Multiobjective optimal allocation problem with probabilistic non-linear cost constraint

Shazia Ghufran*, Saman Khowaja and M. J. Ahsan

Department of Statistics and Operations Research, Aligarh Muslim University, Aligarh, INDIA
*Corresponding Author e-mail id : itsshaziaghufran@gmail.com

Abstract

   This paper considers the optimum compromise allocation in multivariate stratified sampling with non-linear objective function and probabilistic non-linear cost constraint. The probabilistic non-linear cost constraint is converted into equivalent deterministic one by using Chance Constrained programming. A numerical example is presented to illustrate the computational procedure.

Keywords: Multiobjective programming, Multivariate Stratified Sampling, Chance Constrained, Compromise allocation.

Mathematics Subject Classification Primary 62D05.

1. Introduction

   One of the areas of statistics that is most commonly used in all fields of scientific investigation is that of probabilistic sampling. An effective sampling technique is one which provides meaningful knowledge of the important aspects of the population. Stratified sampling is one among the designs of sampling surveys for obtaining such information. This method considers the computation of the stratum sample size, which can be computed by various procedures, but optimum allocation has been found to be a useful approach. In multivariate stratified sampling where more than one characteristic are to be estimated, an allocation which is optimum for one characteristic may not be optimum for other characteristics also. In such situations a compromise criterion is needed to work out a usable allocation which is optimum for all characteristics in some sense. Such an allocation may be called a "Compromise Allocation" because it is based on some compromise criterion. In surveys where several characteristics defined on the population units are highly correlated, the individual optimum allocations for different characteristics may differ relatively little. For such situations Cochran (1977) suggested the use of the character wise average of the individual optimum allocations as a usable compromise allocation. He assumed all the characteristics equally important.
   Several others have studied various criteria for obtaining a usable compromise allocation. Among them are Neyman (1934), Dalenius (1953) and (1957), Ghosh (1958), Yates (1960), Aoyama (1963), Gren (1964) and (1966), Folks and Antle (1965), Hartley (1965), Kokan and Khan (1967), Chatterjee (1972), Ahsan and Khan (1977) and (1982), Chromy (1987), Wywial (1988), Bethel (1989), Kreienbrock (1993), Jahan et al. (1994), Khan et al. (1997), Khan et al. (2003), Ahsan et al. (2005), Kozak (2006), Diaz Garcia and Cortez (2006) and (2008), Ansari et al. (2009) etc. The problem of optimal allocation in stratified sampling is generally stated in two ways. Either one minimizes the cost of survey for a desired precision or the variance of the sample estimate is minimized for a given budget of the survey. Kokan and Khan (1967) formulated the minimization of the cost of the survey for desired precisions on various characters as the following convex programming problem;

$$\left.\begin{array}{l} \underset{n}{Min} \ \sum_{h=1}^{L} c_h n_h, \\ \\ Subject \ \ to \ \sum_{h=1}^{L} \frac{a_{jh}}{n_h} - \sum_{h=1}^{L} \frac{a_{jh}}{N_h} \le k_j, \\ \\ 2 \le n_h \le N_h, \ \ h = 1,..., L. \end{array}\right\} \ j = 1,..., p \tag{1}$$

where L is the number of strata, p is the number of characters to be estimated in the survey and c_h, a_{jh}, k_j and N_h are all positive constants. If the budget of the survey is fixed in advance, say, C, then the multivariate allocation problem is stated to minimize the variances for various characters for a desired precision as the following p convex programming problems.

$$\left.\begin{array}{l} \underset{n}{Min.V} = \sum_{h=1}^{L} \frac{a_{jh}}{n_h} - \sum_{h=1}^{L} \frac{a_{jh}}{N_h} \\ \\ Subject \ to \ \sum_{h=1}^{L} c_h n_h + c_0 \le C \\ \\ and \ \ \ \ 2 \le n_h \le N_h; \ \ h = 1,..., L. \end{array}\right\} \ j = 1,..., p \tag{2}$$

Further, in a survey the costs for enumerating a character in various strata are not known exactly, rather these are being estimated from sample costs. As such the formulated allocation problem should be considered as stochastic programming problem. Stochastic programming problem was first formulated by Dantzig (1955), who suggested a two stage programming technique for its solutions. Later, Charnes and Cooper (1959) developed the chance constrained programming technique in which the chance constraints are converted into equivalent deterministic non-linear constraints.

