ON SOME MATHEMATICAL PROGRAMMING PROBLEMS IN SAMPLE SURVEYS AND RELIABILITY

ABSTRACT

SUBMITTED TO

BABASAHEB BHIMRAO AMBEDKAR UNIVERSITY

(**A CENTRAL UNIVERSITY**)

LUCKNOW

BABASAHEB
BHIMRAO
AMBEDKAR
UNIVERSITY

• LUCKNOW •

प्रज्ञा शील करुणा
ESTABLISHED 1996

FOR THE AWARD OF DEGREE OF

Doctor of Philosophy

IN

APPLIED STATISTICS

Submitted by

Mradula

Under the Supervision Of

Dr. Rahul Varshney

Assistant Professor

DEPARTMENT OF APPLIED STATISTICS

SCHOOL FOR PHYSICAL SCIENCES

BABASAHEB BHIMRAO AMBEDKAR UNIVERSITY

(A CENTRAL UNIVERSITY)

Vidya Vihar, Raebareli Road Lucknow-226 025 (U.P.), India

ENROLMENT NO: 1151/15 **2021**

Abstract

This thesis entitled “On Some Mathematical Programming Problems in Sample Surveys and Reliability" is submitted to the Babasaheb Bhimrao Ambedkar University, Lucknow, India, to supplicate the degree of Doctor of Philosophy in Applied Statistics. It embodies the research work carried out by me in the Department of Applied Statistics, Babasaheb Bhimrao Ambedkar University, Lucknow.

Optimization is a technique to find the best solution among all available solutions for the formulated mathematical programming problems. In the modern world, sample survey and reliability 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. A sample survey is the most efficient, scientific, economical and reliable data collection method for collecting information and suggesting the best estimator to predict unknown population parameters. The objective of the sample survey is to provide these information, about the population under study, with the use of minimum budget and other available resources in feasible time. The problem arises in conducting the sample surveys is to opt a sampling design that either assures the maximum precision of the sample estimates for the given budget of survey or suggest the minimum budget for a given

level of precision on estimates as the case may be. In sampling theory various sampling designs are suggested for different situations and the nature of the population. The emphasis is given, in this thesis, on the formulation of mathematical programming problems of sample surveys and reliability for univariate and multivariate cases. Also, the solutions of mathematical programming problems are suggested by using some suitable optimization techniques.

In stratified sample surveys, the problem of determining the sample sizes from the strata is known as the problem of allocations and several types of allocations i.e. equal, proportional, and optimum allocations are available in literature. Among these, an optimum allocation is considered as the best allocation. In multivariate stratified sampling, where more than one characteristic is defined on each population unit, the optimum allocation for one characteristic will not, in general, be optimum for other characteristics. Thus, the problem of allocation becomes more complex and the problem of sample allocations in multivariate stratified sampling has been discussed by various authors for a long time, starting apparently with Neyman (1934). In such circumstances, a compromise criterion is required, resulting in an optimum allocation for all the characteristics in some sense. Since this allocation is based on a compromise criterion, it is known as a compromise allocation in multivariate stratified sampling. Many authors such as Ali *et al.* (2015), Varshney *et al.* (2015), Haseen *et al.* (2016), Raghav *et al.* (2017), Varshney *et al.* (2017), Ansari *et al.* (2018) and others, worked out for compromise allocation by defining new estimators or used existing estimators in different situations for multivariate case. In this thesis, the problem of obtaining a compromise allocation in various situations is studied in detail. When population means are to be evaluated, Yates (1960) recommended the criterion.

In many decision-making problems, sample surveys, environmental, social, economic, and technical areas have more than one multiple objective functions. It is significant to realize that multiple objectives are often non commensurable and in conflict with each other in

optimization problems. The problem of deriving the statistical information on population characteristics may be formulated as an optimization problem of minimizing the cost of survey subject to the restriction that the loss of precision is within the prescribed limit or alternately minimizing the loss in precision subject to the restriction that the cost of the survey remains within the given budget. Some authors formulated these problems as optimization problems and suggested using various mathematical programming techniques to solve them, such as dynamic programming, lexicographic goal programming, fuzzy programming, genetic algorithm, particle swarm optimization, etc.

