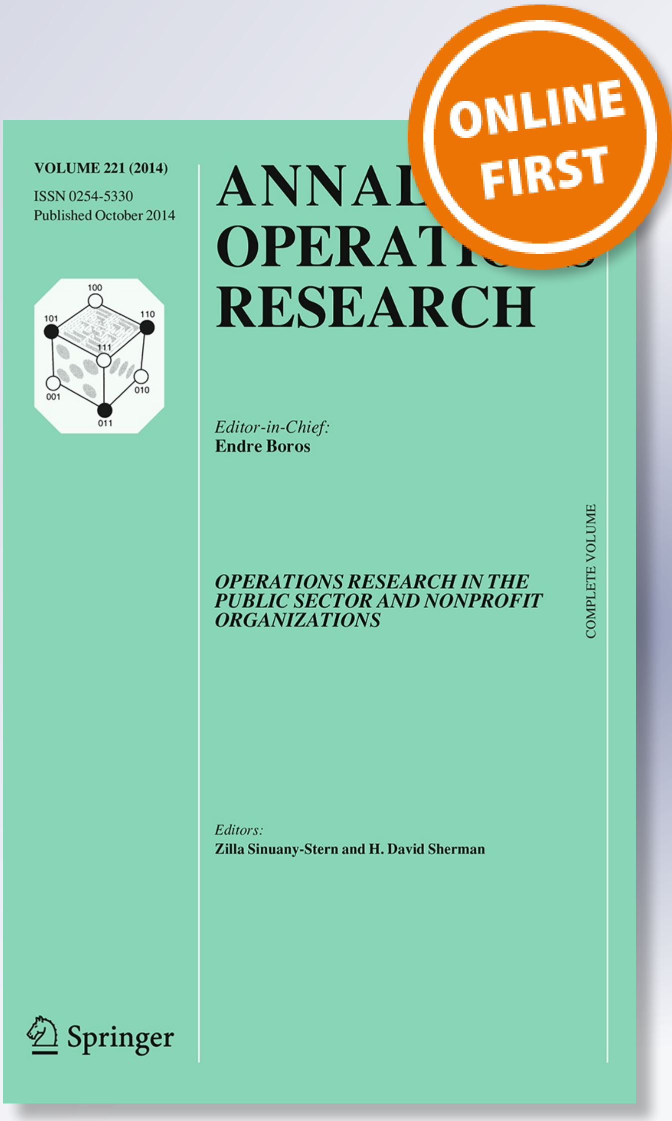
Integer compromise allocation in multivariate stratified surveys

Rahul Varshney, M. G. M. Khan, Ummatul Fatima & M. J. Ahsan

Annals of Operations Research

ISSN 0254-5330

Ann Oper Res
DOI 10.1007/s10479-014-1734-z



Ann Oper Res
DOI 10.1007/s10479-014-1734-z

Integer compromise allocation in multivariate stratified surveys

Rahul Varshney · M. G. M. Khan · Ummatul Fatima · M. J. Ahsan

Abstract In multivariate stratified sampling more than one characteristic are defined on every unit of the population. An optimum allocation which is optimum for one characteristic will generally be far from optimum for others. To resolve this problem, a compromise criterion is needed to work out a usable allocation. In this manuscript, a compromise criterion is discussed and integer compromise allocations are obtained by using goal programming technique. A numerical example is presented to illustrate the computational details, which reveals that the proposed criterion is suitable for working out a usable compromise allocation for multivariate stratified surveys.

Keywords Compromise allocation · Goal programming · Multiobjective nonlinear programming · Multivariate stratified sampling

1 Introduction

The main problem in stratified sampling is the optimum choice of the size of the samples to be selected from various strata either to maximize the precision of the estimate for a fixed

R. Varshney (✉)
Department of Statistics, Amity University, Noida, India
e-mail: itsrahulvarshney@gmail.com

M. G. M. Khan
School of Computing, Information and Mathematical Sciences, The University of the South Pacific, Suva, Fiji
e-mail: khan_mg@usp.ac.fj

U. Fatima · M. J. Ahsan
Department of Statistics and Operations Research, Aligarh Muslim University, Aligarh, India
e-mail: fatimau2011@yahoo.com

M. J. Ahsan
e-mail: mjahsan2001@yahoo.co.in

Published online: 04 October 2014

cost or to minimize the cost of the survey for a fixed precision of the estimate. The sample sizes allocated according to either of the above criteria is termed as "Optimum Allocation". According to Cochran (1977), Tschuprow (1923) was first to give the optimum allocation for a fixed sample size. But his result remained unknown. Neyman (1934) rediscovered it. Later on Stuart (1954) used Cauchy-Schwarz inequality to work out optimum allocation in more general conditions. However, in multivariate populations where more than one characteristic on each selected unit of the population are to be measured, the optimum allocation for one characteristic may be far from optimum for others. In such situations it is almost impossible to find an allocation that is optimum for all the characteristics simultaneously.

