Stochastic nonlinear optimal allocations with Weibull cost function when non-response is observed

Ewere, F.[1, 2], Mbegbu, J. I.[2] and Okon, E. J.[2]

Different cost functions, comprising of linear, quadratic and gamma cost functions, have been utilised based on their functional form and will be obsolete if additional cost is incurred. In this paper, a Weibull cost function when there is non-response is proposed for multivariate stratified sampling surveys. Hansen and Hurwitz (1946) introduced the concept of non-response by subsampling non-respondents in a more careful second attempt, and the variance function was considered to be deterministic. Furthermore, in real-life situations, population parameters are considered to be random and uncertain, and it is advisable to use a stochastic programming approach to model such problems. To optimally distribute sample sizes to strata, the problem will be modelled using the multi-objective stochastic integer nonlinear programming problem (MO-SINLPP) technique. A solution method is suggested using the D1 distance technique and fuzzy programming. The results of the comparative study showed that the compromise optimum allocations outperformed the individual optimum allocation of the different characteristics using the Weibull cost function.

Keywords: Stochastic programming, Weibull cost function, Non-response, Optimal allocation, Fuzzy programming
DOI: https://doi.org/10.63255/01-0127.24/05

1.0 Introduction

Multivariate stratified sampling has been the core of sampling surveys in recent studies and primarily, sampling efficiency is determined by how sample sizes are distributed across strata. In surveys, obtaining information from the complete population rather than just a sample could make decision-making for researchers and policymakers more difficult. One way of improving the precision of estimates without increasing the sample size in a survey is by arranging the entire population into subpopulations called strata. In multivariate surveys, the best allocation for one character in a unit may not be convenient for another in that same unit. In these cases, a compromise that is, in some way, optimal for all the criteria must be found to determine how samples should be distributed throughout various strata. The optimum allocation problem was considered by Cochran (1977) and Sukhatme et al. (1984), for stratified random sampling using a univariate population. Several optimal compromise allocation methods for surveys that are multipurpose have been developed, either by

[1] Corresponding Email: friday.ewere@uniben.edu. The views expressed in this article are those of the authors and do not necessarily reflect the position of the CISON. While efforts are made to ensure the accuracy of this paper, CISON bears no responsibility for errors, omissions, or interpretations. Authors are solely responsible for the accuracy, originality, and ethical standards of their work.
[2] Department of Statistics, University of Benin, Edo State, Nigeria.

minimising the fixed cost for a specified variance or by reducing the coefficient of variation for a given cost.

Stratified sampling surveys present different challenges, one of which is determining the most appropriate stratum size. When allocating sample size to strata, factors such as the cost of collecting data per element in each stratum, variability within each stratum and the population size of each stratum must be considered. Some of these problems are shrouded in different forms of uncertainty. They become stochastic or fuzzy when some of the parameters are random (Alshqaq et al. 2022). These problems may be formulated in areas of random probability or uncertainty of outcome. This can be solved using some well-known methods like the Stochastic Chance Constrained Programming (SCCP) technique, which Charnes and Cooper originally introduced in 1959. The technique provides a robust approach for modelling stochastic control systems and decision-making (Ding and Wang, 2012).

In sample survey, it is expected that information would not be obtained from all the selected units in the sample, but practically, it’s generally not possible due to nonresponse. Nonresponse will exist if the sampled person does not provide the required information in the survey. According to Shafiullah et al. (2014), the degree of nonresponse solely relies on factors like the nature of the population being targeted, the kind of survey considered and the period in which the survey is conducted.

A classical nonresponse concept, designed originally for surveys, which first involved mailing of questionnaires and conducting of interviews second attempt to the subsample of the non-respondents was introduced by Hansen and Hurwitz (1946). Additionally, the estimator of the population mean was obtained, and the expression for its variance was also derived.

El-Badry (1956) modified the Hansen and Hurwitz method by sending multiple waves of questionnaires to the non-responding units to improve the rate of response. Foradori (1961) further expanded El-Badry’s approach to accommodate several sampling designs.

Khan et al. (2008) determined the optimal sample sizes and subsamples into strata when non-response is observed for multivariate stratified sampling survey with Lagrange Multipliers Technique and developed an estimate for stratum mean for the population characteristics that is unbiased and also derived the variance of the Hansen and Hurwitz (1946) estimator, which accounts for the sub-sample of non-respondents. The linear cost function with nonresponse was utilised to solve the nonlinear programming problem (NLPP), with a convex objective function. To further enhance the concept of nonresponse, new estimators were developed by Ahmed et al. (2016) for Scrambled Response on a second attempt.

Optimum compromise allocations with Integer values are recommended since they are widely used for practical implementation. To obtain integer values, Varshney et al. (2011) applied multivariate stratified sampling in the context of non-response by formulating an integer nonlinear programming problem with a linear cost function. A solution method was derived with lexicographic goal programming (LGP) to secure optimal compromise allocations. The procedure compared Cochran's allocation, Chatterjee's allocation, minimised trace and Khan's compromise allocation with the LGP technique. The LGP approach ensured a feasible solution, which other methods may not provide, and proved to be competitive in terms of variability and cost.

Varshney et al. (2012) introduced an optimal sampling design to address non-response when strata weights are unknown. An All Integer Non-Linear Programming Problem (AINLPP) was developed to estimate unknown population means. A solution method was formulated with the goal programming (GP) technique using the linear cost function to obtain sample sizes for strata, and the LINGO optimisation software was used to obtain integer solutions. The proposed compromise criterion using goal programming was effective in determining a compromise allocation in multivariate surveys when travel costs within strata are taken into account and auxiliary information is available.

In real-life scenarios, sampling cost and variance are known to be stochastic and follow probability distributions (Melaku, 1986). Haseen et al. (2014) developed a multivariate stratified sampling technique with non-response, where sampling cost and variance are random, as a multi-objective stochastic programming problem. To achieve the aim, random sampling variances and costs are transformed into a deterministic nonlinear programming problem (NLPP) with chance constraint and a modified E-model. Three different methods were adopted, which are fuzzy programming, D1 distance method and goal programming, to obtain the optimal compromise allocation to the developed problem. The linear cost function, which incorporates non-response, was employed.

Several optimisation techniques in multivariate surveys were examined by Raghav et al. (2014) when non-responses are observed with a linear cost function. These techniques depend on what is known about the population at first, which can be classified as complete, partial, or null information. The problem is addressed using various methods: the value function technique and goal programming when all information is known, the $\in$-constraint method when only partial information is available, and the distance-based method when insufficient information about the variables exists, all aimed at obtaining the compromise allocation of sample sizes under non-response.

The geometric programming technique in multivariate stratified sampling when non-response is observed was utilised by Shafiullah et al. (2015). The allocation problem was derived in two phases; the first phase problem was worked out as a geometric programming problem, and in the second phase, the primal-dual relationship theorem was used. A dual solution was obtained using the formulated geometric programming problem with the help of the quadratic cost function when nonresponse is observed.

Okon and Yahaya (2018) proposed a new gamma cost function in the presence of non-response in a multivariate stratified sampling design. This was further modified by Varshney and Mradula (2019), where they formulated a gamma cost function when non-response is considered and solved it using the LGP method. Contextually, the last known cost approximation in the presence of nonresponse is the gamma cost function proposed by Okon and Yahaya 2018).

Optimum allocation techniques are best formulated with integer bounds to obtain a true optimal solution for any given problem. The cost of a survey plays a critical role in determining the allocation of the sample across different strata. Determining the true functional form of the cost function is of great essence. Problem solving comes with real-life cost, sometimes from hired labour to equipment maintenance, transportation and several other financial constraints. The linear cost function was developed to take care of enumerating cost; the quadratic cost function is considered if transportation cost is significant, and the gamma cost function is relevant if the time taken to obtain data from the units of the population is taken into consideration. The existing cost functions were all proposed based on their functional form. From the literature, if additional cost is incurred in the survey, the existing cost functions become obsolete. Hence, developing a new cost function that will approximate the existing cost functions will be valid in some sense.

Modern electronic devices are currently used in the enumeration of household units during surveys. Advancement in modern technology has made it possible to design complex systems whose safety and operations rely on the reliability of the components that the system is made up of. For instance, a fuse might burn out, a steel column may buckle, or a heat-sensing device or battery might fail. It is widely established that identical components, even when exposed to the same environmental conditions, will fail at different and unpredictable times (Ronald et al. 2007). The linear, quadratic and gamma cost functions may not be suitable if the survey device fails. This failure may affect the enumeration of sampling units and will prompt an additional cost. In this case, a failure cost can be attached per stratum and the time to failure modelled using the Weibull distribution, both at the first attempt and a more careful second attempt.

Making decisions with uncertainty is tasking and unavoidable when dealing with real-life situations. In practice, due to the uncertainty in the population data, the variances as well as the costs should be treated as random variables (Alshaqq et al. 2022 and Mahfouz et al. 2023). Stochastic optimisation, which is also known as probabilistic optimisation, occurs when some or all of the data of the optimisation function follows a probability distribution.

In this light, we suggest a new cost function when nonresponse is observed, known as the Weibull cost function, which approximates the existing cost functions known in the literature. The existing cost functions used by other authors, which are being challenged in this work, are only relevant in the pencil-and-paper era of making questionnaires and will become obsolete with the advent of a technologically based questionnaire system. The Weibull cost function, as well as the variances, will be considered as random variables. The chance constraint and modified E-model technique will be used, respectively, to convert the sampling cost and variance, which are random variables, into a more non-stochastic nonlinear programming problem. The problem will be constructed as an MO-SINLPP.

