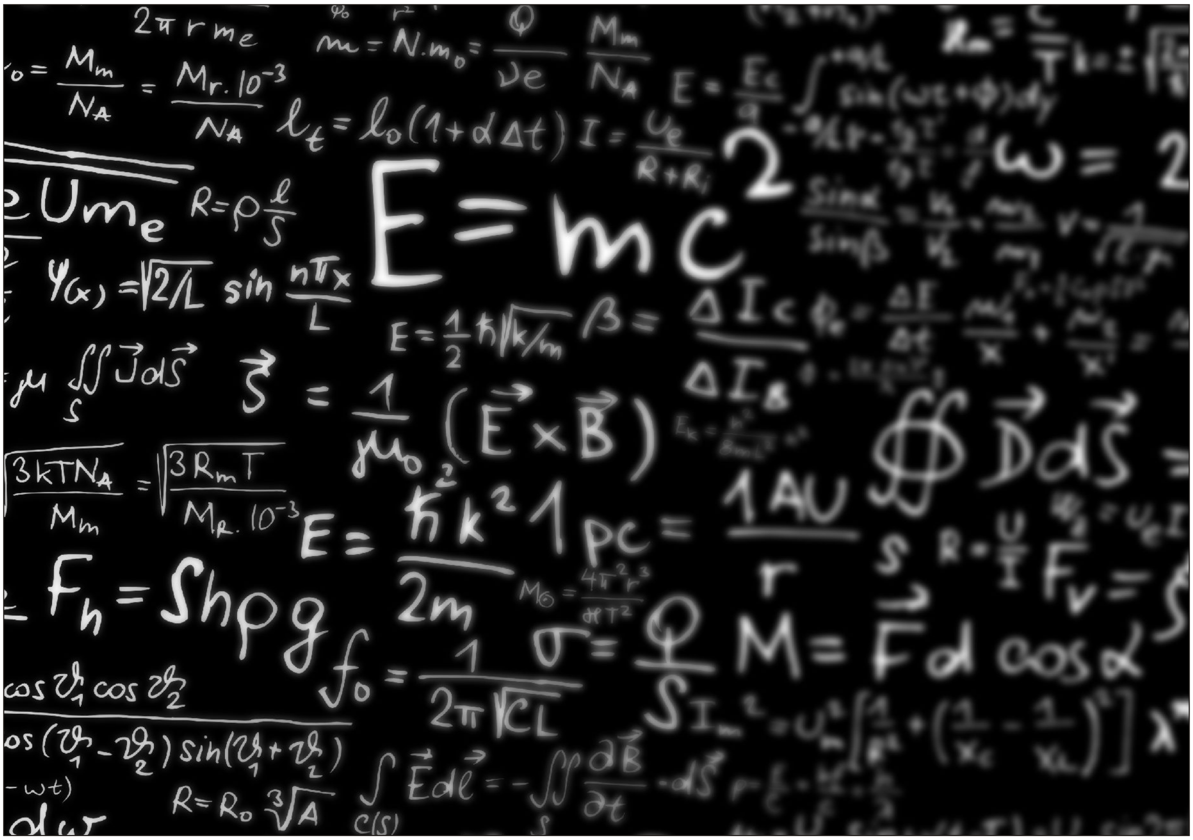

Asian Journal of Mathematics & Statistics

ISSN 1994-5418

science alert

ANSInet
an open access publisher
http://ansinet.com

OPEN ACCESS Asian Journal of Mathematics and Statistics

ISSN 1994-5418
DOI: 10.3923/ajms.2018.1.11

Research Article
A New Approach to Randomized Response Model Using Fuzzy Numbers

[1]Tanveer Ahmad Tarray, [1]Peer Bilal and [2]Muzafar Rasool

[1]Department of Mathematical Sciences, Islamic University of Science and Technology, Awantipora, Pulwama, 192122, Kashmir, India
[2]Department of Computer Science, Islamic University of Science and Technology, Awantipora, Pulwama, 192122, Kashmir, India

Abstract
The crux of this study is to consider a randomized response model using allocation problem in presence of non-response based on model and minimize the variance subject to cost constraint. The costs (measurement costs and total budget of the survey) in the cost constraint are assumed as fuzzy numbers, in particular triangular and trapezoidal fuzzy numbers due to the ease of use. The problem formulated is solved by using Lagrange multipliers technique and the optimum allocation obtained in the form of fuzzy numbers is converted into crisp form using α-cut method at a prescribed value of α. Numerical illustrations are also given in support of the present study and the results are formulated through LINGO.

Key words: Unrelated randomized response technique, optimum allocation, stratified random sampling, sensitive attribute, fuzzy logic

Citation: Tanveer Ahmad Tarray, Peer Bilal and Muzafar Rasool, 2018. A new approach to randomized response model using fuzzy numbers. Asian J. Math. Stat., 11: 1-11.

Corresponding Author: Tanveer Ahmad Tarray, Department of Mathematical Sciences, Islamic University of Science and Technology, Awantipora, Pulwama, 192122, Kashmir, India

Competing Interest: The authors have declared that no competing interest exists.

Data Availability: All relevant data are within the paper and its supporting information files.

INTRODUCTION

The most important things for obtaining data pertaining to human population is the social survey. To measure opinions, attitudes and behaviors that cover a wide band of interests, the social survey has been established as being tremendously practical. The surveys are conducted due to many reasons, non-availability of certain facts/information in the archives being the most understandable and apparent. For instance, if one is interested in knowing crime rate, information about unseen crimes or unreported victimization experience is not available in formal records on crimes. Sometimes the facts about the individuals (in a population) are inaccessible to the investigators for legal reasons. Questionnaires, in particular social surveys, generally consist of many items. Some of the items may be about sensitive/high risk behavior, due to the social stigma carried by them. One problem with research on high-risk behavior is that respondents may consciously or unconsciously provide incorrect information. In psychological surveys, a social desirability bias has been observed as a major cause of distortion in standardized personality measures. Survey researchers have similar concerns about the truth of survey results/findings about such topics as drunk driving, use of marijuana, tax evasion, illicit drug use, induced abortion, shop lifting, child abuse, family disturbances, cheating in exams, HIV/AIDS and sexual behavior. Thus to obtain trustworthy data on such confidential matters, especially the sensitive ones, instead of open surveys alternative procedures are required. Such an alternative procedure known as "randomized response technique (RRT)" was first introduced by Warner[1].

The Randomized response (RR) technique was first introduced by Warner[1] mainly to cut down the possibility of (1) Reduced response rate and (2) Inflated response bias experienced in direct or open survey relating to sensitive issues. Warner[1] himself pointed out how one may get a biased estimate in an open survey when a population consists of individuals bearing a stigmatizing character A or its complement A^C, which may or may not also be stigmatizing. It requires the interviewee to give a "Yes" or "No" answer either to the sensitive question or to its negative depending on the outcome of a randomized device not reported to the interviewer. Greenberg *et al.*[2] derived results for Warner's model in the case of less than completely truthful reporting. Later several modifications in RR technique have been developed by various authors Fox and Tracy[3], Chaudhuri and Mukerjee[4], Singh and Tarray[5-7], Tarray and Singh[8], Singh and Tarray[9], Tarray and Singh[10-13] and Tarray[14].

