Journal of King Saud University – Science 33 (2021) 101448

Contents lists available at ScienceDirect

Journal of King Saud University – Science

journal homepage: www.sciencedirect.com

Original article

On multivariate-multiobjective stratified sampling design under probabilistic environment: A fuzzy programming technique

Abdullah Ali H. Ahmadini [a], Rahul Varshney [b], Mradula [c], Irfan Ali [d,*]

[a] Department of Mathematics, Faculty of Science, Jazan University, Jazan, Saudi Arabia
[b] Department of Statistics, School of Physical & Decision Sciences, Babasaheb Bhimrao Ambedkar University, Lucknow 226 025, India
[c] Department of Statistics, School of Physical & Decision Sciences, Babasaheb Bhimrao Ambedkar University, Lucknow 226 025, India
[d] Department of Statistics & Operations Research, Aligarh Muslim University, Aligarh 202 002, India

ARTICLE INFO

Article history:
Received 18 January 2021
Revised 2 March 2021
Accepted 12 April 2021
Available online 22 April 2021

Keywords:
Multivariate-multiobjective stratified sampling
Stochastic programming
Fuzzy goal programming
Compromise allocation
Gamma cost function

ABSTRACT

In a multivariate stratified sampling design, the individual optimum allocation of one character may not remain optimum to other characteristics. For the solution of such problems, a usable allocation must be required to get precise estimates of the unknown population parameters, which may be near optimum to all characteristics in some sense. The compromise criterion is required to obtain such usable allocation in sampling literature. In this paper, the sample allocation problem is considered as a stochastic nonlinear programming problem and thereafter formulated into a multiobjective programming problem to provide the usable allocation. The formulated problem is solved by using different models of stochastic optimization. Afterwards, the proposed allocation is worked out and compared with some other allocations, which are well defined in sampling, to give a comparative study. Also, the numerical study defines the practical utility of the proposed technique.

1. Introduction

In many real-life practices, the populations may vary in their accessibility. Some parts of the population may be in remote locations, gated buildings or other inaccessible areas. For such situations, the choice of sampling design affects the results of the survey. Then the stratified sampling seems to be the best choice of the sampling technique. For obtaining detailed information about the characteristics of the population, a multivariate stratified sample survey is carried out by splitting the population into L strata. It is assumed that all the p characteristics are defined in each unit of the population. The estimation of unknown population means of p characteristics is required and may be carried out using the Nonlinear Programming Problem (NLPP). Cochran (1977) has been shown that the individual optimum allocation of one character may not remain optimum for others. A compromise criterion may be required to obtain the best allocation, which helps obtain precise information about population parameters. Therefore the allocation based on some compromised criterion is called compromise allocation in multivariate stratified sampling design. Most of the authors (Neyman, 1934; Kokan and Khan, 1967; Chatterjee, 1968; Ahsan, 1975; Khan et al., 1997; Semiz, 2004; Kozak, 2006; Varshney et al., 2012, 2015; Fatima et al., 2014; Muhammad et al., 2015; Muhammad and Husain, 2017; Varshney and Mradula, 2019) discussed the problems of allocation and worked out compromise allocation in multivariate stratified sample surveys. A compromise allocation is obtained either by suggesting different compromise criteria or using the suggested criteria under different conditions, i.e. in the availability of auxiliary information, presence of nonresponse, etcetera. In many sampling designs, the stratum variances are not known in advance but maybe estimable. From the deterministic point of view, such problems may be formulated as NLPPs. However, if the nature of the estimated variances is also considered, it will be an additional restriction to the problem, and therefore the compromise allocation may not be obtained easily. For such situations, the Stochastic Nonlinear Programming Problem (SNLPP) may help to work out the required compromise allocation to obtain sufficient information about population parameters (See (Charnes and Cooper, 1963; Prékopa, 1978; Díaz-García and Garay Tapia, 2007; Kozak and Wang,

* Corresponding author.
E-mail address: irfii.st@amu.ac.in (I. Ali).

Peer review under responsibility of King Saud University.

Production and hosting by Elsevier

https://doi.org/10.1016/j.jksus.2021.101448
1018-3647/

2010; Haseen et al., 2016)). Díaz-García and Ramos-Quiroga (2014) discussed and provided results by solving SNLPPs with a fixed linear cost function. The concept of fuzzy set theory has been discussed by Zadeh (1965) and then Bellman and Zadeh (1970). They utilized the fuzzy approach for dynamic issues. The idea of the fuzzy set was given by Zimmermann (1978) to convert the multiobjective linear programming problem into a single objective linear programming problem. Many authors solved sample allocation problems by using fuzzy programming techniques (See (Gupta et al., 2013; Ali and Hasan, 2013; Varshney et al., 2017; Haq et al., 2020). Fuzzy programming is one of many available optimization techniques that deal with optimization problems under uncertainty. The technique is flexible and thus helps decision-makers have a better understanding of their problems. Such techniques may be applied when situations are not clearly defined and also have uncertainty. The fuzzy programming technique is a more appropriate technique for solving the problem when the data has anykind of uncertainity. Fuzzy programming has been studied and applied recently by several authors in different areas (Elsisi, 2019a, 2019b, 2020; Fakhrzad and Goodarzian, 2019; Elsisi and Soliman, 2020; Fathollahi-Fard et al., 2020; Goodarzian and Hosseini-Nasab, 2021; Lu et al., 2020).

Under the probabilistic environment, the use of the nonlinear cost function is proposed by considering the labour cost as part of the survey's total cost. In this paper, the compromise criterion is suggested for determining the compromise allocation for a multiobjective-multivariate stochastic nonlinear programming problem for the fixed cost in the probabilistic situation. The solution procedure is also given to solve the formulated problem by using an appropriate nonlinear programming technique.

In Section 2, the notations and formulation of the problem are given. The formulated problems' solutions are suggested using two deterministic approaches: modified E-model and chance constraints in Section 3. In the modified E-model formulation, the solution strategy is recommended by utilizing a fuzzy goal programming technique. For the chance constraints model, the solutions are obtained using Lagrange multiplier and integer nonlinear programming techniques. In Section 4, the numerical illustration is discussed by considering the Iris data set, and the data set is obtained by simulation carried out by the software R. The formulated problem is solved through the modified E-model approach and chance constraints technique. The solution procedures are suggested by utilizing fuzzy goal programming problem, Lagrange multiplier method and integer nonlinear programming problem. MATLAB software is used to solve the formulated NLPPs. A comparative study is included by considering some other allocations, as discussed in Section 4.1.1, with the proposed allocation. Finally, the conclusion has been made for using the proposed technique in Section 5.

2. Framework of the problem

Assuming a population of N units that is partitioned into strata of sizes $N_1, N_2, \ldots, N_L$ units such that$\sum_{h=1}^{L} N_h = N$.

For h^{th} stratum, the following notations are introduced as follows:

N_h : Stratum size

$W_h = \frac{N_h}{N}$: Stratum weight

n_h : Sample size

y_{hi}: Observational value of i^{th} stratum unit/stratum sample.

$\bar{Y}_h = \frac{1}{N_h} \sum_{i=1}^{N_h} y_{hi}$: Stratum mean

$\bar{y}_h = \frac{1}{n_h} \sum_{i=1}^{n_h} y_{hi}$: Sample mean

$S_h^2 = \frac{1}{N_h - 1} \sum_{i=1}^{N_h} \left(y_{hi} - \bar{Y}_h \right)^2$: Stratum mean square

$s_h^2 = \frac{1}{n_h - 1} \sum_{i=1}^{n_h} \left(y_{hi} - \bar{y}_h \right)^2$: Sample mean square.

