An integer solution using Branch and Bound Method in Multi-objective stratified sampling design

Mushtaq A. lone, Shakeel. A .Mir, S. Maqbool and M. Ashraf. Bhat

Abstract

This article deals with the determination of optimum allocation of sample sizes using fuzzy goal programming approach in tertiary objective stratified sampling design. In this paper, the degree of satisfaction for each goal is also determined at both integer and non integer optimum solution.

Keywords: stratification, optimal allocation, goal programming, Fuzzy goal programming, Taylors first order approximation, Branch and Bound method.

Introduction

One of the areas of statistics that is most commonly used in all fields of scientific investigation is that of stratified sampling. In stratified sampling the most important consideration is the allocation of sample sizes in each stratum either to minimize the variance subject to cost or minimize cost subject to variance. In stratified sampling the population is divided into certain number of groups called strata and then selecting a simple random sample from each stratum independently. The population characteristics can be inferred with samples from each stratum, exploiting the gain in precision in the estimates, administrative convenience and the flexibility in this design is that different sampling procedure can be used in different sub- populations. The problem of optimally choosing the sample sizes is known as the optimal allocation problem. The problem of optimal allocation in multivariate stratified sampling designs discussed by several authors (see, for example, Gosh (1958), Yates (1960), Hartley (1965), Folks and Antle (1965), Aoyama (1963), Gren (1966), Kokan and Khan(1967), Chatterjee (1972), Bethel (1989), kreienbrock (1993), Khan *et al*. (1997). Ahsan *et al*. (2005), Kozak (2006), Ansari *et al*. (2009), etc). The first mathematical formulation of fuzziness was pioneered by Zadeh (1965). Orlovsky (1980) made a numerous attempts to explore the ability of fuzzy set theory to become a useful tool for adequate mathematical analysis of real world problems. Fuzzy methods have been developed in virtually all branches of decision making problems can be found in Tamiz (1996), Zimmermann (1991) and Ross (1995). Goal programming approach in fuzzy environment has been first introduced by Narashimann(1980).Fuzzy goal programming has been discussed by several authors (see Pal B.B. *et al*(2003), Biwas *et al* (2005), Parra *et al*(2001) etc). In this paper, tertiary objective stratified sampling design is solved using fuzzy goal programming (FGP) approach through LINGO. For practical purpose if the solution is non integer then the FGP is solved using Branch and Bound method instead of rounding the non integer sample sizes to the nearest integral values. However in some situation for small samples the rounded off allocation may become infeasible and non optimal.

Problem formulation

Sometimes it is desirable to divide the population into several sub populations or strata in order to estimate the population parameter. In stratified random sampling, the population of size N divided into L strata of sizes $N_1, N_2, \ldots N_L$. Where $\sum_{i=1}^{L} N_i = N$. If $N_1, N_2, \ldots N_L$ are not known in advance then the strata weights remains unknown. Stratified sampling is useful if the strata weights are known for each stratum. Also, we assume that n_i samples are drawn independently from each stratum. Let $\bar{y}$ denote an unbiased estimate of population mean ($\bar{Y}$), where Y is the characteristic under study. Let $\bar{y}_i$ is an unbiased estimate of the stratum mean $\bar{Y}_i$ such that $\bar{y}_i = \frac{1}{n_i}\sum_{h=1}^{n_i} y_{ih}$ *for all* $(i=1,2,3,...L)$ .Then, $\bar{y}$ given by $\bar{y} = \frac{1}{N}\sum_{i=1}^{L} N_i \bar{y}_i = \sum_{i=1}^{L} W_i \bar{y}_i$

is an unbiased estimate of population mean $\bar{Y}$. The precision of this estimate is measured by the variance of the sample estimate of the population characteristics.

$$V(\bar{y}) = \frac{1}{N}\sum_{i=1}^{L} N_i \bar{y}_i = \frac{1}{N^2}\sum_{i=1}^{L} N_i^2 V(\bar{y}_i) = \sum_{i=1}^{L} W_i^2 S_i^2 x_i = \sum_{i=1}^{L} a_i x_i$$

Since $\sum_{i=1}^{L} \frac{a_i}{N_i}$ is constant and it is sufficient to minimize $V(\bar{y}) = \sum_{i=1}^{L} \frac{a_i}{n_i}$.

