Aligarh Journal of Statistics
Vol. 35(2015),77-90

A Bi-OBJECTIVE COST FUNCTION IN MULTIVARIATE STRATIFIED SURVEYS

Shamsher Khan and M. M. Khalid

ABSTRACT

This paper studies the problem of optimal allocation for multivariate stratified survey as a bi-objective programming problem with the objective to minimize the costs (i.e. measurement and travel) incurred in the survey subject to precision constraint for each characteristic. The unitary cost of measurement and travel are considered as normally distributed random variables. Population variances are assumed to be unknown and replaced by sample variances which are also normally distributed random variables. The precision for each characteristic is specified as multi-choice. To remove the randomness from objective functions, Expected Value Standard Deviation (EVSD) criterion is applied after converting the bi-objective problem into a single objective problem. Chance constraint programming technique is then used for deterministic equivalent of constraints. Thus, the problem of optimal allocation is treated as Stochastic Bi-objective Programming Problem (SBOPP) with multi-choice in right hand side. A numerical illustration is also given for the demonstration of proposed approach solved by Lingo Software.

1. INTRODUCTION

Stratified sampling is the most commonly used sampling design in probability sampling. For stratification past data may be used to divide a heterogeneous population into groups such that the units within each group or strata are alike, Hansen *et al.* (1953).

Before using the stratified sampling the sampler must have the answer to the following questions:

i) How many strata should be there?
ii) What should be the strata boundaries?
iii) How many units are to be selected from each stratum?

The third problem which is known as the allocation problem is considered here with the assumption that the population under study has been already stratified into a number of strata with known strata boundaries and the sampling frame of all the strata are available. An allocation will be the best allocation that can minimize the variance of estimator of the population parameter for a given cost of survey or minimize the cost of the survey for desired precision of the estimate. There are some factors which affect the allocation scheme such as the variance of the population, cost of obtaining an observation from each stratum and the degree of precision. If the population variance is unknown, it can be estimated from a preliminary sample and the estimated variance is used in place of population variance, Sukhatme *et al.* (1984). Diaz-Garcia and Gary-Tapia (2007) worked out with estimated variance in univariate survey when population variance was unknown and replaced by sample variance. They considered the problem of optimal allocation as a non-linear stochastic programming problem. Diaz-Garcia and Gary-Tapia showed that in stratified random sampling the sample variance has asymptotic normal distribution on the basis of the result given by Melaku (1986). Fatima *et al.* (2014) extended this work for multivariate case by formulating the problem of optimal allocation as a multi-objective programming problem and solved it by goal programming technique.

In stratified sampling the population having N units is divided into L non-overlapping and exhaustive groups called strata having $N_1, N_2, N_3, ..., N_L$ units respectively (symbols have their usual meaning as in Cochran (1977), otherwise stateded).These subpopulations are called strata. Let n_h be the size of sample allocated to h^{th} stratum then the general cost function in multivariate survey is given as

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

where c_h is the sampling cost per unit associated to the h^{th} stratum. The term c_0 represents an overhead cost.

If the travel costs between units of a stratum are significant then Beardwood *et al.* (1959) suggested that the total travel cost is better represented by the expression $\sum_{h=1}^{L} t_h \sqrt{n_n}$; where t_h is the travel cost per unit in the h^{th} stratum. This expression is quadratic in n_h .

Some authors worked on the problem of optimal allocation in multivariate surveys with quadratic cost such as Khowaja *et al.* (2012), Ghufran *et al.* (2012) by considering the cost function in the form

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

The cost of measurement, which varies from stratum to stratum during the course of survey and affected due to random causes, such as raining, weather conditions etc; can be considered as random variable. The cost of travel between units is also affected by some factors that are out of control of the sampler such as area of survey, condition of the road, modes of travel etc. Thus the travel costs can also be considered as random variables. Some authors worked out with random costs (measurement and travel) to obtain the optimal allocation in multivariate survey with linear and/or nonlinear cost function such as Bakhshi *et al*. (2010), Javaid *et al.* (2011), Ali *et al.* (2011) by treating the cost as a normally distributed random variable but for the first time the total cost is considered as bi-objective function instead of a quadratic function. Another feature of this paper is the consideration of precision of estimates as multi choice, Khan and Khalid (2013).