Many practical challenges are yet to be solved in surveys and censuses here in Nigeria and Africa at large. In Nigeria, for example, most surveys are not carried out with the correct statistical model in mind. The Weibull cost function suggested in this paper will be a good approximation to the already existing cost functions.

2.0 Methodology

Let represent a finite population which is separated into mutually exclusive and exhaustive strata with sizes such that $\sum_{k=1}^{Q} N_k = N$. A simple random sample without replacement (SRSWOR) of size n_k is drawn independently from each stratum in a way that $\sum_{k=1}^{Q} n_k = n$. If the stratums are first divided into respondents and non-respondents groups, which are mutually exclusive and exhaustive. Let the mutually exclusive strata N_k be divided into N_{k1} and N_{k2}, which represents the respondents and non-respondents group respectively in the k^{th} stratum, where $N_{k2} = N_k - N_{k1}$. The values of N_{k1} and N_{k2} are not known before the sample observations are obtained. Let n_k units be drawn from N_k in the k^{th} stratum and let n_{k1} units be members of the respondents group and the remaining n_{k2} units be members of the non-respondents group, that is $n_{k2} = n_k - n_{k1}$.

To achieve a comprehensive result, a subsample of size r_k is obtained from n_{k2} non-respondents in a more careful second attempt. The necessary information is now obtained from the r_k units and this time around we assume that all the r_k units responded.

Let the subsample sizes at the second attempt be given as

$$r_k = \frac{n_{k2}}{h_k}, \quad k = 1, 2, ...Q \tag{1}$$

Equation (1) above are subsample sizes drawn from n_{k2} non-respondents group of the k^{th} stratum, with sample fraction denoted by $1/h_k$ where $h_k \geq 1$.

The estimator for non-response in the study variable is estimated as;

$$\bar{y}_{jk(w)} = \frac{n_{k1}\bar{y}_{jk1} + n_{k2}\bar{y}_{jk2(r_k)}}{n_k} \tag{2}$$

where $\bar{y}_{jk(w)}$ is the weighted mean, with $E(\bar{y}_{jk(w)}) = \bar{Y}_{jk}$.

The corresponding variance of $\bar{y}_{jk(w)}$ will be obtained as

$$V(\bar{y}_{jk(w)}) = \left(\frac{1}{n_k} - \frac{1}{N_k}\right)S_{jk}^2 + \frac{W_{k2}^2 S_{jk2}^2}{r_k} - \frac{W_{k2} S_{jk2}^2}{n_k} \tag{3}$$

where W_{k2} is the non-respondent weights.

The estimate of total population variance of $\bar{y}_{jk(w)}$ which is unbiased is obtained as

$$V(\bar{y}_{jk}) = \sum_{k=1}^{Q} W_k^2 V(\bar{y}_{jk(w)}) \tag{4}$$

$$V(\bar{y}_{jk}) = \sum_{k=1}^{Q} W_k^2 \left[\left(\frac{1}{n_k} - \frac{1}{N_k}\right)S_{jk}^2 + \frac{W_{k2}^2 S_{jk2}^2}{r_k} - \frac{W_{k2} S_{jk2}^2}{n_k}\right] \tag{5}$$

W_k is the stratum weights. If finite population correction (FPC) is ignored in Equation (5), then the following equation will be obtained: that is,

$$V(\bar{y}_{jk}) = \sum_{k=1}^{Q} \frac{W_k^2\left(S_{jk}^2 - W_{k2} S_{jk2}^2\right)}{n_k} + \sum_{k=1}^{Q} \frac{W_k^2 W_{k2}^2 S_{jk2}^2}{r_k} \tag{6}$$

$\bar{Y}_{jk} = N_k^{-1}\sum_{i=1}^{N_k} y_{jki}$: stratum mean in the k^{th} stratum for j^{th} characteristics;

$S_{jk}^2 = \left(N_k - 1\right)^{-1}\sum_{i=1}^{N_k}\left(y_{jki} - \bar{Y}_{jk}\right)^2$: stratum variance in the k^{th} stratum for j^{th} characteristics;

$S_{jk2}^2 = \left(N_{k2}-1\right)^{-1}\sum_{i=1}^{N_{k2}}\left(y_{jki}-\bar{Y}_{jk2}\right)^2$: stratum variance for non-response group in the k^{th} stratum. $\bar{y}_{jk1} = n_{k1}^{-1}\sum_{i=1}^{n_{k1}} y_{jki}$: sample mean of the first attempt units in the k^{th} stratum.

$\bar{y}_{jk2(r_k)} = r_k^{-1}\sum_{i=1}^{r_k} y_{jki}$: sample mean of units respond on second call in the k^{th} stratum.

2.1 Cost Function with Non-Response Stratified Sampling

The linear cost function if nonresponse is considered is given as

$$C = \sum_{k=1}^{Q} c_{k0}n_k + \sum_{k=1}^{Q} c_{k1}n_{k1} + \sum_{k=1}^{Q} c_{k2}n_{k2} \tag{7}$$

The quantity $W_{k1}n_k$ is used as the expected value since n_{k1} is not known until the first attempt is made. A cost function of the form is expected;

$$C = \sum_{k=1}^{Q}\left(c_{k0}+c_{k1}W_{k1}\right)n_k + \sum_{k=1}^{Q} c_{k2}r_k \tag{8}$$

where C is the total budget, c_{k0} is the cost attributed to selecting the n_k units.

In real life events, when determining survey cost, both measurement unit costs and travel costs within strata are key factors. If travel costs can be substantially measured, the cost function becomes non-linear. A non-linear cost function, which accounts for measurement unit costs and travel costs within strata, provides a better approximation of the survey's actual budget. Beardwood et al. (1959) derived the distance between m randomly dispersed destinations within an area to be asymptotically proportional to $\sqrt{m}$ for large m . Therefore, the cost of visiting the n_k and r_k units in the k^{th} stratum dispersed all over the stratum can be taken as $t_{ko}\sqrt{n_k}$ and $t_{k2}\sqrt{r_k}$ $k = 1, 2, ..., Q$, approximately. t_k is the cost of travelling around selected units in the k^{th} stratum. The quadratic cost function considered is,

$$C = \sum_{k=1}^{Q}\left(c_{k0}+c_{k1}W_{k1}\right)n_k + \sum_{k=1}^{Q} c_{k2}r_k + \sum_{k=1}^{Q} t_{k0}\sqrt{n_k} + \sum_{k=1}^{Q} t_{k2}\sqrt{r_k} \tag{9}$$

where t_{ko} is per unit traveling cost at first attempt, t_{k2} is travel cost for visiting the non-respondent unit within the k^{th} stratum and r_k are samples gotten from the non-respondent unit at second attempt.

If the labour cost, which represents the time taken to obtain data from all selected units, is significant, then the cost function considered by Abubakar and Okon (2018) is given as

$$C=\sum_{k=1}^{Q}\left(c_{k0}+c'_{k1}W_{k1}\right)n_k+\sum_{k=1}^{Q}c'_{k2}r_k+\sum_{k=1}^{Q}t_{k0}\sqrt{n_k}+\sum_{k=1}^{Q}t_{k2}\sqrt{r_k}+\alpha\sum_{k=1}^{Q}\frac{n_k}{\lambda}+\omega\sum_{k=1}^{Q}\frac{r_k}{\lambda^*} \tag{10}$$

where α and ω are the effect of labour cost at the first and second attempt, respectively, with rates λ and λ^*. $c'_{k1}=c_{k1}+u_{k1}$ and $c'_{k2}=c_{k2}+u_{k2}$ are the enumerating and processing unit cost with reward paid to respondents and non-respondents in k strata.

2.2 Weibull Cost Function with Non-Response

In survey enumeration, failure of device is inevitable. Survey devices have tendencies of wearing out with time. However in many complex systems or real life situations, failure rate exhibits a bathtub shape, it shows a form of decrease in the beginning, then it stabilizes and towards the end of its life cycle it enters the wear out phase, at this point the failure rate begins to rise. A constant failure rate means that the component's lifetime T will lead to an exponential distribution, with the rate parameter equal to the constant failure rate λ. The probability distribution function and cumulative distribution function is given as

$$\left.\begin{aligned} f(t)&=\lambda e^{-\lambda t}\\ F(t)&=1-e^{-\lambda t}\end{aligned}\right\} \tag{11}$$

For an irreparable component, the mean time to failure (MTTF) is equal to the expected lifetime

$$E(t)=\int_0^{\infty}tf(t)dt=\frac{1}{\lambda} \tag{12}$$

In some circumstances, the simplifying assumption is not suitable, particularly during the early stages and wear-out phases of a component's lifespan. In such cases, a Weibull distribution will be preferred. Now, if the time to failure for a device affecting sampling units follows a Weibull distribution with state space $t\geq 0$, then the probability distribution function for time to failure is approximately

$$f(t,\theta,\gamma)=\frac{\gamma}{\theta}\left(\frac{t}{\theta}\right)^{\gamma-1}\ell^{-\left(\frac{t}{\theta}\right)^{\gamma}}, \qquad t\geq 0 \tag{13}$$

with $\gamma>0$ as shape parameter and $\theta>0$ scale parameter for the distribution (Karuna and Khan 2017). The equation given in the Equation (15) is a two parameter Weibull distribution.

Proposition 1: If the time to failure is attributed to the sampling units, that is, the selected units of respondents at the first attempt and a more careful second attempt for the subsample of non-respondents, with rates θ_1^* and θ_2^* having an exponential distribution. The sum of independent identically distributed exponential random variables will follow Weibull distributions with parameters (n_k,θ_1^*) and (r_k,θ_2^*) respectively. Summing up to all strata, the Weibull functions will have parameters $\left(\sum_{k=1}^{Q}n_k,\theta_1^*\right)$ and $\left(\sum_{k=1}^{Q}r_k,\theta_2^*\right)$ respectively.