Thus one has to use an allocation that is optimum in some sense for all the characteristics. Such allocations are called compromise allocations in sampling literature, because they are based on some compromised criterion. Various criteria to work out a compromise allocation are available. Dalenius (1957), Ghosh (1958), Yates (1960), Aoyama (1962), Folks and Antle (1965), Chatterjee (1967, 1968), Kokan and Khan (1967), Ahsan (1975–76, 1978), Ahsan and Khan (1977, 1982), Bethel (1985, 1989), Schittkowski (1985–86), Chromy (1987), Kreienbrock (1993), Jahan et al. (1994, 2001), Jahan and Ahsan (1995), Rahim (1995), Khan et al. (1997), Khan et al. (2003, 2008), Singh (2003), Semiz (2004), Díaz-García and Cortez (2006, 2008), Kozak (2006a, b), Ansari et al. (2009, 2011), Varshney et al. (2012) and many others either suggested new compromise criterion or explored the existing criteria under various situations such as the presence of nonresponse, availability of auxiliary information, use of double sampling technique for stratification etcetera.

In the present manuscript the problem of determining the compromise allocation in multivariate stratified random sampling is viewed as a multiobjective integer nonlinear programming problem.

2 An overview of the literature and the proposed approach

In multivariate stratified surveys the individual optimum allocations are of no practical use and we need an allocation that is optimum for all characteristics in some sense. As mentioned in Sect. 1, many authors contributed to obtain a compromise allocation using different compromise criteria. Another approach to this problem is to formulate it as a multiobjective programming problem and convert it into a single objective problem by scalarizing the multiple objectives. Some of the compromise criteria used by these authors are the averaging of individual optimum allocations, minimizing the sum of the relative increases in the variances of estimates due to not using the individual optimum allocations, minimizing the sum of variances of the individual estimates, minimizing the weighted sum of variances of the individual estimates, minimizing the sum of coefficients of variation (CVs) of the individual estimates, minimizing the total cost of survey with tolerance limits on the variances of the estimates, minimizing the total sample size with tolerance limit on the CVs of the estimates, etc. (see Chatterjee 1968; Cochran 1977; Arthanari and Dodge 1981; Sukhatme et al. 1984; Bethel 1989; Khan et al. 2003; Díaz-García and Cortez 2008; Ansari et al. 2011, etc.). Many authors also suggested the criteria by minimizing the variances or CVs of some selected characteristics, such as minimizing the maximum CV of the estimate (see Kozak 2006a). It has been noticed that these criteria works well, especially, to minimize the maximum variance among the estimates. However, most of the criteria failed to minimize the variances of the estimates simultaneously of all characteristics, especially, the characteristics that have smaller variances. It is, therefore, a multiobjective criterion is needed that may minimize the variances of all characteristics simultaneously and hence the sum of variances of all characteristics.

In this paper, the authors have made an attempt to determine a compromise criterion by considering the problem as a multiobjective programming problem that minimizes the CVs of the estimates of the population means of all the characteristics simultaneously.

Furthermore, for practical application of any allocation the integer sample sizes are required, which is not considered by most of the authors. The integer values are obtained by rounded off the continuous values of sample sizes. Rounding off works well in general but it may produce a non-optimum and/or infeasible solution (see Khan et al. 1997) in some cases. To avoid this, the authors formulated the problem of determining the compromise allocation in multivariate stratified sampling as a multiobjective integer nonlinear programming problem (MINLPP) and developed a solution procedure using Goal Programming Technique. For each study variable, the proposed technique introduces a goal variable to ensure that the increase in the variance of its estimate due to not using the individual optimum allocation will not exceed the goal variable. The sum of the goal variables is then minimized under the cost and other restrictions.