Thus in this work, the problem of optimal allocation is formulated as a SBOPP to minimize both the costs simultaneously with chance constraint which has multi-choice in the right hand side. The per unit cost of measurement and the travel cost are considered as independently normally distributed random variables. The population variances are supposed to be unknown and replaced by sample variances which are also random variables with asymptotic normal distribution.

2. FORMULATION OF THE PROBLEM

The problem of optimal allocation to minimize the quadratic cost function for the desired degree of the precision for the estimate of the population parameter for each characteristic can be treated as a mathematical programming problem and stated as follows

$$\begin{aligned} &\text{Minimize } C(n) = c_0 + \sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} t_h \sqrt{n_n} \\ &\text{Subject to } V(\bar{y}_{jst}) \le v_j^{\circ} \ \forall\ j = 1,2,3,\dots,p \\ &\text{and} \quad 2 \le n_h \le N_h ; \end{aligned} \tag{3}$$

We considered that the costs of measurement $c_h\ (h = 1,2,3,\dots,L)$ and the travel costs $t_h\ (h = 1,2,3,\dots,L)$ are normally distributed random variables and the RHS of the j^{th} constraint has k_j number of choices and out of these choices one must be selected by respective constraint with a specified probability p_j°.

Our assumption is the consideartion of quadratic cost function as a bi-objective function then the problem of optimum allocation as a SBOPP will be

$$\text{Minimize } C(n) = \{C_1(n_h), C_2(n_h)\} = \left\{ \sum_{h=1}^{L} c_h n_h, \sum_{h=1}^{L} t_h \sqrt{n_n} \right\}$$

$$\text{Subject to } P\left[\widehat{\text{V}}\text{ar}(\bar{y}_{jst}) \le v_j^\circ \right] \ge p_j^0 \ \forall \ j = 1,2,3,\dots,p \text{ ; and } v_j^0 \in \left\{ v_j^{(1)}, v_j^{(2)}, v_j^{(3)}, \dots, v_j^{(k_j)} \right\}$$

$$2 \le n_h \le N_h \text{ ;} \qquad (4)$$

where $\widehat{\text{V}}\text{ar}(\bar{y}_{jst}) = \sum_{h=1}^{L} \frac{W_h^2 s_{jh}^2}{n_h} - \sum_{h=1}^{L} \frac{W_h^2 s_{jh}^2}{N_h}$

is estimated variance for j^{th} characteristic. The term c_0 is removed because overhead cost is not the part of optimization Kokan (1963).
The expression of variance function $\widehat{\text{V}}\text{ar}(\bar{y}_{jst})$ in the problem (4) indicates that we are using the estimated variance at the place of population variance which is unknown, Diaz-Garcia and Gary-Tapia (2007).

3. SOLUTION METHODS

Suppose that the mean and variance of the normally distributed costs c_h and t_h are as follows

$$\text{mean } E(c_h) = \bar{c}_h \text{ and variance } Var(c_h) = \sigma_{c_h}^2 \text{ ;}$$

$$\text{mean } E(t_h) = \bar{t}_h \text{ and variance } Var(t_h) = \sigma_{t_h}^2 \text{ ;}$$

$$i.e. \ c_h \sim N\left(\bar{c}_h, \sigma_{c_h}^2\right) \text{ and } t_h \sim N\left(\bar{t}_h, \sigma_{t_h}^2\right)$$

If we have the quadratic cost function as given in the equation (2) with random parameters, then optimum allocation can be obtained by solving its deterministic equivalent of the objective function in the form

$$\text{Minimize } k_1 \bar{C} + k_2 \sqrt{\text{Var}\,(\bar{C})} \text{ ;}$$

where k_1 and k_2 are the non-negative constants whose values indicate the relative importance of mean value $\bar{C}$ and standard deviation $\sqrt{\text{Var}(\bar{C})}$ (see Rao (1978)). Without loss of generality we can take $k_1 + k_2 = 1$.