Furthermore,

$\bar{Y} = \frac{1}{N} \sum_{h=1}^{L} \sum_{i=1}^{N_h} y_{hi} = \frac{1}{N} \sum_{h=1}^{L} N_h \bar{Y}_h = \sum_{h=1}^{L} W_h \bar{Y}_h$: describes the overall population mean.

If the estimated value of $\bar{Y}$ is needed, then the stratified sample mean

$$\bar{y}_{st} = \sum_{h=1}^{L} W_h \bar{y}_h,$$

gives an unbiased estimator for $\bar{Y}$with the sampling variance

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

In a multivariate stratified population where p characteristics are given on each population element, then the$p$population means $\bar{Y}_j; j = 1, 2, \ldots, p$ are to be estimated. Since the individual optimum allocation may not be optimum for other characteristics. Let y_{jhi} denotes the value obtained from i^{th}element in h^{th} stratum having j^{th} characteristic and $\bar{Y}_{jh} = \frac{1}{N_h} \sum_{i=1}^{N_h} y_{jhi}$ be the stratum mean of y_{jhi}. Then the sample means for all characteristics in h^{th} stratum are calculated by

$$\bar{y}_{jh} = \frac{1}{n_h} \sum_{i=1}^{n_h} y_{jhi}; j = 1, 2, \ldots, p.$$

For j^{th} characteristic, an unbiased estimate of the overall population mean $\bar{Y}_j$ is given by $\bar{y}_{jst}$ and is expressed by

$$\bar{y}_{jst} = \sum_{h=1}^{L} W_h \bar{y}_{jh}$$

with its sampling variance

$$V(\bar{y}_{j\,st}) = \sum_{h=1}^{L} \frac{W_h^2 S_{jh}^2}{n_h} - \sum_{h=1}^{L} \frac{W_h S_{jh}^2}{N}$$

where S_{jh}^2 is the stratum variance of j^{th} characteristic in h^{th} stratum for$j = 1, 2, \ldots, p$, $h = 1, 2, \ldots, L$ and can be calculated by

$$S_{jh}^2 = \frac{1}{N_h - 1} \sum_{i=1}^{N_h} \left(y_{jhi} - \bar{Y}_{jh} \right)^2.$$

For a multivariate stratified sample survey, the linear cost function may be considered for the overall budget of the survey (Cochran, 1977) and may be expressed as

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

or

$$C - c_0 = C_0 = \sum_{h=1}^{L} c_h n_h,$$

where C_0 denotes the cost to measure all sampling units in all strata, c_h is the per-unit measurement cost of measuring p characteristics on the selected unit in h^{th} stratum, $n_h$is the h^{th} stratum sample size and c_0 is the overhead cost to conduct the survey. If the travel cost within the stratum come into consideration, then the cost function may not have remained linear. Beardwood et al. (1959) suggested the nonlinear cost function for this case. They showed that the distance between n arbitrarily dispersed points is proportional to $\sqrt{n}$. The nonlinear cost function, which includes travel costs, may be expressed as

$$C = c_0 + \sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} t_h \sqrt{n_h},$$

or

$$C_0 = C - c_0 = \sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} t_h \sqrt{n_h} \quad (1)$$

where $t_h\sqrt{n_h}$ is the travel cost incurred for h^{th} stratum. The cost function defined by (1) is quadratic in $\sqrt{n_h}$.

Practically, some other cost factors may be considered, like costs on the reward to respondents, labour costs, etc. Whenever the interviewers want to collect detailed information from the selected respondents, there will be a requirement for more human resources available for a specified time. For these prerequisites, the labour costs may be used for conducting the survey, and therefore the cost function may be expressed as

$$C_0 = C - c_0 = \sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} t_h \sqrt{n_h} + \omega \sum_{h=1}^{L} E(T_h) \quad (2)$$

where ω is unit time labour cost and $\sum_{h=1}^{L} E(T_h)$ is the accumulated labour time to obtain information from all strata. The labour time is available concerning time for sampling units within a stratum and follows an exponential distribution with the rate λ and value $\lambda = 1/(\text{average time})$. To approach n_h units in h^{th} stratum, the labour time has Gamma distribution with parameters (n_h, λ). Subsequently, the distribution of labour time to measure all sampling units within the stratum follows Gamma distribution with $(\sum_{h=1}^{L} n_h, \lambda)$ (Ross, 2009). Hence the expected labour time for all strata may be computed as

$$\begin{aligned}\sum_{h=1}^{L} E(T_h) &= \sum_{h=1}^{L} \left(\int_0^\infty t \lambda e^{-\lambda t} \frac{(\lambda t)^{n_h - 1}}{(n_h - 1)!} dt \right) \\ &= \sum_{h=1}^{L} \left(\frac{1}{(n_h - 1)!} \int_0^\infty t \lambda e^{-\lambda t} (\lambda t)^{n_h - 1} dt \right) \\ &= \sum_{h=1}^{L} \frac{n_h}{\lambda},\end{aligned}$$

for various values of λ (Muhammad and Husain, 2017).

The problem may be formulated in two ways by using a deterministic approach, and the solution is obtained either by minimizing the variance $V(\bar{y}_{j\,st})$ for a fixed cost or by minimizing the cost of the survey with variances of specified limits. Therefore two optimization problems may be described as

$$\begin{aligned}&\text{Minimize} && V(\bar{y}_{j\,st})\,; j = 1, 2, \ldots, p \text{ simultaneously} \\ &\text{subject to} && \sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} \tau_h \sqrt{n_h} + \omega \sum_{h=1}^{L} \tfrac{n_h}{\lambda} \leqslant C_0, \\ &&& 2 \leqslant n_h \leqslant N_h, \\ &\text{and} && n_h \text{ integers};\ h = 1, 2, \ldots, L.\end{aligned} \quad (3)$$

and

$$\begin{aligned}&\text{Minimize} && \sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} \tau_h \sqrt{n_h} + \omega \sum_{h=1}^{L} \tfrac{n_h}{\lambda} \\ &\text{subject to} && V(\bar{y}_{j\,st}) \leqslant V_0^j, \quad j = 1, 2, \ldots, p, \\ &&& 2 \leqslant n_h \leqslant N_h, \\ &\text{and} && n_h \text{ integers};\ h = 1, 2, \ldots, L,\end{aligned} \quad (4)$$

respectively.

If the true values of S_{jh}^2 are unknown, then they may be computed through a starter test or the values of past events (Kozak, 2006).

3. Determination of identical probabilistic sampling variances and cost

If S_{jh}^2 are considered as random variables, then the problems defined in (3) and (4) become SNLPPs. These problems may be converted into their equivalent deterministic problems. Several techniques, like modified E-model, E-model, V-model, chance constraints, etc., are available to solve deterministic problems (Charnes and Cooper, 1963). In this manuscript, modified E-model and chance constraints methods are used to convert the problems into the deterministic form.