This approach provides a compromise allocation which minimizes the total loss in precession of the estimates for not using the individual optimum allocations with respect to different characteristics.

3 Formulation of the problem

Let the population be divided in to L non-overlapping strata of sizes $N_1, N_2, \dots, N_L$ and $W_h = \frac{N_h}{N}$ denotes the stratum weight of hth stratum such that $N = \sum_{h=1}^{L} N_h$. Let n_h be the number of units drawn by simple random sampling without replacement (*SRSWOR*) from the hth stratum consists of N_h units of the population. Let $p \geq 2$, characteristics are defined on each population unit and the estimation of unknown population means $\bar{Y}_j;\ j = 1, 2, \dots, p$ is of interest. It is assumed that the characteristics are independent. The sampling variances $V(\bar{y}_{jst})$ of the estimates $\bar{y}_{jst}$, of the unknown population means $\bar{Y}_j$ are given by

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

where $f_h = \frac{n_h}{N_h}$ is the sampling fraction in the hth stratum.
Ignoring the finite population correction $(1 - f_h)$ we get

$$V(\bar{y}_{jst}) = \sum_{h=1}^{L} \frac{W_h^2 S_{jh}^2}{n_h}; \quad j = 1, 2, \dots, p \tag{1}$$

where $S_{jh}^2 = \frac{1}{N_h - 1} \sum_{i=1}^{N_h} (y_{jhi} - \bar{Y}_{jh})^2;\ j = 1, 2, \dots, p$ are the stratum variances of the jth characteristic in hth stratum, y_{jhi} is the value of the ith unit of the hth stratum for jth characteristic and $\bar{Y}_{jh} = \frac{1}{N_h} \sum_{i=1}^{N_h} y_{jhi}$ is the stratum mean of y_{jhi}.
Furthermore, stratified sample mean for the jth characteristics

$\bar{y}_{jst} = \sum_{h=1}^{L} W_h \bar{y}_{jh};\quad j = 1, 2, \dots, p$ is an unbiased estimate for $\bar{Y}_j = \frac{1}{N} \sum_{h=1}^{L} \sum_{i=1}^{N_h} y_{jhi}$, the overall population mean of the jth characteristic, where $\bar{y}_{jh} = \frac{1}{n_h} \sum_{i=1}^{n_h} y_{jhi}$, is the sample stratum mean for jth characteristic in the hth stratum.

Assume that we know the population variances S_{jh}^2 and means $\bar{Y}_j$. These are unknown in real surveys but can be approximated or known from a recent or preliminary survey (Kozak 2006a).

Let the pre-fixed cost C of the survey is a linear function of n_h given by

$$C = c_0 + \sum_{h=1}^{L} c_h n_h, \tag{2}$$

where c_0 = overhead cost and c_h = per unit cost of measurement of all the characteristics in hth stratum; $h = 1, 2, \ldots, L$.

Then the required compromise allocation will be the solution to the following MINLPP

$$\begin{aligned} &\text{Minimize} && \begin{pmatrix} CV(\bar{y}_{1st}) \\ CV(\bar{y}_{2st}) \\ \vdots \\ CV(\bar{y}_{pst}) \end{pmatrix} \\ &\text{subject to} && \textstyle\sum_{h=1}^{L} c_h n_{h_c} \le C_0 \\ & && 2 \le n_{h_c} \le N_h \\ &\text{and} && n_{h_c} \text{integers}; \quad h = 1, 2, \ldots, L, \end{aligned} \tag{3}$$

where $CV_j = CV(\bar{y}_{jst}) = \sqrt{\frac{V(\bar{y}_{jst})}{\bar{Y}_j^2}}$; $j = 1, 2, \ldots, p$.

Let

$$Z_j = CV_j = \sqrt{\frac{1}{\bar{Y}_j^2} \sum_{h=1}^{L} \frac{W_h^2 S_{jh}^2}{n_{h_c}}} = \sqrt{\sum_{h=1}^{L} \frac{a_{jh}}{n_{h_c}}} \tag{4}$$

where $a_{jh} = \frac{W_h^2 S_{jh}^2}{\bar{Y}_j^2}$; $j = 1, 2, \ldots, p$, $h = 1, 2, \ldots, L$.

Also $C_0 = C - c_0$ and n_{h_c}; $h = 1, 2, \ldots, L$ denote the compromise allocation.