Where $W_i = \frac{N_i}{N}$; $S_i^2 = \frac{1}{N_i - 1}\sum_{h=1}^{N_i}(y_{ih} - \bar{Y}_{ij})^2$; $a_i = W_i^2 S_i^2$ *and* $x_i = \frac{1}{n_i} - \frac{1}{N_i}$.

Also, coefficient of variation is given by

$$(CV) = (CV(\bar{y})) = \frac{SD(\bar{y})}{\bar{Y}} = \frac{\sqrt{V(\bar{y})}}{\bar{Y}} = \left[\sum_{i=1}^{L} \frac{a_i}{n_i}\right]^{1/2} \Big/ \bar{Y}$$

The problem of optimum allocation involves determining the sample sizes that minimize the total variance subjected to sampling cost. Let c_i be the cost per sample in the i^{th} stratum. The sampling cost function is of the form $C = c^o + \sum_{i=1}^{L} c_i n_i$

Where, c^o = Overhead cost and C is the total budget available in the survey. Let $C - c^o = C^*$.

The tertiary objective allocation problem is given below

$$\left.\begin{aligned}
&\text{Minimize } c^{**} = \sum_{i=1}^{L} c_i n_i \\
&\text{Minimize } n^* = \sum_{i=1}^{L} n_i \\
&\text{Minimize } (CV) = \left[\sum_{i=1}^{L} \frac{a_i}{n_i}\right]^{1/2} \Big/ \bar{Y} \\
&\text{Subject to} \\
&V(\bar{y}) = \sum_{i=1}^{L} \frac{a_i}{n_i} \le v^* \\
&2 \le n_i \le N_i \qquad n_i \text{, integer } i = 1,2,3,\ldots,L
\end{aligned}\right\} \qquad (1.1)$$

Where v^* is prefixed variance of the estimate of the population mean. In this tertiary objective problem, the objective is to minimize the cost function, sample sizes and CV subjected to set of constraints variance and non negative restrictions.

Fuzzy goal programming

Fuzzy programming requires no weightings or rankings and generates only a single (efficient) solution. Thus we deal with aspiration intervals rather than single values (such as fuzzy and interval goal programming) have their own attraction. Zimmermann (1983) presented a fuzzy approach to multi-objective linear programming problems. Now, we formulate the fuzzy programming model of problem (1.1) by transforming the objective functions into fuzzy goals by assigning aspiration level to each of them using Zimmermann Max-min approach (1978). The membership function must be described for each fuzzy goal .If $F_t(n) \le g_t$ then

$$\mu_t(n) = \begin{cases} 1 & if \quad F_t(n) \le g_t \\ \dfrac{u_t - F_t(n)}{u_t - g_t} & if \quad g_t \le F_t(n) \le u_t \\ 0 & if \quad F_t(n) \ge u_t \end{cases} \qquad (1.2)$$

If $F_t(n) \ge g_t$ then

$$\mu_t(n) = \begin{cases} 1 & if \quad f_t(n) \ge g_t \\ \dfrac{F_t(n) - l_t}{g_t - l_t} & if \quad l_t \le F_t(n) \le g_t \\ 0 & if \quad F_t(n) \le l_t \end{cases} \qquad (1.3)$$

Where g_t is the aspiration level of the tth objective $F_t(n)$ and u_t and l_t (t= 1, 2 . . . m) are the upper tolerance limit and lower tolerance limit, respectively, for the tth fuzzy goal. Here t=3 the we can write problem (1.1) as

$$\text{Minimize } [\, F_1(n) = \sum_{i=1}^{L} c_i n_i \;, F_2(n) = \sum_{i=1}^{L} n_i \;, F_3(n) = \left[\sum_{i=1}^{L} \frac{a_i}{n_i} \right]^{1/2} \Big/ \bar{Y} \,]$$