3.1. Determination of probabilistic sampling variance through modified E-model

Consider the following formulated SNLPP for j^{th} characteristic which is given as

$$\begin{aligned}&\text{Minimize} && V(\bar{y}_{j\,st}) \\ &\text{subject to} && \sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} \tau_h \sqrt{n_h} + \omega \sum_{h=1}^{L} \tfrac{n_h}{\lambda} \leqslant C_0, \\ &&& 2 \leqslant n_h \leqslant N_h, \\ &\text{and} && n_h \text{ integers};\ h = 1, 2, \ldots, L,\ j = 1, 2, \ldots, p,\end{aligned} \quad (5)$$

where $V(\bar{y}_{j\,st}) = \sum_{h=1}^{L} \frac{W_h^2 S_{jh}^2}{n_h} - \sum_{h=1}^{L} \frac{W_h S_{jh}^2}{N}$ and S_{jh}^2 are random variables.

By considering the limiting distribution of S_{jh}^2 (Melaku, 1968; Díaz-García and Garay Tapia, 2007), define a random variable ζ_h that has an asymptotic $N(E(\zeta_h), V(\zeta_h))$.

For j^{th} characteristic, ζ_{jh} is defined for a multivariate case and is given as

$$\zeta_{jh} = \frac{1}{n_h - 1} \sum_{i=1}^{n_h} \left(y_{jhi} - \bar{Y}_{jh} \right)^2$$

where y_{jhi} denotes the value of the i^{th} unit in h^{th} stratum for j^{th} characteristic and $\bar{Y}_{jh} = N_h^{-1} \sum_{i=1}^{N_h} y_{jhi}$ is stratum mean for h^{th} stratum. The random variable ζ_{jh} has an asymptotic $N(E(\zeta_h), V(\zeta_h))$. These are given as

$$E(\zeta_{jh}) = \frac{n_h}{n_h - 1} S_{jh}^2$$

and

$$V(\zeta_{jh}) = \frac{n_h}{(n_h - 1)^2} [C_{yjh}^4 - (S_{jh}^2)^2]$$

respectively, where C_{yjh}^4 is the fourth mean moment and can be calculated by the following expression

$$C_{yjh}^4 = \frac{1}{N_h} \sum_{i=1}^{N_h} \left(y_{jhi} - \bar{Y}_{jh} \right)^4 ;\ h = 1, 2, \ldots, L,\ j = 1, 2, \ldots, p.$$

Let us define s_{jh}^2 which may be given as

$$s_{jh}^2 = \zeta_h - \frac{n_h}{n_h - 1} \left(\bar{y}_{jhi} - \bar{Y}_{jh} \right)^2,$$

where

$$\frac{n_h}{n_h - 1} \rightarrow 1$$

and

$$\left(\bar{y}_{jhi} - \bar{Y}_{jh} \right)^2 \rightarrow 0 \text{ in probability form.}$$

Then, the sample variance s_{jh}^2 has Normal asymptotical distribution

$$s_{jh}^2 \xrightarrow{a} N(E(\zeta_{jh}),\ V(\zeta_{jh}))$$

It has also seen that the objective function in (5) is a linear function of s_{jh}^2. Therefore, the objective function also follows Normal distribution with mean and variance, which are given as

$$\begin{aligned} E(\hat{V}(\bar{y}_{j\,st})) &= E\left(\sum_{h=1}^{L} \frac{W_h^2 s_{jh}^2}{n_h} - \sum_{h=1}^{L} \frac{W_h s_{jh}^2}{N}\right) \\ &= \sum_{h=1}^{L} \frac{W_h^2 S_{jh}^2}{n_h - 1} - \sum_{h=1}^{L} \frac{W_h}{N}\left(\frac{n_h}{n_h - 1}\right) S_{jh}^2 \end{aligned} \quad (6)$$

and

$$\begin{aligned} V(\hat{V}(\bar{y}_{j\,st})) &= V\left(\sum_{h=1}^{L} \frac{W_h^2 s_{jh}^2}{n_h} - \sum_{h=1}^{L} \frac{W_h s_{jh}^2}{N}\right) \\ &= \sum_{h=1}^{L} \frac{W_h^4}{n_h(n_h-1)^2}\left(C_{yjh}^4 - (S_{jh}^2)^2\right) - \sum_{h=1}^{L} \frac{W_h^2}{N^2}\left[\frac{n_h}{(n_h-1)^2}\left(C_{yjh}^4 - (S_{jh}^2)^2\right)\right] \end{aligned}$$

Therefore by using the modified E-model technique, the objective function may be redefined as

$$f_j(n_h) = k_1 E\left(\hat{V}(\bar{y}_{j\,st})\right) + k_2\sqrt{V(\hat{V}(\bar{y}_{j\,st}))}$$

where k_1 and k_2 are non-negative constants such that $k_1 + k_2 = 1$. The values of k_1 and k_2 will show the existence of the expectation and variance of $\hat{V}(\bar{y}_{j\,st})$. Hence, the equivalent deterministic NLPP to the SNLPP for j^{th} characteristic, given in (5), maybe formulated as

$$\begin{array}{ll} \text{Minimize} & f_j(n_h) \\ \text{subject to} & \sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} \tau_h \sqrt{n_h} + \omega \sum_{h=1}^{L} \frac{n_h}{\lambda} \leqslant C_0, \\ & 2 \leqslant n_h \leqslant N_h, \\ \text{and} & n_h \text{ integers};\ h = 1, 2, \ldots, L,\ j = 1, 2, \ldots, p, \end{array} \quad (8)$$

where

$$\begin{aligned} f_j(n_h) = &\; k_1\left[\sum_{h=1}^{L} \frac{W_h^2 S_{jh}^2}{n_h - 1} - \sum_{h=1}^{L} \frac{W_h S_{jh}^2}{N}\left(\frac{n_h}{n_h - 1}\right)\right] \\ &+ k_2\left[\sum_{h=1}^{L} \frac{W_h^4}{n_h(n_h-1)^2}\left(C_{yjh}^4 - (S_{jh}^2)^2\right) - \sum_{h=1}^{L} \frac{W_h^2}{N^2}\left(\frac{n_h}{(n_h-1)^2}\left(C_{yjh}^4 - (S_{jh}^2)^2\right)\right)\right]^{1/2}. \end{aligned}$$

Since the objective function includes the values of the population variance S_{jh}^2, but these values are unknown in general, in that case, the sample variances s_{jh}^2 may be used. Therefore, the equivalent deterministic NLPP defined in (8) may be given as

$$\begin{array}{ll} \text{Minimize} & f_j(n_h) = k_1\left[\sum_{h=1}^{L} \frac{W_h^2 s_{jh}^2}{n_h - 1} - \sum_{h=1}^{L} \frac{W_h s_{jh}^2}{N}\left(\frac{n_h}{n_h - 1}\right)\right] \\ & + k_2\left[\sum_{h=1}^{L} \frac{W_h^4}{n_h(n_h-1)^2}\left(C_{yjh}^4 - (s_{jh}^2)^2\right) - \sum_{h=1}^{L} \frac{W_h^2}{N^2}\left(\frac{n_h}{(n_h-1)^2}\left(C_{yjh}^4 - (s_{jh}^2)^2\right)\right)\right]^{1/2} \\ \text{subject to} & \sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} \tau_h \sqrt{n_h} + \omega \sum_{h=1}^{L} \frac{n_h}{\lambda} \leqslant C_0, \\ & 2 \leqslant n_h \leqslant N_h, \\ \text{and} & n_h \text{ integers};\ h = 1, 2, \ldots, L,\ j = 1, 2, \ldots, p. \end{array} \quad (9)$$

