PREFACE
PREFACE

This thesis entitled "On Some Optimum Stratified Sampling Designs" is submitted to the Aligarh Muslim University, Aligarh, India, to supplicate the degree of Doctor of Philosophy in Statistics. It embodies the research work carried out by me in the Department of Statistics and Operations research, Aligarh Muslim University, Aligarh.

Sample surveys play an indispensable role in almost every sphere of human activity. In sample surveys the precision suffers when too small a sample is selected and the cost increases when the sample is too large. Since the objective of the sample surveys is to obtain maximum information about population under study at minimum cost, time and other available resources. The basic problem in sample surveys is to choose a sampling design that either assures the maximum precision of the sample estimates for a given cost of the survey or assumes the minimum cost for a given level of precision as the case may be. Such instances in sample surveys may be handled by formulating and solving them as optimization problems. In this thesis an attempt has been made to formulate and solve some problems arising in multivariate stratified sampling.
This thesis consists of five chapters. Chapter 1 is the usual introductory chapter. It provides an introduction to sample surveys, census, various sampling designs and a detailed introduction to multivariate stratified sampling with various sample size allocations. Problem of nonresponse in sample surveys is also introduced. This chapter also gives a brief idea of the application of mathematical programming techniques in sample surveys. The numerical solutions in this thesis are mostly obtained through the optimization software LINGO (2001). Its brief introduction is also presented in this chapter.

In Chapter 2 the problem of determining the integer compromise allocation in multivariate stratified sampling is considered. The problem is formulated as a Multiobjective Integer Nonlinear Programming Problem. Using the value function technique the problem is converted into a single objective problem. Relaxing some of the restrictions a formula for continuous sample sizes is obtained using Lagrange Multipliers Technique. The integer solution is obtained by rounding off the continuous sample sizes to their nearest integer values. In the later part of the chapter the problem of finding a compromise allocation for a multivariate stratified sample survey with a significant travel cost within strata is formulated as a Multiobjective Integer Nonlinear Programming Problem. A solution procedure is proposed
using the goal programming technique. The numerical examples are also presented to illustrate the computational details. This chapter is based on my joint research papers Khowaja et al. (2011a) and Ghufran et al. (2011a) published in the Journals “Communication in Statistics-Simulation and Computation and the “South Pacific Journal of Natural and Applied Sciences” respectively.

Chapter 3 deals with the problem of optimum allocation in multivariate stratified sampling to estimate the \( p \)–overall population means. The objective is to minimize, simultaneously, the coefficients of variation of the estimators of \( p \)–population means under a cost constraint that includes the measurement as well as the travel cost. Usually, some function of the sampling variances of the estimators of the \( p \)–population means is used as an objective function that is to be minimized for a fixed cost given as a linear function of sample allocations. Since the variances are not unit free it is more logical to consider the minimization of some function of squared coefficient of variations as an objective function. Furthermore, the investigators have to approach the sampled units in order to get the observations. This involves some travel cost. Usually this cost is neglected while constructing a cost function. The travel cost may be substantial in
some cases. For example if the strata consist of geographically difficult areas.

The formulated problem of obtaining an optimum compromise allocation turns out to be a Multiobjective All Integer Nonlinear Programming Problem. Additional restrictions are imposed on the sample sizes to avoid oversampling and to ensure the availability of the estimates of the strata variances. In this chapter three different methods are considered to solve the formulated problem viz the Value function method, the $\epsilon$- Constraint method and the Distance based method. Numerical examples are presented to illustrate the computational details of the proposed methods. This chapter is based on my research paper Ghufran et al. (2012a) published in the journal “Communications in Statistics-Simulation and Computation”.

In Chapter 4 the optimum compromise allocation in multivariate stratified sampling with non-linear objective function and probabilistic non-linear cost constraint are considered. The probabilistic non-linear cost constraint is converted into equivalent deterministic one by using Chance Constrained Programming. A numerical example is presented to illustrate the computational details. This work is based on my research paper Ghufran et al. (2011b) published in the “International Journal of Engineering, Science and Technology”.
In Chapter 5 the use of randomized response technique to reduce the rate of non-response in univariate as well as in multivariate stratified random sampling has been explored. This chapter consists of three parts. In the first part, Two-stage stratified Warner’s randomized response model is used to determine the optimum allocation in the presence of non-response. The problem is formulated as a Nonlinear Programming Problem and a method of solution is proposed. Two numerical examples are worked out to illustrate the computational details of the proposed method. In the second part, Two-stage Stratified Warner’s Randomized Response model is used to determine the optimum allocation in the presence of non-response and with a nonlinear cost function which takes care of the travel cost within strata.

Social surveys are often based on questionnaires containing more than one sensitive question, thus the analysis of multiple RR data is of considerable interest. In multivariate stratified surveys with multiple RR data the choice of optimum sample sizes from various strata may be viewed as a Multiobjective Nonlinear Programming Problem (MNLPP). In the third part the randomized response model is extended for multiple sensitive questions. The problem of obtaining a compromise allocation is formulated as MNLPPs with linear and quadratic cost functions separately. The solution to
the formulated problems are achieved through Goal Programming Technique.

This chapter is based on my research paper Ghufran et al. (2012b) published in the “Optimization Letters” and Ghufran et al. (2013) published in the “Journal of Mathematical Modelling and Algorithms in Operations Research”.

A comprehensive list of references, arranged in alphabetical order, is provided at the end of this thesis.