The proposed Weibull cost function with non-response is given as

$$C=\sum_{k=1}^{Q}\left(c_{k0}+c'_{k1}W_{k1}\right)n_k+\sum_{k=1}^{Q}c'_{k2}r_k+\sum_{k=1}^{Q}t_{k0}\sqrt{n_k}+\sum_{k=1}^{Q}t_{k2}\sqrt{r_k}+\sum_{k=1}^{Q}\beta_k\frac{n_k}{\lambda}+\sum_{k=1}^{Q}\beta_k^*\frac{r_k}{\lambda^*}$$
$$\varphi_k\frac{\sum_{k=1}^{Q}n_k}{\theta_1^*}\int_0^\infty\left(\frac{t_1}{\theta_1^*}\right)^{\sum_{k=1}^{Q}n_k-1}\ell^{-\left(\frac{t_1}{\theta_1^*}\right)^{\sum_{k=1}^{Q}n_h}}dt_1+\varphi_k^*\frac{\sum_{k=1}^{Q}r_k}{\theta_2^*}\int_0^\infty\left(\frac{t_2}{\theta_2^*}\right)^{\sum_{k=1}^{Q}r_k-1}\ell^{-\left(\frac{t_2}{\theta_2^*}\right)^{\sum_{k=1}^{Q}r_k}}dt_2 \tag{13}$$

β_k , β_k^* and φ_k , φ_k^* are the cost of a unit time of labour and failure for the respondent and non-respondents units, respectively. In Equation (13) the cost of a unit time of labour is taken over all strata. Let,

$$\left.\begin{aligned}\mathrm{K}_1&=\varphi_k\frac{\sum_{k=1}^{Q}n_k}{\theta_1^*}\int_0^\infty\left(\frac{t_1}{\theta_1^*}\right)^{\sum_{k=1}^{Q}n_k-1}\ell^{-\left(\frac{t_1}{\theta_1^*}\right)^{\sum_{k=1}^{Q}n_k}}dt_1\\ \mathrm{K}_2&=\varphi_k^*\frac{\sum_{k=1}^{Q}r_k}{\theta_2^*}\int_0^\infty\left(\frac{t_2}{\theta_2^*}\right)^{\sum_{k=1}^{Q}r_k-1}\ell^{-\left(\frac{t_2}{\theta_2^*}\right)^{\sum_{k=1}^{Q}r_k}}dt_2\end{aligned}\right\} \tag{14}$$

Then Equation (13) will become

$$C=\sum_{k=1}^{Q}\left(c_{k0}+c'_{k1}W_{k1}\right)n_k+\sum_{k=1}^{Q}c'_{k2}r_k+\sum_{k=1}^{Q}t_{k0}\sqrt{n_k}+\sum_{k=1}^{Q}t_{k2}\sqrt{r_k}+\sum_{k=1}^{Q}\beta_k\frac{n_k}{\lambda}+\sum_{k=1}^{Q}\beta_k^*\frac{r_k}{\lambda^*}+$$
$$\varphi_k\mathrm{K}_1+\varphi_k^*\mathrm{K}_2. \tag{15}$$

from Equation (13), let

$$\left.\begin{aligned}\sum_{k=1}^{Q}\varphi_k E(T_k) &= \sum_{k=1}^{Q}\varphi_k\left[\frac{n_k}{\theta_1^*}\int_0^{\infty}\left(\frac{t_1}{\theta_1^*}\right)^{n_k-1}\ell^{-\left(\frac{t_1}{\theta_1^*}\right)^{n_k}}dt_1\right]\\ \sum_{k=1}^{Q}\varphi_k^* E(T_k^*) &= \sum_{k=1}^{Q}\varphi_k^*\left[\frac{r_k}{\theta_2^*}\int_0^{\infty}\left(\frac{t_2}{\theta_2^*}\right)^{r_k-1}\ell^{-\left(\frac{t_2}{\theta_2^*}\right)^{r_k}}dt_2\right]\end{aligned}\right\} \quad (16)$$

Then the MTTF for Equation (16) will be

$$\left.\begin{aligned}\sum_{k=1}^{Q}\varphi_k E(T_k) &= \sum_{k=1}^{Q}\varphi_k\theta_1^*\Gamma\left(\frac{1}{n_k}+1\right)\\ \sum_{k=1}^{Q}\varphi_k^* E(T_k^*) &= \sum_{h=1}^{Q}\varphi_k^*\theta_2^*\Gamma\left(\frac{1}{r_k}+1\right)\end{aligned}\right\} \quad (17)$$

The cost function in Equation (13) can be rewritten as

$$\begin{aligned}C = &\sum_{k=1}^{Q}\left(c_{k0}+c_{k1}'W_{k1}\right)n_k+\sum_{k=1}^{Q}c_{k2}'r_k+\sum_{k=1}^{Q}t_{k0}\sqrt{n_k}+\sum_{k=1}^{Q}t_{k2}\sqrt{r_k}+\sum_{k=1}^{Q}\beta_k\frac{n_k}{\lambda}+\sum_{k=1}^{Q}\beta_k^*\frac{r_k}{\lambda^*}\\ &+\sum_{k=1}^{Q}\varphi_k\theta_1^*\Gamma\left(\frac{1}{n_k}+1\right)+\sum_{k=1}^{Q}\varphi_k^*\theta_2^*\Gamma\left(\frac{1}{r_k}+1\right)\end{aligned} \quad (18)$$

2.3 Stochastic Form of the Weibull Cost Function with Non-Response

Theorem 1: Assuming the stochastic vector $\tau = (\mathrm{c}_{k0}, c_{k1}', c_{k2}', t_{k0}, t_{k1}, \beta_k, \beta_k^*, \varphi_k, \varphi_k^*, C)$, where $h = 1, 2, ..., L$ and the function $g(n_k, r_k; \tau)$ has the form;

$$\begin{aligned}g(n_k,r_k;\tau) = &\sum_{k=1}^{Q}\left(c_{k0}+c_{k1}'W_{k1}\right)n_k+\sum_{k=1}^{Q}c_{k2}'r_k+\sum_{k=1}^{Q}t_{k0}\sqrt{n_k}+\sum_{k=1}^{Q}t_{k2}\sqrt{r_k}+\sum_{k=1}^{Q}\beta_k\frac{n_k}{\lambda}+\sum_{k=1}^{Q}\beta_k^*\frac{r_k}{\lambda^*}\\ &+\sum_{k=1}^{Q}\varphi_k\theta_1^*\Gamma\left(\frac{1}{n_k}+1\right)+\sum_{k=1}^{Q}\varphi_k^*\theta_2^*\Gamma\left(\frac{1}{r_k}+1\right)-C\end{aligned} \quad (19)$$

If the stochastic vector contains cost that are normally distributed and independent random variables then, $P\left\{g(n_k, r_k; \tau) \le C\right\} \ge p_0$ if and only if $E(g(n_k, r_k; \tau)) + G_\alpha\sqrt{V(g(n_k, r_k; \tau))} \le C$ where p_0 , $0 \le p_0 \le 1$ is a specified probability .

Proof: If $c_{k0} \sim N(\mu_{c_{k0}}, \sigma^2_{c_{k0}})$, $c'_{k1} \sim N(\mu_{c'_{k1}}, \sigma^2_{c'_{k1}})$, $c'_{k2} \sim N(\mu_{c'_{k2}}, \sigma^2_{c'_{k2}})$, $t_{k0} \sim N(\mu_{t_{k0}}, \sigma^2_{t_{k2}})$, $\beta_k \sim N(\mu_{\beta_k}, \sigma^2_{\beta_k})$, $\beta^*_k \sim N(\mu_{\beta^*_k}, \sigma^2_{\beta^*_k})$, $\varphi_k \sim N(\mu_{\varphi_k}, \sigma^2_{\varphi_k})$, and $\varphi^*_k \sim N(\mu_{\varphi^*_k}, \sigma^2_{\varphi^*_k})$

The cost constraint can be written as

$$p\left[\mathrm{g}(n_k, r_k; \tau) \le C\right] \ge p_0 \tag{20}$$

$$\Rightarrow p\left[\frac{\mathrm{g}(n_k, r_k; \tau) - E(\mathrm{g}(n_k, r_k; \tau))}{\sqrt{V(\mathrm{g}(n_k, r_k; \tau))}} \le \frac{C - E(\mathrm{g}(n_k, r_k; \tau))}{\sqrt{V(\mathrm{g}(n_k, r_k; \tau))}}\right] \ge p_0 \tag{21}$$

Equation (21) is a the standard normal variate with $Z \sim N(0,1)$. The standard normal variate is evaluated at z, with $\Phi(z)$ as cumulative density function. If the standard normal variate at which $\Phi(G_\alpha) = p_0$, is represented as G_α. We can write the constraints as

$$\Phi\left[\frac{C - E(\mathrm{g}(n_k, r_k; \tau))}{\sqrt{V(\mathrm{g}(n_k, r_k; \tau))}}\right] \ge \Phi(G_\alpha) \tag{22}$$

Φ is the standardize normal distribution function. The inequality will be satisfied if and only if

$$\left[\frac{C - E(\mathrm{g}(n_k, r_k; \tau))}{\sqrt{V(\mathrm{g}(n_k, r_k; \tau))}}\right] \ge G_\alpha \tag{23}$$