The NLPP given in (9) may be extended as multiobjective-INLPP (MINLPP) for multivariate stratified sampling designs as given as

$$\begin{array}{ll} \text{Minimize} & [f_1(n_h),\ f_2(n_h),\ \ldots, f_p(n_h)] \\ \text{subject to} & \sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} \tau_h \sqrt{n_h} + \omega \sum_{h=1}^{L} \frac{n_h}{\lambda} \leqslant C_0, \\ & 2 \leqslant n_h \leqslant N_h, \\ \text{and} & n_h \text{ integers};\ h = 1, 2, \ldots, L. \end{array} \quad (10)$$

3.1.1. Solution procedure by using the fuzzy goal programming technique

To solve the MINLPP given in (10), the fuzzy goal programming technique may be applied for multivariate sampling design. Since no technique is developed to solve the multiobjective formulation of INLPP, in that case, the problem may be converted into a single objective problem by using a suitable criterion. For such a case, the fuzzy goal programming technique may be used and applied using the following steps.

Stage 1: To get the solution of the MINLPP, a problem of a single objective function is to be required by ignoring the remaining objective functions of other characteristics to work out the optimum solution for each characteristic as an ideal solution.

Stage 2: Step −1 is repeated for all characteristics, and p-optimum solutions are obtained to give the optimum values of objective functions $(f_1, f_2, \ldots, f_p)$.

Stage 3: To compute the payoff matrix, the ideal solutions will give the upper and lower values for each objective function by defining U_j and L_j for j^{th} objective function; $j = 1, 2, \ldots, p$.

These values are computed as

$$U_j = Max\left\{f_1(n_{1h}^*), f_2(n_{2h}^*), \ldots, f_p(n_{ph}^*)\right\}$$

and

$$L_j = Min\left\{f_1(n_{1h}^*), f_2(n_{2h}^*), \ldots, f_p(n_{ph}^*)\right\}$$

where $f_j(n_{jh}^*)$is the optimum value of the objective function for j^{th} characteristic with optimum allocation n_{jh}^*.

Stage 4: The membership function may be defined as

$$\mu_j(n_{jh}) = \begin{cases} 0 & \text{if } f_j(n_{jh}) \geqslant U_j \\ \frac{U_j(n_{jh}) - f_j(n_{jh})}{U_j(n_{jh}) - L_j(n_{jh})} & \text{if } L_j \leqslant f_j(n_{jh}) \leqslant U_j \\ 1 & \text{if } f_j(n_{jh}) \leqslant L_j \end{cases}$$

where $\mu_j(n_{jh})$ is a strictly monotonic decreasing function to the solution n_{jh}, $h = 1, 2, \ldots, L$.

Consider the variable ξ_j which is defined as

$$\xi_j = \frac{U_j(n_{jh}) - f_j(n_{jh})}{U_j(n_{jh}) - L_j(n_{jh})}$$

Stage 5: By the max–min method, we have $Max\,[Min\,(\xi_1,\ \xi_2, \ldots, \xi_p)]$, then

$$\begin{array}{ll} \text{Maximize} & \xi \\ \text{subject to} & \xi_1 \geqslant \xi, \\ & \xi_2 \geqslant \xi, \\ & \vdots \\ & \xi_p \geqslant \xi, \end{array}$$

where $\xi = \underset{j}{Min}\,\{\mu_j(n_{jh});\ j = 1, 2, \ldots, p\}$.

Finally, the mathematical programming formulation for the problem (10) is to be solved by using fuzzy goal programming as follows:

$$\begin{array}{ll} \text{Maximize} & \xi \\ \text{subject to} & f_1 - \xi(U_1 - L_1) \geqslant L_1 \\ & f_2 - \xi(U_2 - L_2) \geqslant L_2 \\ & \vdots \\ & f_p - \xi(U_p - L_p) \geqslant L_p, \\ & \sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} \tau_h \sqrt{n_h} + \omega \sum_{h=1}^{L} \frac{n_h}{\lambda} \leqslant C_0, \\ & \xi \geqslant 0,\ n_h \text{ are integers};\ h = 1, 2, \ldots, L. \end{array} \quad (11)$$

3.2. Determination of probabilistic sampling cost through chance constraints

In this section, the SNLPP is considered for minimizing the total survey cost for a given bound to the estimated variance of the mean. This bound may be specified with tolerance limits for estimated variances of the estimates. This SNLPP may be formulated as:

$$\begin{aligned}
&\text{Minimize} && \textstyle\sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} \tau_h \sqrt{n_h} + \omega \sum_{h=1}^{L} \frac{n_h}{\lambda} \\
&\text{subject to} && P\left[\hat{V}(\bar{y}_{j\,st}) \leqslant V_0^j\right] \geqslant p_0 \,;\ j = 1,\ 2, \ldots, p, \\
& && 2 \leqslant n_h \leqslant N_h, \\
&\text{and} && n_h \text{ integers};\ h = 1,\ 2, \ldots, L,
\end{aligned} \tag{12}$$

where

$$\hat{V}(\bar{y}_{j\,st}) = \sum_{h=1}^{L} \frac{W_h^2 s_{jh}^2}{n_h} - \sum_{h=1}^{L} \frac{W_h s_{jh}^2}{N}.$$

Also, $V_0^j \geqslant 0$ and p_0 is a predetermined probability such that $0 \leqslant p_0 \leqslant 1$.

Since s_{jh}^2 follows an asymptotic $N(E(\zeta_h), V(\zeta_h))$, then the estimated $\hat{V}(\bar{y}_{j\,st})$in (12) also follows asymptotic Normal distribution with mean and variance defined in (6) and (7), respectively. After standardizing the function of the $\hat{V}(\bar{y}_{j\,st})$ in (12), it may be re-expressed as

$$P\left[\frac{\hat{V}(\bar{y}_{j\,st}) - E\{\hat{V}(\bar{y}_{j\,st})\}}{\sqrt{V\{\hat{V}(\bar{y}_{j\,st})\}}} \leqslant \frac{V_0^j - E\{\hat{V}(\bar{y}_{j\,st})\}}{\sqrt{V\{\hat{V}(\bar{y}_{j\,st})\}}}\right] \geqslant p_0$$

where $p_0 = \varphi\left[\frac{V_0^j - E\{\hat{V}(\bar{y}_{j\,st})\}}{\sqrt{V\{\hat{V}(\bar{y}_{j\,st})\}}}\right]$

and $\phi(\cdot)$ represents the function of standard Normal distribution. If e denotes the value of a random variable that follows standard normal distribution such that $\phi(\cdot) = p_0$, with these conditions, the inequality may be expressed as

$$\varphi\left[\frac{V_0^j - E\{\hat{V}(\bar{y}_{j\,st})\}}{\sqrt{V\{\hat{V}(\bar{y}_{j\,st})\}}}\right] \geqslant \varphi(e)$$

Therefore,

$$E\{\hat{V}(\bar{y}_{j\,st})\} + e\sqrt{V\{\hat{V}(\bar{y}_{j\,st})\}} - V_0^j \leqslant 0.$$

