International Journal of Applied Information Systems (IJAIS) – ISSN : 2249-0868
Foundation of Computer Science FCS, New York, USA
Volume 8– No.2, January 2015 – www.ijais.org

A Note on Optimum Allocation with Non-Linear Cost Function

Mohd. Vaseem Ismail, Ph.D.
Jamia Hamdard,
New Delhi, India

Manoj Kr. Sharma
Mewar University
Rajasthan, India

E. A. Khan, Ph.D.
Jamia Hamdard,
New Delhi, India

Kaynat Naseer, Ph.D.
Integral University
Lucknow, India

ABSTRACT
In this paper we consider the optimum allocation for multivariate sampling with non-linear cost function –travel cost. The problem of determining the optimum allocations are formulated as Nonlinear Programming Problems , in which each NLPP has a convex objective function and a non-linear cost constraint. The NLLP's are then solved using Lagrange Multiplier technique and the explicit formula for variance is obtained.

Keywords:
Multivariate sampling, Travel Cost, Optimum Allocation, Nonlinear programming problem.

1. INTRODUCTION
Stratified sampling is the most popular among various sampling designs that are extensively used in sample survey. When a stratified sampling is to be used a sampler has to deal with three basic problems such as (i) the problem of determining the number of strata, (ii) the problem of cutting the stratum boundaries and (iii) the problem of optimum allocation of sample sizes to various strata. In stratified sampling the values of the sample sizes n_h in the respective strata are chosen by the sampler. They may be selected to minimize $V\left(\overline{y}_{st}\right)$ for a specified cost of taking the sample or to minimize the cost for a specified value of $V\left(\overline{y}_{st}\right)$.

The general cost function is of the form

$$\text{Cost} = C = c_o + \sum_{h=1}^{L} c_h n_h{}^{\alpha}$$

Within any stratum the cost is proportional to the size of sample, but the cost per unit c_h may vary from stratum to stratum. The term c_o represents an overhead cost. If travel costs between units are substantial, empirical and mathematical studies suggest that travel costs are better represented by the expression $\sum_{h=1}^{L} t_h \sqrt{n_h}$ if $\alpha = 1/2$ and c_h is replaced by t_h where t_h is the travel cost per unit (Beardwood et al., 1959).

In this paper a method of optimum allocation for multivariate stratified sampling is developed for the non-linear cost function. The problem of determining the optimum allocations are formulated as Nonlinear Programming Problems , in which each NLPP has a convex objective function and a non-linear cost constraint. Several techniques are available for solving these NLPP's , we used Lagrange Multiplier technique to solve the optimum allocation of the value of sample size n_h.

2. FORMULATION OF THE PROBLEM WHEN $C = c_o + \sum_{h=1}^{L} t_h \sqrt{n_h}$ ($\alpha = 1/2$)
In stratified random sampling with a linear cost function, the variance of the estimated mean $\overline{y}_{st}$ is a minimum for a specified cost C, and the cost is a minimum for specified variance $V\left(\overline{y}_{st}\right)$ when $n_h \propto W_h S_h / \sqrt{c_h}$

Suppose that p characteristics are measured on each unit of a population which is partitioned into L strata. Let n_h be the number of units to be drawn with out replacement from the h^{th} stratum $(h = 1,2,..., L)$. For the j^{th} character an unbiased estimate of the population mean $\overline{Y}_j$ is $\overline{y}_{jst}$ whose variance is given by

$$V\left(\overline{y}_{jst}\right) = \sum_{h=1}^{L} W_h^2 S_{hj}^2 X_h, \qquad j = 1,2,...., p \tag{2.1}$$

where

$$W_h = \frac{N_h}{N}, S_{hj}^2 = \frac{1}{N_h - 1} \sum_{i=1}^{N_h} \left(y_{hji} - \overline{Y}_j\right)^2$$

and $X_h = \left(\frac{1}{n_h} - \frac{1}{N_h}\right)$, in usual notations.

Let t_h be the travel cost of enumerating all the p characters on a single unit in the h^{th} stratum. The total cost of survey may be given as

$$C = c_o + \sum_{h=1}^{L} c_h n_h + \sum_{h=1}^{L} t_h \sqrt{n_h} \tag{2.2}$$

Where c_0 is the overhead cost. If we consider that each sample has constant measurement cost then $\sum_{h=1}^{L} c_h n_h = c\sum_{h=1}^{L} n_h = nc$ which can be merged in c_0 then the equation (2.2) reduces to

$$C = c_o + \sum_{h=1}^{L} t_h \sqrt{n_h} \qquad (2.3)$$

For a fixed budget C_o, the problem of determining an optimum allocation may be expressed as the following NLPP:

Minimize $$Z = \sum_{h=1}^{L} W_h^2 S_{hj}^2 X_h$$

Subject to $$\sum_{h=1}^{L} t_h \sqrt{n_h} \le C_0 \qquad (2.4)$$

and

$$1 \le n_h \le N_h; \qquad h = 1, 2, ..., L$$

where $C_o = C - c_o$.