In our case, we have a bi-objective problem with random parameters so the existing technique of single objective mathematical programming problem can not be applied directly. So, our first step is to convert it into a single objective problem with the help of weighted sum method and secondly we remove the randomness by applying the expected value criterion. In order to solve the single objective problem an EVSD criterion developed by Bayoumi *et al.*(2005) is applied. The single objective function by weighted sum method will be

$$\text{Minimize} \quad \begin{aligned} C(n) &= \left\{\lambda C_1(n_h) + (1-\lambda) C_2(n_h)\right\} \\ &= \left\{\lambda\left(\sum_{h=1}^{L} c_h n_h\right) + (1-\lambda)\left(\sum_{h=1}^{L} t_h \sqrt{n_n}\right)\right\} \end{aligned} \tag{5}$$

3.1 Expected Value Standard Deviation (EVSD) Criteria

The deterministic equivalent of the objective function given in equation (5) can be obtained by applying the expected value criterion. The expected value of the objective function will be

$$\begin{aligned} E\left(C(n)\right) &= E\left\{\lambda C_1(n_h) + (1-\lambda) C_2(n_h)\right\} = E\left\{\lambda\left(\sum_{h=1}^{L} c_h n_h\right) + (1-\lambda)\left(\sum_{h=1}^{L} t_h \sqrt{n_n}\right)\right\} \\ &= \lambda\left(\sum_{h=1}^{L} \bar{c}_h n_h\right) + (1-\lambda)\left(\sum_{h=1}^{L} \bar{t}_h \sqrt{n_n}\right) \end{aligned} \tag{6}$$

By using the EVSD criterion the objective function to be minimized will be

$$\text{Minimize}\left\{E\left(C(n)\right) + \sqrt{\text{Var}\left(C(n)\right)}\right\} \tag{7}$$

where the expected value of $C(n)$ is given in equation (6) and the variance function can be calculated as follows

$$\begin{aligned} \text{Var}\left(C(n)\right) &= \text{Var}\left\{\lambda\left(\sum_{h=1}^{L} c_h n_h\right) + (1-\lambda)\left(\sum_{h=1}^{L} t_h \sqrt{n_n}\right)\right\} \\ &= \lambda^2 \text{Var}\left(\sum_{h=1}^{L} c_h n_h\right) + (1-\lambda)^2 \text{Var}\left(\sum_{h=1}^{L} t_h \sqrt{n_n}\right) \\ &\quad + 2\lambda(1-\lambda)\text{Cov}\left(\left(\sum_{h=1}^{L} c_h n_h\right), \left(\sum_{h=1}^{L} t_h \sqrt{n_n}\right)\right) \end{aligned}$$

Consider the variance and covariance terms

$$\mathrm{Var}\left(\sum_{h=1}^{L} c_h n_h\right) = \mathrm{Var}(c_1 n_1 + c_2 n_2 + ... + c_L n_L) = \sum_{h=1}^{L} n_h^2 \sigma_{c_h}^2$$

$$\mathrm{Var}\left(\sum_{h=1}^{L} t_h \sqrt{n_n}\right) = \mathrm{Var}(t_1\sqrt{n_1} + t_2\sqrt{n_2} + ... + t_L\sqrt{n_L}) = \sum_{h=1}^{L} n_h \sigma_{t_h}^2 .$$

The covariance term will vanish as we assume that the costs are independentally distributed random variables. Thus the variance function will take the from as

$$\mathrm{Var}\ (C(n)) = \lambda^2 \left(\sum_{h=1}^{L} n_h^2 \sigma_{c_h}^2\right) + (1-\lambda)^2 \left(\sum_{h=1}^{L} n_h \sigma_{t_h}^2\right) \quad (8)$$

Substituting these values of $E(C(n))$ and $\mathrm{Var}\ (C(n))$, our objective will become

$$\text{Minimize}\left\{E(C(n)) + \sqrt{\mathrm{Var}\ (C(n))}\right\} = \lambda\left(\sum_{h=1}^{L} \bar{c}_h n_h\right) + (1-\lambda)\left(\sum_{h=1}^{L} \bar{t}_h \sqrt{n_n}\right)$$
$$+\sqrt{\lambda^2 \left(\sum_{h=1}^{L} n_h^2 \sigma_{c_h}^2\right) + (1-\lambda)^2 \left(\sum_{h=1}^{L} n_h \sigma_{t_h}^2\right)} \quad ......(9)$$

4. CONVERSION OF CHANCE CONSTRAINT INTO ITS DETERMINISTIC EQUIVALENT AND MULTI CHOICE RHS INTO STANDARD CONSTRAINT

The deterministic equivalent of the constraint having random variables is obtianed by the chance constraint programming technique and the RHS of the constraint is transformed into as standard mathematical programminh problem by using the technique developed by Acharya and Biswal (2011).