The equivalent deterministic NLPP for SNLPP in (12) may be given as

$$\begin{aligned}
&\textit{Minimize} && \textstyle\sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} \tau_h \sqrt{n_h} + \omega \sum_{h=1}^{L} \frac{n_h}{\lambda} \\
&\textit{subject to} && E\{\hat{V}(\bar{y}_{j\,st})\} + e\sqrt{V\{\hat{V}(\bar{y}_{j\,st})\}} - V_0^j \leqslant 0; j = 1,\ 2, \ldots, p,\ 2 \leqslant n_h \leqslant N_h, \\
&\textit{and} && n_h \text{ intgers};\ h = 1, 2, \ldots, L,
\end{aligned} \tag{13}$$

where

$$E\{\hat{V}(\bar{y}_{j\,st})\} + e\sqrt{V\{\hat{V}(\bar{y}_{j\,st})\}} = \left[\sum_{h=1}^{L} \frac{W_h^2 S_{jh}^2}{n_h - 1} - \sum_{h=1}^{L} \frac{W_h S_{jh}^2}{N}\left(\frac{n_h}{n_h - 1}\right)\right]$$
$$+ e\left[\sum_{h=1}^{L} \frac{W_h^4}{n_h(n_h-1)^2}\left(C_{yjh}^4 - \left(S_{jh}^2\right)^2\right) - \sum_{h=1}^{L} \frac{W_h^2}{N^2}\left(\frac{n_h}{(n_h-1)^2}\left(C_{yjh}^4 - \left(S_{jh}^2\right)^2\right)\right)\right]^{1/2} \tag{14}$$

The expression in (14) population variances S_{jh}^2, and these values remain not known in advance. Then s_{jh}^2 maybe substituted in place of S_{jh}^2. Hence the equivalent deterministic NLPP for (12) may be given as

$$\begin{aligned}
&\text{Minimize} && \textstyle\sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} \tau_h \sqrt{n_h} + \omega \sum_{h=1}^{L} \frac{n_h}{\lambda} \\
&\text{subject to} && E\{\hat{V}(\bar{y}_{j\,st})\} + e\sqrt{V\{\hat{V}(\bar{y}_{j\,st})\}} - V_0^j \leqslant 0\,; j = 1,\ 2, \ldots, p, \\
& && 2 \leqslant n_h \leqslant N_h, \\
&\text{and} && n_h \text{ integers};\ h = 1,\ 2, \ldots, L,
\end{aligned} \tag{15}$$

where

$$E\{\hat{V}(\bar{y}_{j\,st})\} + e\sqrt{V\{\hat{V}(\bar{y}_{j\,st})\}} = \left[\sum_{h=1}^{L} \frac{W_h^2 s_{jh}^2}{n_h - 1} - \sum_{h=1}^{L} \frac{W_h s_{jh}^2}{N}\left(\frac{n_h}{n_h - 1}\right)\right]$$
$$+ e\left[\sum_{h=1}^{L} \frac{W_h^4}{n_h(n_h-1)^2}\left(C_{yjh}^4 - \left(s_{jh}^2\right)^2\right) - \sum_{h=1}^{L} \frac{W_h^2}{N^2}\left(\frac{n_h}{(n_h-1)^2}\left(C_{yjh}^4 - \left(s_{jh}^2\right)^2\right)\right)\right]^{1/2} \tag{16}$$

4. Application

A population of size $N = 9500$, with three strata and two characteristics are taken and obtained by simulating 150 observations of Iris data. Iris data set is available in the Software R domain (R Development Core Team, 2018). These observations are divided into three strata where two characteristics (that is, length and width of a leaf of a particular species of flower) are measured on each population unit. The population units for three strata of sizes 3000, 3000 and 3500 are generated by the simulation of Iris data using the software R and the values of s_{jh}^2 and C_{yjh}^4 are computed and reported in Table 1. The values of c_h, τ_h, ω and λ are assumed for numerical illustration accordingly and given in Table 1. The total cost for conducting the survey is taken as 1000 units.

4.1. Solution for modified E-model by fuzzy goal programming technique

Without loss of generality, $k_1 = k_2 = 0.5$ is taken. For the given numeric values, given in Table 1, the formulations of the NLPP for both characteristics are given as

Table 1
Three strata information with two characteristics.

h	N_h	s_{1h}^2	s_{2h}^2	C_{y1h}^4	C_{y2h}^4	c_h	τ_h	ω	λ
1	3000	0.01523817	0.02037975	0.00068061	0.001217395	2	1	100	20
2	3000	0.06898021	0.00957083	0.01434982	0.000268493	3	2	100	20
3	3500	0.16608490	0.01080517	0.08426042	0.000349415	3	3	100	20

$$\begin{array}{lc} \text{Minimize } f_1 = & 0.5\left[\frac{0.00151959}{(n_1-1)}+\frac{0.00687891}{(n_2-1)}+\frac{0.02254338}{(n_3-1)}-\frac{0.00481210n_1}{9500(n_1-1)}-\frac{0.02178320n_2}{9500(n_2-1)}-\frac{0.06118920n_3}{9500(n_3-1)}\right] \\ & +0.5\left[\frac{0.000004459}{(n_1-1)^2n_1}+\frac{0.00009538}{(n_2-1)^2n_2}+\frac{0.00104419}{(n_3-1)^2n_3}-\frac{0.00004472n_1}{9500^2(n_1-1)^2}-\frac{0.00095649n_2}{9500^2(n_2-1)^2}-\frac{0.00769289n_3}{9500^2(n_3-1)^2}\right]^{\frac{1}{2}}, \\ \text{subject to} & 2n_1+3n_2+3n_3+\sqrt{n_1}+2\sqrt{n_2}+3\sqrt{n_3}+5(n_1+n_2+n_3)\leqslant 1000, \\ & 2\leqslant n_h\leqslant N_h, \\ \text{and} & n_h \text{ integers};\ h=1,2,3. \end{array}$$

$$\begin{array}{lc} \text{Minimize } f_2 = & 0.5\left[\frac{0.00203233}{(n_1-1)}+\frac{0.00095443}{(n_2-1)}+\frac{0.00146663}{(n_3-1)}-\frac{0.00643571n_1}{9500(n_1-1)}-\frac{0.00302237n_2}{9500(n_2-1)}-\frac{0.00398085n_3}{9500(n_3-1)}\right] \\ & +0.5\left[\frac{(0.00000797)}{(n_1-1)^2n_1}+\frac{0.00000176}{(n_2-1)^2n_2}+\frac{0.0000042865}{(n_3-1)^2n_3}-\frac{0.00007998n_1}{9500^2(n_1-1)^2}-\frac{0.00001764n_2}{9500^2(n_2-1)^2}-\frac{0.00003158n_3}{9500^2(n_3-1)^2}\right]^{\frac{1}{2}} \\ \text{subject to} & 2n_1+3n_2+3n_3+\sqrt{n_1}+2\sqrt{n_2}+3\sqrt{n_3}+5(n_1+n_2+n_3)\leqslant 1000, \\ & 2\leqslant n_h\leqslant N_h, \\ \text{and} & n_h \text{ integers};\ h=1,2,3. \end{array}$$

characteristics are worked out as given below:

$$n_{11}=18,\ n_{12}=37,\ n_{13}=67,\ f_1=0.00034712,$$

$$n_{21}=54,\ n_{22}=33,\ n_{23}=40,\ f_2=0.00005878.$$

After getting ideal solutions, the payoff matrix may be computed and given in Table 2:

The upper and lower bounds of each objective function may be given as:

$$f_1^l=0.00034712,\ f_1^u=0.00047683,$$

$$f_2^l=0.00005878,\ f_2^u=0.00010331,$$

Therefore the values of L_j and U_j are obtained as

$$L_j=\underset{j}{\text{Min}}\ f_j(n_{jh}^*)=0.000058780 \text{ and } U_j=\underset{j}{\text{Max}}\ f_j(n_{jh}^*) = 0.00047683.$$

Let $\mu_1(n_{1h})$ and $\mu_2(n_{2h})$ are the fuzzy membership function for the functions $f_j(n_{jh})$, $j=1,2$ and they are used for developing a membership function for both characteristics as

$$\mu_1(n_{1h})=\begin{cases}0 & \text{if } f_1(n_{1h})\geqslant 0.00047683\\ \frac{0.00047683-f_1(n_{1h})}{0.00013} & \text{if } 0.00034712\leqslant f_1(n_{1h})\leqslant 0.00047683\\ 1 & \text{if } f_1(n_{1h})\leqslant 0.00034712\end{cases}$$

$$\mu_2(n_{2h})=\begin{cases}0 & \text{if } f_2(n_{2h})\geqslant 0.00010331\\ \frac{0.00010331-f_2(n_{2h})}{0.00004453} & \text{if } 0.000058780\leqslant f_2(n_{2h})\leqslant 0.00010331\\ 1 & \text{if } f_2(n_{2h})\leqslant 0.000058780\end{cases}$$

By using the max–min addition operator, the objective function is revised as

$$\text{Maximize}\left\{7.274210973-\left(\frac{f_1(n_{1h})}{.00013}+\frac{f_2(n_{2h})}{.00004453}\right)\right\}$$

To maximize the above problem with subject to constraints as formulated as

$$\begin{array}{ll} \text{Maximize} & \xi \\ \text{subject to} & f_1-.00012971\xi\geqslant 0.00034712 \\ & f_2-.00004453\xi\geqslant 0.00005878 \\ & \sum_{h=1}^{L}c_hn_h+\sum_{h=1}^{L}\tau_h\sqrt{n_h}+\omega\sum_{h=1}^{L}\frac{n_h}{\lambda}\leqslant C_0, \\ & 2\leqslant n_h\leqslant N_h, \\ \text{and} & \xi\in[0,1],\ n_h \text{ areintegers};\ h=1,2,\dots,L. \end{array}$$

Table 2
The payoff matrix with ideal solutions.

	f_1	f_2
n_{1h}^*	0.00034712	0.00010331
n_{2h}^*	0.00047683	0.00005878

Using MATLAB, the optimal solution to the above problem is obtained as

$$\xi=0.1890,\ n_1=33,\ n_2=35,\ n_3=56$$

with variances $V(\bar{y}_{jst});\ j=1,2$under the proposed allocation given as

$$V(\bar{y}_{1st})=0.000638635,\ V(\bar{y}_{2st})=0.000114102,$$

Therefore, the Trace will be

$$V(\bar{y}_{1st})+V(\bar{y}_{2st})=0.00075274.$$

where,

$$f_1 = 0.5\left[\frac{0.00151959}{(n_1-1)}+\frac{0.00687891}{(n_2-1)}+\frac{0.02254338}{(n_3-1)}-\frac{0.00481210n_1}{9500(n_1-1)}-\frac{0.02178320n_2}{9500(n_2-1)}-\frac{0.06118920n_3}{9500(n_3-1)}\right] +0.5\left[\frac{0.000004459}{(n_1-1)^2n_1}+\frac{0.00009538}{(n_2-1)^2n_2}+\frac{0.00104419}{(n_3-1)^2n_3}-\frac{0.00004472n_1}{9500^2(n_1-1)^2}-\frac{0.00095649n_2}{9500^2(n_2-1)^2}-\frac{0.00769289n_3}{9500^2(n_3-1)^2}\right]^{\frac{1}{2}}$$

$$f_2 = 0.5\left[\frac{0.00203233}{(n_1-1)}+\frac{0.00095443}{(n_2-1)}+\frac{0.00146663}{(n_3-1)}-\frac{0.00643571n_1}{9500(n_1-1)}-\frac{0.00302237n_2}{9500(n_2-1)}-\frac{0.00398085n_3}{9500(n_3-1)}\right] +0.5\left[\frac{(0.00000797)}{(n_1-1)^2n_1}+\frac{0.00000176}{(n_2-1)^2n_2}+\frac{0.0000042865}{(n_3-1)^2n_3}-\frac{0.00007998n_1}{9500^2(n_1-1)^2}-\frac{0.00001764n_2}{9500^2(n_2-1)^2}-\frac{0.00003158n_3}{9500^2(n_3-1)^2}\right]^{\frac{1}{2}}$$

4.1.1. Comparison with other allocations

In this section, a comparative study is carried out where the proposed method is compared with some other well-defined methods of allocation. Some of these methods are as follows:

4.1.1.1. Proportional allocation. For the fixed cost of the survey, the proportional allocation may be obtained by substituting $n_h=nW_h$ in the cost function, and subsequently, stratum-wise allocations, which are rounded off to nearest integers, may be obtained as

$$n_1=40,\ n_2=40,\ n_3=46,$$

also the trace value, under proportional allocation, is computed as 0.0008066.

4.1.1.2. Cochran's average allocation. Cochran (1977) suggested the compromise criterion by taking the average of the individual optimum allocations $n_{jh}^*;h=1,\ 2,\dots,L,\ j=1,\ 2,\dots,p$. These allocations are obtained by solving the individual NLPP for each characteristic; that is, for j^{th} characteristic, the required NLPP may be formulated as

$$\begin{array}{ll} \text{Minimize} & V(\bar{y}_{jst})=\sum_{h=1}^{L}\frac{W_h^2S_{jh}^2}{n_{jh}} \\ \text{subject to} & \sum_{h=1}^{L}c_hn_{jh}+\sum_{h=1}^{L}\tau_h\sqrt{n_{jh}}+\omega\sum_{h=1}^{L}\frac{n_{jh}}{\lambda}\leqslant C_0 \\ \text{and} & n_{jh}\geqslant 0,\ h=1,2,\dots,L,\ j=1,2,\dots,p, \end{array} \quad (19)$$

where $n_{jh}^*=(n_{j1}^*,n_{j2}^*,\dots,n_{jL}^*)$ denotes the individual optimum allocation to the j^{th} characteristic. Therefore Cochran's compromise allocation is computed as

$$n_h=\frac{1}{p}\sum_{j=1}^{p}n_{jh}^*;\ h=1,2,\dots,L.$$

Table 3
Allocations with trace value and relative efficiency.

S.N	Allocations	Allocations $n_1 n_2 n_3 n$	cost	Trace	R.E. w.r.t Proportional Allocation
1	Proportional	40 40 46 126	1007.3	0.0008066	1.00000
2	Cochran	36 35 53 124	995.67	0.0007755	1.04010
3	Sukhatme	25 36 61 122	991.43	0.0007532	1.07089
4	Proposed	33 35 56 124	999.03	0.0007527	1.07155

Table 4
Percentage increase for all characteristics within the variances when the individual optimum for one characteristic is used.