Stratified random sampling is generally obtained by dividing the population into non-overlapping groups called strata and selecting a simple random sample from each stratum. An RR technique using stratified random sampling provides the group characteristics related to each stratum estimator. Also, stratified sample, protect a researcher from the possibility of obtaining a poor sample[15].

A STUDY OF RANDOMIZED RESPONSE TECHNIQUES

The description of the models due to Singh[16] and Kim and Warde[15] are given below:

Singh model: Singh[16] developed randomized response techniques named RRT1 which is given below:

RRT1: In this procedure, each interviewee in a with replacement simple random sample of size n is provided with one randomized response device. It consists of the statement "I belong to the sensitive group" with known probability P, exactly the same probability as used by Warner[1] and the statement "Yes" with probability (1-P). The interviewee is instructed to use the device and report "Yes" or "No" for the random outcome of the sensitive statement according to his/her actual status. Otherwise, it is simply to report the "Yes" statement observed on the randomized response device. The whole procedure is completed by the respondent, unobserved by the interviewer. Then θ_1 the probability of a "Yes" answer in the population is:

$$\theta_1 = P\pi_S + (1-P)$$

An unbiased estimator of π_S due to Singh[16] is given by:

$$\hat{\pi}_1 = \frac{\hat{\theta} - (1-P)}{P}$$

where, $\hat{\theta}$ is the proportion of "Yes" answer in a sample.

The variance of the estimator $\hat{\pi}_1$ is given by:

$$V(\hat{\pi}_1) = \frac{\pi_S(1-\pi_S)}{n} + \frac{(1-P)(1-\pi_S)}{nP}$$

Kim and Warde model: Kim and Warde[15] suggested a stratified randomized response model based on Warner[1] model. Suppose the population is partitioned into strata and a sample is selected by simple random sampling with replacement in each stratum. To get the full benefit from stratification, it is assumed that the number of units in each stratum is known. An individual respondent in the sample of stratum 'i' is instructed to use the randomization device Ri which consists of a sensitive question (S) card with probability Pi and its negative question (S^c) with probability (1-Pi). The respondent should answer the question by "Yes" or "No" without reporting which question card she or he has. A respondent belonging to the sample in different strata will perform different randomization devices, each having different pre-assigned probabilities. Let ni denote the number of units in the sample from stratum i and n denote the total number of units in sample from all stratum so that $n = \sum_{i=1}^{k} n_i$ Under the assumption that these "Yes" or "No" reports are made truthfully and Pi ($\neq$0.5) is set by the researcher, the probability of a "Yes" answer in a stratum i for this procedure is:

$$Z_i = P_i\pi_{Si}+(1-P_i)(1-\pi_{Si}),\ \text{for } (i=1, 2 ..., k)$$

where, Z_i is the proportion of "Yes" answers in a stratum i, π_{Si} is the proportion of respondents with the sensitive trait in a stratum i and P_i is the probability that a respondent in the sample stratum i has a sensitive question (S) card.

The maximum likelihood unbiased estimate of π_{Si} is shown to be:

$$\hat{\pi}_{Si} = \frac{\hat{Z}_i - (1-P_i)}{2P_i - 1},\ \text{for } (i = 1, 2,..., k)$$

where, $\hat{Z}_i$ is the proportion of "Yes" answer in a sample in the stratum i. Since each $\hat{Z}_i$ is a binomial distribution B (ni, Zi) and the selections in different strata are made independently, the maximum likelihood estimate of π_{kw} (which is unbiased) is easily shown to be:

$$\hat{\pi}_{kw} = \sum_{i=1}^{k} w_i \hat{\pi}_{Si} = \sum_{i=1}^{k} w_i \left[\frac{\hat{Z}_i - (1-P_i)}{2P_i - 1}\right]$$

where, N is the number of units in the whole population, Ni is the total number of units in the stratum i and wi = (Ni/N) for (i = 1, 2, ...k) so that $w = \sum_{i=1}^{k} w_i = 1$.

The variance of $\hat{\pi}_{kw}$ is:

$$V(\hat{\pi}_{kw}) = \sum_{i=1}^{k} \frac{w_i^2}{n_i}\left[\pi_{Si}(1-\pi_{Si}) + \frac{P_i(1-P_i)}{(2P_i-1)^2}\right]$$

The optimal (Neyman) allocation of n to n_1, n_2...n_{k-1} and n_k to derive the minimum variance of the $\hat{\pi}_{kw}$ subject to $n = \sum_{i=i}^{k} n_i$ is approximately given by:

$$\frac{n_i}{n} = \frac{w_i\left[\pi_{Si}(1-\pi_{Si}) + P_i(1-P_i)/(2P_i-1)^2\right]^{1/2}}{\sum_{i=1}^{k} w_i\left[\pi_{Si}(1-\pi_{Si}) + P_i(1-P_i)/(2P_i-1)^2\right]^{1/2}}$$

Thus the minimal variance of an estimator $\hat{\pi}_{kw}$ is given by:

$$V(\hat{\pi}_{kw}) = \frac{1}{n}\left[\sum_{i=1}^{k} w_i \left(\pi_{Si}(1-\pi_{Si}) + \frac{P_i(1-P_i)}{(2P_i-1)^2}\right)^{1/2}\right]^2$$

Ki *et al.*[17] envisaged RR technique that applied the same randomization device to every stratum. Stratified random sampling is generally obtained by dividing the population into non-overlapping groups called strata and selecting a simple random sample from each stratum. An RR technique using a stratified sampling gives the group characteristics related to each stratum estimator. Also, stratified samples protect a researcher from the possibility of obtaining a poor sample. For the sake of completeness and convenience to the readers, the descriptions of fuzzy sets, fuzzy numbers, Triangular Fuzzy Number (TFN) and Trapezoidal Fuzzy Number (TrFN) which are reproduced here from Bector and Chandra[18], Mahapatra and Roy[19], Hassanzadeh *et al.*[20] and Aggarwal and Sharma[21].

Fuzzy sets were introduced by Zadeh[22] to represent/manipulate data and information possessing non-statistical uncertainties.