Reliability theory has been extensively developed over the years, since World War II, with a significant contribution by defense personnel. According to the Aeronautical (1994), reliability is the probability that a system will perform satisfactorily for at least a given period when used under stated conditions. Reliability is used to assess the system structures and a significant design measure in many industrial environments such as telecommunication and manufacturing facilities. The design of such hardware systems may be formulated as a reliability optimization problem. The objective of such problems is to either maximize reliability, availability, or performance or minimize cost. In a real-life system, reliability problems data is not always precise due to some uncontrollable and unavoidable situations. Therefore, to handle the imprecise data, the mathematical programming problems may be used to formulate the optimization problems and an appropriate decision may be taken by using appropriate mathematical programming techniques.

The reliability of a system may be maximized subject to the resource constraints, to determine the optimum number of redundant components for each stage, provided that the reliability of each component is known. Recently Kim and Kim (2017), Kakkar *et al.* (2019) and Zhiyuan *et al.* (2019), have obtained different models of reliability and maintenance, explored their best solutions, and applied them to various systems. Reliability has expanded its scope across different sectors in recent years, thus functioning as an integral

element of quality within its structure and development cycle. The real-life problems, such as projects and case studies, may be formulated as mathematical programming problems, and a solution may be worked out by using an appropriate optimization technique.

The aim of this thesis is to explore the formulations of mathematical programming problems in sample surveys and reliability theory and their solutions by using some suitable optimization techniques. It is also aimed to inquire whether optimization techniques have any significant effects on sampling and reliability theory. The discussions in this thesis are included on nonresponse problems, probabilistic environments and boundaries problems under stratified random sampling. For reliability computation, the series-parallel system has been used for the availability and cost constraints under particle swarm optimization technique.

The thesis comprises six chapters. Among them, **Chapter 1** provides a primary reading material to equip the reader for understanding the problems discussed in the following chapters.

Chapter 2 deals with the problem of nonresponse problems which was introduced by Hansen and Hurwitz (1946). The nonresponse problem is solved by introducing the Gamma cost function for the problem and a comparative study, by including some existing compromise criteria with the proposed method, is given. The numerical illustration is also included by using lexicographic goal programming technique for the practical utility of the proposed method. This problem is based on my published research work Varshney and Mradula (2019) and published in the *Journal of Statistical Computation and Simulation.*

In **Chapter 3** a procedure for determining the optimum stratum boundaries and sample size for the sample, with a nonlinear fixed cost function, is explained. The problem is formulated as a nonlinear programming problem. In this chapter the problems are discussed for the population which may follow the distributions like Weibull, Gamma, Log-normal,

etc.

In the **Chapter 4** the use of stochastic programming problem under some probabilistic constraints is discussed for multivariate stratified sample survey. The problem is formulated as a stochastic multiobjective nonlinear programming problem by using a nonlinear cost function. The stochastic nonlinear programming problem is solved using the chance constraint and modified E-model methods, and a solution is suggested by using fuzzy goal programming technique and Lagrange's method. Also a comparative study is included for the practical utility of the method. This problem is based on my accepted research work in the *Journal of King Saud University - Science.*

In **Chapter 5**, the study of the ratio type estimators, where the population mean is unknown but the population median of study variable is known, is carried out and the use of auxiliary information for improving the precision of estimates of the population mean or total is explored. The problem is formulated in two parts,

(i) with linear cost function and solved using the integer programming technique and Lagrange multiplier technique and numerical illustration for the method's practical utility. This part is based on my published research work Mradula *et al.* (2019) and published in the *Journal of Communications in Statistics - Simulation and Computation*.

(ii) With nonlinear cost function and solved it using integer programming problem with MATLAB software.

A comparative study of different other estimators has been made.

In **Chapter 6**, the availability and cost function of a series-parallel system are expressed as a multiobjective model under a fixed time environment. The importance of the solution methods has been discussed with an numerical illustration using particle swarm optimization technique. The proposed method is compared with some other different existing optimization techniques such as genetic algorithm and goal programming technique and the

obtained results explained the effectiveness of the proposed approach in terms of availability and the cost maintenance with system time. The simulation studies are included, by using R software, to evaluate the availability and determining the cost function by using MATLAB.

