}{x_i} - C\right\}\right]_{X^{k(0)}}$$

$$= \left[-\frac{C_1}{(x_1^{k(0)})^2}, -\frac{C_2}{(x_2^{k(0)})^2}, \ldots\ldots, -\frac{C_L}{(x_L^{k(0)})^2}\right].$$

Then

$$\nabla g(X^{k(0)})'\left(X - X^{k(0)}\right) = \sum_{i=1}^{L} \frac{C_i}{x_i^{k(0)}} - \sum_{i=1}^{L} \frac{C_i x_i}{(x_i^{k(0)})^2} .$$

Thus the constraint (4.2.2) linearized at the point $X^{k(0)}$ is

$$G(X) \approx 2\sum_{i=1}^{L}\frac{C_i}{x_i^{k(0)}} - \sum_{i=1}^{L}\frac{C_i x_i}{x_i^{k(0)^2}} - C \le 0. \qquad (4.3.1)$$

We then solve the following LPP:

$$\begin{aligned} &\textit{Minimize} \quad Z_k = \sum_{i=1}^{L} a_{ik} x_i \\ &\textit{Subject to} \quad 2\sum_{i=1}^{L}\frac{C_i}{x_i^{k(0)}} - \sum_{i=1}^{L}\frac{C_i x_i}{(x_i^{k(0)})^2} - C \le 0 \qquad (4.3.2) \\ &\qquad \frac{1}{N_i} \le x_i \le 1, \qquad i = 1,2,\ldots,L. \end{aligned}$$

Denote the solution of LPP (4.3.2) by

$$X^{k(1)} = (x_1^{k(1)}, \ldots, x_L^{k(1)}).$$

At t^{th} iteration we find $X^{k(t)}$ and

$$g(X^{k(t)}) = \sum_{i=1}^{L}\frac{C_i}{x_i^{k(t)}} - C.$$

If $\left|g(X^{k(t)})\right| \le \epsilon_1$ then clearly $X^{k(t)}$ also solves the CPP (4.2.1)-(4.2.3).

Otherwise we linearize the constraint $g(X)$ about the point $X^{k(t)}$ and solve the LPP:

$$\textit{Minimize} \quad Z_k = \sum_{i=1}^{L} a_{ik} x_i$$

$$\textit{Subject to} \quad 2\sum_{i=1}^{L} \frac{C_i}{x_i^{k(l)}} - \sum_{i=1}^{L} \frac{C_i x_i}{(x_i^{k(l)})^2} - C \le 0\,, \qquad l = 0,1,\ldots,t_k^* \qquad (4.3.3)$$

$$\frac{1}{N_i} \le x_i \le 1\,, \qquad i = 1,2,\ldots,L.$$

The process is then repeated until

$\left|g(X^{k(t)})\right| \le \epsilon_1$ say at $t_k^{*\,th}$ iteration. The LPP (4.3.3) for $t_k = t_k^*$ approximates the CPP (4.2.1)-(4.2.3) for $j = k$.

At some stage it is also possible that $\left|g(X^{k(t)})\right| > \epsilon_1$ but $\left|X^{k(t-1)} - X^{k(t)}\right| \le \epsilon_2$. In this case the LPP (4.3.3) does not exactly solve the CPP (4.2.1)-(4.2.3). However, as the point $X^{(t)}$ is getting repeated, we will consider the LPP (4.3.3) to approximate the CPP (4.2.1)-(4.2.3) and take the corresponding t equal to t^*.

Taking $t_0^* = 1$, the following LPPs are now solved for $s = 1,2,\ldots,p$:

$$\text{Minimize } Z_s = \sum_{i=1}^{L} a_{is} x_i$$

$$\text{Subject to } 2\sum_{i=1}^{L} \frac{C_i}{x_i^{s(l)}} - \sum_{i=1}^{L} \frac{C_i x_i}{(x_i^{s(l)})^2} - C \le 0\ , l = 0,1,\ldots,t_k^*; s = 1,2,\ldots,p \quad (4.3.4)$$

$$\frac{1}{N_i} \le x_i \le 1\ , \qquad i = 1,2,\ldots,L.$$

Let the minimum values of Z_s thus found be Z_s^0, $s = 1,2,\ldots,p$ at the corresponding minimal points $X_s^0, s = 1,2,\ldots,p$. The p solutions $X_1^0,\ldots\ldots,X_p^0$ have been obtained by minimizing the individual objective functions subject to the linearized constraints which will give us the aspiration levels being used in Chebyshev goal programming model.

4.4 SOLUTION USING CHEBYSHEV GOAL PROGRAMMING

For obtaining an unique solution **suitable** for all the p objective functions, we use the Chebyshev goal programming technique. The Chebyshev formulation of the multivariate sampling problem (4.2.1)-(4.2.3) is the following LPP:

$$\text{Minimize} \quad \delta$$

$$\text{Subject to} \quad 2\sum_{l=1}^{L}\frac{C_l}{x_l^{s(l)}} - \sum_{l=1}^{L}\frac{C_l x_l}{(x_l^{s(l)})^2} - C \le 0, \; l = 0,1,\dots,t_k^*; s = 1,2,\dots,p$$

$$\sum_{l=1}^{L} a_{ls}x_l - \delta \le Z_s^0, \qquad s = 1,2,\dots,p \qquad (4.4.1)$$

$$\frac{1}{N_l} \le x_l \le 1, \qquad l = 1,2,\dots,L$$

where δ (dummy variable) represents the worst deviation level and Z_s^0, $s = 1,2,\dots,p$ are the aspiration levels.

4.5 ALGORITHM

Let us consider the problem (4.2.1)-(4.2.3).

Set $k = 1$ and $t = 0$.

Step I: If $k > p$, go to Step III. Otherwise find the point $X^{k(t)}$ by solving the LPP (4.3.3).

Step II: If $\left|g(X^{k(t)})\right| \le \epsilon_1$ or $\left|X^{k(t-1)} - X^{k(t)}\right| \le \epsilon_2$ for some t, say t_k^*, where ϵ_1 and ϵ_2 are the suitable tolerance limits, then go to step I with $k = k+1$.

Otherwise go to step I with $t = t+1$.

Step III: Solve LPP (4.3.4) for $s = 1,2,...,p$ to obtain X_s^0, the approximate minimal points for the respective objective functions, with minimum corresponding values of Z_s as Z_s^0.

Step IV: Solve the Chebyshev goal programming model (4.4.1) of the problem (4.2.1)-(4.2.3) to obtain the Chebyshev point X_{ch}^*.

4.6 FUZZY SOLUTION

Like Chebyshev goal programming, the basis of fuzzy programming approach is also to minimize the worst deviation from any goal. For obtaining a fuzzy solution, we first compute for each s ($s = 1,2,...,p$), the maximum and minimum values of the respective objective functions.