It was specifically designed to mathematically represent uncertainty and vagueness and to provide formalized tools for dealing with the imprecision intrinsic to many problems. However, the story of fuzzy logic started much more earlier. To devise a concise theory of logic and later mathematics, Aristotle posited the so-called "Laws of Thought". One of these, the "Law of the Excluded Middle," states that every proposition must either be True (T) or False (F). Even when Parmenides proposed the first version of this law (around 400 Before Christ) there were strong and immediate objections: for example, It was Plato who laid the foundation for what would become fuzzy logic, indicating that there was a third region (beyond T and F) where these opposites "tumbled about." A systematic alternative to the bi-valued logic of Aristotle. Three-valued logic, along with the mathematics to accompany it. The third value can be best be translated as the term "possible," and the numeric value between T and F. Eventually, an entire notation and axiomatic system from which he hoped to derive modern mathematics. Later, four-valued logics, five-valued logics and then declared that in principle there was nothing to prevent the derivation of an infinite-valued logic. three- and infinite-valued logics were the most intriguing, but ultimately settled on a four-valued logic because it seemed to be the most easily adaptable to Aristotelian logic.

The notion of an infinite-valued logic was introduced in Zadeh's seminal work "Fuzzy Sets" where the mathematics of fuzzy set theory and by extension fuzzy logic. This theory proposed making the membership function (or the values F and T) operate over the range of real numbers [0, 1]. New operations for the calculus of logic were proposed and showed to be in principle at least a generalization of classic logic. Fuzzy logic provides an inference morphology that enables approximate human reasoning capabilities to be applied to knowledge-based systems. The theory of fuzzy logic provides a mathematical strength to capture the uncertainties associated with human cognitive processes, such as thinking and reasoning. The conventional approaches to knowledge representation lack the means for representating the meaning of fuzzy concepts. As a consequence, the approaches based on first order logic and classical probability theory do not provide an appropriate conceptual framework for dealing with the representation of commonsense knowledge, since such knowledge is by its nature both lexically imprecise and non-categorical.

There are two main characteristics of fuzzy systems that give them better performance for specific applications.

- Fuzzy systems are suitable for uncertain or approximate reasoning, especially for the system with a mathematical model that is difficult to derive.
- Fuzzy logic allows decision making with estimated values under incomplete or uncertain information

Fuzzy set: A fuzzy set $\tilde{A}$ in a universe of discourse X is defined as the following set of pairs $\tilde{A}\{(x, \mu_{\tilde{A}}(x): x\in X\}$. Here $\mu_{\tilde{A}}(z): X\rightarrow[0, 1]$ is a mapping called the membership function of the fuzzy set $\tilde{A}$ and $\mu_{\tilde{A}}$ is called the membership value or degree of membership of $x\in X$ in the fuzzy set $\tilde{A}$. The larger the value of $\mu_{\tilde{A}}$ the stronger the grade of membership in $\tilde{A}$.

α-Cut: The α-cut for a fuzzy set $\tilde{A}$ is shown by $\tilde{A}_\alpha$ and for $\alpha\in[0,1]$ is defined to be:

$$\tilde{A}_\alpha = \left\{(x \mid \mu_{\tilde{A}}(x) \geq \alpha : x \in X)\right\} \quad (1)$$

where, X is the universal set.

Upper and lower bounds for any α-cut $\tilde{A}_\alpha$ are given by $\tilde{A}_\alpha^U$ and $\tilde{A}_\alpha^L$ respectively.

Fuzzy number: A fuzzy set in R is called a fuzzy number if it satisfies the following conditions:

- A is convex and normal
- A_α is a closed interval for every $\in$ (0, 1]
- Support of is bounded

Triangular fuzzy number (TFN): A fuzzy number $\tilde{A}$ = (p, q, r) is said to be a triangular fuzzy number if its membership function is given by:

$$\mu_{\tilde{A}}(x) = \begin{cases} \dfrac{x-p}{q-p}, & \text{if } p \leq x \leq q, \\ \dfrac{r-x}{r-q}, & \text{if } q \leq x \leq r, \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

Trapezoidal fuzzy number (TrFN): A fuzzy set $\tilde{A}$ = (p, q, r, s) on real numbers R is called a trapezoidal fuzzy number with membership function as follows:

$$\mu_{\tilde{A}}(x) = \begin{cases} 0, & \text{if } x \leq p, \\ \dfrac{x-p}{q-p}, & \text{if } p \leq x \leq q, \\ \dfrac{s-x}{s-r}, & \text{if } r \leq x \leq s, \\ 0, & \text{if } s \leq x \end{cases} \quad (3)$$

PROBLEM FORMULATION

Ki *et al.*[17] suggested a stratified RR technique that applied the same randomization device to every stratum. Stratified random sampling is generally obtained by dividing the population into two over lapping groups called strata and selecting a simple random sample from each stratum. An RR technique using a stratified random sampling gives the group characteristics related to each stratum estimator. Also, stratified sample protect a researcher from the possibility of obtaining a poor sample. Under Ki *et al.*[17] proportional sampling assumption, it may be easy to derive the variance of the proposed estimator; however, it may cause a high cost because of the difficulty in obtaining a proportional sample from some stratum. To rectify this problem, Kim and Warde[15] present a stratified randomized response technique using an optimal allocation which is more efficient than a stratified randomized response technique using a proportional allocation. Singh and Tarray[7] developed a stratified randomized response models designated as SRRT1 which is described below:

SRRT1: In this procedure, an individual respondent in a sample from each stratum is provided with one randomized response device. It consists of the statement "I belong to the sensitive group" with known probability Pi, exactly the same probability as used by Kim and Warde[15] and the statement "Yes" with probability (1-Pi). The interviewee is instructed to use the device and report "Yes" or "No" for the random outcome of the sensitive statement according to his actual status. Otherwise, it is simply requested to report the "Yes" statement observed on the randomized response device. The whole procedure is completed by the respondent, unobserved by the interviewer. A respondent belonging to the sample in different strata will perform different randomization devices, each having different pre-assigned probabilities. The probability of a "Yes" answer in a stratum i for this procedure is:

$$\theta_{1i} = P_i\pi_{Si}+(1-P_i), \quad \text{for } (i = 1, 2..., k) \tag{4}$$

where, θ_{1i} is the proportion of "Yes" answers in a stratum i, π_{Si} is the proportion of respondents with the sensitive trait in a stratum i and Pi is the probability that a respondent in the sample stratum i has a sensitive question.

