

# Compromise Allocation for Combined Ratio Estimates of Population Means of a Multivariate Stratified Population Using Double Sampling in Presence of Non-Response

# Sana Iftekhar, Qazi Mazhar Ali, Mohammad Jameel Ahsan

Department of Statistics & Operations Research, Aligarh Muslim University, Aligarh, India Email: <u>iftekhar.sana54@gmail.com</u>

Received 8 September 2014; revised 16 October 2014; accepted 12 November 2014

Copyright © 2014 by authors and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY). http://creativecommons.org/licenses/by/4.0/

# Abstract

This paper is an attempt to work out a compromise allocation to construct combined ratio estimates under multivariate double sampling design in presence of non-response when the population mean of the auxiliary variable is unknown. The problem has been formulated as a multi-objective integer non-linear programming problem. Two solution procedures are developed using goal programming and fuzzy programming techniques. A numerical example is also worked out to illustrate the computational details. A comparison of the two methods is also carried out.

# **Keywords**

Multivariate Stratified Sampling, Compromise Allocation, Non-Response, Double Sampling Goal Programming, Fuzzy Programming

# **1. Introduction**

Often in sample surveys the main variable is highly correlated to another variable called an auxiliary variable and the data on auxiliary variable are either available or can be easily obtained. In this situation to obtain the estimate of the parameters regarding the main variable the auxiliary information can be used to enhance the precision of the estimate. Ratio and Regression Methods and double sampling technique are some examples. When data are collected on the sampled units of the main variable due to one or the other reason, data for all the se-

How to cite this paper: Iftekhar, S., Ali, Q.M. and Ahsan, M.J. (2014) Compromise Allocation for Combined Ratio Estimates of Population Means of a Multivariate Stratified Population Using Double Sampling in Presence of Non-Response. *Open Journal of Optimization*, **3**, 68-78. <u>http://dx.doi.org/10.4236/ojop.2014.34007</u>

lected units cannot be obtained. This result is an incomplete and less informative sample. This phenomenon is termed as "non response". [1] is the first one to consider this problem. Furthermore, when auxiliary parameters are unknown, they can be estimated from a preliminary large sample. Then a second sample is obtained in which the main and auxiliary, both the variables are measured. Often a second sample is a subsample of the first. In such cases only the main variable is to be measured in the second sample. This technique is called "Double Sampling" or "Two Phase Sampling", [2]-[11] are some who used the auxiliary information in sample surveys. [10] has worked on the problem in which ratio estimator has been considered for population mean under double sampling in presence of non-response for a univariate population.

In the present paper, we considered combined ratio estimators of the population means of a multivariate stratified population using double sampling in presence of non-response. Compromise allocations at first and second phase of double sampling are obtained by formulating the problems as multi-objective integer non-linear programming problems. Solution procedures are developed by using goal programming and fuzzy programming techniques. A numerical example is also worked out to illustrate the computational details. A comparison of the two methods is also carried out.

When auxiliary information is available, the use of Ratio method of estimation is well known in univariate stratified sampling. Formulae are also available to work out optimum allocations to various strata [12]. In multivariate case finding an allocation that gives optimum results for all the characteristics is not possible due to the conflicting nature of the characteristics. Compromise allocation is used in such situations. Furthermore, if the problem of non-response is also there, the situation becomes more complicated. The paper is structured as below:

In Section 2 of the manuscript combined ratio estimates for the population means of the "p" characteristics in presence of non-response using double sampling are constructed. Section 3 formulates the problem of obtaining compromise allocations for phase-I and phase-II of the double sampling as an integer nonlinear programming problem (INLPP). Sections 4 and 5 show that how these INLPP's can be transformed to apply the Goal Programming Technique (GPT) and the Fuzzy Programming Technique (FPT) to solve the transformed problems. Section 6 provides an application of the techniques through a numerical data. In the last Section 8 gives the conclusion and the future work trend for interested readers.

#### 2. The Combined Ratio Estimate in Multivariate Stratified Double Sampling Design in Presence of Non-Response

Consider a multivariate stratified population of size N with L non-overlapping strata of sizes  $N_1, N_2, \dots, N_L$ with  $\sum_{h=1}^{L} N_h = N$ . Let p characteristics be defined on each unit of the population. If  $N_1, N_2, \dots, N_L$  are not known in advance then the strata weights  $W_h = N_h/N, h = 1, 2, \dots, L$  also remain unknown. In such situation double sampling technique may be used to estimate the unknown strata weights. For this a large preliminary simple random sample of size n' is obtained at the first phase of the double sampling, treating the population as unstratified. The number of sampled units  $n'_h; h = 1, 2, \dots, L$  falling in each stratum is recorded. The quantity  $w_h = n'_h/n'$  will give an unbiased estimate of  $W_h$ . Simple random subsamples, without replacement of sizes  $n_h = \mu_h n'_h; h = 1, 2, \dots, L; 0 < \mu_h \le 1$  are then drawn out of  $n'_h$  from each stratum for values of  $\mu_h$  chosen in advance.

