International Journal Of Scientific & Engineering Research, Volume 4, Issue 6, May 2013
ISSN 2229-5518

A Compromise Solution in Multivariate Surveys with stochastic random cost function

Sana Iftekhar, Sanam Haseen, Qazi Mazhar Ali and Abdul Bari

Abstract—In this paper a problem of multivariate stratified sampling for non-linear random cost with certain probability has been formulated using chance constraint method. Here the formulated problem minimizes the coefficient of variation and determines the best compromise allocation. The solution to this formulated problem is calculated via different programming problem viz. lexicographic goal Programming, fuzzy programming, ϵ-constraint approach and a comparative study by these methods has been attempted. An empirical study of the problem has been done at the end of the paper.

Index Terms—Multivariate Stratified Sampling, Coefficient of Variation, Compromise allocation, Non-linear cost, Stochastic Programming, Lexicographic Goal Programming, Fuzzy Programming, ϵ-Constraint approach.

1. Introduction

With the advent of compromise allocation in sampling surveys where multiple characteristics are under study it is well known that Cochran [28] has enlighten with the idea of character wise average of the individual optimum allocation as a promising compromise allocation taking into consideration that all characteristics are equally important.

This problem of obtaining compromise allocation for multiple characteristic under study was experimented by many researchers. Among them are Dalenius [26 & 27], Yates [3], Aoyama [5], Folks and Antle [12], Chatterjee [19 & 20], Huddleston [6], Chromy [11], Bethel [10], Hartley [7], Kokan et al [13], Diaz Garcia and Cortez [8], Khan et. al. [17 & 18] etc.

For any population Coeffecient of Variation (CoV) is expressed as a relative amount of population standard deviation and population mean. According to Ostle [2] coefficient of variation is a special implement for comparing the variation in two series of data which are measured in two different units.

Dantzig [4] was the first who formulated Stochastic Programming Problem (SPP) and suggested a two stage programming technique to solve it. Later, another method for solving SPP by converting the problem into a deterministic non-linear constraint is developed by Charnes & Cooper [1] in 1959.

SPP is a mathematical programming problem that involves *uncertainty*. In SPP the parameters are known or estimated to follow some probability distribution. In a broader sense, SPP is to find solution that is feasible for all most all the possible data simultaneously maximizing objective function which includes the random variables.

Recent work has been done in this field of chance constraint are by Diaz-Garcia [9], Javed and Bakshi [24], Bakhshi [30], Khan et. al [15] and Ghufran et. al. [21] etc.

In this paper the problem of finding compromise allocation in multivariate sampling in case of random variable with normal probability distribution is formulated into non-linear stochastic programming problem and its equivalent deterministic non-linear programming problem. Problem has been solved using different methods. An empirical study for comprehensive detail of different methods used is also being presented.

2. Formulation of the Problem

We consider a multivariate population consisting of N units which is divided into L disjoint strata of sizes $N_1, N_2, \dots, N_L$ such that $N = \sum_{h=1}^{L} N_h$. Suppose that p characteristics $(j = 1,2, \dots, p)$ are measured on each unit of the population. We assume that the strata boundaries are fixed in advanced. Let n_h units be drawn without replacement from the h^{th} stratum $h = 1,2, \dots, L$.

For j^{th} character, an unbiased estimate of the population mean $\bar{Y}_j$ $(j = 1,2, \dots, p)$ denoted by $\bar{y}_{jst}$, has its sampling variance

- *Sana Iftekhar is currently pursuing Ph.D. in Statistics from Aligarh Muslim University, India. E-mail: Iftekhar.sana54@gmail.com.*
- *Sanam Haseen is currently pursuing Ph.D. in Operations Research from Aligarh Muslim University, India. E-mail: sanam.haseen@gmail.com.*
- *Qazi Mazhar Ali is currently Professor in Dept. of Statistics and Operations Research, A.M.U., Aligarh, India.E-mail: qaziali88@gmail.com.*
- *Abdul Bari is currently Professor in Dept. of Statistics and Operations Research, A.M.U., Aligarh, India.E-mail: bariamu2k3@yahoo.co.in.*

