Research Journal of Mathematical and Statistical Sciences ______________________________ISSN 2320–6047
Vol. **3(12),** 16-20, December **(2015)** *Res. J. Mathematical and Statistical Sci.*

Linear / Non Linear plus Fractional Goal Programming (L/NLPFGP) Approach in Stratified Sampling Design

Lone Mushtaq A, Mir Shakeel A.,Singh K.N. and Khan Imran
Division of Agri-stat SKUAST-K, Salimar, Srinagar, INDIA

Available online at: www.isca.in, www.isca.me
Received 2nd October 2015, revised 23rd October 2015, accepted 17th November 2015

Abstract

This article deals with the problem of finding an integer optimal allocation of sample sizes in stratified sampling design as a problem of multi-objective optimization. In this paper tertiary objective stratified sampling design is converted into linear/non linear plus fractional goal programming. Then a fuzzy goal programming approach is used to solve the converted problem through LINGO. When a non integer solution is obtained then Branch and Bound method is used to obtain the integer solution. Numerical illustration is also given for the demonstration of proposed approach.

Keywords: Stratification, optimal allocation, goal programming, fractional programming, Taylors first order approximation, branch and bound method.

Introduction

Sometimes it is desirable to divide the whole population into several sub-populations in order to estimate the population parameter. This is achieved by a technique known as Stratification, in which the population size N is first divided into L strata of sizes $N_1, N_2, \ldots, N_L$ respectively. Strata are non overlapping and they together compromise the whole population i,e $\sum_{i=1}^{L} N_i = N$. There should be homogenity within and heterogeneity between strata. When the strata have been determined, a sample is drawn from each stratum. These drawing are made independently in different strata. If the simple random sampling (SRS) is taken in each stratum, the whole procedure is described as stratified random sampling. Thus the problem in which sample sizes are optimally chosen is known as problem of optimal allocation. The problem of optimal allocation for univariate stratified population was first considered by Neyman[1].

Optimum allocation of sample sizes to the various strata has been stated as a non linear mathematical programming problem in which the objective function is the cost subject to a variance restriction or vice versa. The problem of optimal allocation in stratified sampling designs discussed by several authors (see, for example, Gosh[2], Yates[3], Hartley[4], Folks and Antle[5], Aoyama[6], Gren[7], Kokan and Khan[8], Chatterjee[9], Bethel[10], kreienbrock[11], Khan *et al.*[12]. Ahsan *et al.*[13], Kozak[14], Ansari *et al.*[15], etc). Multiobjective linear plus linear fractional programming problem solutions are found in Hirche[16], Chadha[17],Jain and Lachhwani[18], Schaible[19], Sukhatme *et* al[20] etc. In stratification the Values of N_i must be known. However, in many practical situations, it is usually difficult to perform stratification with respect to the characteristics under study especially because of physical and cost considerations. Generally, the stratification is done according to administrative grouping, geographic regions and on the basis of auxiliary characters. In this paper we propose an algorithm in which the problem of tertiary objective stratified sampling design is converted into linear/non linear plus fractional goal programming.

Problem formulation

Consider a population of size N divided into L strata of sizes $N_1, N_2, \ldots, NL$, such that $\sum_{i=1}^{L} N_i = N$. Also, we assume samples are drawn from each stratum independently such that $\sum_{i=1}^{L} n_i = n$. Since sample mean denoted by $\bar{y}$ is an unbiased estimate of $\bar{Y}$ the population mean. Let $\bar{y}_i$ denotes the sample mean for ith stratum is an unbiased estimate of $\bar{Y}_i$ the stratum mean such that

$$\bar{y}_i = \frac{1}{n_i} \sum_{h=1}^{n_i} y_{ih} \quad \text{for all } (i = 1,2,3,\ldots L) .$$

Then $$\bar{y} = \frac{1}{N} \sum_{i=1}^{L} N_i \bar{y}_i = \sum_{i=1}^{L} W_i \bar{y}_i$$

The precision of this estimate is measured by the variance of the sample estimate of the population characteristics Y .

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

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. 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. c_i = cost per unit in the i^{th} stratum.

Let $C - c^o = C^*$.

The tertiary objective allocation problem is given below

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

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.