For the *jth* characteristics and the *hth* stratum denote by

 $y_{ihi}$  the value of the *ith* population (sample) unit of the main variable.

 $x_{ihi}$  the value of the *ith* population (sample) units of the auxiliary variable.

 $\overline{Y}_{jh} = \frac{1}{N_h} \sum_{i=1}^{N_h} y_{jhi}$  and  $\overline{y}_{jh} = \frac{1}{n_h} \sum_{i=1}^{n_h} y_{jhi}$  the stratum mean and the sample mean respectively for the main

variable.

$$\overline{X}_{jh} = \frac{1}{N_h} \sum_{i=1}^{N_h} X_{jhi}$$
 and  $\overline{x}_{jh} = \frac{1}{n_h} \sum_{i=1}^{n_h} y_{jhi}$  denote the same values for the auxiliary variable.

In double sampling for stratification the combined ratio estimate of the population mean of the *jth* characteristics is given as

$$\overline{y}_{j(CRDS)} = \sum_{h=1}^{L} w_h \overline{r}_{jh}; \quad j = 1, 2, \cdots, p$$
(1)

where "*CR*" and "*DS*" stand for "combined ratio" and "double sampling" respectively. Further,

$$\overline{r}_{jh} = \hat{R}_{jst} \overline{X}_{j}, \qquad \hat{R}_{jst} = \frac{\overline{y}_{jst}}{\overline{x}_{jst}}$$
$$\overline{y}_{js} = \sum_{h=1}^{L} w_h \overline{y}_{jh}$$
$$\overline{x}_{js} = \sum_{h=1}^{L} w_h \overline{x}_{jh}$$

The sampling variance of  $\overline{y}_{j(CRDS)}$  is

$$V(\overline{y}_{j(CRDS)}) = \left(\frac{1}{n'} - \frac{1}{N}\right) S_{yj}^{2} + \frac{1}{n'} \sum_{h=1}^{L} w_{h} \left(\frac{1}{\mu_{h}} - 1\right) S_{rjh}^{2}$$
(2)

 $S_{rih}^2$  in (2) is defined as

$$S_{rjh}^{2} = S_{yjh}^{2} + R_{j}^{2}S_{xjh}^{2} - 2R_{j}S_{x_{j}y_{j}h}$$
(3)

where  $R_i$  are true population ratios given as

$$R_j = \frac{\overline{Y}_j}{\overline{X}_j}$$

 $S_{yjh}^2, S_{xjh}^2$  are the stratum variances of the *jth* characteristics in the *hth* stratum for main variable and auxiliary variables respectively and  $S_{x_jy_jh}$  are the stratum co-variances of the *jth* characteristics in the *hth* stratum for  $j = 1, 2, \dots, p; h = 1, 2, \dots, L$ .

In the presence of non-response, let out of the  $n_h$  units  $n_{h1}$  units respond at the first call and  $n_{h2} = n_h - n_{h1}$  units constitute the non-respondents group. Using [1], a subsample of size  $m_{h2} = k_h^* n_{h2}; 0 < k_h^* \le 1$  out of  $n_{h2}$  is drawn and interviewed with extra efforts. Where  $k_h^*; h = 1, 2, \dots, L$  are fixed in advance.

An combined ratio estimate  $\overline{y}_{i(CRDS)}^*$  of  $\overline{Y}_j$  may be given as

$$\overline{y}_{j(CRDS)}^{*} = \sum_{h=1}^{L} w_h \overline{r}_{jh}^{*}$$
(4)

where

$$\overline{r}_{jh}^* = \frac{n_{h1}\overline{r}_{jh1} + n_{h2}\overline{r}_{jm_{h2}}}{n_h}$$

 $\overline{r}_{jh1}$ ,  $\overline{r}_{jm_{h2}}$  are the sample mean of the ratio estimates for respondents (based on  $n_{h1}$  units) and non-respondents group (based on  $m_{h2}$  units) respectively.