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

where $W_h = \frac{N_h}{N}$ is the stratum weight and $S_{jh}^2 = \frac{1}{N_h-1}\sum_{i=1}^{N_i}\left(y_{jhi} - \bar{Y}_{jh}\right)^2$ is the variance for the j^{th} characteristic in the h^{th} 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 $n_1, n_2, \dots, n_L$ that minimize the variances of various characteristics 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 measurement 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$\sum_{h=1}^{L} t_h\sqrt{n_h}$, where t_h is the travel cost incurred in enumerating a sample unit in the h^{th} stratum.

Assuming this non-linear cost function one should have

$$C = c_0 + \sum_{h=1}^{L} c_h + \sum_{h=1}^{L} t_h\sqrt{n_h} \qquad (2)$$

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

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

Thus the MONLPP of the above problem can be written as

$$\left.\begin{array}{ll} Minimize & \begin{pmatrix}(CoV)_1^2\\ \vdots \\ (CoV)_p^2\end{pmatrix} \\ subject\ to & \sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} t_h\sqrt{n_h} \le C_0 \\ and & 2 \le n_h \le N_h; h = 1,2,\dots,L \end{array}\right\} \qquad (3)$$

where $\sum_{j=1}^{p}(CoV)_j^2 = \begin{pmatrix}(CoV)_1^2\\ \vdots \\ (CoV)_p^2\end{pmatrix}$ and

$$(CoV)_j = CoV(\bar{y}_{jst}); \quad j = 1,2,\dots,p$$

$$= \frac{SD(\bar{y}_{jst})}{\bar{Y}_j} \quad j = 1,2,\dots,p$$

Thus

$$(CoV)_j^2 = \frac{V(\bar{y}_{jst})}{\bar{Y}_j^2} \quad j = 1,2,\dots,p$$

$$= \bar{Y}_j^{-2}\left\{\sum_{h=1}^{L} W_h^2\left(-\frac{n_h}{N_h}\right)\frac{S_{jh}^2}{n_h}\right\} \qquad (4)$$

Thus, the MONLPP may be restated as

$$\left.\begin{array}{ll} Minimize\ Z & = \bar{Y}_j^{-2}\left\{\sum_{h=1}^{L} W_h^2\left(1-\frac{n_h}{N_h}\right)\frac{S_{jh}^2}{n_h}\right\} \\ subject\ to & \sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} t_h\sqrt{n_h} \le C_0 \\ and & 2 \le n_h \le N_h; h = 1,2,\dots,L \end{array}\right\} \qquad (5)$$

For realistic 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 $t_h, h = 1,2,\dots,L$ are independently normally distributed random variables.

Thus the above problem can be written in the following chance constrained programming form as

$$\left.\begin{array}{ll} Minimize\ Z & = \bar{Y}_j^{-2}\left\{\sum_{h=1}^{L} W_h^2\left(1-\frac{n_h}{N_h}\right)\frac{S_{jh}^2}{n_h}\right\} \\ subject\ to & P\left(\sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} t_h\sqrt{n_h} \le C_0\right) \ge p_0 \\ and & 2 \le n_h \le N_h; h = 1,2,\dots,L \end{array}\right\} \qquad (6)$$

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

3. Formulation of Chance Constraint

Assuming the costs c_h and $t_h, h = 1,2,\dots,L$ to be independently normally distributed random variables. The function $\left(\sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} t_h\sqrt{n_h}\right)$ will also be normally distributed with mean $E\left(\sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} t_h\sqrt{n_h}\right)$ and variance $V\left(\sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} t_h\sqrt{n_h}\right)$.