The maximum likelihood estimate of π_{Si} in this procedure will be:

$$\hat{\pi}_{1Si} = \frac{\hat{\theta}_{1i} - (1-P_i)}{P_i}, \quad \text{for } (i = 1, 2,..., k) \tag{5}$$

where, $\hat{\theta}_{1i}$ is the proportion of "Yes" answer in a sample in the stratum i. Since each $\hat{\theta}_{1i}$ is a binomial distribution $B(n_i, \hat{\theta}_{1i})$ and the selections in different strata are made independently, the maximum likelihood estimate of $\pi_S = \sum_{i=1}^{k} w_i\pi_{Si}$ is easily shown to be:

$$\hat{\pi}_{1S} = \sum_{i=1}^{k} w_i\hat{\pi}_{1Si} = \sum_{i=1}^{k} w_i\left[\frac{\hat{\theta}_{1i} - (1-P_i)}{P_i}\right]$$

where they denote N to be the number of units in the whole population, Ni to be the total number of units in the stratum i and wi = (Ni/N) for (i = 1, 2,...k) so that $w = \sum_{i=1}^{k} w_i = 1.$

As each estimator $\hat{\pi}_{1Si}$ is unbiased for π_{Si} the expected value for $\hat{\pi}_S$, the expected value of $\hat{\pi}_{1S}$ is:

$$E(\hat{\pi}_{1S}) = E\left(\sum_{i=1}^{k} w_i\,\hat{\pi}_{1Si}\right) = \sum_{i=1}^{k} w_i\,E\left(\hat{\pi}_{1Si}\right) = \sum_{i=1}^{k} w_i\,\hat{\pi}_{Si} = \pi_S \tag{6}$$

Since each unbiased estimator $\hat{\pi}_{1Si}$ has its own variance, the variance of $\hat{\pi}_{1S}$ is:

$$V(\hat{\pi}_{1S}) = V\left(\sum_{i=1}^{k} w_i\,\hat{\pi}_{1Si}\right) \tag{7}$$

$$= \sum_{i=1}^{k} w_i^2\,V(\hat{\pi}_{1Si}) \tag{8}$$

$$= \sum_{i=1}^{k} \frac{w_i^2}{n_i}\left[\pi_{Si}(1-\pi_{Si}) + \frac{(1-P_i)(1-\pi_{Si})}{P_i}\right] \tag{9}$$

Or:

$$V(\hat{\pi}_S) = \sum_{i=1}^{k} \frac{w_i^2}{n_i^2}\{A_i\} \tag{10}$$

To find the optimum allocation we either maximize the precision for fixed budget or minimize the cost for fixed precision. A linear cost function which is an adequate approximation of the actual cost incurred will be

The linear cost function is:

$$C = C_0 + \sum_{i=1}^{k} c_i n_i \tag{11}$$

where, C_0 is the over head cost, c_i is the per unit cost of measurement in ith stratum, C is the available fixed budget for the survey.

In view of Eq. 4 and 11, the problem of optimum allocation can be formulated as a non linear programming problem (NLPP) for fixed cost as:

$$\left.\begin{aligned} &\text{Minimize } V(\hat{\pi}_S) = \sum_{i=1}^{k} \frac{w_i^2}{n_i} A_i \\ &\text{Subject to } \sum_{i=1}^{k} c_i n_i \le c_0 \\ &1 \le n_i \le N_i \text{ and } n_i \text{ integers}, i = 1, 2,...k \end{aligned}\right\} \tag{12}$$

The restrictions $1 \le n_i$ and $n_i \le N_i$ are placed to have the representation of every stratum in the sample and to avoid the oversampling, respectively.

FUZZY FORMULATION

Generally, real-world situations involve a lot of parameters such as cost and time, whose values are assigned by the decision makers and in the conventional approach, they are required to fix an exact value to the aforementioned parameters. However, decision-makers frequently do not precisely know the value of those parameters. Therefore, in such cases it is better to consider those parameters or coefficients in the decision-making problems as fuzzy numbers. The mathematical modeling of fuzzy concepts was presented by Zadeh[22]. Therefore, the fuzzy formulation of problem (12) with fuzzy cost constraint is given by considering two cases of fuzzy numbers, that is, triangular fuzzy number (TFN) and trapezoidal fuzzy number (TrFN).

For triangular fuzzy number (TFN) we consider:

$$\left.\begin{array}{l}\text{Minimize } \sum_{i=1}^{k} \frac{w_i^2}{n_i} A_i \\ \text{Subject to } \sum_{i=1}^{k} \left(c_i^1, c_i^2, c_i^3\right) n_i \leq \left(c_0^1, c_0^2, c_0^3\right) \\ 1 \leq n_i \leq N_i \text{ and } n_i \text{ integers}, i = 1, 2, \ldots k\end{array}\right\} \quad (13)$$

Where:

$$A_i = \left[\pi_{Si}(1-\pi_{Si}) + \frac{(1-P_i)(1-\pi_{Si})}{P_i}\right] \quad (14)$$

And $\tilde{C}_i = \left(c_i^1, c_i^2, c_i^3\right)$ is triangular fuzzy numbers with membership function:

$$\mu_{\tilde{C}_i}(x) = \begin{cases} \frac{x - c_i^1}{c_i^2 - c_i^1}, & \text{if } c_i^1 \leq x \leq c_i^2 \\ \frac{c_i^3 - x}{c_i^3 - c_i^2}, & \text{if } c_i^2 \leq x \leq c_i^3 \\ 0, & \text{otherwise} \end{cases} \quad (15)$$

Similarly, the membership function for available budget can be expressed as:

$$\mu_{\tilde{C}_i}(x) = \begin{cases} \frac{x - c_0^1}{c_i^2 - c_0^1}, & \text{if } c_0^1 \leq x \leq c_0^2 \\ \frac{c_0^3 - x}{c_0^3 - c_0^2}, & \text{if } c_0^2 \leq x \leq c_0^3 \\ 0, & \text{otherwise} \end{cases} \quad (16)$$

and for trapezoidal fuzzy number (TrFN) we consider:

$$\left.\begin{array}{l}\text{Minimize } \sum_{i=1}^{k} \frac{w_i^2}{n_i} A_i \\ \text{Subject to } \sum_{i=1}^{k} \left(c_i^1, c_i^2, c_i^3, c_i^4\right) n_i \leq \left(c_0^1, c_0^2, c_0^3, c_0^4\right) \\ 1 \leq n_i \leq N_i \text{ and } n_i \text{ integers}, i = 1, 2, \ldots k\end{array}\right\} \quad (17)$$

Where:

$$A_i = \left[\pi_{Si}(1-\pi_{Si}) + \frac{(1-P_i)(1-\pi_{Si})}{P_i}\right] \quad (18)$$

and $\tilde{C}_i = \left(c_i^1, c_i^2, c_i^3, c_i^4\right)$ is trapezoidal fuzzy numbers with membership function:

$$\mu_{\tilde{C}_i}(x) = \begin{cases} 0, & \text{if } x \leq c_i^1, \\ \frac{x - c_i^1}{c_i^2 - c_i^1}, & \text{if } c_i^1 \leq x \leq c_i^2, \\ 1, & \text{if } c_i^2 \leq x \leq c_i^3, \\ \frac{c_i^4 - x}{c_i^4 - c_i^3}, & \text{if } c_i^3 \leq x \leq c_i^4, \\ 0, & \text{if } c_i^4 \leq x \end{cases} \quad (19)$$

Similarly, the membership function for available budget can be expressed as:

$$\mu_{\tilde{C}_0}(x) = \begin{cases} 0, & \text{if } x \leq c_0^1, \\ \frac{x - c_0^1}{c_0^2 - c_0^1}, & \text{if } c_0^1 \leq x \leq c_0^2, \\ 1, & \text{if } c_0^2 \leq x \leq c_0^3, \\ \frac{c_0^4 - x}{c_0^4 - c_0^3}, & \text{if } c_0^3 \leq x \leq c_0^4, \\ 0, & \text{if } c_0^4 \leq x \end{cases} \quad (20)$$

LAGRANGE MULTIPLIERS FORMULATION

Let us now determine the solution of problems (13) by ignoring upper and lower bounds and integer requirements the NLPP with TFNs is solved by Lagrange multipliers technique (LMT).