Diaz-Garcia and Gary-Tapia (2007), considered the univariate case of stratified sampling. Our problem is multivariate, so the deterministic equivalent of the constraints

$$P\left[\hat{\mathrm{V}}\mathrm{ar}(\bar{y}_{jst}) \le v_j^{\circ}\right] \ge p_j^0;\ v_j^0 \in \left\{v_j^{(1)}, v_j^{(2)}, v_j^{(3)}, ..., v_j^{(k_j)}\right\}$$

will be

$$E\left(\hat{\mathrm{V}}\mathrm{ar}(\bar{y}_{jst})\right) + K_j \sqrt{\mathrm{Var}\left(\hat{\mathrm{V}}\mathrm{ar}(\bar{y}_{jst})\right)} \le v_j^{\circ};\ v_j^0 \in \left\{v_j^{(1)}, v_j^{(2)}, v_j^{(3)}, ..., v_j^{(k_j)}\right\} \quad (10)$$

where

$$E\left(\hat{\mathrm{V}}\mathrm{ar}(\bar{y}_{jst})\right)+K_j\sqrt{\mathrm{Var}\left(\hat{\mathrm{V}}\mathrm{ar}(\bar{y}_{jst})\right)}$$
$$=\sum_{h=1}^{L}\frac{W_h^2}{(n_h-1)}S_{jh}^2-\sum_{h=1}^{L}\frac{W_h}{N}\left(\frac{n_h}{n_h-1}\right)S_{jh}^2$$
$$+K_j\left[\sum_{h=1}^{L}\frac{W_h^4}{n_h(n_h-1)^2}\left(C_{jyh}^4-(S_{jh}^2)^2\right)-\sum_{h=1}^{L}\frac{W_h^2}{N^2}\left(\frac{n_h}{(n_h-1)^2}\left(C_{jyh}^4-(S_{jh}^2)^2\right)\right)\right]^2;$$

and $C_{jyh}^4=\frac{1}{N_h}\sum_{i=1}^{N_h}(y_{hij}-\bar{Y}_{hj})^4$ are the fourth moments about stratum means for each characteristic.

The symbol 'K_j' stands for the value of standard normal random variable such that $\Phi(K_j)=p_j^0$, in such a way that the inequality can be established as

$$\Phi\left(\frac{v_j^0-E\left(\hat{\mathrm{V}}\mathrm{ar}(\bar{y}_{jst})\right)}{\sqrt{\mathrm{Var}\left(\hat{\mathrm{V}}\mathrm{ar}(\bar{y}_{jst})\right)}}\right)\geq\Phi(K_j),$$

which holds only if

$$\frac{v_j^0-E\left(\hat{\mathrm{V}}\mathrm{ar}(\bar{y}_{jst})\right)}{\sqrt{\mathrm{Var}\left(\hat{\mathrm{V}}\mathrm{ar}(\bar{y}_{jst})\right)}}\geq K_j\,.$$

As we considered multi-choice precision for each characteristic, so the transformation of the constraints into its standard form will be in the following manner.

Case (i) If $k_j=1$, then the constraint will be same as an ordinary constraint.

Case (ii) If $k_j=2$, then the constraint will be

$$E\left(\hat{\mathrm{V}}\mathrm{ar}(\bar{y}_{jst})\right)+K_j\sqrt{\mathrm{Var}\left(\hat{\mathrm{V}}\mathrm{ar}(\bar{y}_{jst})\right)}\leq\left\{v_j^{(1)},v_j^{(2)}\right\}$$

Out of these two goals, one must be selected. Since the total number of the elements in the set is 2, one binary variable $z_j^{(1)}$ is required.

Introducing binary variable the constraint will be

$$E\left(\hat{\mathrm{V}}\mathrm{ar}(\bar{y}_{jst})\right)+K_j\sqrt{\mathrm{Var}\left(\hat{\mathrm{V}}\mathrm{ar}(\bar{y}_{jst})\right)}\leq z_j^{(1)}v_j^{(1)}+\left(1-z_j^{(1)}\right)v_j^{(2)}$$

$$0\leq z_j^{(1)}\leq 1$$

Case (iii) If $k_j = 3,$ then the constraint will be

$$E\left(\widehat{\text{V}}\text{ar}(\overline{y}_{jst})\right) + K_j\sqrt{\text{Var}\left(\widehat{\text{V}}\text{ar}(\overline{y}_{jst})\right)} \leq \left\{v_j^{(1)}, v_j^{(2)}, v_j^{(3)}\right\}$$

Out of these three goals, one must be selected.