This also implies that

$$\left.\begin{aligned} E\left(\mathrm{g}(n_k, r_k; \tau)\right) = {} & E\left(\sum_{k=1}^{Q} c_{k0} n_k\right) + E\left(\sum_{k=1}^{Q} c'_{k1} W_{k1} n_k\right) + E\left(\sum_{k=1}^{Q} c'_{k2} r_k\right) + \\ & E\left(\sum_{k=1}^{Q} t_{k0}\sqrt{n_k}\right) + E\left(\sum_{k=1}^{Q} t_{k2}\sqrt{r_k}\right) + E\left(\sum_{k=1}^{Q} \beta_k \frac{n_2}{\lambda}\right) + \\ & E\left(\sum_{k=1}^{Q} \beta^*_k \frac{r_k}{\lambda^*}\right) + E\left[\sum_{k=1}^{Q} \varphi_k \theta^*_1 \Gamma\left(\frac{1}{n_k} + 1\right)\right] + E\left[\sum_{k=1}^{Q} \varphi^*_k \theta^*_2 \Gamma\left(\frac{1}{r_k} + 1\right)\right] \end{aligned}\right\} \tag{24}$$

with variance obtained as

$$= \sum_{k=1}^{Q} \begin{pmatrix} n_k \mu_{c_{k0}} + W_{k1} n_k \mu_{c'_{k1}} + r_k \mu_{c'_{k2}} + \mu_{t_{k0}} \sqrt{n_k} + \mu_{t_{k2}} \sqrt{r_k} \\ + \frac{n_k}{\lambda} \mu_{\beta_k} + \frac{r_k}{\lambda^*} \mu_{\beta^*_k} + \theta^*_1 \Gamma\left(\frac{1}{n_k} + 1\right) \mu_{\varphi_k} + \theta^*_2 \Gamma\left(\frac{1}{r_k} + 1\right) \mu_{\varphi^*_k} \end{pmatrix} \tag{25}$$

$$V\left(\mathrm{g}(n_k,r_k;\tau)\right)=\sum_{k=1}^{Q}n_k^2\sigma_{c_{k0}}^2+\sum_{k=1}^{Q}n_k^2W_{k1}^2\sigma_{c'_{k1}}^2+\sum_{k=1}^{Q}r_k^2\sigma_{c'_{k2}}^2+\sum_{k=1}^{Q}n_k\sigma_{t_{k0}}^2+\sum_{k=1}^{Q}r_k\sigma_{t_{k2}}^2+$$
$$\sum_{k=1}^{Q}\frac{n_k^2}{\lambda^2}\sigma_{\beta_k}^2+\sum_{k=1}^{Q}\frac{r_k^2}{\lambda^{*2}}\sigma_{\beta_k^*}^2+\theta_1^{2*}\Gamma\left(\frac{1}{n_k}+1\right)^2\sigma_{\varphi_k}^2+\theta_2^{2*}\Gamma\left(\frac{1}{r_k}+1\right)^2\sigma_{\varphi_k^*}^2 \tag{26}$$

we can write Equation (23) as

$$E(g(n_k,r_k;\tau))+G_\alpha\sqrt{V(g(n_k,r_k;\tau))}\le C \tag{27}$$

Substituting we will obtain the Weibull cost function with nonresponse in its stochastic form as

$$\sum_{k=1}^{Q}\begin{pmatrix}n_k\mu_{c_{k0}}+W_{k1}n_k\mu_{c'_{k1}}+r_k\mu_{c'_{k2}}+\mu_{t_{k0}}\sqrt{n_k}+\mu_{t_{k2}}\sqrt{r_k}\\ +\frac{n_k}{\lambda}\mu_{\beta_k}+\frac{r_k}{\lambda^*}\mu_{\beta_k^*}+\theta_1^*\Gamma\left(\frac{1}{n_k}+1\right)\mu_{\varphi_k}+\theta_2^*\Gamma\left(\frac{1}{r_k}+1\right)\mu_{\varphi_k^*}\end{pmatrix}$$
$$+G_\alpha\sqrt{\begin{matrix}\sum_{k=1}^{Q}n_k^2\sigma_{c_{k0}}^2+\sum_{k=1}^{Q}n_k^2W_{k1}^2\sigma_{c'_{k1}}^2+\sum_{k=1}^{Q}r_h^2\sigma_{c'_{k2}}^2+\sum_{k=1}^{Q}n_k\sigma_{t_{k0}}^2+\sum_{k=1}^{Q}r_k\sigma_{t_{k2}}^2+\\ \sum_{k=1}^{Q}\frac{n_k^2}{\lambda^2}\sigma_{\beta_k}^2+\sum_{k=1}^{Q}\frac{r_k^2}{\lambda^{*2}}\sigma_{\beta_k^*}^2+\theta_1^{2*}\Gamma\left(\frac{1}{n_k}+1\right)^2\sigma_{\varphi_k}^2+\theta_2^{2*}\Gamma\left(\frac{1}{r_k}+1\right)^2\sigma_{\varphi_k^*}^2\end{matrix}}\le C \tag{28}$$

2.4 Equivalent Deterministic Weibull Cost Constraint

Let X be a feasible space of a multi-objective stochastic optimization problem (MOSOP). If the mean and variance constants are unknown it can be replaced with the estimator of mean $\hat{E}(.)$ and $\hat{V}(.)$ as given below

$$X=\begin{cases}n_k,r_k\in R\,|\sum_{k=1}^{Q}\left(\bar{c}_{k0}+\bar{c}'_{k1}W_{k1}\right)n_h+\sum_{k=1}^{Q}\bar{c}'_{k2}r_k+\sum_{k=1}^{Q}\bar{t}_{k0}\sqrt{n_k}+\sum_{k=1}^{Q}\bar{t}_{k2}\sqrt{r_k}+\sum_{k=1}^{Q}\bar{\beta}_h\frac{n_h}{\lambda}+\sum_{k=1}^{Q}\bar{\beta}_h^*\frac{r_h}{\lambda^*}+\\ \qquad\sum_{k=1}^{Q}\bar{\varphi}_k\theta_1^*\Gamma\left(\frac{1}{n_k}+1\right)+\sum_{k=1}^{Q}\bar{\varphi}_k^*\theta_2^*\Gamma\left(\frac{1}{r_k}+1\right)+\\ \qquad G_\alpha\sqrt{\begin{matrix}\sum_{k=1}^{Q}n_k^2\hat{\sigma}_{c_{k0}}^2+\sum_{k=1}^{Q}n_k^2W_{k1}^2\hat{\sigma}_{c'_{k1}}^2+\sum_{k=1}^{Q}r_k^2\hat{\sigma}_{c'_{k2}}^2+\sum_{k=1}^{Q}n_k\hat{\sigma}_{t_{k0}}^2+\sum_{k=1}^{Q}r_k\hat{\sigma}_{t_{k2}}^2+\\ \sum_{k=1}^{Q}\frac{n_k^2}{\lambda^2}\hat{\sigma}_{\beta_k}^2+\sum_{k=1}^{Q}\frac{r_k^2}{\lambda^{*2}}\hat{\sigma}_{\beta_k^*}^2+\theta_1^{2*}\Gamma\left(\frac{1}{n_k}+1\right)\hat{\sigma}_{\varphi_k}^2+\theta_2^{2*}\Gamma\left(\frac{1}{r_h}+1\right)\hat{\sigma}_{\varphi_h^*}^2\end{matrix}}\\ for\ k=1,2...,Q,\ 2\le n_k\le N_k,\ 2\le r_k\le\hat{n}_{k2}\end{cases}$$

3. Modified Expected Value (Modified-E) Model for the Stochastic Sampling Variance

s_{jk}^2 and s_{jk2}^2 are sample variances which are assumed to be random variables. Until a second attempt is made, the sample size n_{k2} of the non-respondent is not known, but its estimates $\hat{n}_{k2} = W_{k1} n_k$, can be used in this regards.

Consider the random variable ξ_{k2} defined by

$$\xi_{k2} = \frac{1}{\hat{n}_{k2} - 1} \sum_{i=1}^{\hat{n}_{k2}} (y_{jki} - \bar{Y}_{jk2})^2, \tag{29}$$

which is the limiting distribution of the sample variances. Having asymptotic normal distribution with mean given as

$$E[\xi_{k2}] = \frac{1}{\hat{n}_{k2} - 1} \sum_{i=1}^{\hat{n}_{k2}} E(y_{jki} - \bar{Y}_{jk2})^2 = \frac{\hat{n}_{k2}}{\hat{n}_{k2} - 1} S_{jk2}^2 \tag{30}$$

Similarly by independence,

$$\text{var}\left(\xi_{k2}\right) = \frac{\hat{n}_{k2}}{\left(\hat{n}_{k2} - 1\right)^2} \text{var}\left(y_{ijk} - \bar{Y}_{jk2}\right) \tag{31}$$

$$= \frac{\hat{n}_{k2}}{\left(\hat{n}_{k2} - 1\right)^2} \left[C_{jk2}^4 - \left(S_{jk2}^2\right)^2 \right] \tag{32}$$

fourth central moment given as C_{jk2}^4 in Equation (32) above can be evaluated as

$$C_{jk2}^4 = \frac{1}{\hat{N}_{k2}} \sum_{k=1}^{\hat{N}_{k2}} \left(y_{ijk} - \bar{Y}_{jk2}\right)^4 \tag{33}$$