Using the results presented in [12]—Sections 5A.2, 12.9 and 13.6 we get  $V(\overline{y}_{j(CRDS)}^*)$  in presence of non-response as

$$V\left(\overline{y}_{j(CRDS)}^{*}\right) = V\left(\overline{y}_{j(CRDS)}\right) + \frac{1}{n'} \sum_{h=1}^{L} w_{h2} \left(\frac{1-k_{h}^{*}}{k_{h}^{*}\mu_{h}}\right) S_{2rjh}^{2}$$

$$= \left(\frac{1}{n'} - \frac{1}{N}\right) S_{yj}^{2} + \frac{1}{n'} \sum_{h=1}^{L} w_{h} \left(\frac{1}{\mu_{h}} - 1\right) S_{rjh}^{2} + \frac{1}{n'} \sum_{h=1}^{L} w_{h2} \left(\frac{1-k_{h}^{*}}{k_{h}^{*}\mu_{h}}\right) S_{2rjh}^{2}$$
(5)

where,

$$S_{2rjh}^2 = S_{2yjh}^2 + R_j^2 S_{2xjh}^2 - 2R_j S_{2x_j y_j h}$$

 $S_{2yjh}^2, S_{2xjh}^2$  are the stratum variances of the *jth* characteristics of the non-respondents in the *hth* stratum for main variable and auxiliary variable respectively.  $S_{2x_jy_jh}$  is the stratum co-variances of the *jth* characteristics of the non-respondents in the *hth* stratum [11].

The total cost of the survey may be given

$$C = c_0 n' + \sum_{h=1}^{L} c_{h1} n_h + \sum_{h=1}^{L} c_{h11} n_{h1} + \sum_{h=1}^{L} c_{h12} m_{h2}$$
(6)

where,

 $c_0$  is the per unit cost of getting information from the preliminary sample n'.

 $c_{h1}$  is the per unit cost of of making the first attempt (Phase I).

 $c_{h11}$  is the per unit cost of processing and analyzing the result of all the *p* characteristics on the  $n_{h1}$  respondents units in the *hth* stratum at Phase I.

 $c_{h12}$  is the per unit cost of measuring and processing the result of all the *p* characteristics on the  $m_{h2}$  subsampled units from non-respondents group in the *hth* stratum at Phase II.

Since  $n_{h1}$  is not known until the first attempt is made, the quantity  $w_{h1}n_h$  may be used as its expected value. The total expected cost  $\hat{C}$  of the survey is then given as

$$\hat{C} = c_0 n' + \sum_{h=1}^{L} (c_{h1} + c_{h11} w_{h1}) n_h + \sum_{h=1}^{L} c_{h12} m_{h2}.$$
(7)

### 3. Formulation of the Problem

In Phase I, we obtain the sample size  $n_h$  in each stratum by minimizing variance given in (5) for fixed cost given in (7). At Phase II subsample size from non-respondents group has been obtained by minimizing the sampling variance in (5) for given cost in (7).

#### 3.1. Formulation of the Problem at Phase I

Expression (5) can be expressed as

$$\left(\overline{y}_{j(CRDS)}^{*}\right) = V_{j} = \sum_{h=1}^{L} \frac{\alpha_{jh}}{n_{h}}; \quad j = 1, 2, \cdots, p$$

$$\tag{8}$$

where the terms independent of  $n_h$  are ignored

and 
$$\alpha_{jh} = \frac{1}{n'} \left( w_h n'_h S^2_{jh} + w_{h2} \left( \frac{\left( 1 - k_h^* \right)}{k_h^*} \right) n'_h S^2_{jh2} \right)$$
 (9)

The cost constraint (7) becomes

$$\sum_{h=1}^{L} (c_{h1} + w_{h1}c_{h11}) n_h \le \hat{C}_0$$
(10)

where

$$\hat{C}_0 = \hat{C} - c_0 n' - \sum_{h=1}^{L} c_{h12} m_{h2}$$

Thus the multi-objective formulation of the problem at Phase I becomes

$$\begin{array}{l}
\text{Minimize } V_{j} = \sum_{h=1}^{L} \frac{\alpha_{jh}}{n_{h}}; j = 1, 2, \cdots, p \text{ simulteniously} \\
\text{Subject to } \sum_{h=1}^{L} \left( c_{h1} + w_{h1} c_{h11} \right) n_{h} \leq \hat{C}_{0} \\
2 \leq n_{h} \leq n'_{h} \\
\text{and } n_{h} \text{ integers}; h = 1, 2, \cdots, L
\end{array}$$

$$(11)$$

(see [11]).

### 3.2. Formulation of the Problem for Phase II

Ignoring the term independent from  $m_{h2}$  in (5), substituting  $k_h^* = m_{h2}/n_{h2}$  and  $v_h = \frac{n_h}{n'_h}$ , for  $j = 1, 2, \dots, p$ , expression (5) can be written as

$$V'_{j} = \sum_{h=1}^{L} \frac{\beta_{jh}}{m_{h2}}; \quad j = 1, 2, \cdots, p$$
(12)

where

$$\beta_{jh} = \frac{1}{n'} \sum_{h=1}^{L} \frac{w_{h2} n_{h2} n'_{h} S^{2}_{jh2}}{n_{h}}$$
(13)