When the constants c_h and $a_{jh}, (h = 1,..., L, j = 1,..., p)$ are fixed, the problem (1) was solved by Kokan and Khan by using an analytical procedure. Prekopa (1995) developed a method from stochastic point of view. The case when sampling variances are random in the constraints (i.e. a_{jh} random in (1)) has been dealt with Diaz-Garcia et al. (2007). Javaid and Bakhshi (2009) considered the case of random costs in (1) and used modified E- model for solving this problem. Bakhshi et al. (2010) find the optimal Sample Numbers in Multivariate Stratified Sampling with a Probabilistic Cost Constraint.

Here we consider the case of a non-linear cost function with random coefficients. The equivalent deterministic model for the problem in (1) is obtained by applying the chance constrained programming technique. The model in (2) with non-linear cost function in constraints is handled by using the modified-E model of Diaz-Garcia et al. (2007).

In the present paper the problem of finding the optimum compromise allocation is formulated as Stochastic Nonlinear Programming Problem (SNLPP) and a method is developed to work out the compromise allocation in a multivariate stratified surveys using the compromise criteria "Minimizing the sum of sampling variances of the estimators of the population parameters of various characteristics". A numerical example is also worked out to illustrate the computational details of the method.

2. Problem formulation

We consider a multivariate population consisting of N units which is divided into L disjoint strata of sizes $N_1, N_2,..., N_L$ such that $N = \sum_{h=1}^{L} N_h$. Suppose that p characteristics $(j = 1,..., p)$ are measured on each unit of the population. We assume that the strata boundaries are fixed in advance. Let n_h units be drawn without replacement from the hth stratum $h = 1,..., L$. For j^{th} character, an unbiased estimate of the population mean $\overline{Y}_j$ $(j = 1,..., p)$, denoted by $\overline{y}_{jst}$, has its sampling variance

$$V(\overline{y}_{jst}) = \sum_{h=1}^{L} \left(\frac{1}{n_h} - \frac{1}{N_h} \right) W_h^2 S_{jh}^2 \ , \ \ j = 1,..., p, \tag{3}$$

where $W_h = \dfrac{N_h}{N}$ is the stratum weight and $S_{jh}^2 = \dfrac{1}{N_h - 1} \sum\limits_{i=1}^{N_i} \left(y_{jhi} - \overline{Y}_{jh} \right)^2$ is the variance for the jth character in the hth stratum. Let C be the upper limit on the total cost of the survey. The problem of optimal sample allocation involves determining the sample sizes $n_1, n_2, \ldots, n_L$ that minimize the variances of various characters under the given sampling budget C. Within any stratum the linear cost function is appropriate when the major item of cost is that of taking the measurements on each unit. If travel costs between units in a given stratum are substantial, empirical and mathematical studies indicate that the costs are better represented by the expression $\sum\limits_{h=1}^{L} t_h \sqrt{n_h}$, where t_h is the travel cost incurred in enumerating a sample unit in the hth stratum. Assuming this non-linear cost function one should have

$$C = c_0 + \sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} t_h \sqrt{n_h} \tag{4}$$

where $c_h ;\ h = 1, 2, \ldots, L$ denote the per unit cost of measurement in the h-th stratum and c_0 is the overhead cost.

The restrictions $2 \le n_h \le N_h ;\ h = 1, 2, \ldots, L$ are introduced to obtain the estimates of the stratum variances and to avoid the problem of oversampling.

Thus the problem (2) with non-linear cost function can be written as

$$\left.\begin{aligned}
& Minimize \quad \sum_{h=1}^{L} \frac{W_h^2 S_{jh}^2}{n_h} - \sum_{h=1}^{L} \frac{W_h^2 S_{jh}^2}{N_h} \\
& Subject\ to \quad \sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} t_h \sqrt{n_h} + c_0 \le C \\
& and \qquad\qquad 2 \le n_h \le N_h ; \quad h = 1, 2, \ldots, L
\end{aligned}\right\} j = 1, 2, \ldots, p \tag{5}$$

Ignoring the term independent of n_h, the allocation problem (5) is reduced to the following p convex programming problems:

$$\left.\begin{aligned}
& Minimize \quad \sum_{h=1}^{L} \frac{W_h^2 S_{jh}^2}{n_h} \\
& Subject\ to \quad \sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} t_h \sqrt{n_h} + c_0 \le C \\
& and \qquad\qquad 2 \le n_h \le N_h ; \quad h = 1, 2, \ldots, L
\end{aligned}\right\} j = 1, 2, \ldots, p \tag{6}$$

In many practical situations the measurement cost c_h and the travel cost t_h in the various strata are not fixed and may be considered as random. Let us assume that c_h and t_h, $h = 1, \ldots, L$ are independently normally distributed random variables. Thus the above problem (6) can be written in the following chance constrained programming form as:

$$\left.\begin{aligned}
& Minimize \quad \sum_{h=1}^{L} \frac{W_h^2 S_{jh}^2}{n_h} \\
& Subject\ to \quad P\left(\sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} t_h \sqrt{n_h} + c_0 \le C \right) \ge p_0 \\
& and \qquad\qquad 2 \le n_h \le N_h ; \quad h = 1, 2, \ldots, L
\end{aligned}\right\} j = 1, 2, \ldots, p \tag{7}$$

where $p_0,\ 0 \le p_0 \le 1$ is a specified probability

3. Solution using Chance Constrained Programming

The costs c_h and t_h, $h = 1,..., L$ have been assumed to be independently normally distributed random variables.