4 The solution using goal programming

Let Z_j^* be the optimum value of Z_j defined by (4) under optimum allocation for the jth characteristic obtained by solving the following all integer nonlinear programming problem (AINLPP)

$$\begin{aligned} &\text{Minimize} && Z_j \\ &\text{subject to} && \textstyle\sum_{h=1}^{L} c_h n_h \le C_0 \\ & && 2 \le n_h \le N_h \\ &\text{and} && n_h \text{ integers}; \quad h = 1, 2, \ldots, L. \end{aligned} \tag{5}$$

Note that the solution to (5) will give the individual optimum allocation for the jth characteristic.

If $\tilde{Z}_j$ denote the value of Z_j for a compromise allocation then $\tilde{Z}_j \geq Z_j^*$ or $\tilde{Z}_j - Z_j^* \geq 0$; $j = 1, 2, \ldots, p$ will give the increase in Z_j due to not using the individual optimum allocation of *j*th characteristic.

Consider the following problem:
"Find n_h such that for *j*th characteristic ($j = 1, 2, \ldots, p$), the increase in the value of the Z_j for each j due to the use of a compromise allocation, instead of its individual optimum allocation, is less than or equal to x_j, where $x_j \geq 0$ ($j = 1, 2, \ldots, p$) are the goal variables" (Schniederjans 1995).
To achieve these goals we must have

$$\tilde{Z}_j - Z_j^* \leq x_j; \quad j = 1, 2, \ldots, p$$

$$\text{or} \quad \tilde{Z}_j - x_j \leq Z_j^*; \quad j = 1, 2, \ldots, p \tag{6}$$

$$\text{or} \quad \sqrt{\sum_{h=1}^{L} \frac{a_{jh}}{n_{h_c}}} - x_j \leq Z_j^*; \quad j = 1, 2, \ldots, p. \tag{7}$$

The total increase in the values of Z_j; $j = 1, 2, \ldots, p$ for not using the individual optimum allocations is given by $\sum_{j=1}^{p} x_j$. Minimizing this total increase will be a suitable compromise criterion to obtain a compromise allocation.
Consequently we have to solve the following goal programming problem (GPP)

$$\begin{aligned}
&\text{Minimize} && \textstyle\sum_{j=1}^{p} x_j \\
&\text{subject to} && \sqrt{\textstyle\sum_{h=1}^{L} \frac{a_{jh}}{n_{h_c}}} - x_j \leq Z_j^*; \quad j = 1, 2, \ldots, p \\
& && \textstyle\sum_{h=1}^{L} c_h n_{h_c} \leq C_0 \\
& && 2 \leq n_{h_c} \leq N_h \\
&\text{and} && x_j \geq 0;\ j = 1, 2, \ldots, p \quad \text{and} \quad n_{h_c} \text{integers}; \quad h = 1, 2, \ldots, L.
\end{aligned} \tag{8}$$

The GPP (8) may be solved by using a mixed integer nonlinear programming technique. However, many optimization softwares are now available. One such software is LINGO (2001). The authors used this software to solve GPP (8). For more information about LINGO one may visit the site: http://www.lindo.com.

5 Some other compromise criteria

For the sake of comparison of the performance of proposed compromise allocation with some other well-known compromised allocations, the following compromise criteria are considered.

5.1 Cochran's average compromise allocation (ACA)

Cochran (1977) suggested the character-wise average of individual optimum allocations of each characteristic as compromise allocation. The allocations are rounded off to the nearest integer. The Cochran's average compromise allocation $n_{h(ACA)}$ is given by

$$n_{h(ACA)} = \frac{1}{p}\sum_{j=1}^{p} n^*_{j,h}; \quad h = 1, 2, \ldots, L, \tag{9}$$

where, $n^*_{j,h}$; $j = 1, 2, \ldots, p$; $h = 1, 2, \ldots, L$ denote the individual optimum allocation for the jth characteristic given by

$$n^*_{j,h} = \frac{C_0 W_h S_{jh}/\sqrt{c_h}}{\sum_{h=1}^{L} W_h S_{jh}\sqrt{c_h}}; \quad h = 1, 2, \ldots, L; \quad j = 1, 2, \ldots, p. \tag{10}$$

5.2 Chatterjee's compromise allocation (CCA)

Chatterjee's compromise allocation, $n_{h(CCA)}$, is obtained by minimizing the sum of the relative increases in the variances of the estimates (Chatterjee 1967), is given as

$$n_{h(CCA)} = \frac{C_0\sqrt{\sum_{j=1}^{p} n^{*2}_{j,h}}}{\sum_{h=1}^{L} c_h\sqrt{\sum_{j=1}^{p} n^{*2}_{j,h}}}; \quad h = 1, 2, \ldots, L, \tag{11}$$

where $n^*_{j,h}$; $h = 1, 2, \ldots, L$; $j = 1, 2, \ldots, p$ are as given by (10).