Subject to (1.4)

$$V(\bar{y}) = \sum_{i=1}^{L} \frac{a_i}{n_i} \le V^*$$

$$2 \le n_i \le N_i \qquad n_i \text{ ,integer } i = 1,2,3,\ldots, L$$

Now, we transform non linear membership functions $\mu_t(n)$ into an equivalent linear membership functions at individual best solution point by using first order Taylor's series as fallows:

$$\mu_t(n) = \mu_t(n_t^*) + [(n_1 - n_{t1}^*)\frac{d\mu_t(n_t^*)}{dn_1} + (n_2 - n_{t2}^*)\frac{d\mu_t(n_t^*)}{dn_2} + \ldots + (n_L - n_{tL}^*)\frac{d\mu_t(n_t^*)}{dn_L}] \quad (1.5)$$

Where n_t^* is the individual best solution.

Thus the NLGPP (1.4) reduces to the following problem as

Minimize $\mu_t(n)$

Subject to

$$V(\bar{y}) = \sum_{i=1}^{L} \frac{a_i}{n_i} \le V^*$$

$$2 \le n_i \le N_i \qquad n_i \text{ , integer } i = 1,2,3,\ldots,L \quad (1.6)$$

The membership function defined in(1.6) has a maximum value one . Thus the aspiration level of defined membership function is unity as given below

$\mu_t(n) + \delta_t^+ = 1$, Where δ_t^+ is the over deviational variable.

Now, Fuzzy goal programming problem (FGP) can be presented as

Minimize δ_t^+

Subjected to

$$V(\bar{y}) = \sum_{i=1}^{L} \frac{a_i}{n_i} \le V^*$$

$$\mu_t(n) + \delta_t^+ = 1 \quad (1.7)$$

$$2 \le n_i \le N_i \qquad n_i \text{ , integer } i = 1,2,3,\ldots,L$$

Numerical illustration

The given data has been taken from Arthanari and Dodge (1993).The population contains 64units, the stratum weights and stratum variance of a population which has been divided into three strata with one characteristic under study is given below in the table 1.

Table 1

i	N_i	$W_i=\frac{N_i}{N}$	S_i^2	$\bar{Y}_i^2$	a_i	C_i
1	16	0.2500	540.0625	62.9375	33.7539	4
2	20	0.3125	14.6737	27.6000	1.4330	1.5
3	28	0.4375	7.2540	14.0714	1.3885	1

Assume that C (available budget) =100 units including c^o and c^o = 30 units (overhead cost).Therefore the total amount available for the survey is C^* = 70 units. Also we assume that the cost of measurement C_i in various strata are $c_1=4, c_2=1.5 \; and \; c_3=1$ for the cost function

$$C = c^o + \sum_{i=1}^{L} c_i n_i$$

After substituting the values of the parameters given in the table (1) above the NLGPP (1.1) is written as when t=3:

$$\left.\begin{aligned}
&Min\;\; F_1(n) = 4n_1 + 1.5n_2 + n_3 \\
&Min\;\; F_2(n) = 4n_1 + 1.5n_2 + n_3 \\
&Min\; F_3(n) = \left[\frac{33.7539}{n_1} + \frac{1.4330}{n_2} + \frac{1.3885}{n_3}\right]^{1/2} / \; 30.52 \\
&subject\,to \\
&\frac{33.7539}{n_1} + \frac{1.4330}{n_2} + \frac{1.3885}{n_3} \le 2.90 \\
&2 \le n_1 \le 16,\;\; 2 \le n_2 \le 20,\;\; 2 \le n_1 \le 2
\end{aligned}\right\} \quad (1.8)$$

In order to find n to satisfy the following fuzzy goals with fuzzy aspiration levels are 70, 23.15 and 0.036 such that $F_1(n) \le 70, F_2(n) \le 23.15 \; and \, F_3(n) \le 0.036$.thus 122, 64, and 0.05 are tolerance limits respectively for the above three goals.