If we observe closely, we can see that

$$s_{jk2}^2 = \xi_k - \frac{\hat{n}_{k2}}{\hat{n}_{k2} - 1} \left(\bar{y}_{jk2} - \bar{Y}_{jk2}\right) \tag{34}$$

where $\frac{\hat{n}_{k2}}{\hat{n}_{k2}-1} \xrightarrow{p} 1$ and $\left(\bar{y}_{jk2}-\bar{Y}_{jk2}\right) \xrightarrow{p} 0$. Then, the sample variances s^2_{jk2} have an asymptotical normal distribution, that is, $s^2_{jk2} \xrightarrow{a}$ $N(E(\xi_k), \text{var}(\xi_k))$. Let,

$$V(\bar{y}_{j(w)}) = \sum_{k=1}^{Q} W_k^2 \left[\frac{s_{jk}^2}{n_k} + \frac{W_{k2}^2 s_{jk2}^2}{r_k} - \frac{W_{k2} s_{jk2}^2}{n_k} \right] \tag{35}$$

The variance function in Equation (35) has a normal distribution, with

$$E\left[\hat{V}(\bar{y}_{j(w)})\right] = \sum_{k=1}^{Q} \left(\frac{W_k^2 s_{yjk}^2}{n_k} + \frac{W_k^2 W_{k2}^2 s_{yjk2}^2}{r_k} - \frac{W_k^2 W_{k2} s_{yjk2}^2}{n_k} \right) \tag{36}$$

$$= \sum_{k=1}^{Q} \left(\frac{W_k^2 E(\xi_k)}{n_k} + \frac{W_k^2 W_{k2}^2 E(\xi_{k2})}{r_k} - \frac{W_k^2 W_{k2} E(\xi_{k2})}{n_k} \right) \tag{37}$$

$$= \sum_{k=1}^{Q} \left(\frac{W_k^2 S_{jk}^2}{(n_k - 1)} + \frac{\hat{n}_{k2} W_k^2 W_{k2}^2 S_{jk2}^2}{r_k (\hat{n}_{k2} - 1)} - \frac{\hat{n}_{k2} W_k^2 W_{k2} S_{jk2}^2}{n_k (\hat{n}_{k2} - 1)} \right) \tag{38}$$

$$E\left[\hat{V}(\bar{y}_{j(w)})\right] = \sum_{k=1}^{Q} \left(\frac{W_k^2 S_{jk}^2}{(n_k - 1)} + \frac{n_k W_k^2 W_{k2}^3 S_{jk2}^2}{r_k (n_k W_{k2} - 1)} - \frac{W_{k2}^2 W_k^2 S_{jk2}^2}{n_k W_{k2} - 1} \right) \tag{39}$$

with,

$$V\left[\hat{V}(\bar{y}_{j(w)})\right] = \sum_{k=1}^{Q} V \left(\frac{W_k^2 s_{jk}^2}{n_k} + \frac{W_k^2 W_{k2}^2 s_{jk2}^2}{r_k} - \frac{W_k^2 W_{k2} s_{jk2}^2}{n_k} \right) \tag{40}$$

$$= \sum_{k=1}^{Q} \left(\frac{W_k^4 V(\xi_k)}{n_k^2} + \frac{W_k^4 W_{k2}^4 V(\xi_{k2})}{r_k^2} - \frac{W_k^4 W_{k2}^2 V(\xi_{k2})}{n_k^2} \right) \tag{41}$$

$$= \sum_{k=1}^{Q} \left(\frac{W_k^4 \left(C_{jk2}^4 - \left(S_{jk2}^2\right)^2 \right)}{n_k (n_k - 1)^2} + \frac{\hat{n}_{k2} W_k^4 W_{k2}^4 \left(C_{jk2}^4 - \left(S_{jk2}^2\right)^2 \right)}{r_k^2 (\hat{n}_{k2} - 1)^2} - \frac{\hat{n}_{k2} W_k^4 W_{k2}^2 \left(C_{jk2}^4 - \left(S_{jk2}^2\right)^2 \right)}{n_k^2 (\hat{n}_{k2} - 1)^2} \right)$$

(42)

$$=\sum_{k=1}^{Q}\left(\frac{W_k^4\left(C_{jk}^4-\left(S_{jk}^2\right)^2\right)}{n_k\left(n_k-1\right)^2}+\frac{n_kW_k^4W_{k2}^5\left(C_{jk2}^4-\left(S_{jk2}^2\right)^2\right)}{r_k^2(n_kW_{k2}-1)^2}-\frac{W_k^4W_{k2}^3\left(C_{jk2}^4-\left(S_{jk2}^2\right)^2\right)}{n_k(n_kW_{k2}-1)^2}\right)$$

(43)

The objective function in the E-model form is written as;

$$f_j(n_k,r_k)=e_1E(\hat{V}(\bar{y}_{j,(w)}))+e_2\sqrt{V(\hat{V}(\bar{y}_{j,(w)}))} \quad (44)$$

Here $e_1, e_2>0$. Roa (2009) suggested that $e_1+e_2=1$. It illustrate the relative significance of the expectation and variance of $V(\bar{y}_{j(w)})$.

Inserting Equation (39) and Equation (43) into Equation (44), we will obtain

$$f_j(n_k,r_k)=e_1\sum_{k=1}^{Q}\left(\frac{W_k^2S_{jk}^2}{(n_k-1)}+\frac{n_kW_k^2W_{k2}^3S_{jk2}^2}{r_k(n_kW_{k2}-1)}-\frac{W_{k2}^2W_k^2S_{jk2}^2}{n_kW_{k2}-1}\right)$$
$$+e_2\sqrt{\sum_{k=1}^{Q}\left(\frac{W_k^4\left(C_{jk}^4-\left(S_{jk}^2\right)^2\right)}{n_k\left(n_k-1\right)^2}+\frac{n_kW_k^4W_{k2}^5\left(C_{jk2}^4-\left(S_{jk2}^2\right)^2\right)}{r_k^2(n_kW_{k2}-1)^2}-\frac{W_k^4W_{k2}^3\left(C_{jk2}^4-\left(S_{jk2}^2\right)^2\right)}{n_k(n_kW_{k2}-1)^2}\right)}$$

(45)

4. Stochastic Individual Allocations

The individual allocation for each characteristic as;

$$\left.\begin{array}{l}
Minimize\ \ f_j(n_{jk}, r_{jk}) = \mathrm{Z}_j = e_1 \sum_{k=1}^{Q}\left(\frac{W_k^2 S_{jk}^2}{\left(n_{jk}-1\right)} + \frac{n_{jk} W_k^2 W_{k2}^3 S_{jk2}^2}{r_{jk}(n_{jk}W_{k2}-1)} - \frac{W_{k2}^2 W_k^2 S_{jk2}^2}{n_{jk}W_{k2}-1}\right) \\
+e_2\sqrt{\sum_{k=1}^{Q}\left(\frac{W_k^4\left(C_{jk}^4-\left(S_{jk}^2\right)^2\right)}{n_{jk}\left(n_{jk}-1\right)^2} + \frac{n_{jk}W_k^4 W_{k2}^5\left(C_{jk2}^4-\left(S_{jk2}^2\right)^2\right)}{r_{jk}^2(n_{jk}W_{k2}-1)^2} - \frac{W_k^4 W_{k2}^3\left(C_{jk2}^4-\left(S_{jk2}^2\right)^2\right)}{n_{jk}(n_{jk}W_{k2}-1)^2}\right)} \\
subject\ to \\
\sum_{k=1}^{Q}\left(\bar{c}_{k0}+\bar{c}'_{k1}W_{k1}\right)n_{jk} + \sum_{k=1}^{Q}\bar{c}'_{k2}r_{jk} + \sum_{k=1}^{Q}\bar{t}_{k0}\sqrt{n_{jk}} + \sum_{k=1}^{Q}\bar{t}_{k2}\sqrt{r_{jk}} + \sum_{k=1}^{Q}\bar{\beta}_k\frac{n_{jk}}{\lambda} + \\
\sum_{k=1}^{Q}\bar{\beta}_k^*\frac{r_{jh}}{\lambda^*} + \sum_{k=1}^{Q}\bar{\varphi}_k\theta_1^*\Gamma\left(\frac{1}{n_{jk}}+1\right) + \sum_{k=1}^{Q}\bar{\varphi}_k^*\theta_2^*\Gamma\left(\frac{1}{r_{jk}}+1\right) + \\
G_\alpha\sqrt{\begin{array}{l}\sum_{k=1}^{Q}n_{jk}^2\hat{\sigma}_{c_{k0}}^2 + \sum_{k=1}^{Q}n_{jk}^2W_{k1}^2\hat{\sigma}_{c'_{k1}}^2 + \sum_{k=1}^{Q}r_{jk}^2\hat{\sigma}_{c'_{k2}}^2 + \sum_{k=1}^{Q}n_{jk}\hat{\sigma}_{t_{k0}}^2 + \sum_{k=1}^{Q}r_{jk}\hat{\sigma}_{t_{k2}}^2 + \\
\sum_{k=1}^{Q}\frac{n_{jk}^2}{\lambda^2}\hat{\sigma}_{\beta_k}^2 + \sum_{k=1}^{Q}\frac{r_{jk}^2}{\lambda^{*2}}\hat{\sigma}_{\beta_k^*}^2 + \theta_1^{2*}\Gamma\left(\frac{1}{n_{jk}}+1\right)\hat{\sigma}_{\varphi_k}^2 + \theta_2^{2*}\Gamma\left(\frac{1}{r_{jk}}+1\right)\hat{\sigma}_{\varphi_k^*}^2\end{array}} \le C \\
2 \le n_{jk} \le N_k,\ 2 \le r_{jk} \le \hat{n}_{k2},\ n_{jk} \text{ and } r_{jk}\ \in \square\ \ \forall\ \ k = 1,2,\ldots,Q
\end{array}\right\}$$

(46)

The stochastic individual allocations for j^{th} $(j = 1,2,\ldots,p)$ characteristics can be obtained by solving the above MO-SINLPP. Z_j^* will be the objective or variance function values of Z_j under the individual allocations.