Let $Z_s(X_j^o) = Z_s^{j0}$, $j = 1,2,...,p$.

Clearly $Z_s^{s0} = Z_s^0 = \min_j Z_s(X_j^0) = L_s$, say.

Denote $\max_j Z_s(X_j^0) = U_s$.

The differences of the maximum and minimum values of Z_s are denoted by $d_s = U_s - L_s$, $s = 1,2,...,p$.

The fuzzy programming formulation of the problem (4.2.1)-(4.2.3) is the following LLP :

$$
\begin{aligned}
&\textit{Minimize} \quad \delta \\
&\textit{Subject to} \quad 2\sum_{l=1}^{L}\frac{C_l}{x_l^{s(l)}} - \sum_{l=1}^{L}\frac{C_l x_l}{(x_l^{s(l)})^2} - C \le 0 \,, \quad l = 0,1,\ldots,t_k^*; s = 1,2,\ldots,p \\
&\sum_{l=1}^{L} a_{ls} x_l - d_s \delta \le Z_s^0 \,, \qquad s = 1,2,\ldots,p \qquad (*) \\
&\frac{1}{N_l} \le x_l \le 1 \,, \qquad l = 1,2,\ldots,L.
\end{aligned}
\qquad (4.6.1)
$$

Comparing (4.4.1) and (4.6.1) it can be noted that the fuzzy programming solution is better than the Chebyshev solution if d_s, the differences between maximum and minimum values of the objective functions, are greater than 1 for all characteristics. The reason behind this is that in this case (i.e. when $d_s > 1$) the constraints (4.6.1.*) in fuzzy programming are less restrictive than the corresponding constraints in Chebyshev problem (4.4.1).

4.7 A NUMERICAL EXAMPLE

Let us consider again the numerical example given in 3.5.1.

By making the transformation $X_i = \frac{100}{x_i}$, the problem (4.2.1)-(4.2.3) is obtained as

$$\begin{aligned} &\textit{Minimize} \quad Z_1 = 0.0036x_1 + 0.0324x_2,\ Z_2 = 0.0081x_1 + 0.0812x_2 \\ &\qquad\qquad \textit{and } Z_3 = 0.0009x_1 + 0.0992x_2 \\ &\textit{Subject to} \quad \frac{3}{x_1} + \frac{4}{x_2} \le 1 \qquad\qquad (4.7.1) \\ &\qquad\qquad 0.5556 \le x_1 \le 100 \\ &\qquad\qquad 0.3704 \le x_2 \le 100. \end{aligned}$$

Let us fix $\in_1$ for the three objective functions be 0.01, 0.006 and 0.08 & $\in_2$ for the three objective functions be 0.005.

The approximated linear programming problems corresponding to the three objective functions Z_1,Z_2 and Z_3, as derived in (4.3.3) are obtained as follows:

$$\begin{aligned} &\textit{Minimize} \quad Z_1 = 0.0036x_1 + 0.0324x_2 \\ &\textit{Subject to} \quad 971.82x_1 + 2915.45x_2 \ge 3139.74 \\ &\qquad\qquad 66.77x_1 + 2915.45x_2 \ge 2342.9 \\ &\qquad\qquad 387.05x_1 + 651.57x_2 \ge 1602.56 \\ &\qquad\qquad 35.69x_1 + 735.7x_2 \ge 1192.12 \\ &\qquad\qquad 126.74x_1 + 154.58x_2 \ge 807.54 \end{aligned}$$

$$13.75x_1 + 205.9x_2 \geq 602.42$$
$$32.21x_1 + 53.99x_2 \geq 390.52$$
$$4.54x_1 + 70.45x_2 \geq 309.52 \quad (4.7.2)$$
$$10.54x_1 + 24.39x_2 \geq 210$$
$$2.28x_1 + 29.95x_2 \geq 171.22$$
$$4.54x_1 + 15.39x_2 \geq 130.76$$
$$1.86x_1 + 17.73x_2 \geq 115.66$$
$$2.78x_1 + 13.54x_2 \geq 1.05$$
$$0.5556 \leq x_1 \leq 100$$
$$0.3704 \leq x_2 \leq 100$$

Minimize $Z_2 = 0.0081x_1 + 0.0812x_2$

Subject to $971.82x_1 + 2915.45x_2 \geq 3139.74$

$$66.77x_1 + 2915.45x_2 \geq 2342.9$$
$$387.05x_1 + 651.57x_2 \geq 1602.56$$
$$35.69x_1 + 735.7x_2 \geq 1192.12$$
$$126.74x_1 + 154.58x_2 \geq 807.54$$
$$13.75x_1 + 205.9x_2 \geq 602.42 \quad (4.7.3)$$
$$32.21x_1 + 53.99x_2 \geq 390.52$$
$$4.54x_1 + 70.45x_2 \geq 309.52$$
$$10.54x_1 + 24.39x_2 \geq 210$$
$$2.28x_1 + 29.95x_2 \geq 171.22$$
$$4.54x_1 + 15.39x_2 \geq 130.76$$
$$1.86x_1 + 17.73x_2 \geq 115.66$$
$$1.99x_1 + 14.59x_2 \geq 101.68$$
$$2.78x_1 + 13.54x_2 \geq 1.05$$
$$0.5556 \leq x_1 \leq 100$$
$$0.3704 \leq x_2 \leq 100$$

$$
\begin{aligned}
\textit{Minimize } & Z_3 = 0.0009x_1 + 0.0992x_2 \\
\textit{Subject to } & 971.82x_1 + 2915.45x_2 \geq 3139.74 \\
& 66.77x_1 + 2915.45x_2 \geq 2342.9 \\
& 0.84x_1 + 2915.45x_2 \geq 2091.54 \\
& 20.64x_1 + 779.58x_2 \geq 1174.24 \\
& 0.33x_1 + 796.34x_2 \geq 1048.74 \\
& 5.69x_1 + 231.71x_2 \geq 591.48 \\
& 0.11x_1 + 238.27x_2 \geq 529.14 \\
& 1.58x_1 + 81.57x_2 \geq 304.84 \\
& 0.05x_1 + 83.88x_2 \geq 273.84 \\
& 0.47x_1 + 37.88x_2 \geq 170.02 \\
& 0.03x_1 + 37.93x_2 \geq 152.34 \\
& 0.5556 \leq x_1 \leq 100 \\
& 0.3704 \leq x_2 \leq 100.
\end{aligned}
\tag{4.7.4}
$$

The solutions X_1^0 X_2^0 and X_3^0 of the three problems (4.7.2), (4.7.3) and (4.7.4) are obtained as:

$X_1^0 = (12.26, 5.24)$ with $Z_1^0 = 0.2138$

$X_2^0 = (14.16, 5.04)$ with $Z_2^0 = 0.5238$

$X_3^0 = (40.63, 3.98)$ with $Z_3^0 = 0.4318$.