Then the function $(\sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} t_h \sqrt{n}_h + c_0)$, will also be normally distributed with mean $E\left(\sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} t_h \sqrt{n}_h + c_0\right)$

and variance $V\left(\sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} t_h \sqrt{n}_h + c_0\right)$.

If $c_h \sim N(\mu_{ch}, \sigma_{ch}^2)$ and $t_h \sim N(\mu_{th}, \sigma_{th}^2)$, then the mean of the function $(\sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} t_h \sqrt{n}_h + c_0)$ is obtained as

$$E\left(\sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} t_h \sqrt{n_h} + c_0\right) = E\left(\sum_{h=1}^{L} c_h n_h\right) + E\left(\sum_{h=1}^{L} t_h \sqrt{n_h}\right) + c_0$$

$$= \sum_{h=1}^{L} n_h E(c_h) + \sum_{h=1}^{L} \sqrt{n_h} E(t_h) + c_0$$

$$= \sum_{h=1}^{L} n_h \mu_{ch} + \sum_{h=1}^{L} \sqrt{n_h} \mu_{th} + c_0 \tag{8}$$

The variance is obtained as

$$V\left(\sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} t_h \sqrt{n_h} + c_0\right) = V\left(\sum_{h=1}^{L} c_h n_h\right) + V\left(\sum_{h=1}^{L} t_h \sqrt{n_h}\right)$$

$$= \sum_{h=1}^{L} n_h^2 V(c_h) + \sum_{h=1}^{L} n_h V(t_h)$$

$$= \sum_{h=1}^{L} n_h^2 \sigma_{ch}^2 + \sum_{h=1}^{L} n_h \sigma_{th}^2 \tag{9}$$

Now let $f(t) = \sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} t_h \sqrt{n}_h + c_0$, then the chance constraint in (7) is given by

$$P\left(f(t) \le C\right) \ge p_0,$$

or $\quad P\left\{\dfrac{f(t) - E(f(t))}{\sqrt{V(f(t))}} \le \dfrac{C - E(f(t))}{\sqrt{V(f(t))}}\right\} \ge p_0,$

where $\left[\dfrac{f(t) - E\{f(t)\}}{\sqrt{V\{f(t)\}}}\right]$ is a standard normal variate with mean zero and variance one. Thus the probability of realizing $\{f(t)\}$ less than or equal to C can be written as

$$P\left(f(t) \le C\right) = \phi\left[\dfrac{C - E\{f(t)\}}{\sqrt{V\{f(t)\}}}\right], \tag{10}$$

where $\phi(z)$ represents the cumulative density function of the standard normal variable evaluated at z. If K_α represents the value of the standard normal variate at which $\phi(K_\alpha) = p_0$, then the constraint (10) can be written as

$$\phi\left[\frac{C - E\{f(t)\}}{\sqrt{V\{f(t)\}}}\right] \geq \phi(K_\alpha). \tag{11}$$

The inequality will be satisfied only if

$$\left[\frac{C - E\{f(t)\}}{\sqrt{V\{f(t)\}}}\right] \geq K_\alpha \ ,$$

or equivalently, $E(f(t)) + K_\alpha \sqrt{V(f(t))} \leq C$.

$$\tag{12}$$

Substituting from (8) and (9) in (12), we get

$$\left(\sum_{h=1}^{L} n_h \mu_{ch} + \sum_{h=1}^{L} \sqrt{n_h}\, \mu_{th} + c_0\right) + K_\alpha \sqrt{\sum_{h=1}^{L} n_h^2 \sigma_{ch}^2 + \sum_{h=1}^{L} n_h \sigma_{th}^2} \leq C \tag{13}$$