5.3 Allocation obtained by minimizing the maximum CV (MaxCV)

The compromise allocation can be obtained by minimizing the maximum CV of the estimates by solving the following AINLPP as suggested by Kozak (2006a)

$$\begin{aligned} &\text{Minimize} && Z_r \\ &\text{subject to} && Z_j \le Z^*_j; \quad j = 1, 2, \ldots, p; \quad j \ne r \\ & && \textstyle\sum_{h=1}^{L} c_h n_h \le C_0 \\ & && 2 \le n_h \le N_h \\ &\text{and} && n_h \text{ integers}; \quad h = 1, 2, \ldots, L, \end{aligned} \tag{12}$$

where $Z_r = CV_r \ge Z_j = CV_j$ for all $j = 1, 2, \ldots, p$.

The resulting compromise allocation is denoted by $n_{h(\text{MaxCV})}$.

6 A numerical illustration

The following numerical example illustrates the Goal Programming approach. The data are from 2002 Agricultural Censuses in Iowa State conducted by National Agricultural Statistics Service, USDA, Washington D.C. cited in Khan et al. (2010) (source: http://www.agcensus.usda.gov/). The 99 counties in the Iowa State are divided into four strata. Two characteristics: (i) The quantity of corn harvested Y_1 and (ii) The quantity of oats harvested Y_2, are of interest. The total amount available for conducting the survey is assumed to be $C = 350$ units with an expected overhead cost $c_0 = 50$ units. This gives $C_0 = C - c_0 = 300$ units. Table 1 gives the strata variances for the two characteristics and per unit cost of measurement with in stratum, c_h; $h = 1, 2, 3$ & 4.
Also $\bar{Y}_1 = 474{,}973.90$ and $\bar{Y}_2 = 1{,}576.25$.

Table 1 Data for four strata with two characteristics

h	N_h	S_{1h}^2	S_{2h}^2	c_h
1	8	29,267,524,195.5	777,174.1	12
2	34	26,079,256,582.8	4,987,812.9	9
3	45	42,362,842,460.8	1,074,510.6	10
4	12	30,728,265,336.9	388,378.5	8

To find the values of Z_1^* and Z_2^*, we have to solve the AINLPP (5) for $j = 1$ & 2.
Using LINGO the solutions for the values given in Table 1 are:

$$\textbf{For } \boldsymbol{j} = 1: \quad n_{1,1}^* = 2, \quad n_{1,2}^* = 9, \quad n_{1,3}^* = 16 \quad \text{and} \quad n_{1,4}^* = 4 \quad \text{with} \quad Z_1^* = 0.0697399.$$
$$\textbf{For } \boldsymbol{j} = 2: \quad n_{2,1}^* = 2, \quad n_{2,2}^* = 18, \quad n_{2,3}^* = 9 \quad \text{and} n_{2,4}^* = 3 \quad \text{with} \quad Z_2^* = 0.157701.$$

Thus the GPP (8) becomes

$$\begin{aligned}
&\text{Minimize} \sum_{j=1}^{2} x_j \\
&\text{subject to} \\
&\sqrt{\frac{0.000847142}{n_{1_c}} + \frac{0.013634624}{n_{2_c}} + \frac{0.038797187}{n_{3_c}} + \frac{0.002001201}{n_{4_c}}} - x_1 \le 0.0697399 \\
&\sqrt{\frac{0.002042574}{n_{1_c}} + \frac{0.236781300}{n_{2_c}} + \frac{0.089354235}{n_{3_c}} + \frac{0.002296662}{n_{4_c}}} - x_2 \le 0.157701 \\
&12n_{1_c} + 9n_{2_c} + 10n_{3_c} + 8n_{4_c} \le 300 \\
&2 \le n_{h_c} \le N_h; \quad h = 1, 2, 3\&4 \\
&\text{and } x_j \ge 0; \quad j = 1\,\&\,2 \quad \text{and} \quad n_{h_c} \text{ integers}; \quad h = 1, 2, 3\,\&\,4.
\end{aligned} \tag{13}$$