The optimal values Z_1^0, Z_2^0 and Z_3^0 will be used as aspiration levels in the Chebyshev goal programming model.

The Chebyshev goal programming model (4.4.1) yields the following LPP:

$$\begin{aligned} &Minimize \quad \delta \\ &Subject\ to \quad 0.0036x_1 + 0.0324x_2 - \delta \le 0.2138 \\ &\qquad 0.0081x_1 + 0.0812x_2 - \delta \le 0.5238 \\ &\qquad 0.0009x_1 + 0.0992x_2 - \delta \le 0.4318 \\ &\qquad 0.5556 \le x_1 \le 100 \\ &\qquad 0.3704 \le x_2 \le 100 \end{aligned}$$

plus the 25 linearized constraints given in (4.7.2), (4.7.3) and (4.7.4).

The Chebyshev point by solving the above problem is $X^*_{ch} = (23.19,\ 4.20)$ with $\delta = 0.0058$. The values of sample sizes n_1 and n_2 are found respectively as 4.31 and 23.81 which rounded to the nearest integers are 4 and 24. The values of the three objective functions (variances) at this point are $Z^1_{ch} = 0.2250, Z^2_{ch} = 0.6409$ and $Z^3_{ch} = 0.4359$.

For obtaining the fuzzy point we find the values of Z_1 at the points X^0_2 and X^0_3, the values of Z_2 at the points X^0_1 and X^0_3 and the values of Z_3 at the points X^0_1 and X^0_2 which are respectively obtained as (0.2142,0.2754), (0.5245,0.6526) and (0.5305,0.5125).

Thus

$L_1 = 0.2138,\ U_1 = 0.0.2754$

$L_2 = 0.0.5238,\ U_2 = 0.6526$

$L_3 = 0.4318, U_3 = 0.5305$

$$d_1 = 0.0616$$
$$d_2 = 0.1288$$
$$d_3 = 0.0987$$

The fuzzy goal programming model (4.6.1) yields the following LPP:

$$\begin{aligned}
&\textit{Minimize} \quad \delta \\
&\textit{Subject to} \quad 0.0036x_1 + 0.0324x_2 - 0.0616\delta \le 0.2138 \\
&\qquad 0.0081x_1 + 0.0812x_2 - 0.1288\delta \le 0.5238 \\
&\qquad 0.0009x_1 + 0.0992x_2 - 0.0987\delta \le 0.4318 \\
&\qquad 0.5556 \le x_1 \le 100 \\
&\qquad 0.3704 \le x_2 \le 100
\end{aligned}$$

plus the 25 linearized constraints given in (4.7.2), (4.7.3) and (4.7.4).

The fuzzy point for the given problem by solving the LPP (4.6.1) is $X^*_{fz} = (22.55, 4.21)$ with $\delta = 0.0607$. The corresponding values of sample sizes n_1 and n_2 are found respectively as 4.43 and 23.76.

It may be remarked that the maximum deviation of the optimum point from the various goals is greater for the fuzzy point as compared to the Chebyshev point. This was expected (since all the

d_j are <<1) as noted in section 4.6. However, after rounding to the nearest integers the solution coincides with that of the rounded solution for Chebyshev method, (i.e. 4,24).

TABLE-4.1

Value of Z_j at the individual optimal points and at the Chebyshev and fuzzy points

	Optimization w.r.t. Z_1	Optimization w.r.t. Z_2	Optimization w.r.t. Z_3	Chebyshev Point	fuzzy point
Rounded n_1 & n_2	(8,19)	(7,20)	(2,25)	(4,24)	(4,24)
Value of Z_1	0.2155	0.2134	0.3096	0.2250	0.2250
Value of Z_2	0.5288	0.5218	0.7299	0.6409	0.6409
Value of Z_3	0.5335	0.5090	0.4419	0.4359	0.4359

The percent increases in the variances for the Chebyshev point (and fuzzy point) as compared to the individual variance minimization points are 104.41%, 122.82%and 0.99% respectively.

4.8 THE CASE OF THE PRESENCE OF BOUNDS ON THE VARIANCES OF SOME CHARACTERS

We now consider the situation where there are tolerance limits on the variances for some of the characteristics. Let the upper limits on the j^{th} variance be given as m_j, $j \in J'$, $J' \subset J = \{1,2,\ldots,p\}$.

Then one requires

$$\sum_{i=1}^{L} a_{ij} x_i \leq m_j, \qquad j \in J'.$$

In this situation, the multiobjective convex programming problem to be solved is

$$\begin{aligned} &\textit{Minimize} \quad Z_j = \sum_{i=1}^{L} a_{ij} x_i, \qquad j \in (J - J') \\ &\textit{Subject to.} \quad \sum_{i=1}^{L} \frac{C_i}{x_i} \leq C \\ &\sum_{i=1}^{L} a_{ij} x_i \leq m_j, \qquad j \in J' \\ &\frac{1}{N_i} \leq x_i \leq 1, \qquad i = 1,2,\ldots,L. \end{aligned} \tag{4.8.1}$$

Let us Consider a population with four strata each of size 150. There are five different characters under study and it is required that the variances of the first, third and fifth characters have the upper tolerance limits 0.70,0.60 and 0.80 respectively. The total field cost is 160 units.

The costs of completely enumerating a unit in the different strata and the coefficients of variance $(a_{ij})$are given in the following table

TABLE-4.2

j / i	a_{ij} 1	2	3	4	5	C_i
1	3.4	5.8	2.4	1.8	2.9	2
2	3.9	1.6	4.8	2.8	5.9	3
3	2.2	4.4	1.0	5.7	3.6	1
4	5.0	2.2	3.9	1.3	4.8	2