The constants $\mu_{ch}, \mu_{th}, \sigma_{ch}$ and σ_{th} in (13) are unknown (by hypothesis). So we will use the estimators of mean

$$E\left(\sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} t_h \sqrt{n_h} + c_0\right) \text{ and variance } V\left(\sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} t_h \sqrt{n_h} + c_0\right) \text{ given by}$$

$$\hat{E}\left(\sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} t_h \sqrt{n_h} + c_0\right) = \sum_{h=1}^{L} n_h \bar{c}_h + \sum_{h=1}^{L} \sqrt{n_h}\, \bar{t}_h \text{ , say} \tag{14}$$

and

$$\hat{V}\left(\sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} t_h \sqrt{n_h} + c_0\right) = \sum_{h=1}^{L} n_h^2 \sigma_{ch}^2 + \sum_{h=1}^{L} n_h \sigma_{th}^2 \text{ , say.} \tag{15}$$

where $\bar{c}_h$, $\bar{t}_h$, σ_{ch}^2 and σ_{th}^2 are the estimated means and variances from the sample.

Thus an equivalent deterministic constraint to the stochastic constraint is given by

$$\left(\sum_{h=1}^{L} \bar{c}_h n_h + \sum_{h=1}^{L} \bar{t}_h \sqrt{n_h} + c_0\right) + K_\alpha \sqrt{\sum_{h=1}^{L} n_h^2 \sigma_{ch}^2 + \sum_{h=1}^{L} n_h \sigma_{th}^2} \leq C \tag{16}$$

Now the problem of allocation in multivariate stratified sample surveys with p-independent characteristics is formulated as a Multiobjective Non Linear Programming Problem (MNLPP). The 'p' objectives are to minimize the individual variances of the estimates of the population means of p-characteristics simultaneously, subject to the non-linear probabilistic cost constraint. The formulated MNLPP is given as

$$
\left.
\begin{array}{ll}
\textit{Minimize} & \begin{pmatrix} V(\bar{y}_{1st}) \\ \vdots \\ V(\bar{y}_{pst}) \end{pmatrix} \\[20pt]
\textit{Subject to} & \hat{E}\left(\sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} t_h \sqrt{n_h} + c_0 \right) + K_\alpha \sqrt{\hat{V}\left(\sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} t_h \sqrt{n_h} \right)} \le C \\[20pt]
& 2 \le n_h \le N_h \\[6pt]
\textit{and} & n_h \ are \ int egers; \quad h = 1, 2, ..., L.
\end{array}
\right\}
\tag{17}
$$

To solve the problem (17) using stochastic programming, we first solve the following p Non Linear Programming Problems (NLPPs) for all the 'p' characteristics separately. The equivalent deterministic non-linear programming problem to the stochastic programming problem is given by

$$
\left.
\begin{array}{ll}
\textit{Minimize} & V_j = \sum_{h=1}^{L} \dfrac{W_h^2 S_{jh}^2}{n_h} \\[20pt]
\textit{Subject to} & \left(\sum_{h=1}^{L} \bar{c}_h n_h + \sum_{h=1}^{L} \bar{t}_h \sqrt{n_h} + c_0 \right) + K_\alpha \sqrt{\left(\sum_{h=1}^{L} \sigma_{ch}^2 n_h^2 + \sum_{h=1}^{L} \sigma_{th}^2 n_h \right)} \le C \\[20pt]
& 2 \le n_h \le N_h \\[6pt]
\textit{and} & n_h \ are \ int egers; \quad h = 1, 2, ..., L.
\end{array}
\right\} \ j = 1, 2, ..., p.
\tag{18}
$$

Let $\underline{n}^*_{jh} = \left(n^*_{j1}, n^*_{j2}, ..., n^*_{jL} \right)$ denote the solution to the j-th NLPP in (18) with V_j^* as the value of the objective function given by

$$
V_j^* = \sum_{h=1}^{L} \dfrac{W_h^2 S_{jh}^2}{n_{jh}}; \quad j = 1, 2, ..., p
\tag{19}
$$

A reasonable criterion to workout a compromise allocation may be to "Minimize the sum of the variances $V_j; \ j = 1, 2, ..., p$".

We may express the multi-objective NLPP (17) using (19) with the above compromise criterion as the following single objective NLPP as

$$
\left.
\begin{array}{ll}
\textit{Minimize} & \sum_{j=1}^{L} V_j \\[20pt]
\textit{Subject to} & \left(\sum_{h=1}^{L} \bar{c}_h n_h + \sum_{h=1}^{L} \bar{t}_h \sqrt{n_h} + c_0 \right) + K_\alpha \sqrt{\left(\sum_{h=1}^{L} \sigma_{ch}^2 n_h^2 + \sum_{h=1}^{L} \sigma_{th}^2 n_h \right)} \le C \\[20pt]
& 2 \le n_h \le N_h \\[6pt]
\textit{and} & n_h \ are \ int egers; \quad h = 1, 2, ..., L.
\end{array}
\right\}
\tag{20}
$$

4. Numerical Illustration

In the table below the stratum sizes, stratum weights, stratum standard deviations, measurement costs, and the travel costs within stratum are given for four different characteristics under study in a population stratified in five strata. The data are mainly from Chatterjee (1968). The values of strata sizes are added assuming the population size as 6000. The total budget of the survey is assumed to be 1500 units with an overhead cost $c_0 = 300$ units.