The research work presented in Chapters 2 to 6 of this thesis is based on my following published/accepted/communicated research papers.

1. VARSHNEY, R. AND MRADULA. (2019). Optimum allocation in multivariate stratified sampling design in the presence of nonresponse with Gamma cost function. *Journal of Statistical Computation and Simulation.* **89(13)**, 2454–2467.

2. MRADULA., YADAV, S.K., VARSHNEY, R., AND DUBE, M. (2019). Efficient estimation of population mean under stratified random sampling with linear cost function. *Communications in Statistics - Simulation and Computation.*
DOI: https://doi.org/10.1080/03610918.2019.1643479.

3. AHMADINI, A .A. H., MRADULA., VARSHNEY, R., AND ALI, I. On multivariate-multiobjective stratified sampling design under probabilistic environment: a fuzzy programming approach. *Journal of King Saud University - Science.* (Accepted)

4. YADAV, S.K., MRADULA., AND VARSHNEY, R. Efficient estimation of population mean under stratified random sampling using non linear cost function. (Communicated)

5. MRADULA., VARSHNEY, R., AND DUBE, M. Determining optimum strata boundaries using auxiliary variable with nonlinear cost function. (Communicated)

6. MRADULA., VARSHNEY. R., AND ALI, I. On availability and cost function for maintenance time of continuous operating series-parallel systems using particle swarm optimization. (Communicated)

Bibliography

Aeronautical, R. I. (1994). *Reliability engineering*. Prentice Hall, Englewood Cliffs, NJ.

Ali, I., Shafiullah., and Haseen, S. (2015). Geometric programming approach in multivariate stratified sample surveys in case of non-response. *Electronic Journal of Applied Statistical Analysis*, 8 (1), 28–43.

Ansari, A. H., Varshney, R., and Ahsan, M. (2018). Compromise mixed allocation in multivariate stratified sampling using dynamic programming technique. *Journal of Advanced Statistics*, 3 (4), 71–79.

Hansen, M. H. and Hurwitz, W. N. (1946). The problem of non-response in sample surveys. *Journal of the American Statistical Association*, 41 (236), 517–529.

Haseen, S., Ali, I., and Bari, A. (2016). Multiobjective stochastic multivariate stratified sampling in presence of nonresponse. *Communications in Statistics - Simulation and Computation*, 45 (8), 2810–2826.

Kakkar, M. K., Bhatti, J., Malhotra, R., Kaur, M., and Goyal, D. (2019). Availability analysis of an industrial system under the provision of replacement of a unit using genetic algorithm. *International Journal of Innovative Technology and Exploring Engineering*, 9 (1), 1236–1241.

Kim, H. and Kim, P. (2017). Reliability redundancy allocation problem considering optimal redundancy strategy using parallel genetic algorithm. *Reliability Engineering and System Safety*, 159 (C), 153–160.

Mradula, Yadav, S. K., Varshney, R., and Dube, M. (2019). Efficient estimation of population mean under stratified random sampling with linear cost function. *Communications in Statistics - Simulation and Computation*. DOI: https://doi.org/10.1080/03610918.2019.1643479.

Neyman, J. (1934). On the two different aspects of the representative method: the method of stratified sampling and the method of purposive selection. *Journal of the Royal Statistical Society*, 97 (4), 558–625.

Raghav, Y. S., Khan, M. F., and Khalil, S. (2017). Sample allocation problem in multiobjective multivariate stratified sample survey under two stage randomized response model. *Journal of Mathematical and Computational Science*, 7 (6), 1074–1089.

Varshney, R., Khan, M. G. M., Fatima, U., and Ahsan, M. J. (2015). Integer compromise allocation in multivariate stratified surveys. *Annals of Operations Research*, 226 (1), 659–668.

Varshney, R., Gupta, S., and Ali, I. (2017). An optimum multivariate multiobjective stratified sampling design: fuzzy programming approach. *Pakistan Journal of Statistics and Operation Research*, 13 (4), 829–855.

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

Zhiyuan, O., Yu, L., Sheng-Jia, R., and Tao, J. (2019). An improved particle swarm optimization algorithm for reliability redundancy allocation problem with mixed redundancy strategy and heterogeneous components. *Reliability Engineering and System Safety*, 181 (C), 62–74.