The multiobjective convex programming formulation of the above problem is as follows

$$Minimize\ V_2 = \frac{5.8}{n_1} + \frac{1.6}{n_2} + \frac{4.4}{n_3} + \frac{2.2}{n_4},\ V_4 = \frac{1.8}{n_1} + \frac{2.8}{n_2} + \frac{5.7}{n_3} + \frac{1.3}{n_4}$$

$$Subject\ to\ 2n_1 + 3n_2 + n_3 + 2n_4 \le 160$$

$$\frac{3.4}{n_1} + \frac{3.9}{n_2} + \frac{2.2}{n_3} + \frac{5.0}{n_4} \le 0.70$$

$$\frac{2.4}{n_1} + \frac{4.8}{n_2} + \frac{1.0}{n_3} + \frac{3.9}{n_4} \le 0.60 \qquad (4.8.2)$$

$$\frac{2.9}{n_1} + \frac{5.9}{n_2} + \frac{3.6}{n_3} + \frac{4.8}{n_4} \le 0.80$$

$$and \qquad 1 \le n_i \le 150,\ i = 1,2,3,4.$$

Using n_i for $\frac{1}{x_i}$, the problem (4.8.2) reduces to the following form:

$$Minimize\ V_2 = 5.8x_1 + 1.6x_2 + 4.4x_3 + 2.2x_4, V_2 = 1.8x_1 + 2.8x_2 + 5.7x_3 + 1.3x_4$$

$$Subject\ to\ \frac{2}{x_1} + \frac{3}{x_2} + \frac{1}{x_3} + \frac{2}{x_4} \le 160$$

$$3.4x_1 + 3.9x_2 + 2.2x_3 + 5.0x_4 \le 0.70 \qquad (4.8.3)$$

$$2.4x_1 + 4.8x_2 + x_3 + 3.9x_4 \le 0.60$$

$$2.9x_1 + 5.9x_2 + 3.6x_3 + 4.8x_4 \le 0.80$$

$$0.0067 \le x_i \le 1, i = 1,2,3,4.$$

The solutions X_2^0 and X_4^0 by solving LPPs (4.3.4) for $s = 2,4$ are obtained as

$$X_2^0 = (0.0467, 0.0544, 0.0357, 0.0490) \& V_2^0 = 0.6464$$

$$X_4^0 = (0.0555, 0.0505, 0.0358, 0.0463) \& V_4^0 = 0.5055$$

The optimal values of V_2^0 and V_4^0 are used as aspiration levels in the Chebyshev model.

The Chebyshev point by solving the LPP (4.4.1) is $X_{ch}^* = (0.0519, 0.0519, 0.0379, 0.0452)$ with $\delta = 0.0078$. The values of sample sizes n_1, n_2, n_3 and n_4 are found respectively as $19.2678, 19.2678, 26.3852$ and 22.1239 which round to the nearest integers are $19, 19, 26$ and 22.

The solution is being summarized in the following table:

TABLE-4.3

	Opt. w.r.t. V_2	**Opt.w.r.t V_4**	**Cheb.point.**
Rounded values of sample sizes	(21,18,28,20)	(18,20,30,22)	(19,19,26,22)
Valueof V_2	0.6322	0.6486	0.6587
Valueof V_4	0.5099	0.4891	0.5204

The percent increases in the variances for the Chebyshev point as compared to the individual variance minimization points are 104.19%, and 106.40%.

CHAPTER-V

OPTIMAL ESTIMATION OF MEANS OF SEVERAL VARIABLES USING MULTIVARIATE AUXILIARY INFORMATION UNDER STRATIFIED SAMPLING

5.1 INTRODUCTION

Most of the sample surveys are devoted to collect information on several variables simultaneously. The usual problem in multipurpose surveys is to estimate the population means or totals of several variables simultaneously by using a number of auxiliary variables the information on which may be available through the past census data or it may be collected through diverting a part of the survey budget. In a land survey, for instance the estimates of the total number of agricultural labourers, literates and schedule casts for a certain administrative block may be easily available through past census data and the information on the variables such as the number of households, number of male workers and number of cultivators of the villages may not be readily available but may be known through diverting a part of the survey budget to it.

The problem of estimation of the population mean (or total) of a single survey variable in the situation where population means (or totals) of several auxiliary variables are known has been considered by several authors including Olkin (1958), Raj (1965), Srivastava

(1965, 1966), Rao and Mudholkar (1967), Singh (1967), Srivastava (1971), Tripathi (1970, 1976, 1987) and Mukherjee et. al. (1987).

The use of information on several auxiliary variables for estimating the population means of more than one principal variables has also been considered by several authors. Tripathi and Khattree (1989) discussed the estimation of means of principal variables $y_1,...,y_p$ under simple random sampling, in the situations where means of auxiliary variables $x_1,...,x_q$ are known. Further, Tripathi (1989) extended the result to the case of two occasions. Tripathi and Chaubey (1993) have considered the problem of obtaining the optimum probabilities of selection based on $x_1,...,x_q$ in pps sampling for estimating the means of $y_1,...,y_p$. Recently, Tripathi and Chaubey (2000) discussed the problem of estimating the mean of a vector variable $\underline{y}=(y_1,...,y_p)'$ based on a general sampling design and on the knowledge of means of several variables $\underline{x}=(x_1,...,x_q)'$ for a finite population. They also gave the criterion of preference of one estimation procedure over the others in a quite general form stronger than customary criteria.

In this chapter, we discuss the estimation of finite population mean vector $(\overline{Y}_1,...,\overline{Y}_p)=\underline{\overline{Y}}'$ of the principal variables $(Y_1,...,Y_p)=\underline{Y}'$, under *stratified sampling design*, in the situations where mean vector $(\overline{X}_1,...,\overline{X}_q)=\underline{\overline{X}}'$ of the auxiliary variables $(X_1,...,X_q)=\underline{X}'$ is known.

5.2 NOTATION

Consider a finite population $U = \{1,2,...,N\}$. The population is divided into L strata.

Let

y_{ijh} = the value of i^{th} unit for j^{th} estimation character in the h^{th} stratum.

and

x_{ikh} = The value of i^{th} unit for k^{th} auxiliary character in the h^{th} stratum.

$(j = 1,2,...,p;\ k = 1,2,...,q;\ h = 1,2,...,L)$.

Let $\underline{y}_{ih}$ be the observed value of the vector of estimation variables $y_1,...,y_p$ on the i^{th} unit in the h^{th} stratum and similarly let $\underline{x}_{ih}$ be the observed value of the vector of auxiliary variables $x_1,...,x_q$ on the i^{th} unit in the h^{th} stratum.

The population mean vectors of the estimation variables and of the auxiliary variables in the h^{th} stratum are given respectively as

$$\overline{\underline{Y}}_h = \frac{1}{N_h}\sum_{i=1}^{N_h}\underline{y}_{ih}$$

and $\underline{X}_h = \frac{1}{N_h}\sum_{i=1}^{N_h}\underline{x}_{ih}$.

Denote by $\overline{\underline{Y}} = \sum_{h=1}^{L} W_h \overline{\underline{Y}}_h$

and $\overline{\underline{X}} = \sum_{h=1}^{L} W_h \overline{\underline{X}}_h$.