Table 1

Values of N_h, W_h, c_h, t_h and S_{jh} for five strata and four characteristics

h	N_h	W_h	S_{jh}			
			S_{1h}	S_{2h}	S_{3h}	S_{4h}
1	1500	0.25	28	206	38	120
2	1920	0.32	24	133	26	184
3	1260	0.21	32	48	44	173
4	480	0.08	54	37	78	92
5	840	0.14	67	9	76	117

In this problem $c_1, c_2, c_3, c_4, c_5, t_1, t_2, t_3, t_4$ and t_5 are independently normally distributed random variables with known means and standard deviations

$$E(c_1) = 1,\ E(c_2) = 1,\ E(c_3) = 1.5, E(c_4) = 1.5 \ and \ E(c_5) = 2$$
$$E(t_1) = 0.5,\ E(t_2) = 0.5,\ E(t_3) = 1, E(t_4) = 1 \ and \ E(t_5) = 1.5$$
$$V(c_1) = 0.25,\ V(c_2) = 0.25,\ V(c_3) = 0.35,\ V(c_4) = 0.35 \ and \ V(c_5) = 0.45.$$
$$V(t_1) = 0.125,\ V(t_2) = 0.125,\ V(t_3) = 0.175,\ V(t_4) = 0.175 \ and \ V(t_5) = 0.225.$$

Using the values given in Table 1 the NLPP (18) and their optimal solutions $\underline{n}_j^*$; $j = 1, 2, 3, 4$ with the corresponding values of V_j^* are listed below. These values are obtained by software LINGO.

For $j = 1$

$$
\left.
\begin{aligned}
&Minimize \quad \frac{49}{n_1} + \frac{58.9824}{n_2} + \frac{45.1584}{n_3} + \frac{18.6624}{n_4} + \frac{87.9844}{n_5} \\
&Subject\ to \ \left(1n_1 + 1n_2 + 1.5n_3 + 1.5n_4 + 2n_5 + 0.5\sqrt{n_1} + 0.5\sqrt{n_2} + 1\sqrt{n_3} + 1\sqrt{n_4} + 1.5\sqrt{n_5} + 300\right) \\
&\qquad + 2.33 \sqrt{\frac{\left(0.25n_1^2 + 0.25n_2^2 + 0.35n_3^2 + 0.35n_4^2 + 0.45n_5^2\right) +}{\left(0.125n_1 + 0.125n_2 + 0.175n_3 + 0.175n_4 + 0.225n_5\right)}} \le 1500 \\
&and \qquad\qquad 2 \le n_1 \le 1500 \\
&\qquad\qquad\qquad 2 \le n_2 \le 1920 \\
&\qquad\qquad\qquad 2 \le n_3 \le 1260 \\
&\qquad\qquad\qquad 2 \le n_4 \le 480 \\
&\qquad\qquad\qquad 2 \le n_5 \le 840
\end{aligned}
\right\} \qquad (21)
$$

The optimum allocation $\underline{n}_1^* = \left(n_{11}^*, n_{12}^*, n_{13}^*, n_{14}^*, n_{15}^*\right)$ is

$n_{11}^* = 132.999,\ n_{12}^* = 143.2324,\ n_{13}^* = 107.7228,\ n_{14}^* = 72.3840,\ n_{15}^* = 127.6964.$

The corresponding value of the variance ignoring finite population correction (fpc) is $V_1^* = 2.148212$.

For j = 2

$$\text{Minimize} \quad \frac{2652.25}{n_1} + \frac{1811.3536}{n_2} + \frac{101.6064}{n_3} + \frac{8.7616}{n_4} + \frac{1.5876}{n_5}$$

$$\text{Subject to} \quad \left(1n_1 + 1n_2 + 1.5n_3 + 1.5n_4 + 2n_5 + 0.5\sqrt{n_1} + 0.5\sqrt{n_2} + 1\sqrt{n_3} + 1\sqrt{n_4} + 1.5\sqrt{n_5} + 300\right)$$

$$+2.33 \sqrt{\frac{\left(0.25n_1^2 + 0.25n_2^2 + 0.35n_3^2 + 0.35n_4^2 + 0.45n_5^2\right)+}{\left(0.125n_1 + 0.125n_2 + 0.175n_3 + 0.175n_4 + 0.225n_5\right)}} \leq 1500 \qquad (22)$$

$$\text{and} \qquad 2 \leq n_1 \leq 1500$$
$$2 \leq n_2 \leq 1920$$
$$2 \leq n_3 \leq 1260$$
$$2 \leq n_4 \leq 480$$
$$2 \leq n_5 \leq 840$$

