ISSN 1684-8403
Journal of Statistics
Volume 24, 2017. pp. 112-119

Conversion of Tertiary Objective Stratified Sampling Design into Fractional Goal Programming

Mushtaq Ahmad Lone[1], *Shakeel Ahmad Mir*[2] *and Imran Khan*[3]

Abstract

The aim of this article is to determine an optimal allocation of sample sizes in Stratified Sampling design, when a tertiary objective Stratified Sampling design is converted into bi-objective fractional goal programming. When a non-integer solution is obtained after solving the bi-objective fractional goal programming through LINGO then Branch and Bound method provides an integer solution.

Keywords

Optimal allocation, Goal programming, Nonlinear programming, Fractional goal programming. Stratification and Branch and Bound method.

1. 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 consisting of N units is first divided into L sub-populations of sizes $N_1, N_2, \ldots, N_L$ units respectively.

[1] Division of Agriculture Statistics, SKUAST-Kashmir, Shalimar, Srinagar, Jammu and Kashmir, India-190025.
Email: lonemushtaq11@gmail.com
[2] Division of Agriculture Statistics, SKUAST-Kashmir, Shalimar, Srinagar, Jammu and Kashmir, India-190025.
Email: mir_98@msn.com
3 Division of Agriculture Statistics, SKUAST-Kashmir, Shalimar, Srinagar, Jammu and Kashmir, India-190025.
Email: imrankhan@skuastkashmir.ac.in

These sub-populations are non-overlapping and together they compromise the whole population i.e. $\sum_{i=1}^{L} N_i = N$. These sub-populations are called strata. These strata are so formed that they are homogeneous within and heterogeneous between. When the strata have been determined, a sample is drawn from each stratum, the drawing being 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. In stratification the Values of N_i must be known. However, in many practical situations, it is usually difficult to stratify with respect to the study variable under consideration 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 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. The problem of optimally choosing the sample sizes is known as the optimal allocation problem. The problem of optimal allocation in Stratified Sampling designs discussed by several authors (see, for example, Ahsan et al. (2005), Ansari et al. (2009), Aoyama (1963), Bethel (1989), Chatterjee (1972), Ghosh (1958), Gren (1966), Hartley (1965), Khan et al. (1997), Kozak (2006), Kreienbrock (1993), Singh and Tarray (2014), Yates (1960) etc). It has been seen that a lot of optimization problems from engineering, natural resources and economics require the optimization of a ratio between physical and/or economic functions and where the objective functions appear as a ratio or quotient of other functions, constitute a fractional programming problem. Von Neumann (1945) used generalized fractional programming (GFP) problem in modeling an expanding economy. Algorithms for solving linear fractional programming problems are well known in in Carnes and Cooper, (1962) and Isbell and Marlow, (1956).

In this article, a tertiary objective Stratified Sampling design is converted into bi-objective fractional programming problem and a goal programming approach is used to solve the programming problem through LINGO software. In practical situations, if the solution is non-integer then rounding of the non-integer sample sizes to the nearest integral values may sometime provide an infeasible and non-optimal solution. Thus in order to get the integer allocation instead of Rounding of non-integer values to the nearest integer, it is better to use Branch and Bound method. The Branch and Bound method consists of two strategies, alternatively followed, till the desired solution is obtained. One strategy consists in branching a problem into two sub problems, and the other in solving each of the two sub problems to obtain the minimum or suitable lower bound of the objective function.

2. Problem formulation

Consider a population of size N divided into L strata of sizes $N_1, N_2, \ldots N_L$ and N_i units in the ith stratum such that $\sum_{i=1}^{L} N_i = N$. 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

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

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.

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

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

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

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

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

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.

3. Fractional goal programming approach

For Multi-Objective Non-Linear Programming Problem (MONLPP) the proposed approach can be outlined and illustrated through an example as given below:

Step 1: Find the individual solution of each objective function subject to the system constraints.

Step 2: Find the maximum value of $F_t(n)$ says it $F_t(n^*)$ .where $F_t(n)$ is the t^{th} objective function.

Step 3: Divide all of the remaining $F_{t-1}(n)$ by $F_t(n^*)$.

Step 4: we get fractional programming say $\xi_t(n)$.

Step 5: Form the goal programming model.

Step 6: solve Step 5 through LINGO.

Step 7: if the solution is integer.
Step 8: stop. Otherwise,
Step 9: use Branch and Bound Method.

4. Numerical illustration

The population contains 64 units, the stratum weights and stratum variance of a population which has been divided into three strata with one characteristic under study is given below.

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 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 NLPP eq. (1.1) is written as when t=3:

$$\left.\begin{aligned}
&Min \;\; F_1(n) = 4n_1 + 1.5n_2 + n_3 \\
&Min \;\; F_2(n) = 4n_1 + 1.5n_2 + n_3 \\
&Min \; F_3(n) = \left[\frac{33.7539}{n_1} + \frac{1.4330}{n_2} + \frac{1.3885}{n_3}\right]^{1/2} \Big/ \; 30.52 \\
&subject \;\; to \\
&\frac{33.7539}{n_1} + \frac{1.4330}{n_2} + \frac{1.3885}{n_3} \le 2.90 \\
&2 \le n_1 \le 16, \;\; 2 \le n_2 \le 20, \;\; 2 \le n_1 \le 20
\end{aligned}\right\} \qquad (1.2)$$

Solving eq. (1.2) through LINGO using step 1, we get $F_1(n) = 70, \; F_2(n) = 23.15, and \; F_3(n) = 0.036$.