Consider a random sample of size n from a finite population U. On each of the sample unit, the measurement for p estimation variables $y_1,\ldots,y_p$ and the q auxiliary variables $x_1,\ldots,x_q$ are obtained as

$$\begin{pmatrix} y_{1_1},\ldots,y_{p_1} \\ y_{1_2},\ldots,y_{p_2} \\ \vdots \quad\quad \vdots \\ y_{1_n},\ldots,y_{p_n} \end{pmatrix} \text{ and } \begin{pmatrix} x_{1_1},\ldots,x_{q_1} \\ x_{1_2},\ldots,x_{q_2} \\ \vdots \quad\quad \vdots \\ x_{1_n},\ldots,x_{q_n} \end{pmatrix}.$$

Let the population be stratified into L strata and denote by $\underline{y}_{ih}$ the vector of sample values of estimation variables on the i^{th} unit in the h^{th} stratum, $i = 1,2,\ldots,n_h; h = 1,2,\ldots,L$ and denote by $\underline{x}_{ih}$ the vector of sample values of auxiliary variables on the i^{th} unit in the h^{th} stratum, $i = 1,2,\ldots,n_h; h = 1,2,\ldots,L$.

The customary unbiased estimators of $\underline{\overline{Y}}_h$ and $\underline{\overline{X}}_h$ are given by

$$\hat{\underline{\overline{Y}}}_h = \frac{1}{n_h}\sum_{i=1}^{n_h} \underline{y}_{ih}$$

and $$\hat{\underline{\overline{X}}}_h = \frac{1}{n_h}\sum_{i=1}^{n_h} \underline{x}_{ih}.$$

Denote by $$\hat{\underline{\overline{Y}}} = \sum_{h=1}^{L} W_h \hat{\underline{\overline{Y}}}_h$$

and $$\hat{\underline{\overline{X}}} = \sum_{h=1}^{L} W_h \hat{\underline{\overline{X}}}_h.$$

5.3 THE PROPOSED CLASS OF ESTIMATORS

For h^{th} stratum, let us define

$$\tilde{\underline{\overline{Y}}}_h = \hat{\underline{\overline{Y}}}_h + T_h\left(\underline{\overline{X}}_h - \hat{\underline{\overline{X}}}_h\right),\ h = 1,2,\ldots,L,$$

where $\hat{\underline{\overline{Y}}}_h = \left(\hat{\overline{Y}}_{h1},\ldots,\hat{\overline{Y}}_{hp}\right)'$ and $\hat{\underline{\overline{X}}}_h = \left(\hat{\overline{X}}_{h1},\ldots,\hat{\overline{X}}_{hk}\right)'$ are the customary unbiased estimators of $\underline{\overline{Y}}_h$ and $\underline{\overline{X}}_h$ respectively, and $T_h = \left(t_{jk}^{h}\right)$ is a $p \times q$ matrix of statistics.

The class of estimators for the vector of population mean $\underline{\overline{Y}}$ may be defined as

$$\tilde{\overline{\underline{Y}}}_{(st)} = \sum_{h=1}^{L} W_h \left[\hat{\overline{\underline{Y}}}_h + T_h \left(\underline{\overline{X}}_h - \hat{\overline{\underline{X}}}_h\right)\right] \tag{5.3.1}$$

where $T_h = \begin{pmatrix} t_{11}^h & t_{12}^h \cdots t_{1q}^h \\ \vdots & \vdots \\ t_{p1}^h & t_{p2}^h \cdots t_{pq}^h \end{pmatrix}_{p\times q}$, and t_{jk}^h are suitably chosen statistics such that their means exits. It may be noted that parallel to random sampling case several interesting estimators may be generated from $\tilde{\overline{\underline{Y}}}_{(st)}$ for specific choices of T_h.

We will consider only the class of estimators (5.3.1) when T_h is a pre-specified non-random matrix.

5.4 CRITERION OF OPTIMIZATION

For fixed T_h, $\tilde{\overline{\underline{Y}}}_{(st)}$ is unbiased for $\underline{\overline{Y}}$ and its MSE matrix $M\left(\tilde{\overline{\underline{Y}}}_{(st)}\right)$ is obtained below. We have

$$\left(\tilde{\overline{\underline{Y}}}_{(st)} - \underline{\overline{Y}}\right) = \sum_{h=1}^{L} W_h \left[\left(\hat{\overline{\underline{Y}}}_h - \underline{\overline{Y}}_h\right) - T_h \left(\hat{\overline{\underline{X}}}_h - \underline{\overline{X}}_h\right)\right].$$

On squaring both sides,

$$\left(\tilde{\bar{\underline{Y}}}_{(st)} - \bar{\underline{Y}}\right)^2 = \sum_{h=1}^{L} W_h^2 \left[\left(\hat{\bar{\underline{Y}}}_h - \bar{\underline{Y}}_h\right)^2 + T_h^2\left(\hat{\bar{\underline{X}}}_h - \bar{\underline{X}}_h\right)^2 - 2T_h\left(\hat{\bar{\underline{Y}}}_h - \bar{\underline{Y}}_h\right)\left(\hat{\bar{\underline{X}}}_h - \bar{\underline{X}}_h\right)\right].$$

Taking expectation on both sides, we have

$$E\left(\tilde{\bar{\underline{Y}}}_{(st)} - \bar{\underline{Y}}\right)^2 = \sum_{h=1}^{L} W_h^2 \begin{bmatrix} E\left(\hat{\bar{\underline{Y}}}_h - \bar{\underline{Y}}_h\right)\left(\hat{\bar{\underline{Y}}}_h - \bar{\underline{Y}}_h\right)' + T_h^2 E\left(\hat{\bar{\underline{X}}}_h - \bar{\underline{X}}_h\right)\left(\hat{\bar{\underline{X}}}_h - \bar{\underline{X}}_h\right)' \\ -2T_h E\left(\hat{\bar{\underline{Y}}}_h - \bar{\underline{Y}}_h\right)\left(\hat{\bar{\underline{X}}}_h - \bar{\underline{X}}_h\right)' \end{bmatrix}$$

Or

$$M\left(\tilde{\bar{\underline{Y}}}_{(st)}\right) = \sum_{h=1}^{L} W_h^2 \left[V_{yy} + T_h V_{xx} T_h' - T_h C_{yx}' - C_{yx} T_h'\right], \qquad (5.4.1)$$

where

$$V_{yy} = E\left(\hat{\bar{\underline{Y}}}_h - \bar{\underline{Y}}_h\right)\left(\hat{\bar{\underline{Y}}}_h - \bar{\underline{Y}}_h\right)'$$

$$V_{xx} = E\left(\hat{\bar{\underline{X}}}_h - \bar{\underline{X}}_h\right)\left(\hat{\bar{\underline{X}}}_h - \bar{\underline{X}}_h\right)'$$

and $C_{yx} = E\left(\hat{\bar{\underline{Y}}}_h - \bar{\underline{Y}}_h\right)\left(\hat{\bar{\underline{X}}}_h - \bar{\underline{X}}_h\right)'$.