The optimum allocation $\underline{n}_2^* = \left(n_{21}^*, n_{22}^*, n_{23}^*, n_{24}^*, n_{25}^*\right)$ is

$n_{21}^* = 303.1810,\ n_{22}^* = .259.2840,\ n_{23}^* = 60.5848,\ n_{24}^* = 18.3975,\ n_{25}^* = 6.6782.$

The corresponding value of the variance ignoring fpc is $V_2^* = 18.12507$.

For j = 3

$$\text{Minimize} \quad \frac{90.25}{n_1} + \frac{69.2224}{n_2} + \frac{85.3776}{n_3} + \frac{38.9376}{n_4} + \frac{113.2096}{n_5}$$

$$\text{Subject to} \quad \left(1n_1 + 1n_2 + 1.5n_3 + 1.5n_4 + 2n_5 + 0.5\sqrt{n_1} + 0.5\sqrt{n_2} + 1\sqrt{n_3} + 1\sqrt{n_4} + 1.5\sqrt{n_5} + 300\right)$$

$$+2.33 \sqrt{\frac{\left(0.25n_1^2 + 0.25n_2^2 + 0.35n_3^2 + 0.35n_4^2 + 0.45n_5^2\right)+}{\left(0.125n_1 + 0.125n_2 + 0.175n_3 + 0.175n_4 + 0.225n_5\right)}} \leq 1500 \qquad (23)$$

$$\text{and} \qquad 2 \leq n_1 \leq 1500$$
$$2 \leq n_2 \leq 1920$$
$$2 \leq n_3 \leq 1260$$
$$2 \leq n_4 \leq 480$$
$$2 \leq n_5 \leq 840$$

The optimum allocation $\underline{n}_3^* = \left(n_{31}^*, n_{32}^*, n_{33}^*, n_{34}^*, n_{35}^*\right)$ is

$n_{31}^* = 142.0023,\ n_{32}^* = 126.7286,\ n_{33}^* = 117.2123,\ n_{34}^* = 82.6231,\ n_{35}^* = 117.3308$

The corresponding value of the variance ignoring fpc is $V_3^* = 3.346324$.

For $j = 4$

$$\left. \begin{aligned} Minimize \quad & \frac{900}{n_1} + \frac{3466.8544}{n_2} + \frac{1319.8689}{n_3} + \frac{54.1696}{n_4} + \frac{268.3044}{n_5} \\ Subject\ to \quad & \left(1n_1 + 1n_2 + 1.5n_3 + 1.5n_4 + 2n_5 + 0.5\sqrt{n_1} + 0.5\sqrt{n_2} + 1\sqrt{n_3} + 1\sqrt{n_4} + 1.5\sqrt{n_5} + 300\right) \\ & + 2.33 \sqrt{\frac{\left(0.25n_1^2 + 0.25n_2^2 + 0.35n_3^2 + 0.35n_4^2 + 0.45n_5^2\right) +}{\left(0.125n_1 + 0.125n_2 + 0.175n_3 + 0.175n_4 + 0.225n_5\right)}} \leq 1500 \\ and \quad & 2 \leq n_1 \leq 1500 \\ & 2 \leq n_2 \leq 1920 \\ & 2 \leq n_3 \leq 1260 \\ & 2 \leq n_4 \leq 480 \\ & 2 \leq n_5 \leq 840 \end{aligned} \right\} \quad (24)$$

The optimum allocation $\underline{n}_4^* = \left(n_{41}^*, n_{42}^*, n_{43}^*, n_{44}^*, n_{45}^*\right)$ is

$n_{41}^* = 139.7336$, $n_{42}^* = 246.2649$, $n_{43}^* = 139.3793$, $n_{44}^* = 31.8239$, $n_{45}^* = 59.5315$.

The corresponding value of the variance ignoring fpc is $V_4^* = 36.19729$.