We define the membership functions for the three fuzzy goals as:

$$\mu_1(n)=\begin{cases}1 & if \quad F_t(n)\le 70\\ \dfrac{122-F_t(n)}{122-70} & if \quad 70\le F_t(n)\le 12\\ 0 & if \quad F_t(n)\ge 122\end{cases} \tag{1.9}$$

$$\mu_2(n)=\begin{cases}1 & if \quad F_t(n)\le 23.15\\ \dfrac{64-F_t(n)}{64-23.15} & if \quad 23.15\le F_t(n)\le 64.\\ 0 & if \quad F_t(n)\ge 64\end{cases} \tag{2.0}$$

$$\mu_t(n)=\begin{cases}1 & if \quad F_t(n)\le 0.036\\ \dfrac{0.05-F_t(n)}{0.05-0.036} & if \quad 0.036_t\le F_t(n)\le 0.05\\ 0 & if \quad F_t(n)\ge 0.05\end{cases} \tag{2.1}$$

Transform membership function by using first order Taylor's series about points n_1^* = (14.29, 4.80, 5.80), n_2^* = (16, 3.60, 3.54) and n_3^* = (16,20,28) for $\mu_1(n)$, $\mu_2(n)$ *and* $\mu_3(n)$ respectively.

$$\left.\begin{aligned}\mu_1(n)&=\mu_1(n_1^*)=-0.08n_1-0.03n_2-0.02n_3+2.4\\ \mu_2(n)&=\mu_1(n_2^*)=-0.02n_1-0.02n_2-0.02n_3+1.463\\ \mu_3(n)&=\mu_1(n_3^*)=0.104n_1+0.0028n_2+0.0014n_3+0.7589\end{aligned}\right\} \tag{2.2}$$

Using (2.2) Fuzzy goal programming problem (FGP) can be presented as

Minimize $\delta_1^+ + \delta_2^+ + \delta_3^+$

Subject to

$$\left.\begin{aligned}&-0.08n_1-0.03n_2-0.02n_3+2.4+\delta_1^+=1\\ &-0.02n_1-0.02n_2-0.02n_3+1.463+\delta_2^+=1\\ &0.104n_1+0.0028n_2+0.0014n_3-0.7589+\delta_3^+=1\\ &\frac{33.7539}{n_1}+\frac{1.4330}{n_2}+\frac{1.3885}{n_3}\le 2.90\\ &2\le n_1\le 16,\ 2\le n_2\le 20,\ \ 2\le n_1\le 2\end{aligned}\right\} \tag{2.3}$$

After solving FGPP (2.3) through LINGO software, we get

$n_1 = 16, n_2 = 3.43, n_3 = 3.73, \delta_1^+ = 0.058, , \delta_2^+ = 0, .033 and\ \delta_3^+ = 0.08$. Thus the obtained solution is non integer. In order to get the integer values NLGPP (2.3) is solved using Branch and Bound method. Various nodes of branch and Bound Method for NLGPP (2.3) are presented below in figure (a).