5. Compromise Solution with Fuzzy Programming

Consider $\bar{\xi}_j$ to be the solutions obtained individually for the problem in Equation (46), that is, $\bar{\xi}_j = (n_1,\ldots,n_Q,r_1,\ldots,r_Q)$. Let $\bar{\xi}_j^B$ be the best solution obtained individually for the j^{th} characteristics subject to the cost constraints.

$$Z_j^B = Z(\bar{\xi}_j^B) = \min Z_j(\bar{\xi}_j);\ j = 1,2,\ldots,p$$

The corresponding fuzzy goal is of the form

$$Z_j(\bar{\xi}_j) \ge \mathrm{Z}_j^B;\ j = 1,2,\ldots,p$$

Next, is by constructing a payoff matrix using the individual solution which is best with the variance values at each best solution.

$$\begin{array}{c} \\ \bar{\xi}_1 \\ \bar{\xi}_2 \\ \vdots \\ \bar{\xi}_p \end{array} \begin{array}{c} \begin{array}{cccc} Z_1(\bar{\xi}) & Z_2(\bar{\xi}) & \cdots & Z_p(\bar{\xi}) \end{array} \\ \begin{pmatrix} Z_1(\bar{\xi}_1^B) & Z_2(\bar{\xi}_1) & \ldots & Z_p(\bar{\xi}_1) \\ Z_1(\bar{\xi}_2) & Z_2(\bar{\xi}_2^B) & \ldots & Z_p(\bar{\xi}_2) \\ \vdots & \vdots & \ddots & \vdots \\ Z_1(\bar{\xi}_p) & Z_2(\bar{\xi}_p) & \ldots & Z_p(\bar{\xi}_p) \end{pmatrix} \end{array}$$

The minimum value of each column of $Z_j(\bar{\xi})$ ($j=1,2,...,p$) gives the lower tolerance limit and the maximum value gives the least acceptable level of achievement. Let Z_j^B and Z_j^L represent the upper and the lower tolerance limit respectively. Now, let $d_j = Z_j^L - Z_j^B$.

Then the solution method using fuzzy programming with MO-SINLPP in Equation (46) is given a

$$\left.\begin{array}{l}
\textit{Minimize } \rho \\
\textit{subject to} \\
e_1 \sum_{k=1}^{Q} \left(\frac{W_k^2 S_{jk}^2}{(n_{kca}-1)} + \frac{n_{kca} W_k^2 W_{k2}^3 S_{jk2}^2}{r_{kca}(n_{kca} W_{k2}-1)} - \frac{W_{k2}^2 W_k^2 S_{jk2}^2}{n_{kca} W_{k2}-1} \right) + \\
e_2 \sqrt{\sum_{k=1}^{Q} \left(\frac{W_k^4 \left(C_{jk}^4 - \left(S_{jk}^2\right)^2 \right)}{n_{kca}\left(n_{kca}-1\right)^2} + \frac{n_{kca} W_k^4 W_{k2}^5 \left(C_{jk2}^4 - \left(S_{jk2}^2\right)^2 \right)}{r_{kca}^2 (n_{kca} W_{k2}-1)^2} - \frac{W_k^4 W_{k2}^3 \left(C_{jk2}^4 - \left(S_{jk2}^2\right)^2 \right)}{n_{kca}(n_{kca} W_{k2}-1)^2} \right)} - \rho d_j \le Z_j^B \\
\sum_{k=1}^{Q} \left(\bar{c}_{k0} + \bar{c}'_{k1} W_{k1} \right) n_{kca} + \sum_{k=1}^{Q} \bar{c}'_{k2} r_{kca} + \sum_{k=1}^{Q} \bar{t}_{k0} \sqrt{n_{kca}} + \sum_{k=1}^{Q} \bar{t}_{k2} \sqrt{r_{kca}} + \sum_{k=1}^{Q} \bar{\beta}_k \frac{n_{kca}}{\lambda} + \\
\sum_{k=1}^{Q} \bar{\beta}_k^* \frac{r_{kca}}{\lambda^*} + \sum_{k=1}^{Q} \bar{\varphi}_k \theta_1^* \Gamma\left(\frac{1}{n_{kca}} + 1 \right) + \sum_{k=1}^{Q} \bar{\varphi}_k^* \theta_2^* \Gamma\left(\frac{1}{r_{kca}} + 1 \right) + \\
G_\alpha \sqrt{\begin{array}{l} \sum_{k=1}^{Q} n_{kca}^2 \hat{\sigma}_{c_{k0}}^2 + \sum_{k=1}^{Q} n_{kca}^2 W_{k1}^2 \hat{\sigma}_{c'_{k1}}^2 + \sum_{k=1}^{Q} r_{kca}^2 \hat{\sigma}_{c'_{k2}}^2 + \sum_{k=1}^{Q} n_{kca} \hat{\sigma}_{t_{k0}}^2 + \sum_{k=1}^{Q} r_{kca} \hat{\sigma}_{t_{k2}}^2 + \\ \sum_{k=1}^{Q} \frac{n_{kca}^2}{\lambda^2} \hat{\sigma}_{\beta_k}^2 + \sum_{k=1}^{Q} \frac{r_{kca}^2}{\lambda^{*2}} \hat{\sigma}_{\beta_k^*}^2 + \theta_1^{2*} \Gamma\left(\frac{1}{n_{kca}} + 1 \right) \hat{\sigma}_{\varphi_k}^2 + \theta_2^{2*} \Gamma\left(\frac{1}{r_{kca}} + 1 \right) \hat{\sigma}_{\varphi_k^*}^2 \end{array}} \le C \\
\rho \ge 0;\ 2 \le n_{kca} \le N_k,\ 2 \le r_{kca} \le \hat{n}_{k2},\ n_{kca} \text{ and } r_{kca} \in \mathbb{Z}\ \forall\ k = 1,2,...,Q
\end{array}\right\}$$

(47)

The worst deviation level is represented by the decision variable ρ. "ca" represents compromise allocation.

6. D_1 -Distance Technique

A pre-emptive-based method for solving multi-objective optimisation problems by ranking the goals according to the order of importance is often called Lexicographic Goal Programming (LGP). In certain situations, solutions that are considered best are chosen over better pre-emptive structures. D_1 distance technique chooses a proper pre-emptive structure which in turn provides are better.

If we consider p characteristic variances, we can rank or form a hierarchy of goals for each characteristic variance. If preference be given to the variance function representing the first characteristics, we may obtain a compromise solution in the form, in some sense as $n_{kca}^{(1)} = n_{1ca}^{(1)}, n_{2ca}^{(1)}, n_{3ca}^{(1)}, ..., n_{Qca}^{(1)}$. Correspondingly, we produce results for all the lexicographic goal programming problems with priority structures and determine the optimal solution for each one. The solutions that correspond to the second character variance will be $n_{kca}^{(2)} = n_{1ca}^{(2)}, n_{2ca}^{(2)}, n_{3ca}^{(2)}, ..., n_{Qca}^{(2)}$ if we do this continuously we will obtain $n_{kca}^{(R)} = n_{1ca}^{(R)}, n_{2ca}^{(R)}, n_{3ca}^{(R)}, ..., n_{Qca}^{(R)}$.

If we consider the above solution, the ideal solution, which corresponds to different pre-emptive structures, are clearly stated Table 1. The solution can be described as

$$n_{kca}^{*} = \left\{\max\left(n_{1ca}^{(1)}, n_{1ca}^{(2)}, ..., n_{1ca}^{(R)}\right), \max\left(n_{2ca}^{(1)}, n_{2ca}^{(2)}, ..., n_{2ca}^{(R)}\right), ..., \max\left(n_{Qca}^{(R)}, n_{Qca}^{(R)}\right), ..., n_{Qca}^{(R)}\right\} = n_{1ca}^{*}, n_{2ca}^{*}, ..., n_{Qca}^{*}$$

and

$$r_{kca}^{*} = \left\{\max\left(r_{1ca}^{(1)}, r_{1ca}^{(2)}, ..., r_{1ca}^{(R)}\right), \max\left(r_{2ca}^{(1)}, r_{2ca}^{(2)}, ..., r_{2ca}^{(R)}\right), ..., \max\left(r_{Qca}^{(R)}, r_{Qca}^{(R)}\right), ..., r_{Qca}^{(R)}\right\} = r_{1}^{*}, r_{2}^{*}, ..., r_{Q}^{*}$$

Ideal solutions are not obtainable in practice. The best compromise solution is chosen from the solution with the shortest distance to the ideal solution. The pre-emptive pattern that corresponds is established as the best pre-emptive pattern from the pre-emptive patterns considered.