The cost constraint becomes

$$\sum_{h=1}^{L} c_{h12} m_{h2} \le \hat{C}_{0}'$$

where

$$\hat{C}_{0}' = \hat{C} - c_{0}n' - \sum_{h=1}^{L} (c_{h1} + w_{h1}c_{h11})n_{h}$$
(14)

Then the multi-objective formulation of the problem at Phase II becomes

$$\begin{array}{l}
\text{Minimize } V_j' = \sum_{h=1}^{L} \frac{\beta_{jh}}{m_{h2}}; \ j = 1, 2, \cdots, p; \text{ simultaneously} \\
\text{Subject to } \sum_{h=1}^{L} c_{h12} m_{h2} \leq \hat{C}_0' \\
2 \leq m_{h2} \leq n_{h2} \\
\text{and } m_{h2} \text{ integers; } h = 1, 2, \cdots, L
\end{array} \right\}.$$
(15)

# 4. Formulation as a Goal Programming Problem

## 4.1. Phase I

Let  $V_j^*$  be the optimal value of  $V_j$  under optimum allocation for the *jth* characteristics obtained by solving the following integer non-linear programming for all the  $j = 1, 2, \dots, p$  characteristics separately.

$$\begin{array}{l}
\text{Minimize } V_{j} = \sum_{h=1}^{L} \frac{\alpha_{jh}}{n_{h}} \\
\text{Subject to } \sum_{h=1}^{L} \left( c_{h1} + w_{h1}c_{h11} \right) n_{h} \leq \hat{C}_{0} \\
2 \leq n_{h} \leq n'_{h} \\
\text{and } n_{h} \text{ integers; } h = 1, 2, \cdots, L
\end{array} \right\}$$
(16)

Further let

$$\tilde{V}_{j} = \tilde{V}_{j}\left(n_{1}, n_{2}, \cdots, n_{L}\right) = \sum_{h=1}^{L} \frac{\alpha_{jh}}{n_{h}}$$

$$\tag{17}$$

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

Obviously  $\tilde{V}_j \ge V_j^*$  and  $\tilde{V}_j - V_j^* \ge 0$ ;  $j = 1, 2, \dots, p$  will give the increase in the variances due to not using the individual optimum allocation for *jth* characteristics.

Let  $d_j \ge 0$  denote the tolerance limit specified for  $(\tilde{V}_j \ge V_j^*)$ ;  $j = 1, 2, \dots, p$ .

We have  $\tilde{V}_{j} - V_{j}^{*} \le d_{j}; j = 1, 2, \dots, p$ 

or 
$$V_j - d_j \le V_j^*; j = 1, 2, \cdots, p$$
  
or  $\sum_{h=1}^{L} \frac{\alpha_{jh}}{n_h} - d_j \le V_j^*; j = 1, 2, \cdots, p$  (18)

A suitable compromise criterion to work out a compromise allocation at phase-I will then be to minimize the sum of deviations  $d_i$ . Therefore the Goal Programming problem at phase-I may be given as

$$\begin{array}{c}
\text{Minimize } \sum_{j=1}^{p} d_{j} \\
\text{Subject to } \sum_{h=1}^{L} \frac{\alpha_{jh}}{n_{h}} - d_{j} \leq V_{j}^{*} \\
\sum_{h=1}^{L} \left( c_{h1} + w_{h1}c_{h11} \right) n_{h} \leq \hat{C}_{0} \\
2 \leq n_{h} \leq n'_{h} \\
d_{j} \geq 0 \\
\text{and } n_{h} \text{ integers} \\
h = 1, 2, \cdots, L; j = 1, 2, \cdots, p
\end{array}$$

$$(19)$$

(See [13]). Where  $d_j \ge 0$ ;  $j = 1, 2, \dots, p$  are the goal variables.

The goal is now to minimize the sum of deviations from the respective optimum variances.

#### 4.2. Phase II

Similarly, at phase II Goal Programming formulation of the problem (15) will be

$$\begin{array}{c}
\text{Minimize } \sum_{j=1}^{p} d'_{j} \\
\text{Subject to } \sum_{h=1}^{L} \frac{\beta_{jh}}{m_{h2}} - d'_{j} \leq V_{j}^{*\prime} \\
\sum_{h=1}^{L} c_{h12}m_{h2} \leq \hat{C}'_{0} \\
2 \leq m_{h2} \leq n_{h2} \\
d'_{j} \geq 0 \\
\text{and } m_{h2} \text{ integers; } h = 1, 2, \cdots, L; j = 1, 2, \cdots, p
\end{array}$$

$$(20)$$

# 5. Formulation as a Fuzzy Programming Problem

#### 5.1. Phase I

To obtain Fuzzy solution we first compute maximum value  $U_k$  and minimum value  $L_k$  for each characteristic. where

$$U_{k} = \max V_{j}\left(\binom{n^{*}}{h,j}\right)L_{k} = \min V_{j}\left(\binom{n^{*}}{h,j}\right); \ j,k = 1, 2, \cdots, p$$
(21)

where  $\binom{n_{h,j}^*}{j}$  denote the optimum allocation for the *jth* characteristics and the maximum and minimum are for all  $V_j$ , among their values for a particular j = k.