Since $2^1 < k_j < 2^2$, two binary variables $z_j^{(1)}$ and $z_j^{(2)}$ are required. So 3 can be expressed as

$$\binom{2}{2}+\binom{2}{1} \text{ or } \binom{2}{1}+\binom{2}{0}$$

Hence there will be restriction on remaining one (i.e., $4-3$) term by introducing an additional constraint in the problem (10).

In this case two models are formulated with the help of two binary variables $z_j^{(1)}$ and $z_j^{(2)}$ in this manner.

Model (a)

$$E\left(\widehat{\text{V}}\text{ar}(\overline{y}_{jst})\right) + K_j\sqrt{\text{Var}\left(\widehat{\text{V}}\text{ar}(\overline{y}_{jst})\right)}$$

$$\leq \left(1-z_i^{(1)}\right)\left(1-z_i^{(2)}\right)v_j^{(1)} + \left(1-z_i^{(1)}\right)z_i^{(2)}v_j^{(2)} + z_i^{(1)}\left(1-z_i^{(2)}\right)v_j^{(3)}$$

$$z_j^{(1)} + z_j^{(2)} \leq 1;\, 0 \leq z_j^{(1)} \leq 1;\, 0 \leq z_j^{(2)} \leq 1.$$

Model (b)

$$E\left(\widehat{\text{V}}\text{ar}(\overline{y}_{jst})\right) + K_j\sqrt{\text{Var}\left(\widehat{\text{V}}\text{ar}(\overline{y}_{jst})\right)}$$

$$\leq \left(1-z_j^{(1)}\right)z_j^{(2)}v_j^{(1)} + z_j^{(1)}\left(1-z_j^{(2)}\right)v_j^{(2)} + z_j^{(1)}z_j^{(2)}v_j^{(3)}$$

$$z_j^{(1)} + z_j^{(2)} \geq 1;\, 0 \leq z_j^{(1)} \leq 1;\, 0 \leq z_j^{(2)} \leq 1;$$

Finally, by using EVSD criteria for the objective function and by applying transformation techniques discussed above, we get the following problem to be solved

$$\text{Minimize}\left\{E\left(C(n)\right) + \sqrt{\text{Var }\left(C(n)\right)}\right\} = \lambda\left(\sum_{h=1}^{L}\overline{c}_h n_h\right) + (1-\lambda)\left(\sum_{h=1}^{L}\overline{t}_h\sqrt{n_n}\right) + \sqrt{\lambda^2\left(\sum_{h=1}^{L}n_h^2\sigma_{c_h}^2\right) + (1-\lambda)^2\left(\sum_{h=1}^{L}n_h\sigma_{t_h}^2\right)}\Bigg\}$$

Subject to

$$E\left(\widehat{\mathrm{V}}\mathrm{ar}(\overline{y}_{jst})\right)+K_j\sqrt{\mathrm{Var}\left(\widehat{\mathrm{V}}\mathrm{ar}(\overline{y}_{jst})\right)}\le\left\{v_j^{(k_j)}\right\};$$

and $2\le n_h\le N_h$; n_h must be integer. (11)

Where k_j is the number of choice for the precision of j^{th} characteristic. For each value of λ, we get efficient solutions.

It is also assumed that the population variance is unknown and replaced by the sample variance. So the constraint in problem (10) will be

$$\sum_{h=1}^{L}\frac{W_h^2}{(n_h-1)}s_{jh}^2-\sum_{h=1}^{L}\frac{W_h}{N}\left(\frac{n_h}{n_h-1}\right)s_{jh}^2$$

$$+K_j\left[\sum_{h=1}^{L}\frac{W_h^4}{n_h(n_h-1)^2}\left(C_{jyh}^4-(s_{jh}^2)^2\right)-\sum_{h=1}^{L}\frac{W_h^2}{N^2}\left(\frac{n_h}{(n_h-1)^2}\left(C_{jyh}^4-(s_{jh}^2)^2\right)\right)\right]^2\le\left\{v_j^{(k_j)}\right\}.$$