Now, we consider the following criteria of preference given by Tripathi & Chaubey (2000):

Let $\underline{M}(\underline{Z}_y) = E\left[(\underline{Z}_y - \underline{\bar{Y}})(\underline{Z}_y - \underline{\bar{Y}})'\right]$ denote the mean square error (MSE) matrix of an estimator $\underline{Z}_y$ of $\underline{\bar{Y}}$.

C.P. (1) : An estimator $\underline{Z}_y$ is said to be better than another estimators $\underline{Z}'_y$ of $\underline{\bar{Y}}$ if and only if $M(\underline{Z}'_y) - M(\underline{Z}_y)$ is non negative definite whatever be the value of $\underline{y}_1, \ldots, \underline{y}_N$.

C.P. (2): Let $C = \{\underline{Z}_y\}$ be a class of estimators of $\underline{\bar{Y}}$. An estimator $\underline{Z}_{oy} \in C$ is said to be optimum for $\underline{\bar{Y}}$ in the class C if and only if $M(\underline{Z}_y) - (\underline{Z}_{oy})$ is non-negative definite (n.n.d.) for all $\underline{Z}_y (\neq \underline{Z}_{oy})$ in the class C and for all possible values of $\underline{y}_1, \ldots, \underline{y}_N$.

We will find the optimum value of T_h in (5.3.1) under the criterion C.P. (2).

5.5 OPTIMUM CHOICE OF T_h

For obtaining the optimum choice of T_h, we differentiate (5.4.1) w.r.t. T_h and equate to zero.

$$\frac{\partial\left(M\left(\underline{\tilde{\bar{Y}}}_{(st)}\right)\right)}{\partial T_h} = \sum_{h=1}^{L} W_h^2 \left[-2C'_{yx} + 2T_h' V_{xx}\right] = 0$$

$$\Rightarrow -2C_{yx}^{'} + 2T_h^{'} V_{xx} = 0$$

$$2T_h^{'} V_{xx} = 2\, C_{yx}^{'}$$

$$T_h^{opt} = C_{yx} V_{xx}^{-1}. \qquad (5.5.1)$$

Substituting the optimum value of T_h in (5.4.2), we have

$$M\left(\tilde{\bar{\underline{Y}}}_{(st)}\right) = \sum_{h=1}^{L} W_h^2 \begin{bmatrix} V_{yy} + C_{yx} V_{xx}^{-1} V_{xx} C_{yx}^{'} V_{xx}^{-1} - C_{yx} V_{xx}^{-1} C_{yx}^{'} \\ - C_{yx} V_{xx}^{-1} C_{yx}^{'} \end{bmatrix}$$

$$= \sum_{h=1}^{L} W_h^2 \left[V_{yy} + C_{yx} V_{xx}^{-1} C_{yx}^{'} - 2C_{yx} V_{xx}^{-1} C_{yx}^{'} \right]$$

$$= \sum_{h=1}^{L} W_h^2 \left[V_{yy} - C_{yx} V_{xx}^{-1} C_{yx}^{'} \right].$$

Hence, optimum MSE Matrix of $\bar{\underline{Y}}$ is given by

$$M\left(\tilde{\bar{\underline{Y}}}_{(st)}^{opt}\right) = \sum_{h=1}^{L} W_h^2 \left[V_{yy} - C_{yx} V_{xx}^{-1} C_{yx}^{'} \right]. \qquad (5.5.2)$$

Now, consider the difference

$$M\left(\tilde{\tilde{\underline{Y}}}_{(st)}\right)-M\left(\tilde{\tilde{\underline{Y}}}_{(st)}^{opt}\right)=\sum_{h=1}^{L}W_h^2\left[V_{yy}+T_hV_{xx}T_h'-T_hC_{yx}'-C_{yx}T_h'\right]$$
$$-\sum_{h=1}^{L}W_h^2\left[V_{yy}-C_{yx}V_{xx}^{-1}C_{yx}'\right]$$

$$=\sum_{h=1}^{L}W_h^2\left[T_hV_{xx}T_h'-T_h\,C_{yx}'-C_{yx}\,T_h'+C_{yx}V_{xx}^{-1}C_{yx}'\right]$$

$$=\sum_{h=1}^{L}W_h^2\begin{bmatrix}T_hV_{xx}T_h'-T_h^{opt}V_{xx}T_h^{opt'}+T_h^{opt}V_{xx}T_h^{opt'}\\-T_hC_{yx}'-C_{yx}T_h'+C_{yx}V_{xx}^{-1}C_{yx}'\end{bmatrix}$$

$$=\sum_{h=1}^{L}W_h^2\begin{bmatrix}T_hV_{xx}T_h'-T_h^{opt}V_{xx}T_h^{opt'}+T_h^{opt}C_{yx}'-T_h\,C_{yx}'\\-C_{yx}T_h'+C_{yx}T_h^{opt'}\end{bmatrix}$$

$$=\sum_{h=1}^{L}W_h^2\begin{bmatrix}T_hV_{xx}T_h'-T_h^{opt}V_{xx}T_h^{opt'}-\left(T_h-T_h^{opt}\right)C_{yx}'\\-C_{yx}\left(T_h-T_h^{opt}\right)'\end{bmatrix}$$

$$=\sum_{h=1}^{L}W_h^2\begin{bmatrix}\left(T_h-T_h^{opt}\right)V_{xx}\left(T_h-T_h^{opt}\right)'+\left(T_h-T_h^{opt}\right)\left(V_{xx}T_h^{opt}-C_{yx}'\right)\\+\left(T_h^{opt}V_{xx}-C_{yx}\right)\left(T_h-T_h^{opt}\right)'\end{bmatrix}. \quad (5.5.3)$$

Since the first term on the RHS of (5.5.3) is non-negative definite (n.n.d.), the difference on the LHS for $T_h \neq T_h^{opt}$ can be made n.n.d. if and only if

$$T_h = C_{yx}\,V_{xx}^{-1}. \quad (5.5.4)$$

Hence the optimum choice of T_h w.r.t. the criterion C.P.(2) is as given in (5.5.4).