The difference of the maximum value  $U_k$  and minimum values  $L_k$  are denoted by  $d_k = U_k - L_k; k = 1, 2, \dots, p$ .

The Fuzzy Programming Problem (FPP) corresponding to the (11) at phase I is given by the following NLPP

$$\begin{array}{c}
\text{Minimize}\,\delta\\
\text{Subject to}\,\sum_{h=1}^{L} \frac{\alpha_{jh}}{n_{h}} - \delta\left(d_{k}\right) \leq V_{j}^{*}\\
\sum_{h=1}^{L} \left(c_{h1} + w_{h1}c_{h11}\right)n_{h} \leq \hat{C}_{0}\\
2 \leq n_{h} \leq n_{h}'\\
d_{k} \geq 0\\
\text{and } n_{h} \text{ integers}\\
h = 1, 2, \cdots, L; \, j, k = 1, 2, \cdots, p
\end{array}$$

$$(22)$$

where  $\delta \ge 0$  is the decision variable representing the worst deviation level.

#### 5.2. Phase II

Similarly, the Fuzzy Programming Problem corresponding to the (15) at phase II is given by the following NLPP

$$\begin{array}{c}
\text{Minimize} \delta'\\
\text{Subject to} \sum_{h=1}^{L} \frac{\beta_{jh}}{m_{h2}} - \delta'(d'_{k}) \leq V_{j}^{*\prime}\\
\sum_{h=1}^{L} c_{h12}m_{h2} \leq \hat{C}_{0}'\\
2 \leq m_{h2} \leq n_{h2}\\
d'_{k} \geq 0\\
\text{and } m_{h2} \text{ integers}\\
h = 1, 2, \cdots, L; \ j, k = 1, 2, \cdots, p
\end{array}$$

$$(23)$$

where  $\delta' \ge 0$  is the decision variable representing the worst deviation level.

The NLPPs may be solved by using the optimization software [14]. For further information about LINGO one may visit the site: <u>http://www.lindo.com</u>.

#### **6. A Numerical Example**

The data in **Table 1** use are from [15]. A population of size N = 3850 is divided into four strata. Two characteristic  $Y_1$  and  $Y_2$  are defined on each unit of the population. The values of  $X_1$  and  $X_2$  are used as the auxiliary information corresponding on the main variable  $Y_1$  and  $Y_2$ . The authors have assumed the values for  $R_1 = 1.48$  and  $R_2 = 0.843$ . **Table 2** shows the other data. Each stratum is divided into respondents and non-respondents as shown in **Table 2**.

It is assumed that  $v_h$  and  $k_h^*$  are known and the preliminary sample size n' = 1000.

In the last column of Table 2, l = 1 is for respondents group and l = 2 is for non-respondents group.

The total cost for the survey is taken as C 3000 units. Out of which 750 units are for the preliminary sample n', 1900 units are for phase-I and 350 units are for phase-II.

Tal	ble	1.	D	ata	for	four	strata	and	two	charact	terist	ics.
-----	-----	----	---	-----	-----	------	--------	-----	-----	---------	--------	------

							<i>j</i> = 1			<i>j</i> = 2		
h	$W_h$	$\mathcal{V}_h$	$k_{\scriptscriptstyle h}^*$	$C_{h1}$	$C_{h11}$	C <sub>h12</sub>	$S_{y_{1h}}^{2}$	$S_{x_{1h}}^{2}$	$S^{_2}_{_{x_1y_{1h}}}$	$S_{y_{2h}}^{2}$	$S_{_{x_{2h}}}^{2}$	$S^{2}_{_{x_{2}y_{2h}}}$
1	0.32	0.4	0.5	1	2	3	784	242	341	1444	481	628
2	0.21	0.5	0.6	1	3	4	576	192	250	676	255	294
3	0.27	0.6	0.7	1	4	5	1024	341	445	1936	645	842
4	0.20	0.65	0.75	1	5	6	2916	972	1268	6084	2028	2645