Bayoumi *et al.*(2005) also suggested that if we apply expected value standard deviation criterion for SBOPP and convert the bi-objective problem into single objective problem, then we get the set of non-dominated and efficient soltutions by gradually increasing the value of λ. In this way, our objective will be

$$\text{Min C}(n)=\lambda\left(\sum_{h=1}^{L}\overline{c}_h n_h+\sqrt{\left(\sum_{h=1}^{L}n_h^2\sigma_{c_h}^2\right)}\right)$$

$$+(1-\lambda)\left(\sum_{h=1}^{L}\overline{t}_h\sqrt{n_n}+\sqrt{\left(\sum_{h=1}^{L}n_h\sigma_{t_h}^2\right)}\right)$$

Subject to

$$\sum_{h=1}^{L}\frac{W_h^2}{(n_h-1)}s_{jh}^2-\sum_{h=1}^{L}\frac{W_h}{N}\left(\frac{n_h}{n_h-1}\right)s_{jh}^2$$

$$+K_j\left[\sum_{h=1}^{L}\frac{W_h^4}{n_h(n_h-1)^2}\left(C_{jyh}^4-(s_{jh}^2)^2\right)-\sum_{h=1}^{L}\frac{W_h^2}{N^2}\left(\frac{n_h}{(n_h-1)^2}\left(C_{jyh}^4-(s_{jh}^2)^2\right)\right)\right]^2\le\left\{v_j^{(k_j)}\right\}$$

and $2\le n_h\le N_h$; n_h must be integer. (12)

5. NUMERICAL ILLUSTRATION

For the purpose of numerical illustration, the data given in table 1, are taken from Diaz-Garcia and Gary-Tapia (2007), and modified to as our requirement i.e. for multivariate survey. Suppose we have three characters under study and the whole population is divided into four strata.

We assume that the costs are random variables and unknown population variances are replaced by sample variance.

The mean and variance of measurement costs $c_h\ (h=1,2,3,...,L)$ are assumed as follows: $E(c_1)=25$, $E(c_2)=23$, $E(c_3)=28$, $E(c_4)=30$

$$\text{Var}(c_1)=30\,,\ \text{Var}(c_2)=25\,,\ \text{Var}(c_3)=34\,,\ \text{Var}(c_4)=32\,.$$

Table 1: Stratum Weights, Sample Variances, fourth moments for the three characters under study

h	N_h	W_h	s_{1h}^2	s_{2h}^2	s_{3h}^2	C_{1h}^4	C_{2h}^4	C_{3h}^4
1	2500	0.24	0.1694	0.1969	0.1496	0.0884	0.0848	0.0799
2	2300	0.22	8.4317	7.7431	7.3417	330.4106	310.6041	320.1604
3	2800	0.26	0.0972	0.0792	0.0827	0.0319	0.0400	0.0391
4	3000	0.28	3.8590	4.5809	4.8950	34.1001	36.2314	35.0099

Similarly the mean and varaiance of travel costs $t_h\ (h=1,2,3,...,L)$ are assumed as follows

$$E(t_1)=15\,,\ E(t_2)=13\,,\ E(t_3)=18\,,\ E(t_4)=20 \text{ and}$$

$$\text{Var}(t_1)=20\,,\ \text{Var}(t_2)=25\,,\ \text{Var}(t_3)=24\,,\ \text{Var}(t_4)=15$$

We also assume the multi choice nature of precision i.e. each characteristic has a set of choices with acceptable precisions. These choices for each characteristic will be as follows

$$V(\bar{y}_{1st})\le v_1^{\circ},\ v_1^{\circ}\in\left\{v_1^{(1)},v_1^{(2)}\right\}=\{0.015,0.016\};$$

$$V(\bar{y}_{2st})\le v_2^{\circ},\ v_2^{\circ}\in\left\{v_2^{(1)},v_2^{(2)},v_2^{(3)}\right\}=\{0.018,0.019,0.020\};$$

$$V(\bar{y}_{3st})\le v_3^{\circ},v_3^{\circ}\in\left\{v_3^{(1)},v_3^{(2)},v_3^{(3)}\right\}=\{0.014,0.017,0.018\}$$

The constraint satisfying probability $p_j^0=0.99\ \forall\ j=1,2,3.$

It is supposed that we need a sampling plan to minimize both the costs simultaneously and which ensures that the estimate for each characteristic can take only one value from the set of variance specified as choices for each characteristic.