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

$$E\left(\sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} t_h\sqrt{n_h}\right) = E\left(\sum_{h=1}^{L} c_h n_h\right) + E\left(\sum_{h=1}^{L} t_h\sqrt{n_h}\right)$$
$$= \sum_{h=1}^{L} n_h E(c_h) + \sum_{h=1}^{L} \sqrt{n_h} E(t_h)$$
$$= \sum_{h=1}^{L} n_h \mu_{c_h} + \sum_{h=1}^{L} \sqrt{n_h}\, \mu_{t_h} \qquad (7)$$

And the variance is obtained as

$$V\left(\sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} t_h\sqrt{n_h}\right) = V\left(\sum_{h=1}^{L} c_h n_h\right) + V\left(\sum_{h=1}^{L} t_h\sqrt{n_h}\right)$$
$$= \sum_{h=1}^{L} n_h^2 V(c_h) + \sum_{h=1}^{L} n_h V(t_h)$$
$$= \sum_{h=1}^{L} n_h^2\, \sigma_{c_h}^2 + \sum_{h=1}^{L} n_h\, \sigma_{th}^2 \qquad (8)$$

Now let $f(k) = \left(\sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} t_h\sqrt{n_h}\right)$, then the chance constraint in (8) is given by

$$P(f(k) \le C) \ge p_0$$

$$\text{means } P\left\{\frac{f(k) - E(f(k))}{\sqrt{V(f(k))}} \le \frac{C - E(f(k))}{\sqrt{V(f(k))}}\right\} \ge p_0,$$

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

$$P(f(k) \le C) = \emptyset\left[\frac{C - E(f(k))}{\sqrt{V(f(k))}}\right], \qquad (9)$$

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

$$\emptyset\left[\frac{C - E(f(k))}{\sqrt{V(f(k))}}\right] \ge \emptyset(K_\alpha) \qquad (10)$$

The inequality will be satisfied only if

$$\emptyset\left[\frac{C - E(f(k))}{\sqrt{V(f(k))}}\right] \ge K_\alpha, \qquad (11)$$

or equivalently,

$$E(f(k)) + K_\alpha\sqrt{V(f(k))} \le C \qquad (12)$$

Substituting from (7) and (8) in (12), we get

$$\left(\sum_{h=1}^{L} n_h\, \mu_{c_h} + \sum_{h=1}^{L} \sqrt{n_h}\, \mu_{t_h}\right) + K_\alpha\sqrt{\sum_{h=1}^{L} n_h^2\, \sigma_{c_h}^2 + \sum_{h=1}^{L} n_h\, \sigma_{th}^2} \le C$$

The constants μ_{c_h}, μ_{t_h}, σ_{ch} and σ_{th} in (13) are unknown (by hypothesis). So we will use the estimator of mean $E\left(\sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} t_h\sqrt{n_h}\right)$ and variance $V\left(\sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} t_h\sqrt{n_h}\right)$ is given below

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

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

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

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

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

Now the problem of allocation in multivariate stratified sample surveys with p-independent characteristics is formulated as a MOINLPP. 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 MOINLPP is given as

$$\left.\begin{array}{l} Minimize\ Z\ \begin{pmatrix}(CoV)_1^2\\ \vdots \\ (CoV)_p^2\end{pmatrix} \\ subject\ to\ \ \hat{E}\left(\sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} t_h\sqrt{n_h}\right) + K_\alpha\sqrt{\hat{V}\left(\sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} t_h\sqrt{n_h}\right)} \le C \\ \qquad 2 \le n_h \le N_h \\ and \quad n_h\ are\ integers;\ \ h = 1,2,\ldots,L \end{array}\right\} (17)$$

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

$$\left.\begin{array}{l} \text{Minimize } Z = (CoV)_j = \bar{Y}_j^{-2}\left\{\sum_{h=1}^{L} W_h^2\left(1-\frac{n_h}{N_h}\right)\frac{S_{jh}^2}{n_h}\right\} \\ \text{subject to } \left(\sum_{h=1}^{L}\bar{c}_h n_h + \sum_{h=1}^{L}\bar{t}_h\sqrt{n_h}\right) + K_\alpha\sqrt{\left(\sum_{h=1}^{L}\sigma_{ch}^2 n_h^2 + \sum_{h=1}^{L}\sigma_{t_h}^2 n_h\right)} \le C \\ \qquad 2 \le n_h \le N_h \\ \text{and } \quad n_h \text{ are integers}; \quad j = 1,2,\dots p; \quad h = 1,2,\dots,L \end{array}\right\} \quad (18)$$

4. Solution of the problem using different methods

The MOINLPP (18) can be solved using different methods for finding compromise allocations at different context.