Table 2. Data for groups of respondents and non-respondents.											
h	Group	$S^2$	$\mathbf{S}^2$	$\mathbf{S}^2$	$\mathbf{S}^2$	$\mathbf{C}^2$	$\mathbf{S}^2$	$W_{hl}$			
	Group	<i>y</i> <sub>1h</sub>	<i>x</i> <sub>1<i>h</i></sub>	$\mathcal{O}_{x_1y_{1h}}$	y <sub>2h</sub>	<i>D</i> <sub><i>x</i><sub>2<i>h</i></sub></sub>	$D_{x_2y_{2h}}$	<i>l</i> = 1, 2			
1	Respondent	361.06	103.16	157.04	767.82	255.76	333.93	$w_{11} = 0.70$			
1	Non-respondent	310.55	88.73	135.07	454.76	151.48	197.93	$w_{12} = 0.30$			
2	Respondent	373.79	124.60	162.24	449.92	169.72	195.67	$w_{21} = 0.80$			
2	Non-respondent	326.29	108.76	141.62	353.81	133.46	153.88	$w_{22} = 0.20$			
2	Respondent	930.15	309.75	404.22	1272.88	424.07	553.60	$w_{31} = 0.75$			
3	Non-respondent	560.28	186.85	243.48	1165.98	388.46	507.10	$w_{_{32}} = 0.25$			
4	Respondent	2355.98	785.33	1024.48	2690.53	896.84	1169.70	$w_{41} = 0.72$			
4	Non-respondent	1013.08	337.69	440.53	2403.55	801.18	1044.93	$w_{42} = 0.28$			

Using estimated values of strata weights the values of  $n'_h = w_h n'; h = 1, 2, \dots, L$  are obtained as

$$n'_1 = 320, n'_2 = 210, n'_3 = 270, n'_4 = 200$$
 with  $\sum_{h=1}^4 n'_h = 1000$ 

# 6.1. Computation of Compromise Allocation Using Goal Programming Technique (GPT)

# 6.1.1. Individual Optimum Allocation (Phase I)

Using data from Table 1 and Table 2, we compute the individual optimum allocation for each characteristic by using NLPP (11) will be the solution to:

For j = 1

Minimize 
$$V_1 = \frac{39.16074}{n_1} + \frac{13.82271}{n_2} + \frac{55.71960}{n_3} + \frac{62.80995}{n_4}$$
  
Subject to  
 $2.4n_1 + 3.4n_2 + 4n_3 + 4.6n_4 \le 1900$   
 $2 \le n_1 \le 320$   
 $2 \le n_2 \le 210$   
 $2 \le n_3 \le 270$   
 $2 \le n_4 \le 200$   
and  $n_h$  are integers  $h = 1, 2, \dots, L$ 

Using optimization software LINGO we get the optimal solution as

$$(n_{h,1}^*) = (159,80,146,144), \quad V_1^* = 1.236899$$

For j = 2

Minimize 
$$V_2 = \frac{109.6218}{n_1} + \frac{21.00712}{n_2} + \frac{85.26067}{n_3} + \frac{151.4584}{n_4}$$
  
Subject to  
 $2.4n_1 + 3.4n_2 + 4n_3 + 4.6n_4 \le 1900$   
 $2 \le n_1 \le 320$   
 $2 \le n_2 \le 210$   
 $2 \le n_3 \le 270$   
 $2 \le n_4 \le 200$   
and  $n_h$  are integers  $h = 1, 2, \dots, L$ 

Using optimization software LINGO we get the optimal solution as

$$\binom{n^*}{n_{h,2}} = (183, 69, 126, 157), \quad V_2^* = 2.544852.$$

**6.1.2. Compromise Solution Using Goal Programming (Phase I)** Using data from **Table 1** and **Table 2** the Goal Programming Problem (19) can be formulated as

$$\begin{aligned} \text{Minimize} d_1 + d_2 \\ \text{Subject to} \\ \frac{39.16074}{n_1} + \frac{13.82271}{n_2} + \frac{55.71960}{n_3} + \frac{62.80995}{n_4} - d_1 \leq 1.236899 \\ \frac{109.6218}{n_1} + \frac{21.00712}{n_2} + \frac{85.26067}{n_3} + \frac{151.4584}{n_4} - d_2 \leq 2.544852 \\ 2.4n_1 + 3.4n_2 + 4n_3 + 4.6n_4 \leq 1900 \\ 2 \leq n_1 \leq 320 \\ 2 \leq n_2 \leq 210 \\ 2 \leq n_3 \leq 270 \\ 2 \leq n_4 \leq 200 \\ d_1 \geq 0, d_2 \geq 0 \end{aligned}$$

and  $n_h$  are integers  $h = 1, 2, \dots, L$ 

Using optimization software LINGO we get the optimal solution as

$$n_1^* = 175, \quad n_2^* = 72, \quad n_3^* = 134, \quad n_4^* = 152,$$

with  $d_1^* = 0.79$  and  $d_2^* = 0.603434$  and the optimum value of the objective function  $d_1^* + d_2^* = 0.0139343$ .