Using LINGO, the proposed compromise allocation is found to be

$$n_{1_c}^* = 2, \quad n_{2_c}^* = 15, \quad n_{3_c}^* = 12 \quad \text{and} \quad n_{4_c}^* = 2$$

with $x_1^* = 0.004867374$ and $x_2^* = 0.001676615$ and the optimum value of the objective function $x_1^* + x_2^* = 0.00654399$, and $Z_1 = CV_1 = 0.074607272$, $Z_2 = CV_2 = 0.159377615$ and $CV_1 + CV_2 = 0.233984887$.
For the data given in Table 1, the individual optimum allocations $n_{j,h}^*$; $h = 1, 2, \ldots, L$; $j = 1, 2, \ldots, p$ using formula (10) are given as:

$$\begin{aligned}
&\text{For } j = 1: \quad n_{1,1}^* = 2.09960, \quad n_{1,2}^* = 9.72632, \quad n_{1,3}^* = 15.56497, \quad n_{1,4}^* = 3.95229 \\
\text{and} \quad &\text{for } j = 2: \quad n_{2,1}^* = 1.45114, \quad n_{2,2}^* = 18.04109, \quad n_{2,3}^* = 10.51400, \quad n_{2,4}^* = 1.88458.
\end{aligned}$$

Using (9), the Cochran's average compromise allocation rounded off to the nearest integer is:

$$n_{1(ACA)} = 2, n_{2(ACA)} = 14, \quad n_{3(ACA)} = 13, \quad n_{4(ACA)} = 3$$

with $Z_1 = CV_1 = 0.071055883$, $Z_2 = CV_2 = 0.159916208$ and $CV_1 + CV_2 = 0.230972091$.

Table 2 Various allocations with sum of CVs and the cost incurred

S. no.	Allocations	n_1	n_2	n_3	n_4	Sum of CVs	Cost incurred
1	Proportional	3	11	14	4	0.240006425	307
2	Cochran's	2	14	13	3	0.230972091	304
3	Chatterjee's	2	14	13	3	0.230972091	304
4	Minimizing Maximum CV	2	9	16	4	0.252740241	297
5	Proposed allocation	2	15	12	2	0.233984887	295

Using the above value of $n^*_{j,h}$, the Chatterjee's compromise allocation given by (11), rounded off to nearest integer is:

$$n_{1(CCA)} = 2, n_{2(CCA)} = 14, \quad n_{3(CCA)} = 13, \quad n_{4(CCA)} = 3$$

with $Z_1 = CV_1 = 0.071055883$, $Z_2 = CV_2 = 0.159916208$ and $CV_1 + CV_2 = 0.230972091$.
The allocation $n_{h(\text{MaxCV})}$ obtained by minimizing the maximum CV (MaxCV) is obtained by solving the following AINLPP:

$$\begin{aligned}
&\text{Minimize } Z_2 = \sqrt{\frac{0.002042574}{n_{1_c}} + \frac{0.236781300}{n_{2_c}} + \frac{0.089354235}{n_{3_c}} + \frac{0.002296662}{n_{4_c}}}\\
&\text{subject to}\\
&\sqrt{\frac{0.000847142}{n_{1_c}} + \frac{0.013634624}{n_{2_c}} + \frac{0.038797187}{n_{3_c}} + \frac{0.002001201}{n_{4_c}}} \le 0.0697399\\
&12n_{1_c} + 9n_{2_c} + 10n_{3_c} + 8n_{4_c} \le 300\\
&2 \le n_{h_c} \le N_h; \quad h = 1, 2, 3\&4\\
&\text{and } \ n_{h_c} \text{ integers}; \quad h = 1, 2, 3 \ \& \ 4, \ \text{ where } n_{h_c} = n_{h(\text{MaxCV})}.
\end{aligned} \tag{14}$$