4.1 Lexicographic goal programming

Lexicographic goal programming (Díaz-García and Cortez [8]) requires complete information to find solution of hierarchical order arranged according to the importance of the variances. If there are p variances of the estimates $\overline{y}_{jh(w)}$ of $\overline{Y}_j$ arranged in order of their importance and that $(V_1, V_2, \dots, V_n)$ is the arrangement in lexicographic order of importance that is first characteristic is the most important one while the p^{th} is the least important.

At the first stage of the solution the following MOINLPP (18) for $j = 1$ has been obtained. Let $(CoV)_1^*$be the optimal value of the objective function $(CoV)_1$ and $d_1 \ge 0$ is such that$(CoV)_1 - (CoV)_1^* \le d_1$.

At the second stage of the solution the INLPP

$$\left.\begin{array}{l} \text{Minimize } \overline{Y}_2^{-2}\left\{\sum_{h=1}^{L} W_h^2\left(1-\frac{n_h}{N_h}\right)\frac{S_{jh}^2}{n_h}\right\} + d_1 \\ \text{subject to } Y\left\{\sum_{h=1}^{L} W_h^2\left(1-\frac{n_h}{N_h}\right)\frac{S_{jh}^2}{n_h}\right\} - d_1 \le (CoV)_1^* \\ \left(\sum_{h=1}^{L}\bar{c}_h n_h + \sum_{h=1}^{L}\bar{t}_h\sqrt{n_h}\right) + K_\alpha\sqrt{\left(\sum_{h=1}^{L}\sigma_{ch}^2 n_h^2 + \sum_{h=1}^{L}\sigma_{ch}^2 n_h\right)} \le C \\ d_1 \ge 0; \quad j = 1,2,\dots p, \\ \qquad 2 \le n_h \le N_h \\ \text{and } \quad n_h \text{ are integers}; \quad h = 1,2,\dots,L. \end{array}\right\} \quad (19)$$

is obtained.

Similarly at the last stage ($p^{th}\, stage$) of the solution the INLPP to be solved would be

$$\left.\begin{array}{l} \text{Minimize } \overline{Y}_p^{-2}\left\{\sum_{h=1}^{L} W_h^2\left(1-\frac{n_h}{N_h}\right)\frac{S_{jh}^2}{n_h}\right\} + \sum_{j=1}^{p} d_p \\ \text{subject to } Y\left\{\sum_{h=1}^{L} W_h^2\left(1-\frac{n_h}{N_h}\right)\frac{S_{jh}^2}{n_h}\right\} - d_j \le (CoV)_j^* \\ \left(\sum_{h=1}^{L}\bar{c}_h n_h + \sum_{h=1}^{L}\bar{t}_h\sqrt{n_h}\right) + K_\alpha\sqrt{\left(\sum_{h=1}^{L}\sigma_{ch}^2 n_h^2 + \sum_{h=1}^{L}\sigma_{ch}^2 n_h\right)} \le C \\ d_j \ge 0; \quad j = 1,2,\dots p, \\ \qquad 2 \le n_h \le N_h \\ \text{and } \quad n_h \text{ are integers}; \quad h = 1,2,\dots,L. \end{array}\right\} \quad (20)$$

where $d_j \ge 0; j = 1,2,\dots,p-1$ are goals variables whose values are to be determined such that the total increase in the coefficient of variation is minimized and $(CoV)_j^*$ denote the variance under individual optimal allocation for j^{th} characteristic, $j = 1,2,\dots,p$.

It is to be noted that between (19) and (20) there are $(p-3)$ more stages.