After substituting these values in the problem given in (11) we solve the resulting problem with the help of LINGO software package (2011). By gradually increasing the value of λ , we get the results given in table 2.

Table 2: Total cost of the survey for different values of λ

λ	n_1	n_2	n_3	n_4	z_1^1	z_2^1	z_2^2	z_3^1	z_3^2	Cost C(n)
0.0	08	67	06	145	0	1	0	1	0	498.21
0.1	11	72	07	132	0	1	0	1	0	1103.91
0.2	11	70	09	131	0	1	0	1	0	1739.15
0.3	11	70	09	131	0	1	0	0	1	2380.21
0.4	12	73	09	128	0	1	0	1	0	3022.78
0.5	12	73	09	128	0	1	0	0	0	3666.05
0.6	12	73	09	128	0	1	0	1	0	4309.74
0.7	12	73	09	128	0	1	0	0	0	4953.68
0.8	12	73	09	128	0	1	0	0	1	5597.77
0.9	12	73	09	128	0	1	0	1	0	6241.96
1.0	12	73	09	128	0	1	0	1	0	6886.22

Table 3. Total cost of the survey for different values of k_1 and k_2

k_1	k_2	n_1	n_2	n_3	n_4	z_1^1	z_2^1	z_2^2	z_3^1	z_3^2	Cost C(n)
0.0	1.0	20	76	18	120	0	1	0	0	1	795.45
0.1	0.9	16	73	10	125	0	1	0	0	0	1380.86
0.2	0.8	14	73	10	126	0	1	0	1	0	1954.87
0.3	0.7	14	71	09	128	0	1	0	1	0	2526.32
0.4	0.6	12	73	09	128	0	1	0	1	0	3096.40
0.5	0.5	12	73	09	128	0	1	0	1	0	3666.05
0.6	0.4	11	70	09	131	0	1	0	0	0	4235.43
0.7	0.3	11	70	09	131	0	1	0	1	0	4803.68
0.8	0.2	11	70	09	131	0	1	0	0	0	5371.93
0.9	0.1	11	70	09	131	0	1	0	0	1	5940.18
1.0	0.0	11	70	09	131	0	1	0	0	0	6508.43

If we consider the single objective (quadratic cost function as given in equation (2)) at the place of bi-objective cost then the objective will be to minimize the function

$$C(n) = k_1\bar{C} + k_2\sqrt{\text{Var}\,(\bar{C})} \quad = k_1\left(\sum_{h=1}^{L}\bar{c}_h n_h + \sum_{h=1}^{L}\bar{t}_h\sqrt{n_n}\right) + k_2\left(\sqrt{\left(\sum_{h=1}^{L} n_h^2\sigma_{c_h}^2 + \sum_{h=1}^{L} n_h\sigma_{t_h}^2\right)}\right)$$

Subject to the constraints given in equation (11).

After solving the allocation problem with this objective function, at different values of k_1 and k_2, we get the results given in Table 3.

DISCUSSION

From the results presented in Table 2; it can be observed that the total cost is increasing with the increment in λ and if we are giving equal weights (i.e. $\lambda = 0.5$) to both the objectives (measurement cost and travel cost) then by applying the EVSD criteria the incurred cost is 3666.05. In the similar manner if we consider the single objective (i.e. quadratic cost function) then for different values of k_1 and k_2 the cost is inreasing as k_1 increases and at $k_1 = k_2 = 0.5$; the incurred cost is 3666.05. This result shows that if we are giving equal preference to the mean value of cost function and the standard deviation of the cost function then the cost is same as for bi-objective consideration for equal weights. One important thing to be noticed here is that in the bi-objective case we are giving weights to the objective functions and in the single objective case we are giving the weights to the mean value and the standard deviation of the cost function. Thus, these results indicate that if the costs are random then bi-objective cost function can be used in the place of quadratic cost function.

On the other hand, regarding the precision of the estimate, we have the allocation scheme as:

$$n_1 = 12,\ n_2 = 73,\ n_3 = 9,\ n_4 = 128$$

in both cases i.e., in single objective at $\lambda = 0.5$ and bi-objective at $k_1 = k_2 = 0.5$. The values of variances with this allocation scheme are

$$\bar{y}_{1st} = 0.015\,,\ \bar{y}_{2st} = 0.018\,,\ \bar{y}_{3st} = 0.014\,.$$

These variances ensure that the estimate of each characteristic takes exactly one value from the set of choices specified as the precision.