Using LINGO the solution to the AINLPP (14) is given as:

$$n_{1(MaxCV)} = 2, \quad n_{2(MaxCV)} = 9, \quad n_{3(MaxCV)} = 16, \quad n_{4(MaxCV)} = 4$$

with $Z_1 = CV_1 = 0.069739899$, $Z_2 = CV_2 = 0.183000342$ and $CV_1 + CV_2 = 0.252740241$.
Also the proportional allocation is given as:

$$n_{h(prop)} = \frac{C_0 W_h}{\sum_{h=1}^{L} c_h W_h}; \quad h = 1, 2, \ldots, L \quad \text{(see Sukhatme et al. 1984).}$$

For the data of Table 1 we get

$$n_{1(prop)} = 3, \quad n_{2(prop)} = 11, \quad n_{3(prop)} = 14, \quad n_{4(prop)} = 4$$

with $Z_1 = CV_1 = 0.069234527$, $Z_2 = CV_2 = 0.170771899$ and sum of CVs $= 0.240006425$.
These allocations, sum of CVs and the cost incurred are given in Table 2.

7 Conclusion

Usually the trace of the variance-covariance matrix is used as a measure of performance of a compromise allocation (see Sukhatme et al. 1984). As the trace is the sum of individual variances that usually differ in their units of measurement to add the variances are not mathematically feasible.

In this paper we considered the sum of the CVs as a measure of the performance of a compromise allocation because they are unit free and can be added.

The cost incurred given in the last column of Table 2 reveals that Proportional, Cochran's and Chatterjee's allocations are infeasible because they violate the cost constraints.

Among the feasible compromise allocations the proposed allocation gives the minimum value of the sum of CVs hence it can be said that its performance is best as regards the sum of CVs is concerned.

Acknowledgments The authors are grateful to the Editors and the learned Reviewers for their valuable comments and suggestions that helped to revise the manuscript in its present form. The author M. J. Ahsan is thankful for the financial assistance provided under 'Emeritus Fellowship' by the University Grants Commission, Govt. of India to carry out this research.

References

Ahsan, M. J. (1975–76). A procedure for the problem of optimum allocation in multivariate stratified random sampling. *Aligarh Bulletin of Mathematics*, *5–6*, 37–42.

Ahsan, M. J. (1978). Allocation problem in multivariate stratified random sampling. *Journal of Indian Statistical Association*, *16*, 1–5.

Ahsan, M. J., & Khan, S. U. (1977). Optimum allocation in multivariate stratified random sampling using prior information. *Journal of Indian Statistical Association*, *15*, 57–67.

Ahsan, M. J., & Khan, S. U. (1982). Optimum allocation in multivariate stratified random sampling with overhead cost. *Metrika*, *29*, 71–78.

Ansari, A. H., Najmussehar, & Ahsan, M. J. (2009). On multiple response stratified random sampling design. *Journal of Statistics Sciences*, *1*(1), 45–54.

Ansari, A. H., Varshney, R., Najmussehar, & Ahsan, M. J. (2011). An optimum multivariate-multiobjective stratified sampling design. *Metron*, *LXIX*(3), 227–250.

Aoyama, H. (1962). Stratified random sampling with optimum allocation for multivariate population. *Annals of the Institute of Statistical Mathematics*, *14*, 251–258.

Arthanari, T. S., & Dodge, Y. (1981). *Mathematical Programming in Statistics*. New York: Wiley.

Bethel, J. (1985). An optimum allocation algorithm for multivariate surveys. *Proceedings of American Statistical Association Survey Research Methods Section*, 209–212.

Bethel, J. (1989). Sample allocation in multivariate surveys. *Survey Methodology*, *15*, 47–57.

Chatterjee, S. (1967). A note on optimum allocation. *Scandinavian Actuarial Journal*, *50*, 40–44.

Chatterjee, S. (1968). Multivariate stratified surveys. *Journal of the American Statistical Association*, *63*, 530–534.

Chromy, J. R. (1987). Design optimization with multiple objectives. *Proceedings of American Statistical Association Survey Research Methods Section*, 194–199.

Cochran, W. G. (1977). *Sampling techniques*. New York: Wiley.

Dalenius, T. (1957). *Sampling in Sweden: Contributions to the methods and theories of sample survey practice*. Stockholm: Almqvist and Wiksell.

Díaz-García, J. A., & Cortez, L. U. (2006). Optimum allocation in multivariate stratified sampling: Multi-objective programming. Comunicación Técnica No. I-06-07/28-03-2006 (PE/CIMAT).