Using the computed values of V_j^*; $j = 1, 2, 3, 4$ and the compromise criterion conjectured in section 3, the Stochastic Programming Problem given in (20) may be expressed as:

$$\left. \begin{aligned} Minimize \quad & \frac{1377.5081}{n_1} + \frac{2449.5358}{n_2} + \frac{740.9373}{n_3} + \frac{35.7884}{n_4} + \frac{56.2973}{n_5} \\ Subject\ to \quad & \left(1n_1 + 1n_2 + 1.5n_3 + 1.5n_4 + 2n_5 + 0.5\sqrt{n_1} + 0.5\sqrt{n_2} + 1\sqrt{n_3} + 1\sqrt{n_4} + 1.5\sqrt{n_5} + 300\right) \\ & + 2.33 \sqrt{\frac{\left(0.25n_1^2 + 0.25n_2^2 + 0.35n_3^2 + 0.35n_4^2 + 0.45n_5^2\right) +}{\left(0.125n_1 + 0.125n_2 + 0.175n_3 + 0.175n_4 + 0.225n_5\right)}} \leq 1500 \\ and \quad & 2 \leq n_1 \leq 1500 \\ & 2 \leq n_2 \leq 1920 \\ & 2 \leq n_3 \leq 1260 \\ & 2 \leq n_4 \leq 480 \\ & 2 \leq n_5 \leq 840 \end{aligned} \right\} \quad (25)$$

The optimum compromise allocation which is the solution to the SNLPP (25) given by the software LINGO is:

$n_{1c}^* = 194.9769$, $n_{2c}^* = 228.4192$, $n_{3c}^* = 113.5010$, $n_{4c}^* = 34.5660$, $n_{5c}^* = 57.7153$.

With the corresponding value of the objective function as 67.92505.

5. Discussion

In multivariate stratified surveys when the use of individual optimum allocations is not possible, we need to work out an allocation that is optimum for all characteristics in some sense that is, using a compromise criterion to work out an allocation. This paper is a profound study of an optimum compromise allocation with cost constraint function as random parameters with certain probabilities. After rounding off to the nearest integer value we get the optimum compromise allocation obtained by optimization software LINGO as

$$n_{1c}^* = 195,\ n_{2c}^* = 228,\ n_{3c}^* = 114,\ n_{4c}^* = 35,\ n_{5c}^* = 58.$$

6. Conclusion

In this paper, we focused on multiobjective nonlinear programming problem involving random variables.The solution to the formulated problem is worked out using the compromise criterion "Minimizing the sum of sampling variances of the estimators of the population parameters of various characteristics. The proposed approach can be extended to solve multiobjective chance constrained bilevel programming problems with continuous random variables in chance constraints and/or in objective function.

References

Ahsan, M. J. and Khan, S. U. 1977. Optimum allocation in multivariate stratified random sampling using prior information, *Journal of Indian Statistical Association*, Vol. 15, pp. 57-67.

Ahsan, M. J. and Khan, S. U. 1982. Optimum allocation in multivariate stratified random sampling with overhead cost, *Metrika*, Vol. 29, pp. 71-78.

Ahsan, M. J., Najmussehar and Khan, M. G. M. 2005. Mixed allocation in stratified sampling, *Aligarh Journal of Statistics*, Vol. 25, pp. 87-97.

Ansari, A. H., Najmussehar and Ahsan, M. J. 2009. On multiple response stratified random sampling design, *International Journal of Statistical Sciences*, Kolkata, India, Vol. 1, No. 1, pp. 1-11.

Aoyama, H. (1963). Stratified random sampling with optimum allocation for multivariate populations, *Annals of the Institute of Statistical Mathematics*, Vol. 14, pp. 251-258.

Bakhshi, Z. H. Khan, M.F. and Ahmad, Q.S. 2010. Optimal sample numbers in multivariate stratified sampling with a probabilistic cost constraint. *International journal of Mathematics and Applied Statistics,* Vol. 1, No. 2, pp. 111-120.

Bethal, J. 1989. Sample allocation in multivariate surveys, *Survey Methodology*, Vol. 15, pp. 47 – 57.

Charnes, A. and Cooper, W.W. 1959. Chance constrained programming. *Management Science,* Vol. 5, pp. 73-79.

Chromy, J. R. 1987. Design optimization with multiple objectives. *Proceedings of the Survey Research Methods section*, American Statistical Association, pp. 194–199.

Chatterjee, S. 1968. Multivariate stratified surveys, *J. Amer. Stat. Assoc.,* Vol. 63, pp. 530-534.

Chatterjee, S. 1972. A study of optimum allocation in multivariate stratified surveys. *Skandinavisk Actuarietidskrift*, Vol. 55, pp. 73-80.

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

Dalenius, T. 1953. The multivariate sampling problem. *Skandinavisk Actuarietidskrift*, Vol. 36, pp. 92-102

Dalenius, T. 1957. *Sampling in Sweden.* Contributions to the Methods and Theories of Sample Survey Practice, Almqvist and Wicksell, Stockholm.