6. CONCLUSION

The problem of optimal allocation to minimize both the costs (measurement and travel) simultaneously is formulated as a bi-objective programming problem. The unitary cost of measurement and travel are considered as random variables so both the objectives become stochastic and the problem of optimal allocation reduces to as a stochastic bi-objective programming problem. Firstly the bi-objective problem is converted into a single objective problem and then EVSD criterion is applied to solve the problem. A numerical illustration is also given for the purpose of demonstration. A comparison of the obtained results is done with the results of single objective quadratic cost function and we conclude that the problem of optimum allocation to minimize the quadratic cost can be treated as a bi-objective programming problem. In future, this approach can be applied in the existing literature of quadratic cost function in order to derive the managerial insights.

REFERENCES

Acharya, S. and Biswal, M.P. (2011). Solving probabilistic programming problems involving multi-choice parameters. *OPSEARCH*, 48: 217-235.

Ali, I., Raghav, Y.S., Bari, A. (2011).Compromise allocation in multivariate stratified surveys with stochastic quadratic cost function. *Journal of Statistical Computation and Simulation*, 83(5):962-976.

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, 1(2):111–120.

Bayoumi, B. I., El-Sawy, A. A., Baseley, N. L., Yousef, I.K., Widya, A.M. (2005). Determining the efficient solutions for bi-criteria programming problems with random variables in both the objective functions and the constraints. *J. KSIAM*, 9(1): 99-110.

Beardwood , J. , Halton , J. H. , Hammersley , J. M. (1959). The shortest path through many points. *Proc. Cambridge Phil. Soc.*, 55: 299 – 327.

Cochran, W.G. (1977). *Sampling Techniques (Third Edition).* John Wiley and Sons, New York.

Díaz-García, J.A. and Garay-Tapia, Ma. M. (2007). Optimum allocation in stratified surveys: Stochastic programming. *Computational Statistics & Data Analysis*, 51: 3016-3026.

Fatima, U., Varshney, R., Siddiqui, N., Ahsan, M.J. (2014). On compromise mixed allocation in multivariate stratified sampling with random parameters. *J Math Model Algor*, 13(4), 523-536.

Ghufran, S., Khowaja, S., Ahsan, M.J. (2012).Optimum multivariate stratified sampling designs with travel cost: A multi-objective integer nonlinear programming approach. *Communications in Statistics - Simulation and Computation*,41(5):598-610.

Hansen, M. H., Hurwitz, W.N. and Madow, W.G. (1953). *Sample Survey Methods and Theory*, Volume I, New York: John Wiley and Sons, Inc..

Javaid, S., Bakhshi Z. H., Khalid, M.M. (2011). Use of stochastic programming for sample allocation in two-stage stratified sampling. *International Journal of Agricultural and Statistical Sciences*, 7(2): 379-392.

Khan, S. and Khalid, M.M. (2013). Multi-choice for precision in multivariate stratified surveys: A compromise solution. *International Journal of Operations Research*, 10(4):171-181.

Khowaja, S., Ghufran, S., Ahsan, M.J. (2012). Multi-objective optimization for optimum allocation in multivariate stratified sampling with quadratic cost. *Journal of Statistical Computation and Simulation*, 82(12):1789-1798.

Kokan, A. R. (1963). Optimum allocation in multivariate surveys. *J. Roy. Statist. Soc. Ser., A*, 126: 557-565.

LINGO (2011). User’s Guide, Lindo Systems Inc., Chicago, IL, USA.

Melaku, A. (1986). Asymptotic normality of the optimal allocation in multivariate stratified random sampling. *Sankhyā: The Indian Journal of Statistics, Series B*, 48(2): 224-232.

Rao, S.S. (1978). *Optimization: Theory and Applications*, John & Wiley.

Sukhatme, P.V., Sukhatme, B.V., S. Sukhatme, and Ashok, C. (1984). *Sampling Theory of Surveys with Applications*, 3rd ed., Iowa State University Press and Indian Society of Agricultural Statistics, Ames, IW and New Delhi.

Shamsher Khan* and M. M. Khalid
Department of Statistics & Operations Research
Aligarh Muslim University,
Aligarh-202002, India.

E-mail**: shamsherstats@gmail.com*

Received: 15.11.2014

Revised: 03.01.2015