BIBLIOGRAPHY

Aghezzaf, B. (2001) On generalized convexity and duality in multiobjective programming problems *Advanced Modeling and Optimization*, **3**, (1), 7-15

Aghezzaf, B. and **Ouaderhman, T.** (2000) An interior multiobjective linear programming algorithm *Advanced Modeling and Optimization*, **2**, (1), 41-52

Ahsan, M.J. and **Khan, S.U.** (1977) Optimum allocation in multivariate stratified random sampling using prior information *Jour Ind. Stat. Assoc* , **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

Aoyama, H. (1963) Stratified random sampling with optimum allocation for multivariate populations *Ann. Inst Stat. Math* , **14**, 251-258

Arvanitis, L.G. and **Afonja, B.** (1971) Use of the generalized variance and the gradient projection method in multivariate stratified sampling *Biometrics*, **27**, 110-127

Bellman, R.E. (1957) Dynamic Programming. *Princeton University Press, Princeton*

Bethel, J. (1985) An optimum allocation algorithm for multivariate survey *Proceedings of the Survey Research Section, ASA*, 204-212

Bethel, J. (1989). Sample allocation in multivariate surveys. *Survey of Methodology*, **15**, (1), 47-57.

Bowley, A.L. (1926). Measurements of precision attained in sampling. *Bull. Inst. Internat. Stat.*, **22**, 1-62.

Chaddha, R.L., Hardgrave, W.W., Hudson, D.J., Segal, M. and **Suurballe, J.W.** (1971). Allocation of total sample size when only the stratum means are of interest. *Technometrices*, **13**, (4), 817-831.

Chakraborty, M. and **Dubey, O.P.** (2001). Goal programming with quadratic preferences-An interactive approach. *IJOMAS*, **17**, (1), 25-34.

Chakraborty, M. and **Sinha, A** (1995). Multiobjective transportation problem: A goal programming approach. *Industrial Engineering Journal*. **24**, (7), 17-22.

Charnes, A. and **Cooper, W.W.** (1961). Management models and industrial applications of linear programming. *John Wiley and Sons, New York, Kluwer Publishers, Boston.*

Charnes, A. and **Cooper, W.W.** (1977). Goal programming and multiple objective optimizations, Part I. *European Jour. Oper. Res.*, **1**, 39-54.

Chatterjee, S. (1967). A note on optimum allocation. *Skand. Akt.* **50**, 40-44.

Chatterjee, S. (1968). Multivariate stratified surveys. *Jour. Amer. Stat. Assoc.* **63**, 530- 534.

Chernayak, O.I. and **Chornous, G.** (2000). Optimal allocation in stratified sampling with a non-linear cost function *Theory of Stochastic Processes*, **6** (22) 3-4, 6-17

Chernayak, O.I. and **Starytskyy, A.** (1998) Optimal allocation in stratified sampling with convex cost function *Visnyk of kviv University. Economics*, (39), 42-46, *(In Ukrainian)*.

Chromy, J.R. (1987) Design optimization with multiple objectives *Proceedings of the Survey Research Section, ASA*, 194-199

Cochran, W.G. (1940) The estimation of the yields of the cereal experiments by sampling for the ratio of grain to total produce *Jour Agricul. Sc.*, **30**, 262-275

Cochran, W.G. (1942) Sampling theory when the sampling units are of unequal sizes *Amer. Stat. Assoc.*, **37**, 199-212

Cochran, W.G. (1963) Sampling Techniques, (2nd ed) *Wiley, New York*

Cochran, W.G. (1977) Sampling Techniques (3rd ed) *John Wiley and Sons, New York*

Dalenius, T. (1953) Multivariate sampling problem *Skand. Akt.*, **36**, 92-122

Dalenius, T. (1957) Sampling in sweden *Almqvist and Wiksell, Stockholm.*

Das, A.K. (1988) Contribution to the theory of sampling strategies based on auxiliary information *Ph.D thesis submitted to B C.K.V., Mohanpur, Nadia, West Bengal, India*

Das, A.K. and **Tripathi, T.P.** (1980) Sampling strategies for population mean when the coefficient of variation of an auxiliary character is known *Sankhya,* Ser C, **42**, 76-86

Ericsion, W.A. (1965) Optimum stratified sampling using prior information *Jour. Amer, Stat. Assoc.*, **60**, 750-772

Geary, R.C. (1949) Sampling methods applied to Irish agraicultural statistics *Technical Series,* Sept 1949

Ghosh, S.P. (1958) A note on stratified random sampling with multiple characters *Cal. Stat Assoc Bull.*, **8**, 81-89

Gupta, S. and **Chakraborty, M.** (1997) Multiobjective linear programming A Fuzzy programming approach *IJOMAS,* **13**, (2), 207-214

Han-Lin Li and **Chian-Son Yu** (2000) A fuzzy multibojective program with quasiconcave membership functions and fuzzy coefficients *Fuzzy Sets and Systems,* **109**, 59-81

Hansen, M.H. and **Hurwitz, W.N.** (1943) On the theory of sampling from finite populations *Ann. Math. Stat* , **14**, 333-362

Hansen, M.H and **Hurwitz, W.N.** (1951) Response errors in surveys *Jour. Amer. Stat. Assoc, 46, 147-190.*

Hansen, M.H., Hurwitz, W.N. and Madow, W.G. (1953) Sample survey methods and theory *John Wiley and Sons, New Yorks.*

Huddleston, H.F., Claypool, P.L. and **Hocking, R.R.** (1970) Optimal sample allocation to strata using convex programming *Applied Statistics,* **19**, 273-278

Ignizio, J P. and **Cavalier, T. M.** (1994). Linear Programming. *Prentice Hall International, New Jersey.*

Ignizio, J.P. (1976). Goal programming and extensions. *Lexington, Mass: Heath (Lexington books).*

Ignizio, J.P. (1983). Generalized goal programming. *Special Issue of Computers and Operations Research,* **10**, (4).

Ignizio, J.P. (1985). Introduction to linear Goal programming. *Beverly Hills: Sage.*

Jahan, N., Khan, M.G. and **Ahsan, M.J.** (1994). A generalized compromise allocation. *Jour. Ind. Stat. Assoc.,* **32**, 95-101.

Kassem, M.A.E. and **Ammar, E.E.** (1996). A parametric study of multiobjective NLP problems with fuzzy parameters in the objective functions. *Fuzzy Sets and Systems,* **80**, 187-196.

Kelly, J.E. (1960). The cutting-plane method for solving convex programs. *Jour. Soc. Indust. Appl. Math.,* **8**, 703-712.

Khan, E. A (1997). On use of mathematical programming techniques in some optimization problems arising in stratified sample surveys. *Ph.D. thesis submitted to Aligarh Muslim University, Aligarh, India.*

Khan, S.U. and **Bari, A.** (1977). A procedure for integer solution to some allocation problem. *Pure and Applied Mathematika Sciences,* **5**, 25-32.

Khan, S.U. and **Tripathi, T.P.** (1967). The use of multi-auxiliary information in double sampling. *Jour. Ind. Stat. Assoc.,* **5**, 42-48.