#### 6.1.3. Individual Optimum Allocation (Phase II)

As in Section 6.1.1 for the given data the individual optimum allocations for each the two characteristics using NLPP (15) are:

For j = 1

$$(m_{h,1}^*) = (21, 12, 19, 24) \quad V_1^{*'} = 0.7599099$$

For j = 2

$$\binom{m_{h,2}^*}{}=(21,9,19,26)$$
  $V_2^{*'}=1.742418.$ 

**6.1.4. Compromise Solution Using Goal Programming (Phase II)** For the given data, as in Section 6.1.2 Goal Programming Problem (20) gives the following optimal solution

$$m_{12}^* = 22, \quad m_{22}^* = 10, \quad m_{32}^* = 20, \quad m_{42}^* = 24,$$

with  $d_1^{*'} = 0.002751545$  and  $d_2^{*'} = 0.002866150$  and the optimum value of the objective function  $d_1^{*'} + d_2^{*'} = 0.005617695$ .

#### 6.2. Computations of Compromise Solution Using Fuzzy Programming Technique (FPT)

#### 6.2.1. Compromise Solution Using Fuzzy Programming (Phase I)

To obtain fuzzy solution we first obtained the maximum value and minimum value as given in (21) for each characteristic by using individual optimum allocation worked out in Section 6.1.1

$$U_1 = 1.256605$$
  $L_1 = 1.236899$   
 $U_2 = 2.58685$   $L_2 = 2.544852$ 

and 
$$d_1 = 0.019706$$
  $d_2 = 0.042954$ .

After computing the optimum allocation and optimum variances for two characteristics the compromise optimal solution for the above problem can be obtained by solving the given Fuzzy Programming Problem (FPP) of (23)

# $\begin{array}{l} \text{Minimize } \delta \\ \text{Subject to} \\ \\ \frac{39.16074}{n_1} + \frac{13.82271}{n_2} + \frac{55.71960}{n_3} + \frac{62.80995}{n_4} - \delta\left(0.019706\right) \leq 1.236899 \\ \frac{109.6218}{n_1} + \frac{21.00712}{n_2} + \frac{85.26067}{n_3} + \frac{151.4584}{n_4} - \delta\left(0.042954\right) \leq 2.544852 \\ 2.4n_1 + 3.4n_2 + 4n_3 + 4.6n_4 \leq 1900 \\ 2 \leq n_1 \leq 320 \\ 2 \leq n_2 \leq 210 \\ 2 \leq n_3 \leq 270 \\ 2 \leq n_4 \leq 200 \\ \delta \geq 0 \\ \text{and} n_b \text{ are integers } h = 1, 2, \cdots, L \end{array}$

Using optimization software LINGO we get the optimal solution as

$$n_1^* = 171, n_2^* = 75, n_3^* = 135, n_4^* = 151$$
 with  $\delta = 0.2486958$ 

#### 6.2.2. Compromise Solution Using Fuzzy Programming (Phase II)

Similarly, using data from Table 1 and Table 2 the Fuzzy Programming Problem (23) gives the following optimal solution

$$m_{12}^* = 22$$
,  $m_{22}^* = 10$ ,  $m_{32}^* = 20$ ,  $m_{42}^* = 24$  with  $\delta' = 0.2423372$ .

#### 7. Summary of the Results

- 1 - DI

In the following results obtained using Goal Programming Technique and Fuzzy Programming Technique are summarized.

#### 8. Conclusions

**Table 3** and **Table 4** show the values of the variance of the combined ratio estimates of the population means at Phase-I and Phase-II respectively, for the two characteristics. The figures show that both the approaches the Goal Programming Approach and the Fuzzy Programming Approach give almost same results. However, at Phase-I the Goal Programming Approach is slightly more precise in terms of the trace value (See [16]).

The Goal Programming and Fuzzy Programming technique and some other techniques like Dynamic Programming and Separable Programming can be used to solve a wide variety of mathematical programming problems. These techniques may be of great help in solving multivariate sampling problem also. Like determining

rable 5. Compromise solution at Phase 1.											
Tashniquas		Allo	cations		Varia	ances	Trace	Cost incurred			
Techniques	$n_1$	<i>n</i> <sub>2</sub>	<i>n</i> <sub>3</sub>	$n_{4}$	$V_1$	$V_2$	$V_{1} + V_{2}$	С			
GPT	175	72	134	152	1.244799	2.550886	3.795685	1900			
FPT	171	75	135	151	1.242011	2.555754	3.797765	1900			

#### S. Iftekhar et al.

Table 4. Compromise solution at Phase II.												
Tachniques		Alloc	ations		Varia	inces	Trace	Cost incurred				
rechniques -	<i>m</i> <sub>12</sub>	<i>m</i> <sub>22</sub>	<i>m</i> <sub>32</sub>	<i>m</i> <sub>42</sub>	$V_1'$	$V_2'$	$V_1' + V_2'$					
GPT	22	10	20	24	0.7626614	1.745284	2.507945	350				
FPT	22	10	20	24	0.7626614	1.745284	2.507945	350				

the number of strata, strata boundaries and compromise allocations in multivariate stratified sampling. Little work has been done to solve the above mentioned optimization problems in real life situations. For example when the estimates of the population parameters used in formulating the problems are themselves treated as random variables with assumed or known distributions. In such cases the formulated problems becomes a multivariate stochastic programming. Further, apart from a linear cost function, nonlinear functions may be used that may include travel cost, labour cost, rewards to the respondent and incentives to the investigators etc. Interested researchers may expose these situations.