4.2 Fuzzy Programming

When the optimal solution is not a crisp solution, instead a compromise solution is required for the problem. The problem is required to be formulated into a fuzzy programming problem (Haseen et. al [23]).

It has already been considered in section 4.1 that $(CoV)_j^*$ be the optimal value of $(CoV)_j$ obtained by solving the MOINLPP (18).

Further let

$$\widetilde{CoV}_j = \widetilde{CoV}_j(n_1, n_2, \dots, n_h, \dots, n_L) \quad (21)$$

denote the value of the coefficient of variation under the compromise allocation, where $n_h; h = 1,2,\dots,L$ are to be worked out.

Obviously $\widetilde{CoV}_j \ge CoV_j^*$ and $\widetilde{CoV}_j - CoV_j^* \ge 0; j = 1,2,\dots,p$ will give the increase in the variance due to not using the individual optimum allocation for j^{th} characteristic.

To obtain a fuzzy solution, we first compute the maximum value U_k and the minimum value L_k, for each $k = 1,2,\dots,p$.

Now,

$$L_k = \min_j Z_k\left(n_{h,j}^*\right) \qquad U_k = \max_j Z_k\left(n_{h,j}^*\right)$$

where $n_{h,j}^*$ denote the optimum allocation for the j^{th} characteristic in four strata.

The differences of the maximum and minimum values of Z_k are denoted by $d_k = U_k - L_k, k = 1, 2, \ldots, p$.

The fuzzy programming formulation of the MOINLPP in (18) is given by the following INLPP:

$$\left.\begin{aligned} &\text{Minimize} \quad \delta \\ &\text{subject to} \quad \overline{Y}_k^{-2}\left\{\sum_{h=1}^{L} W_h^2\left(1-\frac{n_h}{N_h}\right)\frac{S_{jh}^2}{n_h}\right\} - \delta d_k \le CoV_k^* \\ &\left(\sum_{h=1}^{L}\bar{c}_h n_h + \sum_{h=1}^{L}\bar{t}_h\sqrt{n_h}\right) + K_\alpha\sqrt{\left(\sum_{h=1}^{L}\sigma_{ch}^2 n_h^2 + \sum_{h=1}^{L}\sigma_{ch}^2 n_h\right)} \le C \\ &2 \le n_h \le N_h \\ &\textit{and} \quad n_h \textit{ are integers}; \quad k = 1,2,\ldots p; \; h = 1,2,\ldots,L. \end{aligned}\right\} \quad (22)$$

where δ ≥ 0 is the decision variable representing the worst deviation level.

The fuzzy programming may be solved using the optimization software LINGO-13 [14].

4.3 The ϵ - Constraint Approach

The ϵ-constraint method was introduced by Haimes et.al [29]. It was used when partial information about the characteristics is available. In their method they selected one objective and set added all other objectives into constraints after setting an upper bound to each of them. This method only needs to identify the most important characteristic for obtaining the compromise allocation.

Under this approach we express the problem for obtaining the integer compromise allocation as

$$\left.\begin{aligned} &\textit{Minimize } (CoV)_k^2 \\ &\textit{subject to } (CoV)_j^2 \le (CoV)_j^{*2} \\ &\left(\sum_{h=1}^{L}\bar{c}_h n_h + \sum_{h=1}^{L}\bar{t}_h\sqrt{n_h}\right) + K_\alpha\sqrt{\left(\sum_{h=1}^{L}\sigma_{ch}^2 n_h^2 + \sum_{h=1}^{L}\sigma_{ch}^2 n_h\right)} \le C \\ &2 \le n_h \le N_h \\ &\textit{and} \quad n_h \textit{ are integers}; \quad j = 1,2,\ldots p, \quad h = 1,2,\ldots,L. \end{aligned}\right\} \quad (23)$$

where the k^{th} characteristic, $k \in \{1, 2, \ldots, p\}$, is assumed to be most important and $(CoV)_j^{*2}$ is a predetermined bound for the *p-1* remaining coefficients of variation *j=1,2, …, p; j ≠ k*.