Table 1: Computation of Ideal Solution

Pre-emptive Structure	n_{1ca}	n_{2ca}	$\cdots$	n_{Qca}	r_{1ca}	$\cdots$	r_{Qca}

$Z^{(1)}$	$n_{1ca}^{(1)}$	$n_{2ca}^{(1)}$	$\cdots$	$n_{Qca}^{(1)}$	$r_{1ca}^{(1)}$	$\cdots$	$r_{Qca}^{(1)}$
$Z^{(2)}$	$n_{1ca}^{(2)}$	$n_{2ca}^{(2)}$	$\cdots$	$n_{Qca}^{(2)}$	$r_{1ca}^{(2)}$	$\cdots$	$r_{Qca}^{(2)}$
$\vdots$	$\vdots$	$\vdots$	$\cdots$		$\vdots$	$\cdots$	
$Z^{(R)}$	$n_{1ca}^{(R)}$	$n_{2ca}^{(R)}$	$\cdots$	$n_{Qca}^{(R)}$	$r_{1ca}^{(R)}$	$\cdots$	$r_{Qca}^{(R)}$
Ideal Solution	n_{1ca}^{*}	n_{2ca}^{*}	$\cdots$	n_{Qca}^{*}	r_{1ca}^{*}	$\cdots$	r_{Qca}^{*}

Table 2: Distance Calculation

	$n_{1ca}\ldots n_{Qca}$	$r_{1ca}\ldots r_{Qca}$	$(\mathrm{D}_1)^r$
$Z^{(1)}$	$\left\lvert n_{1ca}^{*}-n_{1ca}^{(1)}\right\rvert\ldots\left\lvert n_{Qca}^{*}-n_{Qca}^{(1)}\right\rvert$	$\left\lvert r_{1ca}^{*}-r_{1ca}^{(1)}\right\rvert\ldots\left\lvert r_{Qca}^{*}-r_{Qca}^{(1)}\right\rvert$	$\sum_{k=1}^{Q}\left\lvert\left(n_{kca}^{*}-n_{kca}^{(1)}\right)+\left(r_{1ca}^{*}-r_{1ca}^{(1)}\right)\right\rvert$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$Z^{(r)}$	$\left\lvert n_{1ca}^{*}-n_{1ca}^{(r)}\right\rvert\ldots\left\lvert n_{Qca}^{*}-n_{Qca}^{(r)}\right\rvert$	$\left\lvert r_{1ca}^{*}-r_{1ca}^{(r)}\right\rvert\ldots\left\lvert r_{Qca}^{*}-r_{Qca}^{(r)}\right\rvert$	$\sum_{k=1}^{Q}\left\lvert\left(n_{kca}^{*}-n_{kca}^{(r)}\right)+\left(r_{1ca}^{*}-r_{1ca}^{(r)}\right)\right\rvert$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$Z^{(R)}$	$\left\lvert n_{1ca}^{*}-n_{1ca}^{(R)}\right\rvert\ldots\left\lvert n_{Qca}^{*}-n_{Qca}^{(R)}\right\rvert$	$\left\lvert r_{1ca}^{*}-r_{1ca}^{(R)}\right\rvert\ldots\left\lvert r_{Qca}^{*}-r_{Qca}^{(R)}\right\rvert$	$\sum_{k=1}^{Q}\left\lvert\left(n_{kca}^{*}-n_{kca}^{(R)}\right)+\left(r_{1ca}^{*}-r_{1ca}^{(R)}\right)\right\rvert$

The goal programming technique defined below will be used to obtain the compromise allocation

$$
\left.\begin{array}{l}
\min_{1\le r\le R} \sum_{h=1}^{L} (d_{kr}^{+}, d_{kr}^{-}) \\
\textit{subject to} \\
\qquad e_1 \sum_{k=1}^{Q} \left(\frac{W_k^2 S_{jk}^2}{(n_{kca}-1)} + \frac{n_{kca} W_k^2 W_{k2}^3 S_{jk2}^2}{r_{kca}(n_{kca}W_{k2}-1)} - \frac{W_{k2}^2 W_k^2 S_{jk2}^2}{n_{kca}W_{k2}-1} \right) + \\
e_2 \sqrt{ \sum_{k=1}^{Q} \left(\frac{W_k^4 \left(C_{jk}^4 - \left(S_{jk}^2\right)^2 \right)}{n_{kca}\left(n_{kca}-1\right)^2} + \frac{n_{kca} W_k^4 W_{k2}^5 \left(C_{jk2}^4 - \left(S_{jk2}^2\right)^2 \right)}{r_{kca}^2 (n_{kca}W_{k2}-1)^2} - \frac{W_k^4 W_{k2}^3 \left(C_{jk2}^4 - \left(S_{jk2}^2\right)^2 \right)}{n_{kca}(n_{kca}W_{k2}-1)^2} \right) } + d_{kr}^{-} - d_{kr}^{+} \le Z_j^{(R)} \\
\sum_{k=1}^{Q} \left(\bar{c}_{k0} + \bar{c}'_{k1} W_{k1} \right) n_{kca} + \sum_{k=1}^{Q} \bar{c}'_{k2} r_{kca} + \sum_{k=1}^{Q} \bar{t}_{k0} \sqrt{n_{kca}} + \sum_{k=1}^{Q} \bar{t}_{k2} \sqrt{r_{kca}} + \sum_{k=1}^{Q} \bar{\beta}_k \frac{n_{kca}}{\lambda} + \\
\qquad \sum_{k=1}^{Q} \bar{\beta}_k^{*} \frac{r_{kca}}{\lambda^{*}} + \sum_{k=1}^{Q} \bar{\varphi}_k \theta_1^{*} \Gamma\left(\frac{1}{n_{kca}} + 1 \right) + \sum_{k=1}^{Q} \bar{\varphi}_k^{*} \theta_2^{*} \Gamma\left(\frac{1}{r_{kca}} + 1 \right) + \\
G_\alpha \sqrt{ \begin{array}{l} \sum_{k=1}^{Q} n_{kca}^2 \hat{\sigma}_{c_{k0}}^2 + \sum_{k=1}^{Q} n_{kca}^2 W_{k1}^2 \hat{\sigma}_{c'_{k1}}^2 + \sum_{k=1}^{Q} r_{kca}^2 \hat{\sigma}_{c'_{k2}}^2 + \sum_{k=1}^{Q} n_{kca} \hat{\sigma}_{t_{k0}}^2 + \sum_{k=1}^{Q} r_{kca} \hat{\sigma}_{t_{k2}}^2 + \\ \sum_{k=1}^{Q} \frac{n_{kca}^2}{\lambda^2} \hat{\sigma}_{\beta_k}^2 + \sum_{k=1}^{Q} \frac{r_{kca}^2}{\lambda^{*2}} \hat{\sigma}_{\beta_k^{*}}^2 + \theta_1^{2*} \Gamma\left(\frac{1}{n_{kca}} + 1 \right) \hat{\sigma}_{\varphi_k}^2 + \theta_2^{2*} \Gamma\left(\frac{1}{r_{kca}} + 1 \right) \hat{\sigma}_{\varphi_k^{*}}^2 \end{array} } \le C \\
\left(n_{kca}^{*} - n_{kca}^{(r)} \right) + \left(r_{1ca}^{*} - r_{1ca}^{(r)} \right) + d_{kr}^{-} - d_{kr}^{+} = 0 \\
d_{kr}^{+} \ge 0, d_{kr}^{+} \ge 0, \quad 1 \ge r \ge R \\
2 \le n_{kca} \le N_k,\ 2 \le r_{kca} \le \hat{n}_{k2},\ n_{kca} \text{ and } r_{kca} \textit{ are } \text{integers } \forall\ \ k = 1, 2, \ldots, Q
\end{array}\right\} \tag{48}
$$

7. Numerical Illustrations

A simulation study will be carried out to show a section of the computational procedure. The R software is used to simulate data for different characteristics see Reddy and Khan (2020), while the LINGO-20 software for optimisation problems will be employed to obtain numerical solutions to the integer nonlinear programming problems. LINGO is an efficient and comprehensive modelling tool built to solve linear and nonlinear optimisation problems. The version used is LINGO-20 with the user's guide (2017), or visit http://www.lindo.com.

Let $u_{k1} = 5$ unit and $u_{k2} = 5$ unit represent the reward given to the respondents and the sub-sample of the non-respondents in a much more careful second attempt. Let λ and λ^* be the reciprocal of the average time taken to obtain data from the selected units be (say 30min, 35min, 40min, etc on average for an individual. Let also θ and θ^* be the reciprocal of average time to failure in enumerating the selected units be (say 30min, 35min, 40min etc. per unit), with the total fixed cost $C_0 = 6000$ units. Let the expected cost per unit time of labour and failure taken over all strata be $E(\beta_k, \beta_k^*) = \underset{\forall k}{E}(10,15)$ and $E(\varphi_k, \varphi_k^*) = \underset{\forall k}{E}(10,20)$ with $V(\beta_k, \beta_k^*) = \underset{\forall k}{V}(2.5,4.5)$ and $V(\varphi_k, \varphi_k^*) = \underset{\forall k}{V}(2.5,6.0)$. Let the percentage of non-response in each stratum be represented by 20, 30, 26 and 28. The objective $j = 1, 2$ is carried out in four strata, that is, $k = 1,\ 2,\ 3,\ 4$.

The R software is used to generate data for two different characteristics. If the population size of N = 1544 is considered to be divided into four strata, then we can obtain the population means, the fourth central moments and standard deviations for all characteristics that are specified on each unit through R- R-simulation (Reddy and Khan 2020). Let 99% probability satisfy the chance constraint required for the problem. Then we obtain G_α is such a way that $\phi(G_\alpha) = 0.99$. At 99% confidence limit, the corresponding value of the standard normal variate G_α is 2.33 from tables. The information on the simulation study is given in Table 1.

Table 1: Data obtained through R-simulation

k	N_k	W_k	S_{1k}^2	S_{2k}^2	C_{1k}^4	C_{2k}^4	W_{k1}
1	410	0.2655	97.8546	58.8913	24904.31	9399.36	0.80
2	496	0.3212	103.4125	66.1549	29691.97	13817.88	0.70
3	321	0.2079	96.2571	67.8723	29455.70	12849.46	0.74
4	317	0.2053	107.2381	59.0672	33676.30	10242.08	0.72

Table 2: Data for non-response obtained through R-simulation

k	N_k	W_k	S_{1k2}^2	S_{2k2}^2	C_{1k2}^4	C_{2k2}^4	W_{k1}	W_{k2}
1	410	0.2655	110.9813	73.6280	29587.08	12693.93	0.80	0.20
2	496	0.3212	109.9064	64.0851	29395.70	13159.13	0.70	0.30
3	321	0.2079	79.3220	52.2284	17700.00	7013.00	0.74	0.26
4	317	0.2053	134.7390	51.6128	56857.70	7177.71	0.72	0.28

Table 3: Survey cost values.