Linear / non linear plus fractional goal programming problem (L/NLPFGPP): For tertiary objective non linear programming problem (TONLPP) the proposed approach can be outlined as given below:

Step 1: Find the maximum value of $F_t(n)$ say it $F_t(n^*)$. if t=3, then our objectives are $F_1(n), F_2(n), and\ F_3(n)$ subject to set of constraints. Let $F_1(n)$=cost has maximum optimal value.

Step 2: Divide each objective individually by $F_t(n^*)$ and all of the remaining $F_{t-1}(n)$ .e,g we write $\xi_{t1}(n) = \frac{F_2(n)}{F_1(n)}$ and $\xi_{t2}(n) = \frac{F_3(n)}{F_1(n)}$ fractional programming with two objectives.
Step 3: we get fractional programming.
Step 4: Add those terms (objectives) to Step 3 which are not in fractional programming separately.
Step 5: we obtain (L/NLPFGPP). e, g $\lambda_1(n) = \xi_{t1}(n) + F_3(n)$ and $\lambda_2(n) = \xi_{t2}(n) + F_2(n)$ with two objectives subjected to set of constraints.
Step 6: Define the membership function for t^{th} objective.
Step 7: Transform membership function bys using first order Taylor's approximation.
Step 8: Form the Fuzzy goal programming model
Step 9: Find n^* after solving transformed (FGPP).
Step 10: if solution is non integer then use Branch and Bound method to get integer solution.
Step 11: End.

Numerical illustration

The given data has been taken from Arthanari and Dodge[21].The population contains 64 units, the stratum weights and stratum variance of a population divided into three strata with one characteristic under study is given below in the table- 1.

Table-1
Data for three strata

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

Assume that C (available budget) =100 units including c^o and c^o = 30 units (overhead cost).Therefore the available amount for the survey is C^*= 70 units. Also, for various strata the cost of measurement C_i can be assumed as $c_1 = 4, c_2 = 1.5$ 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 NLPP (1) is written as when t=3:

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

Solving (2) through LINGO using step 1, we get $F_1(n) = 70$, $F_2(n) = 23.15$, *and* $F_3(n) = 0.036$

Finally after using steps 2, 3, 4 and step 5. We get (L/NLPFGPP) in the following form

$$Min\ \lambda_1(n) = F_3(n) + \xi_1(n)$$

$$Min\ \ \lambda_2(n) = F_2(n) + \xi_2(n)$$

$$\left.\begin{aligned}
&subject\ to\\
&\frac{33.7539}{n_1} + \frac{1.4330}{n_2} + \frac{1.3885}{n_3} \le 2.90\\
&2 \le n_1 \le 16,\ \ 2 \le n_2 \le 20,\ \ 2 \le n_1 \le 2
\end{aligned}\right\} \quad (3)$$

Thus the transformed problem (3) has two goals. These goals can be attained by solving (3) for each single objective i, e goal (1) = $k_1^{"} = 0.3645$ and goal (2) = $k_2^{"} = 16.59$.

In order to find n to satisfy the following fuzzy goals with fuzzy aspiration levels are 0.3645 and 16.59such that $\lambda_1(n) \le 0.3645$, *and* $\lambda_3(n) \le 16.59$ .thus 0.6117 and 64 are tolerance limits respectively for the above two goals.

For the two fuzzy goals the membership functions can be defined as:

$$\mu_1(n) = \begin{cases} 1 & if \quad \lambda_1(n) \le\ 0.3645\\ \dfrac{0.6117 - \lambda_1(n)}{0.6117 - 0.3645} & if \qquad 0.3645 \le \lambda_1(n) \le 0.6117\\ 0 & if\ \lambda_1(n) \ge 0.6117 \end{cases} \quad (4)$$

$$\mu_2(n) = \begin{cases} 1 & if \quad \lambda_1(n) \le 16.59\\ \dfrac{64 - \lambda_1(n)}{64 - 16.59} & if \qquad 16.59 \le \lambda_1(n) \le 64\\ 0 & if\ \lambda_1(n) \ge 64 \end{cases} \quad (5)$$

Using first order Taylor's series transform membership function about points $n_1^* = (16, 3.85, 3.32)$, and $n_2^* = (11.77, 2.425, 2.39)$ for $\mu_1(n)$, *and* $\mu_2(n)$ respectively.