# **Acknowledgements**

The authors are thankful to the Editor for his valuable remarks and suggestions that helped us a lot in improving the standard of the paper. This research work is partially supported by the UGC grant of Emeritus Fellowship to the author Mohammed Jameel Ahsan for which he is grateful to UGC.

#### References

- [1] Hansen, M.H. and Hurwitz, W.N. (1946) The Problem of Non-Response in Sample Surveys. *Journal of the American Statistical Association*, **41**, 517-529. <u>http://dx.doi.org/10.1080/01621459.1946.10501894</u>
- [2] Rao, P.S.R.S. (1986) Ratio Estimation with Sub-Sampling the Non-Respondents. Survey Methodology, 12, 217-230.
- [3] Rao, P.S.R.S. (1987) Ratio and Regression Estimates with Sub-Sampling the Non-Respondents. *Special Contributed Session of the International Statistical Association Meeting*, 2-16 September 1987, Tokyo.
- [4] Khare, B.B. and Srivastva, S. (1993) Estimation of Population Mean Using Auxiliary Character in Presence of Non-Response. *National Academy Science Letters*, 16, 111-114.
- [5] Khare, B.B. and Srivastva, S. (1997) Transformed Ratio Type Estimators for the Population Mean in the Presence of Non-Response. *Communications in Statistics—Theory and Methods*, 26, 1779-1791. http://dx.doi.org/10.1080/03610929708832012
- [6] Raiffa, H. and Schlaifer, R. (1961) Applied Statistical Decision Theory. Graduate School of Business Administration, Harvard University, Boston.
- [7] Ericson, W.A. (1965) Optimum Stratified Sampling Using Prior Information. *Journal of the American Statistical Association*, 60, 750-771. <u>http://dx.doi.org/10.1080/01621459.1965.10480825</u>
- [8] Ahsan, M.J. and Khan, S.U. (1982) Optimum Allocation in Multivariate Stratified Random Sampling with Overhead Cost. *Metrika*, 29, 71-78. <u>http://dx.doi.org/10.1007/BF01893366</u>
- [9] Dayal, S. (1985) Allocation in Sample Using Values of Auxiliary Characteristics. *Journal of Statistical Planning and Inference*, 11, 321-328. <u>http://dx.doi.org/10.1016/0378-3758(85)90037-0</u>
- [10] Khan, M.G.M., Maiti, T. and Ahsan, M.J. (2010) An Optimal Multivariate Stratified Sampling Design Using Auxiliary Information: An Integer Solution Using Goal Programming Approach. *Journal of Official Statistics*, 26, 695-708.
- [11] Varshney, R., Najmussehar and Ahsan, M.J. (2011) An Optimum Multivariate Stratified Double Sampling Design in Non-Response. *Optimization Letters*, 6, 993-1008.
- [12] Cochran, W.G. (1977) Sampling Techniques. 3rd Edition, John Wiley& Sons, New York.
- Schniederjans, M.J. (1995) Goal Programming: Methodology and Applications. Kluwer, Dordrecht. <u>http://dx.doi.org/10.1007/978-1-4615-2229-4</u>
- [14] Lingo User's Guide (2013) Lingo-User's Guide. LINDO SYSTEM INC., Chicago.
- [15] Haseen, S., Iftekhar, S., Ahsan, M.J. and Bari, A. (2012) A Fuzzy Approach for Solving Double Sampling Design in Presence of Non-Response. *International Journal of Engineering Science and Technology*, 4, 2542-2551.
- [16] Sukhatme, P.V., Sukhatme, B.V., Sukhatme, S. and Asok, C. (1984) Sampling Theory of Surveys with Applications. 3rd Edition, Iowa State University Press, Iowa and Indian Society of Agricultural Statistics, New Delhi.



IIIIII II

 $\checkmark$ 

Scientific Research Publishing (SCIRP) is one of the largest Open Access journal publishers. It is currently publishing more than 200 open access, online, peer-reviewed journals covering a wide range of academic disciplines. SCIRP serves the worldwide academic communities and contributes to the progress and application of science with its publication.

Other selected journals from SCIRP are listed as below. Submit your manuscript to us via either submit@scirp.org or Online Submission Portal.