Percentage increment within the variances			
Characteristics	1	2	
1	0	0.31866816	0.014920487
2	0.55578140	0	0.229092349

For the given numerical values, as given in Table 1, the compromise allocation suggested by Cochran (1977) is worked out as $n_1 = 36,\ n_2 = 35,\ n_3 = 53$. The variances of both characteristics are calculated as $V^*(\bar{y}_{1st}) = 0.000664098$ and $V^*(\bar{y}_{2st}) = 0.000111389$ respectively. Therefore the trace value is 0.00077549.

4.1.1.3. Sukhatme's compromise allocation. This compromise allocation is obtained by optimizing the Trace of the variance–covariance matrix of the estimator. The solution to the following NLPP will give the desired compromise allocation Sukhatme et al. (1984), and the formulation of the NLPP is given as

$$\begin{array}{ll} \text{Minimize} & \sum_{j=1}^{p} V(\bar{y}_{jst}) = \sum_{j=1}^{p}\sum_{h=1}^{L}\frac{W_h^2 S_{jh}^2}{n_h} \\ \text{subject to} & \sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L}\tau_h\sqrt{n_h} + \omega\sum_{h=1}^{L}\frac{n_h}{\lambda} \leqslant C_0 \\ \text{and} & n_h \geqslant 0,\ h = 1,2,\ldots,L,\ j = 1,2,\ldots,p. \end{array}$$

On the substitution of numerical values from Table 1, the solution is obtained as $n_1 = 25,\ n_2 = 36,\ n_3 = 61$ and therefore, the trace value is calculated as 0.000753271.

In Table 3, it appears that the proposed allocation gives the least trace value compared to the values obtained by other allocations. Furthermore, the relative efficiency of the proposed allocation to proportional allocation is maximum among the others. Table 4 shows the percentage increase in both characteristics' variances when the individual optimum allocation of one characteristic is used for both characteristics, and the proposed allocation is utilized. The proposed allocation provides lesser values of percentage increase in the variances with respect to individual allocations. Table 5 shows the percentage increase in both characteristics' variances when other allocations are used instead of individual allocations. For the proposed allocation, these values are minimum in comparison to others. Therefore it may be claimed that the suggested allocation may be regarded as the best allocation.

Based on the above discussion, the proposed allocation works well in comparison to other allocations.

Table 6
Allocations with the incurred cost.

S. No	Allocations	$n_1 n_2\ n_3 n$	Cost
1	Lagrange multiplier (non integer)	12.97 52.74 61.33 127	1045.10
2	Lagrange multiplier (rounded)	13 53 61 127	1044.59
3	Lagrange multiplier (integer)	13 53 62 128	1052.78
4	Stochastic (non integer)	18.66 36.95 66.30 122	997.630
5	Stochastic (rounded)	19 35 56 124	997.890
6	Stochastic (integer)	19 38 65 122	997.870

4.2. Solution by chance constraints

When the cost of carrying out a sample survey is high and a specified limit on the variances are given, then this method may be used. With the specified values of $V_0^1 = 0.00080$ and $V_0^2 = 0.00040$, the value of e is 2.3263 such that $\phi(e) = p_0 = 0.99$. The equivalent deterministic problem of SNLPP may be given as

$$\begin{array}{ll} \text{Minimize} & \sum_{h=1}^{3} c_h n_h + \sum_{h=1}^{3}\tau_h\sqrt{n_h} + \omega\sum_{h=1}^{3}\frac{n_h}{\lambda} \\ \text{subject to} & \sum_{h=1}^{3}\frac{W_h^2 s_{1h}^2}{(n_h-1)} - \sum_{h=1}^{3}\frac{W_h s_{1h}^2}{9500}\left(\frac{n_h}{n_h-1}\right) \\ & +2.3263\left[\sum_{h=1}^{3}\frac{W_h^4}{n_h(n_h-1)^2}\left(C_{y1h}^4 - (s_{1h}^2)^2\right) - \sum_{h=1}^{3}\frac{W_h^2}{(9500)^2}\left(\frac{n_h}{(n_h-1)^2}\left(C_{y1h}^4 - (s_{1h}^2)^2\right)\right)\right]^{1/2} \leqslant 0.00080, \\ & \sum_{h=1}^{3}\frac{W_h^2 s_{2h}^2}{(n_h-1)} - \sum_{h=1}^{3}\frac{W_h s_{2h}^2}{9500}\left(\frac{n_h}{n_h-1}\right) \\ & +2.3263\left[\sum_{h=1}^{3}\frac{W_h^4}{n_h(n_h-1)^2}\left(C_{y2h}^4 - (s_{2h}^2)^2\right) - \sum_{h=1}^{3}\frac{W_h^2}{(9500)^2}\left(\frac{n_h}{(n_h-1)^2}\left(C_{y2h}^4 - (s_{2h}^2)^2\right)\right)\right]^{1/2} \leqslant 0.00040, \\ & 2 \leqslant n_h \leqslant N_h, \\ \text{and} & n_h \text{ integers};\ h = 1,2,3. \end{array} \tag{20}$$

The NLPP (20) solutions are obtained using MATLAB by the Lagrange multiplier technique and INLPP technique. The solutions are reported in Table 6.

Table 6 shows the Lagrange multiplier technique and INLPP technique to minimize the survey's total cost. If the continuous solution to the NLPP (20) is considered, then the nonlinear programming technique is preferable to that of the Lagrange multiplier technique. If the continuous solution is adjusted off to the closest whole number, then the nonlinear programming technique provides the survey's minimum cost. Furthermore, if integer restriction is a must, then the use of a nonlinear programming technique is advisable.

5. Conclusion

In general, stratum variances' true values may not be known in advance but may be estimated. In this way, the problem is defined as a multivariate-multiobjective SNLPP in this paper. The formulated SNLPPs may be converted into their deterministic form using a modified E- model and chance constraints techniques. The formulated problems' solutions may be computed either by minimizing the sampling variances for a fixed cost or minimizing the cost

Table 5
Percentage increment when optimization is done about the criteria.

Percentage increment within the variances of distinct characteristics beneath assorted criteria				
Characteristics	Proportional	Cochran	Sukhatme	Proposed
1	15.36391	9.441428	2.40957	0.014920487
2	3.225985	7.91908	27.73566	0.229092349

for the fixed precision value of the variances of estimates. For numerical illustration, the data are generated by conducting simulation using R, and the formulated NLPPs may be solved by using MATLAB. The proposed compromise allocation provides the best outcomes for the given numerical application than that obtained by other compromise criteria, as discussed in this paper. Furthermore, for large scale investigations, it is vital to select the appropriate method for attaining the study's objectives.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

All the authors are very thankful to the Editor in Chief and the anonymous reviewers who helped improve the paper's quality and presentation substantially.