It is to be noted that the choice of k^{th} characteristic and the lower limits $(CoV)_j^{*2}$ represent the evaluator's subjective preferences, and so if there were no solution to the (23), this would mean that at least one of the limits of $(CoV)_j^2$ had been set too low and must be revised. For further information one can refer to Ríos, et. al. [25].

5. A Numerical Example

To implement it practically, we use the data are obtained from the 2002 Agriculture Censuses in Iowa State conducted by National Agricultural Statistics Service, USDA, Washington DC. The 99 counties in Iowa are divided into four strata. Two characteristic are defined, first one is the quantity of corn harvested X_1 and second the quantity of oats harvested X_2 are given below

Table 1: Data for four strata and two characteristics

h	N_h	W_h	S_{1h}^2	S_{1h}^2
1	8	0.0808	21601503189.8	1154134.2
2	34	0.3434	19734615816.7	7056074.8
3	45	0.4545	27129658750.0	2082871.3
4	12	0.1212	17258237358.5	732004.9

Also $\overline{X}_1$ and $\overline{X}_2$ are assumed to be known as $\overline{X}_1$=474973.90 and $\overline{X}_2$=1576.25.

The data has been taken from Ghufran et. al.[22] and Kozok [16].

It is of course untrue in real survey. In practice some approximations of these parameters are used; they can be known from a recent or preliminary survey (Kozak (2006)).

The total amounts available for the survey is $C_0 = C - c_0$=200 units, where c_0=50 units is the expected over cost, and C=250 units is the total budget of the survey.

In this problem $c_1, c_2, c_3, c_4, t_1, t_2, t_3, t_4$ are independently normally distributed random variables with assumed means and standard deviations which are given below:

$$E(c_1) = 15, E(c_2) = 7, E(c_3) = 5, E(c_4) = 9,$$

$$E(t_1) = 10, E(t_2) = 5, E(t_3) = 2, E(t_4) = 6$$

$$V(c_1) = 3.75, V(c_2) = 1.75, V(c_3) = 1.25, V(c_4) = 2.25,$$

$$V(t_1) = 2.5, V(t_2) = 1.25, V(t_3) = 0.5, V(t_4) = 1.5$$

Using the values given in table 1, the MOINLPP 20 and their optimal solutions $n_{h,j}^*$, $j = 1, 2;\; h = 1, 2, 3, 4$ with the corresponding values of $(CoV)_j^*$ are given below. These values are being obtained by using software LINGO-13 [14].

For $j = 1$ the optimum allocation is $n_{1,1}^* = 2$, $n_{1,2}^* = 5$, $n_{1,3}^* = 9$, $n_{1,4}^* = 2$, the corresponding individual objective value is $(CoV)_1^* = 0.00467051$

For $j = 2$ the optimum allocation is $n^*_{2,1} = 2$, $n^*_{2,2} = 7$, $n^*_{2,3} = 6$, $n^*_{2,4} = 2$, the corresponding individual objective value is $(CoV)^*_2 = 0.0659471$

6. Discussion

Table 3 and 4 gives a comprehensive detail of the optimum compromise allocations and their objective values using different methods in different context when the costs are considered to be independent and normally distributed. For instance, we can see through the table that the coefficient of variation and optimal compromise allocation found using Lexicographic programming is the best of all the results found by other methods. But Lexicographic goal programming is used when complete information about the data is available on the other hand when we have only partial information about the data we can use the ϵ-constraint Approach to find the best compromise allocations. The problem has also been solved using fuzzy programming that has an optimum compromise allocation but with greater value of coefficient of variation.

Table 2: Compromise allocation obtained by different methods

Compromise Allocation	Lexicographic Programming	Fuzzy Programming	ϵ-constraint Approach	
		delta = 0.5625201	priority $j = 1$	priority $j = 2$
n_1	2	2	2	2
n_2	7	6	7	5
n_3	6	7	6	9
n_4	2	2	2	2
Total	17	17	17	18

Table 3: Corresponding value of the objective function

Coefficient of Variation	Lexicographic Programming	Fuzzy Programming	ϵ-constraint Approach	
			priority $j = 1$	priority $j = 2$
CoV_1	0.005461061	0.005115211	0.00546 1061	0.004670 51
CoV_2	0.0659471	0.06979772	0.06594 71	0.075463 46
Total	0.071408161	0.074912931	0.07140 8161	0.080133 97

References

[1] A. Charnes, W.W. Cooper, "Chance constrained programming", *Management Science*, pp. 73-79, 1959.