Díaz-García, J. A., & Cortez, L. U. (2008). Multi-objective optimisation for optimum allocation in multivariate stratified sampling. *Survey Methodology*, *34*(2), 215–222.

Folks, J. L., & Antle, C. E. (1965). Optimum allocation of sampling units to strata when there are R responses of interest. *Journal of the American Statistical Association*, *60*, 225–233.

Ghosh, S. P. (1958). A note on stratified random sampling with multiple characters. *Calcutta Statistical Association Bulletin*, *8*, 81–89.

Jahan, N., & Ahsan, M. J. (1995). Optimum allocation using separable programming. *Dhaka University Journal of Sciences*, *43*(1), 157–164.
Jahan, N., Khan, M. G. M., & Ahsan, M. J. (1994). A generalized compromise allocation. *Journal of the Indian Statistical Association*, *32*, 95–101.
Jahan, N., Khan, M. G. M., & Ahsan, M. J. (2001). Optimum compromise allocation using dynamic programming. *Dhaka University Journal of Sciences*, *49*(2), 197–202.
Khan, M. G. M., Ahsan, M. J., & Jahan, N. (1997). Compromise allocation in multivariate stratified sampling: An integer solution. *Naval Research Logistics*, *44*, 69–79.
Khan, M. G. M., Khan, E. A., & Ahsan, M. J. (2003). An optimal multivariate stratified sampling design using dynamic programming. *Australian & New Zealand Journal of Statistics*, *45*(1), 107–113.
Khan, M. G. M., Khan, E. A., & Ahsan, M. J. (2008). Optimum allocation in multivariate stratified sampling in presence of non-response. *Journal of Indian Society of Agricultural Statistics*, *62*(1), 42–48.
Khan, M. G. M., Maiti, T., & 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*(4), 695–708.
Kokan, A. R., & Khan, S. U. (1967). Optimum allocation in multivariate surveys: An analytical solution. *Journal of Royal Statistical Society*, *B29*(1), 115–125.
Kozak, M. (2006a). On sample allocation in multivariate surveys. *Communications in Statistics Simulation and Computation*, *35*, 901–910.
Kozak, M. (2006b). Multivariate sample allocation: Application of random search method. *Statistics in Transition*, *7*(4), 889–900.
Kreienbrock, L. (1993). Generalized measures of dispersion to solve the allocation problem in multivariate stratified random sampling. *Communications in Statistics Theory and Methods*, *22*(1), 219–239.
LINGO (2001). *LINGO-User's Guide*. Published by LINDO SYSTEM INC., 1415, North Dayton Street, Chicago, Illinois, 60622, USA.
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 the Royal Statistical Society*, *97*(4), 558–625.
Rahim, M. A. (1995). Use of distance function to optimize sample allocation in multivariate surveys: A new perspective. *Proceedings of American Statistical Association Survey Research Methods Section*, 365–369.
Schittkowski, K. (1985-86). NLPQL: A FORTRAN subroutine solving constrained nonlinear programming problems. *Annals of Operations Research*, *5*, 485–500.
Schniederjans, M. J. (1995). *Goal programming: Methodology and applications*. Dordrecht: Kluwer Academic Publishers.
Semiz, M. (2004). Determination of compromise integer strata sample sizes using goal programming. *Hacettepe Journal of Mathematics and Statistics*, *33*, 91–96.
Singh, S. (2003). *Advanced sampling theory with applications: How Michael 'selected' Amy*. Dordrecht: Kluwer Academic Publishers.
Stuart, A. (1954). A simple presentation of optimum sampling results. *Journal of the Royal Statistical Society*, *B16*(2), 239–241.
Sukhatme, P. V., Sukhatme, B. V., Sukhatme, S., & Asok, C. (1984). *Sampling theory of surveys with applications*. Iowa State University Press, Iowa and Indian Society of Agricultural Statistics, New Delhi.
Tschuprow, A. A. (1923). On the mathematical expectation of the moments of frequency distributions in the case of correlated observations. *Metron*, *2*, 461–493.
Varshney, R., Najmussehar, & Ahsan, M. J. (2012). An optimum multivariate stratified double sampling design in presence of non-response. *Optimization Letters*, *6*(5), 993–1008.
Yates, F. (1960). *Sampling methods for censuses and surveys*. London: Charles Griffin and Co., Ltd.