The maximum value is $F_1(n) = 70$ using step 2. After using step 3, we get

$$\frac{F_2(n)}{F_1(n)} = \xi_1(n) \;\; and \;\; \frac{F_3(n)}{F_1(n)} = \xi_2(n)$$

Thus the Fractional Goal Programming Problem (FGPP) can be represented as,

$$\left.\begin{array}{l} Min \quad \xi_1(n) \\ Min \quad \xi_2(n) \\ subject \;\; to \\ \dfrac{33.7539}{n_1} + \dfrac{1.4330}{n_2} + \dfrac{1.3885}{n_3} \le 2.90 \\ 2 \le n_1 \le 16, \;\; 2 \le n_2 \le 20, \;\;\; 2 \le n_1 \le 2 \end{array}\right\} \quad (1.3)$$

Thus the transformed problem eq. (1.3) has two goals. These goals can be attained by solving eq. (1.3) for each single objective i, e goal (1) = $k_1^{**} = 0.3170$and goal (2) = $k_2^{**} = 0.0002997$.

In order to determine the allocation of sample sizes, obtained by solving the Nonlinear Goal Programming Problem (NLGPP) given below. Also, it is assumed that both goals are equally important and same weight is assigned.

$$\left.\begin{array}{l} Min \;\; Z = 0.5(\delta_1^+ + \delta_1^-) + 0.5(\delta_2^+ + \delta_2^-) \\ subject \;\; to \\ (n_1 + n_2 + n_3)/(4n_1 + 1.5n_2 + n_3) - \delta_1^+ + \delta_1^- \le 0.3170 \\ \left\{ \left(\dfrac{33.7539}{n_1} + \dfrac{1.4330}{n_2} + \dfrac{1.3885}{n_3} \right)^{1/2} / \; 30.52 \right\} / (4n_1 + 1.5n_2 + n_3) - \delta_2^+ + \delta_2^- <= 0.0002997 \\ \dfrac{33.7539}{n_1} + \dfrac{1.4330}{n_2} + \dfrac{1.3885}{n_3} \le 2.90 \\ 2 \le n_1 \le 16, \;\; 2 \le n_2 \le 20, \;\;\; 2 \le n_1 \le 28 \end{array}\right\} \quad (1.4)$$

After solving FGPP eq. (1.4) through LINGO software, we get $n_1 = 16, n_2 = 3.86, n_3 = 3.32, \delta_1^+ = 0.018, \, , \delta_2^+ = 0, \;\; \delta_3^+ = 0.00008 \;\; and \;\; \delta_3^+ = 0$ Thus the obtained solution is noninteger. In order to get the integer values NLGPP eq.

(1.4) is solved using Branch and Bound method. The integer optimal solution using Branch and Bound method is given by $n_1 = 16, n_2 = 5, and\, n_3 = 3.$

5. Conclusion

This paper concludes that how a tertiary objective stratified sampling design is converted into bi-objective fractional programming problem which is solved through a goal programming approach. In practical situations, rounding of non-integer allocation to the nearest integral values may sometime provide an infeasible and non-optimal solution. In order to get an integer allocation then the developed algorithm shows that how one gets an integer optimum allocation of sample sizes for a conversed programming problem. This approach can be easily applied for Multi- objective conversed programming problems.

References

1. Ahsan, M. J., Najmussehar and Khan, M. G. M. (2005). Mixed allocation in Stratified Sampling. *Aligrah Journal of Statistics Sciences*, **25**, 87-97.
2. 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.
3. Aoyama H., (1963). Stratified Random Sampling with optimum allocation for multivariate populations. *Annals of the Institute of Statistical Mathematics,* **14**, 251–258.
4. Bethel J. (1989). Sample Allocation in multivariate surveys. *Survey Methodology*, **15**, 40-57.
5. Carnes and Cooper, W. W. (1962). Programming with Linear Fractional Functional. *Naval Research Logistics Quartely,* **9**, 181-186.
6. Chatterjee, S. (1972). A study of optimum allocation in Multivariate Stratified surveys. *Skandinavisk Actuarietidskrift*, **55**, 73-80.
7. Ghosh, S. P. (1958). A note on Stratified Random Sampling with multiple characters. *Calcutta Statistical Bulletin*, **8**, 81-89.
8. Gren, J. (1966). Some Application of Non-linear Programming in Sampling Methods. *Przeglad Statystyczny*, **13**, 203–217.
9. 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.*

10. Isbell, J. R. and Marlow, W. H. (1956). Atrition games, *Naval Research Logistics Quartely*, **3**, 1-99.
11. Khan. M. G. M., Ahsan, M. J. and Jahan, N. (1997). Compromise allocation in Multivariate Stratified Sampling: An integer solution. *Naval Research Logistics*, **44**, 69-79.
12. Kozok M., (2006). On sample allocation in multivariate surveys. *Communication in Statistics-Simulation and Computation*, **35**, 901-910.
13. 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.
14. Singh, H. P. and Tarray, T. A. (2014). A Stratified Mangat and Singh's optional randomized response model using proportional and optimal allocation. *Statistica*, **73(1)**, 61-79.
15. Von Neumann, J. (1945). A model of general economic equilibrium. *Review Economic Studies*, **13**, 1–9.
16. Yates, F. (1960). *Sampling Methods for Censuses and Surveys*. (3rd ed.) Charles Griffin, London.