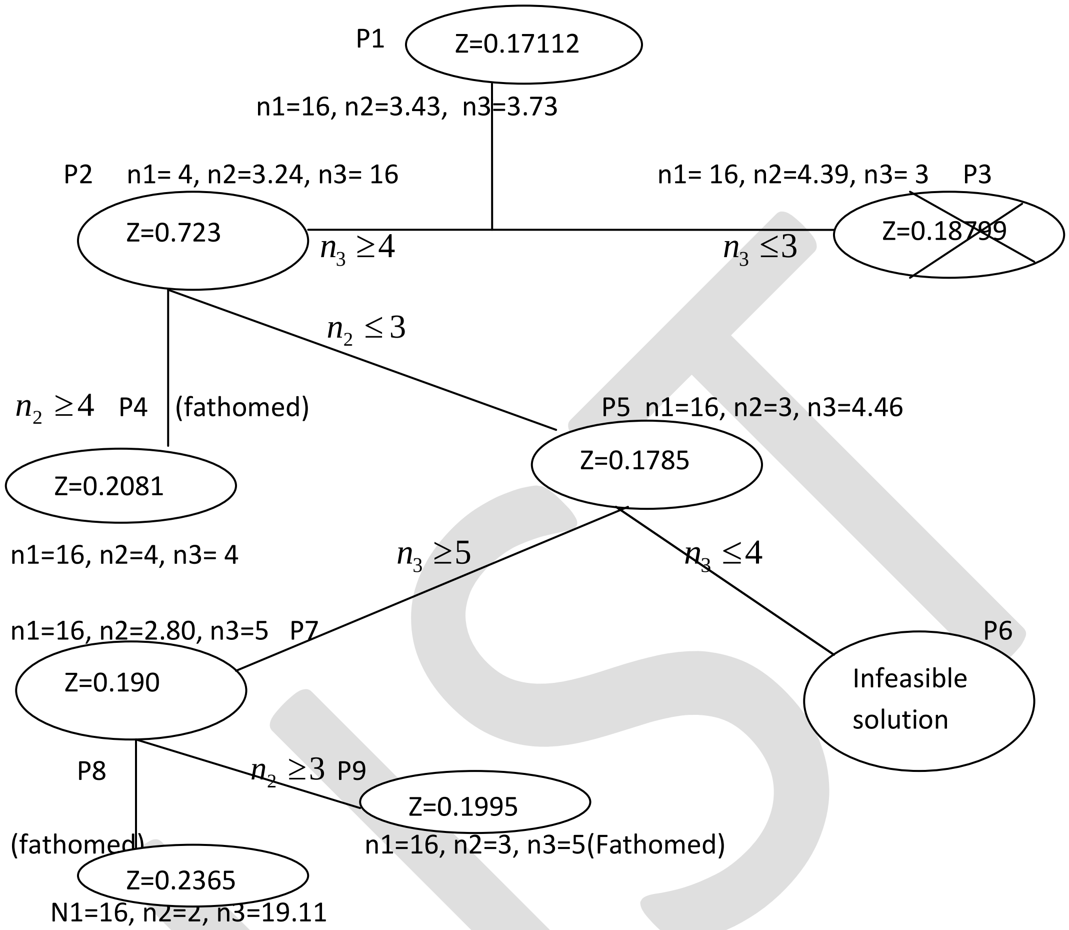

Figure (a): Various nodes of branch and Bound Method for NLGPP (2.3)

Since $n_1, n_2 = 3 and n_3$ are required to be integers, we branch problem P_1into two sub problems P_2 and P_3 by introducing the constraints $n_3 \geq 4$ and $n_3 \leq 3$ respectively indicated by the value $n_3 = 3.73$ which lies between 3 and 4. This process of replacing a problem by two sub problems is called branching. The solution of these two sub problems can be obtained using LINGO software a shown in figure (a) above. We could branch from either one. Thus we choose P2 because it has the more nearly optimal value of the objective function. Again P2 is divided into two branches P4 and P5.Node P_4 is fathomed with integer value and we are left with P_5, thus P_5 is divided into two branches P_6 and $P_{7.}$ P_6 is infeasible and finally we branch P_7 into two sub problems P_8 and P_9 respectively. Thus P_9 is fathomed with integer value and P_8 is fathomed.

Thus the integer optimal solution using Branch and Bound method is given by $n_1 = 16, n_2 = 3, and\ n_3 = 5$. At non integer solution $(n_1 = 16, n_2 = 3.43, n_3 = 3.73)$ the membership function values of $\mu_1(n),\ \mu_2(n)\, and\, \mu_3(n)$ indicates that goals $F_1(n), F_2(n)\ and\ F_3(n)$ are satisfied 94%, 99% and 92% respectively for the obtained solution. Now the membership function value at integer solution indicates that goals are satisfied 93%, 98% and 92% respectively for the obtained solution.