$$\begin{aligned}
&\mu_1(n) = \mu_1(n_1^*) = 0.0198n_1 - 0.0025n_2 - 0.0474n_3 + 0.9354\\
&\mu_2(n) = \mu_1(n_2^*) = -0.021n_1 - 0.0210n_2 - 0.0210n_3 + 1.3483
\end{aligned} \quad (6)$$

Thus Fuzzy goal programming problem (FGP) can be presented as

Minimize $\mu_t(n)$

Subject to

$$\left.\begin{aligned}
&V(\bar{y}) = \sum_{i=1}^{L} \frac{a_i}{n_i} \le V^*\\
&2 \le n_i \le N_i \qquad n_i \text{ , integer} \quad i = 1,2,3,...,\ L
\end{aligned}\right\} \quad (7)$$

The membership function defined in (7) 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

$$\left.\begin{aligned}
&V(\bar{y}) = \sum_{i=1}^{L} \frac{a_i}{n_i} \le V^*\\
&\mu_t(n) + \delta_t^+ = 1\\
&2 \le n_i \le N_i \qquad n_i \text{ , integer} \quad i = 1,2,3,..., L
\end{aligned}\right\} \quad (8)$$

Using (8) the (FGP) for (3) can be presented as

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

Subject to

$$\left.\begin{aligned}
&-0.0198n_1 - 0.0025n_2 - 0.0474n_3 + 0.9354 + \delta_1^+ = 1\\
&-0.021n_1 - 0.0210n_2 - 0.0210n_3 + 1.3483 + \delta_2^+ = 1\\
&\frac{33.7539}{n_1} + \frac{1.4330}{n_2} + \frac{1.3885}{n_3} \le 2.90\\
&2 \le n_1 \le 16,\ \ 2 \le n_2 \le 20,\ \ 2 \le n_1 \le 2
\end{aligned}\right\} \quad (9)$$

After solving FGPP (9) through LINGO[22] software, we get $n_1 = 15.30, n_2 = 3.52$, and $n_3 = 4.85$. Thus the obtained solution is non integer. In order to get the integer values NLGPP (9) is solved using Branch and Bound method. Various nodes of branch and Bound Method for NLGPP (9) are presented below in figure -1

In this method problem, P1 is divided into two branches P2 and P3.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, thus we choose P5 as it has more nearly optimal value function. Similarly P5 is divided into two branches P6 and P7, thus we choose P7 and is divided into again two branches P8 and P9. At node P9 we obtain integer value and the problem becomes fathomed. The integer optimal solution using Branch and Bound method is given by $n_1 = 15, n_2 = 4, and\ n_3 = 5$.

Conclusion

In this paper teiary objective stratified sampling design is formulated as an linear / non linear plus fractional goal programming (L/NLPFGP) is solved using fuzzy goal programming approach through LINGO. An integer optimal allocation is obtained using Branch and Bound method after solving the fuzzy goal programming problem.