The Lagrangian function may be:

$$\varphi(n_h, \lambda) = \sum_{i=1}^{k} \frac{w_i^2}{n_i}\{A_i\} + \lambda\left[\sum_{i=1}^{k}\left(c_i^{(1)}, c_i^{(2)}, c_i^{(3)}\right) n_i - \left(c_0^{(1)}, c_0^{(2)}, c_0^{(3)}\right)\right] \quad (21)$$

Differentiating Eq. 21 with respect to n_i and λ and equating to zero, we get the following sets of equations:

$$\frac{\nabla\varphi}{\nabla n_i} = 0 \Rightarrow n_i = \frac{w_i^2}{n_h^2}\{A_i\} + \lambda\left(c_i^{(1)}, c_i^{(2)}, c_i^{(3)}\right) \quad (22)$$

Or:

$$n_i = \frac{1}{\sqrt{\lambda}} \frac{\sqrt{\{A_i\}}}{\left(c_i^{(1)}, c_i^{(2)}, c_i^{(3)}\right)} \quad (23)$$

Also:

$$\frac{\nabla\varphi}{\nabla\lambda} = \left\{\sum_{i=1}^{k}\left(c_i^{(1)}, c_i^{(2)}, c_i^{(3)}\right)n_i - \left(c_0^{(1)}, c_0^{(2)}, c_0^{(3)}\right)\right\} = 0 \quad (24)$$

Which gives:

$$\sum_{i=1}^{k} w_i\left(c_i^{(1)}, c_i^{(2)}, c_i^{(3)}\right)\sqrt{\frac{\{A_i\}}{\lambda\left(c_i^{(1)}, c_i^{(2)}, c_i^{(3)}\right)}} - \left(c_0^{(1)}, c_0^{(2)}, c_0^{(3)}\right) = 0 \quad (25)$$

Or:

$$\frac{1}{\sqrt{\lambda}} = \frac{\left(c_0^{(1)}, c_0^{(2)}, c_0^{(3)}\right)}{\sum_{i=1}^{k} w_i\sqrt{\{A_i\}\left(c_i^{(1)}, c_i^{(2)}, c_i^{(3)}\right)}} \quad (26)$$

Substituting Eq. 23 in Eq. 26, we have:

$$n_i^* = \frac{\left(c_0^{(1)}, c_0^{(2)}, c_0^{(3)}\right), w_i\sqrt{\frac{\{A_i\}}{\left(c_i^{(1)}, c_i^{(2)}, c_i^{(3)}\right)}}}{\sum_{i=1}^{k} w_i\sqrt{\{A_i\}\left(c_i^{(1)}, c_i^{(2)}, c_i^{(3)}\right)}} \quad (27)$$

In similar manner, the optimum allocation of NLPP Eq. 17 with trapezoidal fuzzy number can be obtained as follows:

$$n_i^* = \frac{\left(c_0^{(1)}, c_0^{(2)}, c_0^{(3)}, c_0^{(4)}\right), w_i\sqrt{\frac{\{A_i\}}{\left(c_i^{(1)}, c_i^{(2)}, c_i^{(3)}, c_i^{(4)}\right)}}}{\sum_{i=1}^{k} w_i\sqrt{\{A_i\}\left(c_i^{(1)}, c_i^{(2)}, c_i^{(3)}, c_i^{(4)}\right)}} \quad (28)$$

To convert fuzzy allocations into a crisp allocation by-cut method.

PROCEDURE FOR CONVERSATION OF FUZZY NUMBERS

The fuzzy allocations into a crisp allocation by α-cut method let $\tilde{A} = (p, q, r)$ be a TFN. An α-cut for $\tilde{A}$, $\tilde{A}_\alpha$ computed as:

$$\alpha = \frac{x-p}{q-p} \Rightarrow \tilde{A}_\alpha^L = x = (q-p)\,\alpha + p$$

and:

$$\alpha = \frac{r-x}{r-q} \Rightarrow \tilde{A}_\alpha^U = x = r - (r-q)\,\alpha \quad (29)$$

where, $\tilde{A}_\alpha = \left[\tilde{A}_\alpha^L, \tilde{A}_\alpha^U\right]$ is the corresponding α-cut as shown in Fig. 1. The allocation obtained in Eq. 27 is in the form of triangular fuzzy number, therefore by using Eq. 29 the equivalent crisp allocation is given by:

$$n_i^* = \frac{\left(c_0^{(3)} - \left(c_0^{(3)} - c_0^{(2)}\right)\right), w_i\sqrt{\frac{\{A_i\}}{\left(c_h^{(1)} + \left(c_h^{(2)} - c_h^{(1)}\right)\right)}}}{\sum_{i=1}^{k} w_i\sqrt{\{A_i\}\left(c_h^{(1)} + \left(c_h^{(2)} - c_h^{(1)}\right)\right)}} \quad (30)$$

Similarly, let $\tilde{A}$ = (p, q, r, s) be a TrFN. An α-cut for $\tilde{A}$, $\tilde{A}_\alpha$ computed as:

$$\alpha = \frac{x-p}{q-p} \Rightarrow \tilde{A}_\alpha^L = x = (q-p)\alpha + p$$

And:

$$\alpha = \frac{s-x}{s-q} \Rightarrow \tilde{A}_\alpha^U = x = s - (s-q)\alpha \quad (31)$$

where, $\tilde{A}_\alpha = \left[\tilde{A}_\alpha^L, \tilde{A}_\alpha^U\right]$ is the corresponding α-cut as shown in Fig. 2. The allocation obtained in Eq. 30 is in the form of triangular fuzzy number, therefore by using Eq. 31 the equivalent crisp allocation is given by:

$$n_i^* = \frac{\left(c_0^{(4)} - \left(c_0^{(4)} - c_0^{(3)}\right)\right), w_i\sqrt{\frac{\{A_i\}}{\left(c_h^{(1)} + \left(c_h^{(2)} - c_h^{(1)}\right)\right)}}}{\sum_{i=1}^{k} w_i\sqrt{\{A_i\}\left(c_h^{(1)} + \left(c_h^{(2)} - c_h^{(1)}\right)\right)}} \quad (32)$$