Khorramshahgol, R. and **Hooshiari, A.** (1991) Three shortcomings of goal programming and their solution *Jour. Inform Optim. Sc* **12**, (3), 459-566

Kish, L. (1965) Survey sampling *John Wiley.*

Kokan, A.R. (1963) Optimum allocation in multivariate surveys *Jour Roy Stat Soc , Sec* A, **126**, 557-565

Kokan, A.R. and **Khan, S.U.** (1967) Optimum allocation in multivariate surveys an analytical solution *Jour Roy Stat Soc ,* Ser B, **29**, 115-125

Malhotra, N. and **Arora, S.R.** (1999) An algorithm for solving Bilevel programming problem-Goal programming approach *IJOMAS*, **15**, (3), 235-244

March, J.G. and **Simon, H.A.** (1958) Organizations *New York John Wily*

Mohan, C. and **Nguyen, H.T.** (1999) Preference level interactive method for solving multiobjecitve Fuzzy programming problems *Asia-Pacific Journal of Operational Research,* **16**, 63-86

Mukherjee, R., Rao, T.J. and **Vijayan, K.** (1987) Regression type estimators using multi-auxiliary information *Australian Jour. Stat.,* **29**, 244-254

Murthy, M.N. (1964) Product method of estimation *Sankhya,* Ser A, **26**, 294-307

Murthy, M.N. (1967) Sampling Theory and Methods *Statistical Publishing Society, Calcutta*

Nandi, S.B. and **Aich, A.B.** (1995). Optimal stratified sampling an information theoretic approach. *Second International Triennial Calcutta Symposium on Probability and Statistics.*

Narasimhan, R. (1980). Goal programming in Fuzzy environment. *Decision Science,* **11**, 326.

Neyman, J. (1934). On the two different aspects of representative method: The method of stratified sampling and the method of purposive selection. *Jour. Roy. Stat.* Soc., **97**, 558-606.

Neyman, J. (1938). Contributions to the theory of sampling human populations. *Jour. Amer. Stat. Assoc.,* **33**, 101-116.

Olkin, I. (1958). Multivariate ratio estimation for finite populations. *Biometrika,* **45**, 154-165.

Omule, S.A.Y. (1985). Optimum design in multivariate stratified sampling. *Biom. Jour.,* **27**, (8), 907-912.

Peter, J.H. and **Bucher** (1940). The 1940 section sample survey of crop aggregates in Indiana and Iowa, U.S. Dept. of Agriculture.

Rahim, M.A. (1994). Sample allocation in multivariate stratified design: an alternative to convex programming. *Abstracts for Survey Research Methods Joint Statistical Meetings.*

Raj, D. (1965). On a method of using multi-auxiliary information in sample surveys. *Jour. Amer. Stat. Assoc.,* **60**, 270-277.

Raj, D. (1968). Sampling Theory. *Tata McGraw Hill.*

Rao, P.S.R.S. and **Mudholkar, G.S.** (1967). Generalized multivariate estimators for the mean of a finite population.

Jour. Amer. Stat. Assoc., **62**, 1009-1012

Robson, D.S. (1957) Applications of multivariate polykays to the theory of unbiased ratio-type estimators *Jour. Amer. Stat Assoc* , **52**, 511-522

Roy, B. (1971) Problems and methods with multiple functions *Mathematical Programming*, **1**, 239-266

Simon, H.A. (1957) Administrative Behavior *New York· Macmillan*

Singh, M.P. (1967) Multivariate product method of estimation for the finite population *Jour. Ind Soc. Agricul. Stat* , **19**, 1-10

Soland, R.M. (1967) Optimum multivariate stratified sampling with prior information *Skand Akt* , **5**, 31-39

Soyibo, A. (1985) Goal programming methods and applications A survey *Jour Inform Optim. Sc* , **6**, (3), 247-264

Srivastava, S.K. (1965) An estimate of the mean of a finite population using several auxiliary character *Jour. Ind. Stat Assoc* , **3**, 189-194

Srivastava, S.K. (1966) On ratio and linear regression method of estimation with several auxiliary variables *Jour. Ind Stat Assoc* , **4**, 66-72

Srivastava, S.K. (1971) A generalized estimator for the mean of a finite population using multiauxiliary information *Jour. Amer Stat Assoc* , **66**, 404-407

Srivastava, S.K. (1980) A class of estimators using auxiliary information in sample surveys *Canadian Jour. Stat.*, **8**, 253-254

Sukhatme, P.V., Sukhatme, B.V., Sukhatme, S. and **Ashok, C.** (1984) Sampling theory of surveys with applications *Iowa State University Press, U S.A.*

Tripathi, T.P. (1970) Contributions to the sampling theory using multivariate information *Ph D thesis submitted to Punjabi University, Patiala, India*

Tripathi, T.P. (1973) Double sampling for inclusion probabilities and regression method of estimation *Jour. Ind. Stat. Assoc* , **10**, 33-46

Tripathi, T.P. (1976) On double sampling for multivariate ratio and difference method of estimation *Jour. Ind. Stat. Assoc* **33**, 33-54

Tripathi, T.P. (1987) A class of estimators for population mean using multivariate auxiliary information under general sampling designs *Aligarh Jour Stat.*, 7, 49-62

Tripathi, T.P. (1989) Optimum estimation of mean vector for dynamic population *Invited paper in the proceedings of the International Symposium on Optimization and Statistics, held at A.M.U. Aligarh,* 8-21 Dec 1989

Tripathi, T.P. and **Chaubey, Y.P.** (1993) Optimum probabilities of the selection in PPS sampling based on super population model and multivariate information *Tech. Report No. 2/93,* Stat-Math, ISI, Calcutta

Tripathi, T.P. and **Chaubey, Y.P.** (2000) Estimators for finite population mean vector based on multivariate auxiliary information *Current development in survey sampling, Editor A K P C Swain, The Modern Book Depot Bhubaneshwar, India*

Tripathi, T.P. and **Khattree, R.** (1989) Simultaneous estimation of several means using multivariate auxiliary information *Tech Report No 13/89 Stat-Math Unit, Indian Statistical Institute, Calcutta, India*

Watson, D.J. (1937) The estimation of lead areas *Jour Agricul Sc*, **27**, 474

Yates, F. (1960) Sampling methods for censuses and surveys, (2nd ed) *Charles Griffin and Co Ltd London*

Zimmermann, H.J. (1978) Fuzzy programming and linear programming with several objective functions *Fuzzy Sets and Systems*, **1**, 45-55

Zimmermann, H.J. (1981) Fuzzy mathematical programming *Computers and operations research*, **4**, 291-298

Zimmermann, H.J. (1985) Fuzzy sets and theory and its applications *Boston Kluwer*