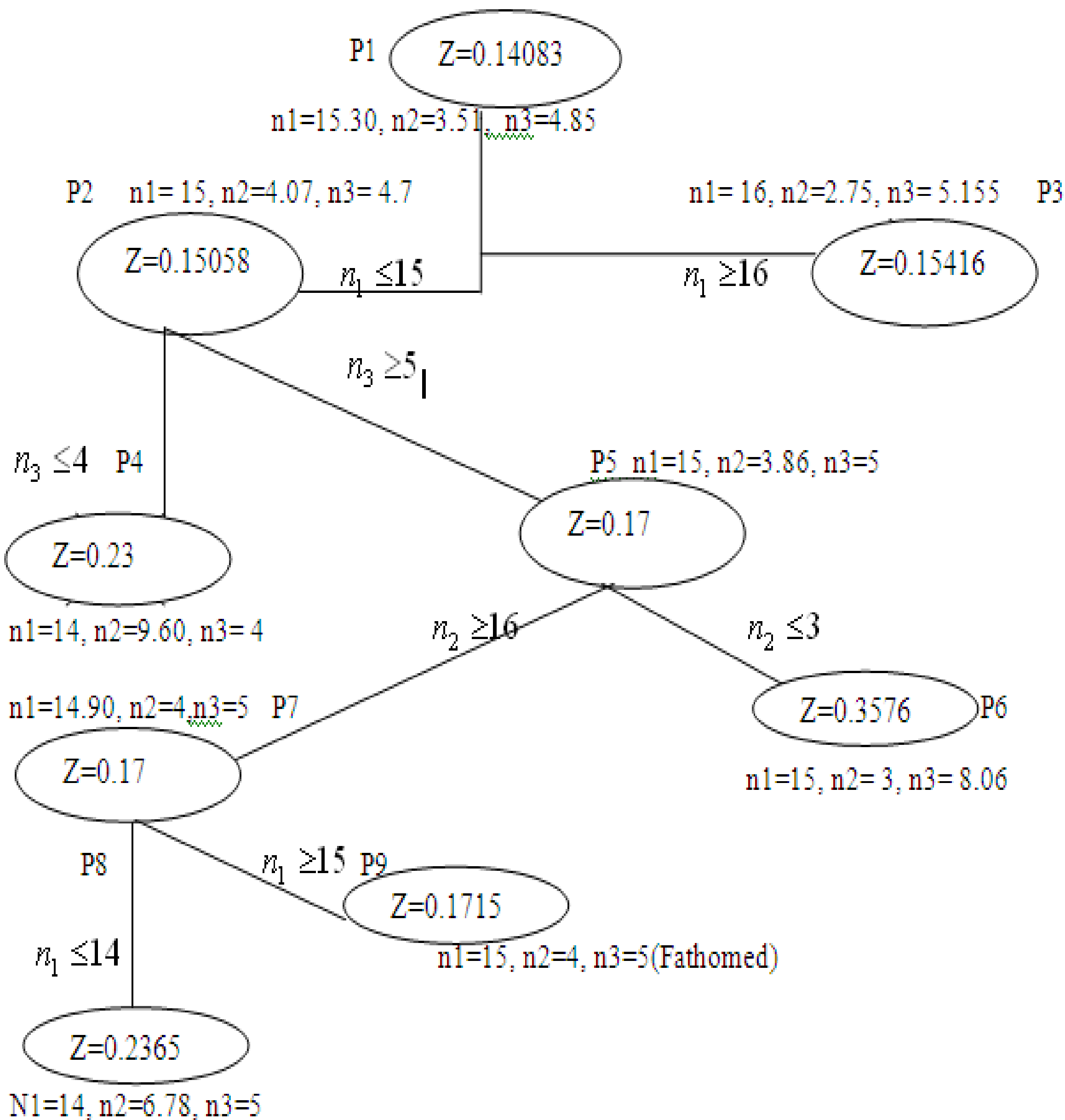

Figure-1
Various nodes of branch and Bound Method for NLGPP (9)

References

1. Neyman J., On the two Different aspects of the representative method: The method of stratified sampling and the method of purposive selection. *Journal of Royal Statistical Society* .,**97(4)**, 558-625 **(1934)**

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

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

4. Hartley H.O., Multiple purpose optimum allocation in stratified sampling, *Proc. Amer. Statist. Assoc., Social Statist.*, 258-261 **(1965)**

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

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

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

8. Kokan A.R. and Khan S.U., Optimum allocation in multivariate surveys: An analytical solution, *Journal of Royal Statistical Society*, B, **29,** 115-125 **(1967)**

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

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

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

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

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

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

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

16. Hirche J., A note on programming problems with linear plus linear fractional objective functions, *European Journal of Operation Research*, **26(1-2),** 49-64 **(1984)**

17. Chadha S.S., Dual of the sum of linear and linear fractional program. *European Journal of Operation Research*, **67,**136-139 **(1993)**

18. Jain S. and lachhwani K., Sum of linear and fractional multiobjective programming problem under Fuzzy rules constraints. *Australian journal of basic and Applied Science*, **4(2),** 105-108 **(2008)**

19. Schaible S., A note on the sum of linear and linear fractional functions, *Naval Research Logistic Quarterly,***24,** 961-963 **(1977)**

20. Sukhatme et al., Sampling theory of Surveys with Applications, *Iowa State University Press, Iowa, U. S. A. and Indian Society of Agricultural Statistics, New Delhi, India*, **(1984)**

21. Arthanari T.S. and Dodge Y., Mathematical programming in statistics, *New York: John Willy*, **(1993)**

22. LINGO, LINGO User’s Guide, LINDO Systems Inc.,1415, North Dayton Street, Chicago, Illinois, 60622, USA **(2013)**