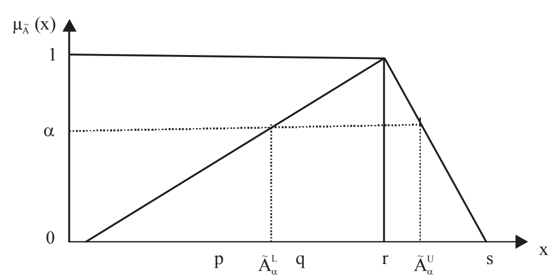

Fig. 1: Triangular fuzzy number with an α-cut

The allocations obtained by Eq. 30 and Eq. 32 provide the solution to NLPP Eq. 13 and 17 if it satisfies the restriction $1 \le n_i \le N_h$, i = 1, 2,..., k. The allocations obtained in Eq. 30 and 32 may not be integer allocations, so to get integer allocations, round off the allocations to the nearest integer values. After rounding off we have to be careful in rechecking that the round-off values satisfy the cost constraint. Now we further discuss equal and proportional allocations as follows:

Equal allocation: In this method, the total sample size is divided equally among all the strata, that is, for the th Stratum:

$$n_i = \frac{n}{k} \tag{33}$$

where, can be obtained from the cost constraint equation as follows:

$$\sum_{i=1}^{k} w_i \sqrt{\{A_i\}\left(c_h^{(1)} + \left(c_h^{(2)} - c_h^{(1)}\right)\right)}\, n_i = \left(c_0^{(4)} - \left(c_0^{(4)} - c_0^{(3)}\right)\right) \tag{34}$$

$$n_i \propto w_i$$

Or:

$$n_i = n w_i \tag{35}$$

Now substituting the value of n_i in Eq. 34, we get:

$$n_i = \frac{N\left(c_0^{(4)} - \left(c_0^{(4)} - c_0^{(3)}\right)\right)}{\sum_{i=1}^{k} \sqrt{\left(c_h^{(1)} + \left(c_h^{(2)} - c_h^{(1)}\right)\right)} N_h} \tag{36}$$

Proportional allocation: This allocation was originally proposed by Bowley[23]. This procedure of allocation is very common in practice because of its simplicity. When no other information except N_i, the total number of units in the i^{th} stratum, is available, the allocation of a given sample of size n to different strata is done in proportion to their sizes, that is, in the i^{th} stratum:

$$n_i = n \frac{N_i}{N} \tag{37}$$

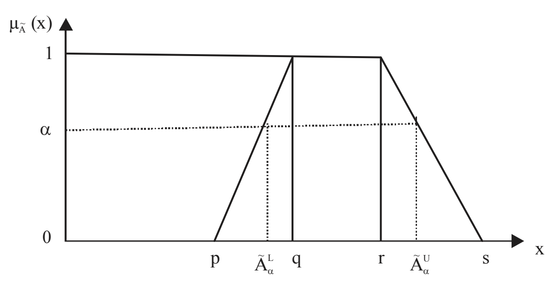

Fig. 2: Trapezoidal fuzzy number with an α-cut

NUMERICAL ILLUSTRATION

A hypothetical example is given to illustrate the computational details of the proposed problem. Let us suppose the population size is 1000 with total available budget of the survey as TFNs and TrFNS are (3500, 4000, 4800) and (3500, 4000, 4400, 4600) units, respectively. The other required relevant information is given in Table 1. By using the value of Table 1, it was computed the values of A_i which is given in Table 2.

After substituting all the values from Table 1 and 2 in Eq. 13, the required FNLLP is given as:

$$\left.\begin{aligned} &\text{Minimize } V(\hat{\pi}_S) = \frac{0.02495672}{n_1} + \frac{0.13331246}{n_2} \\ &\text{Subject to } (1, 2, 4)\, n_1 + (18, 20, 24) n_2 \le (3500, 4000, 4800) \\ &\qquad 1 \le n_1 \le 300 \\ &\qquad 1 \le n_2 \le 700 \end{aligned}\right\} \tag{38}$$

The required optimum allocations for problem (Eq. 13) obtained by substituting the values from Table 1 and 2 in Eq. 30 at α = 0.5 will be:

$$n_1 = \frac{(4800 - 800\alpha)0.3\sqrt{(0.2772969)/(\alpha+1)}}{0.3\sqrt{(0.2772969)(\alpha+1)} + 0.7\sqrt{(0.2716829)(2\alpha+18)}}$$

$$n_2 = \frac{(4800 - 800\alpha)0.7\sqrt{(0.2716829)/(2\alpha+18)}}{0.3\sqrt{(0.2772969)(\alpha+1)} + 0.7\sqrt{(0.2716829)(2\alpha+18)}}$$

In similar manner, optimum allocation for problem (Eq. 17) obtained by substituting the values from Table 1 and 2 in Eq. 32 at α = 0.55 will be:

$$n_1 = \frac{(3750)0.3\sqrt{(0.2772969)/(\alpha+1)}}{0.3\sqrt{(0.2772969)(\alpha+1)} + 0.7\sqrt{(0.2716829)/(2\alpha+18)}}$$

$$n_2 = \frac{(4400 - 200\alpha)0.7\sqrt{(0.2716829)/(2\alpha+18)}}{0.3\sqrt{(0.2772969)(\alpha+1)} + 0.7\sqrt{(0.2716829)(2\alpha+18)}}$$

The values of X_i, A_i and $A_i w_i^2$ are calculated as given in Table 2.

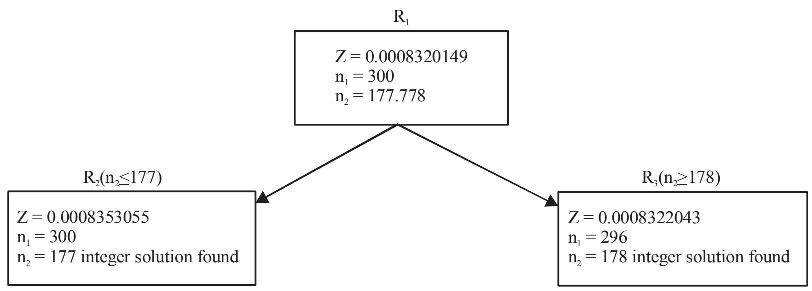

Fig. 3: Various nodes of NLPP

Table 1: Stratified population with two strata

Stratum (i)	T_i	w_i	π_Y	π_{Si}	P_i	(c_h^1, c_h^2, c_h^3)	(c_0^1, c_0^2, c_0^3)
1	0.495	0.3	0.91	0.48	0.9	(1,2,4)	(1,2,4,7)
2	0.95	0.7	0.91	0.53	0.1	(18,20,24)	(18,20,24,26)

Table 2: Calculated values of A_i and $A_i w_i^2$

Stratum (i)	X_i	A_i	$A_i w_i^2$
1	0.501715	0.272969	0.0249572
2	0.5471	0.2716829	0.1331246