[2] B. Ostle,"Statistics in Research Basic Concepts and Techniques for Research Workers. 1st ed. Ames", IA: Iowa State College Press, 1954.

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

[4] G.B. Dantzig, "Linear programming under uncertainty", *Management Science*, pp. 3-4, 1955.

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

[6] H. F. Huddleston, P. L. Claypool, R. R. Hocking., "Optimal Sample Allocation to Strata Using Convex Programming", *Journal of the Royal Statistical Society. Series C*, Vol. 19, pp. 273-278, 1970.

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

[8] J. A. D. Garcia, L. U. Cortez, " Optimum allocation in multivariate stratified sampling: multi-objective programming", *Comunicacion Tecnica*, I-06-07/28-03-206 (PE/CIMAT), 2006.

[9] J. A. D. Garcia, M. M. G Tapia, "Optimum allocation in Stratified surveys: Stochastic Programming", *Computational Statistics and Data Analysis*, Vol. 51, No. 6, pp. 3016-3026, 2007.

[10] J. Bethel, "Sample allocation in multivariate surveys", *Survey Methodology*, Vol. 15, pp. 47 – 57, 1989.

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

[12] J.L. Folks and C.E. Antle, "Optimum allocation of sampling units to the strata when there are r responses of interest", *Journal of American Statistical Association*, Vol. 60, No. 309, pp. 225 – 233, 1965.

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

[14] Lindo Systems Inc., "LINGO User's Guide", Lindo Sys-tems Inc., Chicago, 2001.

[15] M. F. Khan, I. Ali, Y. S. Raghav, A. Bari, "Allocation in Multivariate Stratified Surveys with Non-Linear Random Cost Function", *American*

Journal of Operations Research, Vol. 2, pp. 100-105, 2012.

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

[17] M.G.M. Khan, E.A. Khan, M.J. Ahsan, "An optimal multivariate stratified sampling design using dynamic programming", *Australian and New Zealand J. Statist.*, Vol. 45, pp. 107 – 113, 2003.

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

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

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

[21] S. Ghufran, S. Khowaja, M. J. Ahsan, "Multiobjective optimal allocation problem with probabilistic non-linear cost constraint", *International Journal of Engineering, Science and Technology*, MultiCraft, Vol. 3, pp. 135-145, 2011a.

[22] S. Ghufran, S. Khowaja, M. J. Ahsan, "Optimum Multivariate Stratified Sampling Designs with travel Cost: A Multiobjective Integer Nonlinear Programming Approach", *Communications in Statistics-Simulation and Computation*, Vol. 41, No. 5, pp. 598-610, 2011b.

[23] S. Haseen, S. Iftekhar, M. J. Ahsan and A. Bari, "A fuzzy approach for solving double sampling design in presence of non-response", *International Journal of Engineering Science and Technology*, Vol. 4, No. 06, pp. 2542 – 2551, 2012.

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

[25] S. Ríos, S. Ríos Insua, M. J. Ríos Insua, "Procesos de Decisión Multicriterio", *Madrid: EUDEMA*, 1989.

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

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

[28] W.G. Cochran, "Sampling Techniques", 3rd ed., John Wiley, New York, 1977.

[29] Y. Y. Haimes, L. S. Lasdon, D. A. Wismer, "On a bicriterion formulation of the problems of integrated system dentification and system optimization", *IEEE Transactions on Systems, Man, and Cybernetics*, 1, pp. 296–297, 1971.

[30] Z. H. Bakhshi, M.F. Khan, Q.S. Ahmad, "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, 2010.