Dantzig, G.B. 1955. Linear programming under uncertainty, *Management Science*, Vol. 1, pp. 3-4.

Folks, J.L. and Antle, C.E. 1965. Optimum allocation of sampling units to the strata when there are r responses of interest, *Journal of American Statistical Association*, Vol. 60, pp. 225 - 233.

Garcia, J. A. D. and Cortez, L. U. 2006. Optimum allocation in multivariate stratified sampling: multi-objective programming, *Comunicacion Tecnica No. I-06-07/28-03-206 (PE/CIMAT)*.

Garcia, J. A. D., and Tapia, M. M. G. 2007. Optimum allocation in Stratified surveys: Stochastic Programming, *Computational Statistics and Data Analysis*, Vol. 51, pp. 3016-3026.

Garcia, J. A. D. and Cortez, L. U. 2008. Multi-objective optimization for optimum allocation in multivariate stratified sampling, *Survey Methodology*, Vol. 34, No. 2, pp. 215-222.

Ghosh, S. P. 1958. A note on stratified random sampling with multiple characters, *Calcutta Statistical Association Bulletin*, Vol. 8, pp. 81-89.

Gren, J. 1964. Some methods of sample allocation in multivariate stratified sampling. *Przeglad Statystyczny*, Vol. 11, pp. 361-369 (in Polish).

Gren, J. 1966. Some application of non-linear programming in sampling methods. *Przeglad Statystyczny*, Vol. 13, pp. 203-217 (in Polish).

Hartley, H. O. 1965. Multiple purpose optimum allocation in stratified sampling. *Proc. Amer. Statist. Assoc., Social Statist. Sec.*, pp. 258-261.

Jahan, N., Khan, M. G. M. and Ahsan, M. J., 1994. A generalized compromise allocation, *Journal of the Indian Statistical Association*, Vol. 32, pp. 95-101.

Javed, S. Bakhshi, Z. H. and Khalid, M. M. 2009. Optimum allocation in Stratified Sampling with random costs. *International Review of Pure and Applied Mathematics,* Vol. 5, No. 2, pp. 363-370

Khan, M.G.M., Ahsan, M.J. and Jahan, N. 1997. Compromise Allocation in Multivariate Stratified Sampling: An Integer Solution, *Naval Research Logistics,* Vol. 44, pp. 69 – 79.

Khan, M.G.M., Khan, E.A. and Ahsan, M.J. 2003. An optimal multivariate stratified sampling design using dynamic programming, *Australian andNew Zealand J. Statist.*, Vol. 45, No. 1, pp. 107 – 113.

Kokan, A.R. and Khan, S.U. 1967. Optimum allocation in multivariate surveys: An analytical solution, *Journal of Royal Statistical Society,* Ser. B, Vol. 29, pp. 115 – 125.

Kozok, M. 2006. On sample allocation in multivariate surveys, *Communication in Statistics-Simulation and Computation*, Vol. 35, pp. 901-910.

Kreienbrock, L. 1993. Generalized measures of dispersion to solve the allocation problem in multivariate stratified random sampling, *Communication in Statistics-Theory and Methods,* Vol. 22, No. 1, pp. 219-239.

Neyman, J. 1934. On the two different aspects of the representative method: The method of stratified sampling and the method of purposive selection, *Journal of Royal Statistical Society*, Vol. 97, 558-625.

Prekopa, A. 1995. Stochastic Programming, *Kluwer Academic Publishers*, Series Mathematics and its Applications.

Wywial, J. 1988. Minimizing the spectral radius of means vector from sample variance-covariance matrix sample allocation between strata. *Prace Naukowe Akademii Ekonomicznej we Wroclawiu, Wroclaw, Poland*, Vol. 404, 223-235 (in Polish).

Yates, F. 1960. *Sampling Methods for Censuses and Surveys,* 3rd ed., Charles Griffin and Co. Ltd., London.

Biographical notes

Shazia Ghufran received M.Sc and M.Phil in Statistics from Aligarh Muslim University, Aligarh, India, Presently she is pursuing Ph.D from department of Statistics and Operations research Aligarh Muslim University, Aligarh, India.

Saman Khowaja received M.Sc and M.Phil in Statistics from Aligarh Muslim University, Aligarh, India, Presently she is pursuing Ph.D from department of Statistics and Operations research Aligarh Muslim University, Aligarh, India.

Prof. M.J. Ahsan He is an Professor in the Department of Statistics and Operations Research Muslim University, Aligarh, India; He has more than 30 years of experience in teaching and research.