Table 3: Calculated values of optimum allocation and variance

Case	n_1	n_2	Variance
LMT (optimum allocation)			
TFN	318.15	271.00	0.000569678
TrFN	206.4617	201.76	0.000780694

Applying the α-cut and LMT, the optimum allocation after is obtained and summarized in Table 3 for both the cases i.e., case of TFN and case of TrFN with variance as[24-25]:

Case-I:

$$\left.\begin{aligned} &\text{Minimize } V(\hat{\pi}_S) = \frac{0.02495672}{n_1} + \frac{0.1331246}{n_2} \\ &\text{sujbect to } (1)n_1 + (18)n_2 \le (3500) \\ &\qquad 1 \le n_1 \le 300 \\ &\qquad 1 \le n_2 \le 700 \end{aligned}\right\}$$

Using the above minimization problem, we get optimal solution as n_1 = 300, n_2 = 177.778 and optimal value is Minimize V($\hat{\pi}_S$) = 0.0008320149.

Since n_1 and n_2 are required to be the integers, we branch problem R_1 into two sub problems R_2 and R_3 by introducing the constraints $n_2 \le 177$ and $n_2 \ge 178$ respectively indicated by the value n_1 = 300 and n_2 = 177 and n_1 =296 and n_2 = 178. Hence the solution is treated as optimal. The optimal value is n_1 = 296 and n_2 = 178 and optimal solution is to minimize V ($\hat{\pi}_S$) = 0.0008320149. It may be noted that the optimal integer values are same as obtained by rounding the n_i to the nearest integer. Let us suppose V ($\hat{\pi}_S$) = Z, the various nodes for the NLPP utilizing case-I, are presented in Fig. 3.

Case-II:

$$\left.\begin{aligned} &\text{Minimize } V(\hat{\pi}_S) = \frac{0.02495672}{n_1} + \frac{0.1331246}{n_2} \\ &\text{sujbect to } (2)n_1 + (20)n_2 \le (4000) \\ &\qquad 1 \le n_1 \le 300 \\ &\qquad 1 \le n_2 \le 700 \end{aligned}\right\}$$

Using the above minimization problem, we get optimal solution as n_1 = 240.86, n_2 = 175.91 and optimal value is Minimize V ($\hat{\pi}_S$) = 0.0008603746.

Since n_1 and n_2 are required to be the integers, so problem R_1 is further branched into sub problems R_2; R_2; R_4 and R_5 with additional constraints as $n_1 \le 240$; $n_1 \ge 241$; $n_2 \le 175$ and $n_2 \ge 176$; respectively. Problems R_2, R_4 and R_5 stand fathomed as the optimal solution in each case is integral in n_1 and n_2. Problem R_3 has been further branched into sub problems R_4 and R_5 with additional constraints as $n_1 \le 175$ and $n_1 \ge 176$; respectively which suggests that R_6 is fathomed and R_7 has no feasible solution. The optimal value is n_1 = 240 and n_2 = 136 and optimal solution is to Minimize V ($\hat{\pi}_S$) = 0.0008603761. Let us suppose V ($\hat{\pi}_S$) = Z, the various nodes for the NLPP utilizing case-II, are presented in Fig. 4.

Case-III:

$$\left.\begin{aligned} &\text{Minimize } V(\hat{\pi}_S) = \frac{0.02495672}{n_1} + \frac{0.1331246}{n_2} \\ &\text{sujbect to } (4)n_1 + (24)n_2 \le (4800) \\ &\qquad 1 \le n_1 \le 300 \\ &\qquad 1 \le n_2 \le 700 \end{aligned}\right\}$$

Using the above minimization problem, we get optimal solution as n_1 = 180.25, n_2 = 170 and optimal value is Minimize V ($\hat{\pi}_S$) = 0.0009217340.