References

Ahsan, M.J., 1975-1976. A procedure for the problem of optimum allocation in multivariate stratified random sampling stratified random sampling. Aligarh. Bull. Math. 5-6, 37–42. DOI:10.1080/01621459.1968.11009271.
Ali, I., Hasan, S.S., 2013. Integer fuzzy programming approach in bi-objective selective maintenance allocation problem. J. Math. Model. Algo. 13 (2), 113–124. https://doi.org/10.1007/s10852-013-9220-1.
Beardwood, J., Halton, J.H., Hammersley, J.M., 1959. The shortest path through many points. Math. Proc. Cambridge 55, 299–327. https://doi.org/10.1017/S0305004100034095.
Bellman, R.E., Zadeh, L.A., 1970. Decision-making in a fuzzy environment. Manage. Sci. 17 (4), 141–164. https://doi.org/10.1287/mnsc.17.4.B141.
Charnes, A., Cooper, W.W., 1963. Deterministic equivalents for optimizing and satisficing under chance constraints. Oper. Res. 11 (1), 18–39. https://doi.org/10.1287/opre.11.1.18.
Chatterjee, S., 1968. Multivariate stratified surveys. J. Am. Stat. Assoc. 63, 530–534.
Cochran, W.G., 1977. Sampling Techniques. John Wiley, New York.
Díaz-García, J.A., Garay Tapia, M.M., 2007. Optimum allocation in stratified surveys: stochastic programming. Comput. Stat. Data Anal. 51, 3016–3026.
Díaz-García, J.A., Ramos-Quiroga, R., 2014. Stochastic optimal design in multivariate stratifies sampling. Optimization 63 (11), 1665–1688. https://doi.org/10.1080/02331934.2012.728218.
Elsisi, M., 2019a. Future search algorithm for optimization. Evol. Intell. 12 (1), 21–31.
Elsisi, M., 2019b. Design of neural network predictive controller based on imperialist competitive algorithm for automatic voltage regulator. Neural. Comput. Appl. 31 (9), 5017–5027.
Elsisi, M., 2020. New variable structure control based on different meta-heuristics algorithms for frequency regulation considering nonlinearities effects. Int. Trans. Electr. Energy Syst. 30, (7) e12428.
Elsisi, M., Soliman, M., 2020. Optimal design of robust resilient automatic voltage regulators. ISA Trans. 108, 257–268.
Fakhrzad, M.B., Goodarzian, F., 2019. A fuzzy multiobjective programming approach to develop a green closed-loop supply chain network design problem under uncertainty: modifications of imperialist competitive algorithm. RAIRO Oper. Res. 53 (3), 963–990.
Fathollahi-Fard, A.M., Ahmadi, A., Goodarzian, F., Cheikhrouhou, N., 2020. A bi-objective home healthcare routing and scheduling problem considering patients' satisfaction in a fuzzy environment. Appl. Soft Comput. 93,. https://doi.org/10.1016/j.asoc.2020.106385 106385.
Fatima, U., Varshney, R., Najmussehar, S., Ahsan, M.J., 2014. On compromise mixed allocation in multivariate stratified sampling with random parameters. J. Math. Model. Algor. 13 (4), 523–536. https://doi.org/10.1007/s10852-013-9242-8.
Goodarzian, F., Hosseini-Nasab, H., 2021. Applying a fuzzy multiobjective model for a production–distribution network design problem by using a novel self-adoptive evolutionary algorithm. IJSS: O&L. 8 (1), 1–22.
Gupta, N., Ali, I., Bari, A., 2013. Fuzzy goal programming approach in selective maintenance reliability model. Pak. J. Stat. Oper. Res. 9(3), 321–331. DOI:10.18187/pjsor.v9i3.654.
Haq, A., Ali, I., Varshney, R., 2020. Compromise allocation problem in multivariate stratified sampling with flexible fuzzy goals. J Stat Comput. Simul. 90 (9), 1557–1569.
Haseen, S., Ali, I., Bari, A., 2016. Multiobjective stochastic multivariate stratified sampling in presence of nonresponse. Commun. Stat. - Simul. Comput. 45(8), 2810–2826. . 2014.926173.
Khan, M.G.M., Ahsan, M.J., Jahan, N., 1997. Compromise allocation in multivariate stratified sampling: an integer solution. Nav. Res. Logist. 44 (1), 69–79.
Kokan, A.R., Khan, S.U., 1967. Optimum allocation in multivariate surveys: an analytical solution. J. R. Stat. Soc. B. 29 (1), 115–125.
Kozak, M., 2006. On sample allocation in multivariate surveys. Commun. Stat. - Theory Methods 35 (4), 901–910. https://doi.org/10.1080/03610910600880286.
Kozak, M., Wang, H.Y., 2010. On stochastic optimization in sample allocation among strata. Metron LXVIII (1), 95–103.
Lu, Y., Lang, M., Sun, Y., Li, S., 2020. A fuzzy intercontinental road-rail multimodal routing model with time and train capacity uncertainty and fuzzy programming approaches. IEEE Access 8, 27532–27548.
Melaku, A., 1968. Asymptotic normality of the optimal allocation in multivariate stratified random sampling. Sankhya 48, 224–232.
Muhammad, Y.S., Husain, I., 2017. Trade off between cost and variance for a multiobjective compromise allocation in stratified random sampling. Commun. Stat.-Theory Methods. 46(6), 2655–2666. DOI:10.1080/03610926.2015.1040507.
Muhammad, Y.S., Shabir, J., Husain, I., Abd-el. Moemen, M., 2015. Multiobjective compromise allocation in multivariate stratified sampling using extended lexicographic goal programming with Gamma cost function. J. Math. Model. Algorithms 14 (3), 255–265. https://doi.org/10.1007/s10852-014-9270-z.
Neyman, J., 1934. On the two different aspects of the representative method: the method of stratified sampling and the method of purposive selection. J. Royal Stat. Soc. 97 (4), 558–625. https://doi.org/10.2307/2342192.
Prékopa, A., 1978. The use of stochastic programming for the solution of the some problems in statistics and probability. Technical Summary Report #1834, University of Wisconsin Madison, Mathematical Research Center, Madison.
R Development Core Team. R., 2018. A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org.
Ross, S.M., 2009. Introduction to Probability and Statistics for Engineers and Scientists. Associated Press. ISBN:978-0-12-370483-2.
Semiz, M., 2004. Determination of compromise integer strata sample sizes using goal programming. J. Math. Stat. 33, 91–96.
Sukhatme, P. V., Sukhatme, B.V., Sukhatme, S., Asok, C., 1984. Sampling theory of surveys with applications. Iowa State University Press, Iowa, USA and Indian Society of Agricultural Statistics, New Delhi, India.
Varshney, R., Najmussehar., Ahsan, M.J., 2012. Estimation of more than one parameters in stratified sampling with fixed budget. Math. Method. Oper. Res. 75(2), 185–197. DOI:10.1007/s00186 -012-0380-y.
Varshney, R., Khan M.G.M., Fatima, U., Ahsan, M.J., 2015. Integer compromise allocation in multivariate stratified surveys. Ann. Oper. Res. 226(1): 659–668. DOI:10.1007/s10479-014-1734-z.
Varshney, R., Gupta, S., Ali, I., 2017. An optimum multivariate-multiobjective stratified sampling design: Fuzzy programming approach. Pak. J. Stat. Oper. Res. 13 (4), 829–855. https://doi.org/10.18187/pjsor.v13i4.1834.
Varshney, R., Mradula, 2019. Optimum allocation in multivariate stratified sampling design in the presence of nonresponse with Gamma cost function. J. Stat. Comput. Simul. 89 (13), 2454–2467. https://doi.org/10.1080/00949655.2019.1620747.
Zadeh, L.A., 1965. Fuzzy Sets. Inf. Control. 8, 338–390.
Zimmermann, H.J., 1978. Fuzzy programming and linear programming with several objective functions. Fuzzy Sets Syst. 1, 45–55.