Conclusion

This paper concludes that when a non integer solution exists after solving the FGPP then Branch and Bound method is used to obtain the integer solution of the FGPP. It can be easily seen from the non integer solution obtained above if we rounding off allocations as n_3 approaches to 4. But using Branch and Bound method $n_3 = 5$. Also the satisfaction percentage can be easily seen for each objective at both integer and non integer optimum allocation.

References

Kokan, A. R. and Khan, S. U. (1967) .Optimum allocation in multivariate surveys: An analytical solution.Journal of Royal Statistical Society, Ser. B, 29, 115125.

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

Yates, F. (1960).Sampling Methods for Censuses and Surveys. 3rd ed. London: Charles Griffin.

Aoyama H., **(**1963). Stratified Random Sampling with Optimum Allocation for Multivariate Populations, *Annals of the institute of statistical Mathematics***.** 14:251–258.

Hartley H.O.,(1965). Multiple purpose optimum allocation in stratified sampling, in : *Proceedings of the American Statistical Association, Social Statistics Section. Alexandria, VA: American Statistical Association, 258-261.*

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

Gren , J. (1966) . Some Application of Non-linear Programming in Sampling Methods, *Przeglad Statystyczny*, **13**:203–217

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

Bethel J.,(1989). Sample Allocation in Multivariate Surveys. *Survey Methodology*, **15**, 40-57 **.**

Kreienbrock, L .(1993). Generalized measures of dispersion to solve the allocation problem in multivariate stratified random sampling . Cmmunication in Statistics : theory and methods 22(1) : 219-239.

Khan. M.G.M., and Ahsan, M. J., and Jahan, N.(1997). Compromise allocation in multivariate stratified sampling : An integer solution .Naval Research Logistics 44:69-79.

Ahsan, M. J., Najmussehar and Khan ,M. G.M. (2005). Mixed allocation in stratified sampling .Aligrah Journal of Statistics sciences 25:87-97.

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

Ansari A.H., Najmussehar and Ahsan M.J. (2009). On Multiple Response Stratified Random Sampling Design, *International Journal of Statistical Sciences*, **1(1)**, 45-54.

Arthanari ,T.s.,\and Dodge, Y. (1993).Mathematical programming in statistics.New York: John Willy.

Lingo 13.0, LINDO *inc.ltd.*

Zadeh,L.A.(1965). Fuzzy sets. *Information and Control*, **8:338-353.**

Orlovsky, S. A. (1980). On Formulation of A General Fuzzy Mathematical programming problem. *Fuzzy Sets and Systems*, 3: 311-321.

Tamiz, M.(1996). Multi-Objective programming and goal programming theories and applications. *Germany : Springer-Verlag.*

Zimmermann, H. j. (1991) Fuzzy SetTheory and Its Applications. *(2nd rev.ed). Boston: Kulwer*.

Ross, T. J. 1995. Fuzzy logic with engineering Applications. *New York: Mcgraw-Hill*.

Narashimann.R.(1980). On fuzzy Goal Programming-Some constraints. *Decision sciences* **11**: 532-538.

Pal, B.B.,Monitra, B. N and Maulik .U.(2003). A Goal programming Procedure for Fuzzy Multiobjective Linear fractional Programming Problem. *Fuzzy Sets and Systems* **139**: 391-405.

Biswas, A and Pal, B.B. (2005). Application of Fuzzy Goal Programming technique to Land Use planning in Agricultural System . *Omega* 33: 391-398.

Parra, M.A., Terol, A. B., and Uria, M. V. R. (2001). A Fuzzy Goal Programming Approach to Portfolio selection. *European Journal of Operational Research* 133: 287-297.

Zimmermann HJ.(1978). Using fuzzy sets in Operation research, *European journal of Operational research*,13: 201-206.

Zimmermann HJ.(1983).fuzzy programming and linear programming with several objective functions. *Fuzzy sets and systems* 1:45-55.