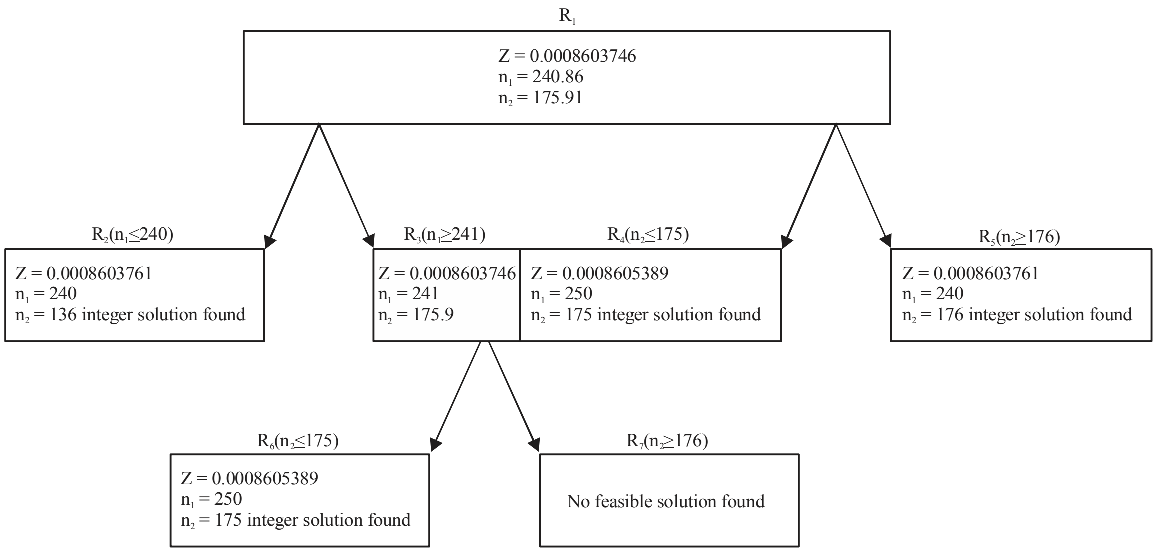

Fig. 4: Various nodes of NLPP

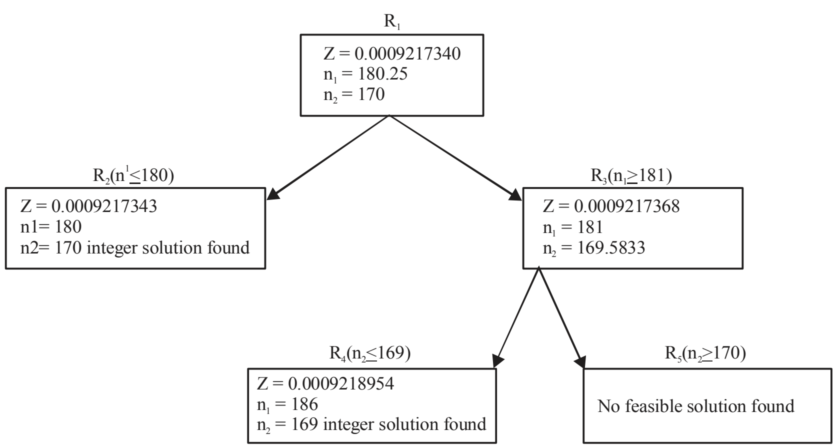

Fig. 5: Various nodes of NLPP

Since n_1 and n_2 are required to be the integers, so problem R_1 is further branched into sub problems R_2 and R_3 with additional constraints as $n_1 \leq 180$; $n_1 \geq 181$, respectively. Problems R_2 stand fathomed as the optimal solution in each case is integral in n_1 and n_2. Problem R_3 has been further branched into sub problems R_4 and R_5 with additional constraints as $n_2 \leq 169$ and $n_2 \geq 170$; respectively. R_4 is fathomed and R_5 has no feasible solution. Hence the solution is treated as optimal. The optimal value is $n_1 = 180.25$ and $n_2 = 170$ and optimal solution is to Minimize $V(\hat{\pi}_S) = 0.0009217340$. Let us suppose $V(\hat{\pi}_S) = Z$, the various nodes for the NLPP utilizing case-III, are presented in Fig. 5.

In both the three cases we find that the optimal value $n_1 = 296$ and $n_2 = 178$ and optimal solution is to Minimize $V(\hat{\pi}_S) = 0.0008320149$.

DISCUSSION

A stratified randomized response method assists to solve the limitations of randomized response that is the loss of individual characteristics of the respondents. The optimum allocation problem for two-stage stratified random sampling based on Singh and Tarray[7] model with fuzzy costs is formulated as a problem of fuzzy nonlinear programming problem. The problem is then solved by using Lagrange multipliers technique for obtaining optimum allocation. The optimum allocation obtained in the form of fuzzy numbers is converted into an equivalent crisp number by using α-cut method at a prescribed value of α.

For practical purposes we need integer sample sizes. Therefore, in instead of rounding off the continuous solution, we have obtained integer solution, by formulating the problem as fuzzy integer nonlinear programming problem and obtained the integer solution by LINGO software.

REFERENCES

1. Warner, S.L., 1965. Randomized response: A survey technique for eliminating evasive answer bias. J. Am. Statist. Assoc., 60: 63-69.
2. Greenberg, B.G., A.L.A. Abul-Ela, W.R. Simmons and D.G. Horvitz, 1969. The unrelated question randomized response model: Theoretical framework. J. Am. Statist. Assoc., 64: 520-539.
3. Fox, J.A. and P.E. Tracy, 1986. Randomized Response: A Method for Sensitive Surveys. SEGE Publications, Newbury Park, CA., ISBN-13: 9780803923096, Pages: 80.
4. Chaudhuri, A. and R. Mukerjee, 1988. Randomized Response: Theory and Techniques. Marcel Dekker, New York.
5. Singh, H.P. and T.A. Tarray, 2012. A stratified unknown repeated trials in randomized response sampling. Commun. Statist. Applic. Methods, 19: 751-759.
6. Singh, H.P. and T.A. Tarray, 2013. An alternative to Kim and Warde's mixed randomized response technique. Stataistica, 73: 379-402.
7. Singh, H.P. and T.A. Tarray, 2014. An improvement over Kim and Elam stratified unrelated question randomized response mode using neyman allocation. Sankhya B, 77: 91-107.
8. Tarray, T.A. and H.P. Singh, 2015. A randomized response model for estimating a rare sensitive attribute in stratified sampling using Poisson distribution. Model Assist. Statist. Applic., 10: 345-360.
9. Singh, H.P. and T.A. Tarray, 2016. An improved BAR-LEV, bobovitch and boukai randomized response model using moments ratios of scrambling variable. Hacettepe J. Math. Statist., 45: 593-608.
10. Tarray, T.A. and H.P. Singh, 2017. A stratified unrelated question randomized response model using Neyman allocation. Commun. Statist.-Theory Methods, 46: 17-27.
11. Tarray, T.A. and H.P. Singh, 2015. Some improved additive randomized response models utilizing higher order moments ratios of scrambling variable. Model Assist. Statist. Applic., 10: 361-383.
12. Tarray, T.A. and H.P. Singh, 2016. New procedures of estimating proportion and sensitivity using randomized response in a dichotomous finite population. J. Modern Applied Statist. Methods, 15: 635-669.
13. Tarray, T.A. and H.P. Singh, 2017. A survey technique for estimating the proportion and sensitivity in a stratified dichotomous finite population. Statist. Applic., 15: 173-191.
14. Tarray, T.A., 2017. Scrutinize on Stratified Randomized Response Technique. GRIN Verlag, Munich, ISBN: 9783668554580.
15. Kim, J.M. and W.D. Warde, 2004. A stratified Warner's randomized response model. J. Statist. Plann. Inference, 120: 155-165.
16. Singh, S., 1993. An alternative to Warner's randomized response technique. Statistica, 53: 67-71.
17. Ki, H.H., K.Y. Jun and Y.L. Hwa, 1994. A stratified randomized response technique. Korean J. Applied Statist., 7: 141-147.
18. Bector, C.R. and S. Chandra, 2005. Fuzzy Mathematical Programming and Fuzzy Matrix Games. Vol. 169, Springer, Berlin, Germany.
19. Mahapatra, G.S. and T.K. Roy, 2006. Fuzzy multi-objective mathematical programming on reliability optimization model. Applied Math. Comput., 174: 643-659.
20. Hassanzadeh, R., I. Mahdavi, N. Mahdavi-Amiri and A. Tajdin, 2012. An α-cut approach for fuzzy product and its use in computing solutions of fully fuzzy linear systems. Proceedings of the 2012 International Conference on Industrial Engineering and Operations Management, July 3-6, 2012, Istanbul, Turkey, pp: 2047-2056.
21. Aggarwal, S. and U. Sharma, 2013. Fully fuzzy multi-choice multi-objective linear programming solution via deviation degree. Int. J. Pure Applied Sci. Technol., 19: 49-64.
22. Zadeh, L.A., 1965. Fuzzy sets. Inform. Control, 8: 338-353.
23. Bowley, A.L., 1926. Measurement of the precision attained in sampling. Bull. Int. Statsit. Inst., 22: 1-62.
24. Naik, P.A. and K.R. Pardasani, 2014. Finite element model to study effect of Na^+/K^+ Pump and Na^+/Ca^{2+} exchanger on calcium distribution in oocytes in presence of buffers. Asian J. Math. Stat., 7: 21-28.
25. Naik, P.A. and K.R. Pardasani, 2013. Finite element model to study effect of buffers in presence of voltage gated Ca^{2+} channels on calcium distribution in oocytes for one dimensional unsteady state case. Int. J. Modern Biol. Med., 4: 190-203.