$E(c_{k0})$	$E(c'_{k1})$	$E(c'_{k2})$	$E(t_{k0})$	$E(t_{k2})$	$V(c_{k0})$	$V(c'_{k1})$	$V(c'_{k2})$	$V(t_{k0})$	$V(t_{k2})$
1.0	2.0	6.0	0.5	1.5	0.25	0.25	1.25	0.15	0.30
1.0	3.0	7.0	0.5	1.5	0.25	0.50	1.50	0.15	0.30
1.0	5.0	9.0	0.5	1.5	0.25	0.75	1.75	0.15	0.30
1.0	6.0	10.0	0.5	1.5	0.25	1.00	2.00	0.15	0.30

Table 4: Individual optimum allocation (IOA) values

	Individual Optimal at:	
	$j=1$	$j=2$
Z_1^*	0.0082347	0.5330028
Z_2^*	0.8768263	0.0064172
n_{j1}	210	167
n_{j2}	252	252
n_{j3}	47	11
n_{j4}	109	200
r_{j1}	2	3
r_{j2}	6	10
r_{j3}	2	2
r_{j4}	18	2

Table 5: Results for ideal solution

Preemptive Structure	n_{1ca}	n_{2ca}	n_{3ca}	n_{4ca}	r_{1ca}	r_{2ca}	r_{3ca}	r_{4ca}
$\left(Z_1^{(1)}, Z_2^{(2)}\right)$	187	156	57	13	13	7	3	6
$\left(Z_2^{(2)}, Z_1^{(1)}\right)$	195	163	40	54	14	6	3	15
Ideal Solution	195	163	57	54	14	7	3	15

Table 6: D_1 distance of compromise solution

Preemptive levels	Distances of compromise solutions								Sum of Distances
	n_{1ca}	n_{2ca}	n_{3ca}	n_{4ca}	r_{1ca}	r_{2ca}	r_{3ca}	r_{4ca}	
$\left(Z_1^{(1)}, Z_2^{(2)}\right)$	0	0	17	0	0	1	0	0	18
$\left(Z_2^{(2)}, Z_1^{(1)}\right)$	8	7	0	41	1	0	0	9	66

Table 7: Compromise allocation obtained using different methods

	Fuzzy	D_1 Distance
$\hat{Z}_1$	0.06411706	0.05567108
$\hat{Z}_2$	0.13992440	0.11234679
n_{1ca}	198	195
n_{2ca}	159	165
n_{3ca}	22	57
n_{4ca}	31	54
r_{1ca}	12	14
r_{2ca}	3	7
r_{3ca}	1	3
r_{4ca}	7	16

Table 8: Trace and cost incurred.

Approach	$\hat{Z}_1$	$\hat{Z}_2$	Trace	Cost incurred
D_1 Distance	0.06411706	0.13992440	0.20404146	4733.969
Fuzzy	0.05567108	0.11234679	0.16801787	4414.410
IOA $j=1$	0.00823470	0.53300280	0.54123750	5992.618
IOA $j=2$	0.00641717	0.87682630	0.88324347	5728.483

Tables 1 and 2, shows data obtained using the R-software (Stratified-R) for two characteristics in four strata. Table 3 shows the assumed survey cost values. Table 4 shows the optimum allocation obtained individually for $j=1$ and $j=2$, respectively. In this case, the optimum allocation of $j=1$ is used for $j=2$ and vice versa. The ideal solution of the D_1 distance technique is presented in Table 5 with distances from the evaluated ideal solution presented in Table 6. The compromise allocations using the D_1 distance method and fuzzy programming is given in Table 7. The traces and costs incurred are presented in Table 8. The traces here represent the total variance with respect to all the variables in the variance-covariance matrices, which are estimates of finite population means of the characteristics considered. Since the characteristics under study are assumed to be independent the covariances are assumed to be zero. The trace value of the different approaches is given in Table 8. This paper, have implemented the Weibull cost function which approximates every existing cost function when nonresponse is considered. The gamma cost function was also sufficiently modified. We observed closely that cost incurred did not surpass the total budgeted allocation given for the survey and also the compromise allocation technique outperformed the individual optimum allocations using the proposed Weibull cost function when nonresponse is considered.

8. Conclusion

The proposed Weibull cost function when nonresponse is observed approximates the linear, quadratic and gamma cost functions. The compromise allocation using the integer nonlinear programming problem (INLPP) with proposed Weibull cost function has solved the issue observed in stratified sampling survey when more than one variable is considered for optimum allocation. Furthermore, two compromise allocation methods were compared with the individual allocations. The compromised allocation method outperformed the individual allocation with lower trace value and less cost incurred as shown in Table 8. The Weibull cost function will fit into any convex optimization problem.

References

Abubakar, Y. and Okon, E. J. (2018). Optimum compromise allocation in multivariate stratified sampling design with gamma cost function: A case of non-response. *Journal of the Nigerian Association of Mathematical Physics*, 48, 93 - 102.

Ahmed, S., Shabbir, J., and Hafeez, W. (2016). Hansen and Hurwitz estimator with scrambled response on second call in stratified random sampling. *Journal of Statistics Applications and Probability*, 5(2), 357-369.

Alshqaq, S. S. A., Ahmadini A. A. and Ali I. (2022). Nonlinear stochastic multiobjective optimization problem in multivariate stratified sampling design. *Hindawi-Mathematical Problems in Engineering*, Volume 2022, 1-16.

Beardwood, J., Halton, J. H., and Hammersley J. M. (1959). The shortest path through many points. *Mathematical Proceedings of the Cambridge Philosophical Society*, 55; 299–327.

Charnes, A. and Cooper, W. W. (1959) "Chance-constrained programming," *Management Science*, 6(1), 73–79.

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

Ding X. and Wang C. (2011). A novel algorithm of stochastic chance-constrained linear programming and its application. *Hindawi-Mathematical Problems in Engineering*, Volume 2012, 1-17.

El-Badry, M. A. (1956). A sampling procedure for mailed questionnaires. *Journal of the American Statistical Association*, 51, 209-227.

Foradori, G. T. (1961). *Some non-response sampling theory for two stage designs.* Mimeograph Series No. 297, North Carolina State University, Raleigh.

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

Haseen, S., Ali, I., and Bari, A. (2014). multi-objective stochastic multivariate stratified sampling in presence of non-response. *Communications in Statistics-Simulation and Computation*, 45(8), 2810-2826.

Karuna G. Reddy and Khan M.G.M. (2018). optimal stratification in stratified designs using weibull-distributed auxiliary information. *Communications in Statistics - Theory and Methods*, 1–17.

Khan, M. G. M., Khan, E. A., and Ahsan, M. J. (2008). optimum allocation in multivariate stratified sampling in presence of non-response. *Journal of the Indian Society of Agricultural Statistics*, 62(1), 42-48.

LINGO-User's Guide (2017). LINGO-User's Guide. LINDO SYSTEM INC., USA.

Mahfouz, M. I., Rashwan M. M. and Khadr Zeinab A. (2023). Optimal stochastic allocation in multivariate stratified sampling. *Mathematics and Statistics*, 11(4), 676 – 684.

Melaku A. (1986). Asymptotic normality of the optimal allocation in multivariate stratified random sampling. *Indian Journal of Statistics. Series B*, 48. 224–232.

Raghav, Y. S., Ali, I., and Bari, A. (2014). Multi-Objective Nonlinear Programming Problem Approach in Multivariate Stratified Sample Surveys in Case of Non-Response. *Journal of Statistical Computation and Simulation*, 84(1); 22-36.

Reddy K. G. and Khan, M. G. M. (2020) "stratifyR: an R package for optimal stratification and sample allocation for univariate populations," *Australian & New Zealand Journal of Statistics*, 62(3), 383–405.

Rao S. (2009) *Engineering optimization theory and practice*. Wiley, Hoboken, New Jersey

Ronald, E. W., Raymond, H. M., Sharon, L. M. and Keying, Y. (2007). *Probability and statistics for engineers and scientist*. 8th Edition. Upper Saddle River, New Jersey: Pearson Education, Inc.

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

Shafiullah, Khan, M. F., and Ali, I. (2014). Fuzzy geometric programming in multivariate stratified sample surveys in presence of non-response with quadratic cost function. *American Journal of Operations Research*, 4, 173-188.

Sukhatme, P.V., Sukhatme, B.V., Sukhatme, S. and Asok, C. (1984). *Sampling theory of surveys with applications*. Iowa State University Press, Ames and Indian Society of Agricultural Statistics, New Delhi

Varshney R, and Mradula (2019). Optimum allocation in multivariate stratified sampling design in the presence of nonresponse with gamma cost function. *Journal Statistical Computing and Simulation*, 89(13), 2454–2467.

Varshney, R., Ahsan, M. J. and Khan, M. G. M. (2011). An optimum multivariate stratified sampling design with non-response: a lexicographic goal programming approach. *Journal of Mathematical Modeling and Algorithms*, 65, 291-296.

Varshney, R., Najmussehar, and Ahsan, M. J. (2012). estimation of more than one parameters in stratified sampling with fixed budget. *Mathematical Method and Operations Research*, 75(2), 185–197.