The optimum choice of n_h for an individual characteristic can be determined by minimizing the variance in (2.1) for the given cost in (2.3), or by minimizing the cost for fixed variance. We can use Lagrange multipliers technique to determine the optimum value of n_h.

The Lagrange function ϕ is defined as

$$\phi(n_h, \lambda) = \sum_{h=1}^{L}\left(\frac{1}{n_h} - \frac{1}{N_h}\right) W_h^2 S_h^2 + \lambda\left(\sum_{h=1}^{L} t_h \sqrt{n_h} - C_o\right) \qquad (2.5)$$

where λ is a Lagrange multiplier.

The necessary conditions for the solution of the problem are

$$\frac{\partial \phi}{\partial n_h} = -\sum_{h=1}^{L} \frac{1}{n_h^2} W_h^2 S_h^2 + \frac{1}{2}\lambda \sum_{h=1}^{L} \frac{t_h}{\sqrt{n_h}} = 0 \qquad (2.6)$$

$$\frac{\partial \phi}{\partial \lambda} = \sum_{h=1}^{L} t_h \sqrt{n_h} - C_o = 0 \qquad (2.7)$$

Solving (2.6) we get

$$n_h = \left(\frac{2W_h^2 S_h^2}{\lambda t_h}\right)^{2/3} \qquad (2.8)$$

$$\sum_{h=1}^{L} n_h = n \qquad (2.9)$$

(2.9) gives

$$\lambda = \frac{2}{n^{3/2}}\left\{\sum_{h=1}^{L}\left(\frac{W_h^2 S_h^2}{t_h}\right)^{2/3}\right\}^{3/2} \qquad (2.10)$$

Putting the value of λ in (2.8) we get

$$n_h = n\left(\frac{W_h^2 S_h^2}{t_h}\right)^{2/3} \sum_{h=1}^{L}\left(\frac{t_h}{W_h^2 S_h^2}\right)^{2/3} \qquad (2.11)$$

This gives

$$V\left(\bar{y}_{st}\right) = \frac{1}{n}\left\{\sum_{h=1}^{L}\left(W_h S_h t_h\right)^{2/3} \sum_{h=1}^{L}\left(\frac{W_h^2 S_h^2}{t_h}\right)^{2/3}\right\} \qquad (2.12)$$

Ignoring the term $\frac{1}{N_h}$

3. FORMULATION OF THE PROBLEM

when

$$C = c_o + \sum_{h=1}^{L} c_h n_h^2 \quad (\alpha = 1)$$

Using Lagrange function we get

$$n_h = \left(\frac{W_h^2 S_h^2}{\lambda C_h}\right)^{1/2} \qquad (3.1)$$

$$\lambda = \frac{1}{n^2}\left[\sum_{h=1}^{L}\left(\frac{W_h^2 S_h^2}{C_h}\right)^{1/2}\right]^2 \qquad (3.2)$$

$$V\left(\bar{y}_{st}\right) = \frac{1}{n}\sum_{h=1}^{L}\left(W_h^2 S_h^2 C_h\right)^{1/2} \sum_{h=1}^{L}\left(\frac{W_h^2 S_h^2}{C_h}\right)^{1/2} \qquad (3.3)$$

4. FORMULATION OF THE PROBLEM WHEN $C = c_o + \sum_{h=1}^{L} c_h n_h^2$ ($\alpha = 2$)

Using Lagrange function we get

$$n_h = n\left(\frac{W_h^2 S_h^2}{c_h}\right)^{1/3} \sum_{h=1}^{L}\left(\frac{c_h}{W_h^2 S_h^2}\right)^{1/3} \qquad (4.1)$$

$$\lambda = \frac{1}{2n^3}\left[\sum_{h=1}^{L}\left(\frac{W_h^2 S_h^2}{c_h}\right)^{1/3}\right]^{1/3} \qquad (4.2)$$

$$V\left(\bar{y}_{st}\right) = \frac{1}{n}\left\{\sum_{h=1}^{L}\left(W_h^4 S_h^4 c_h\right)^{1/3} \sum_{h=1}^{L}\left(\frac{W_h^2 S_h^2}{c_h}\right)^{1/3}\right\} \qquad (4.3)$$

5. FORMULATION OF THE PROBLEM

WHEN $C = c_o + \sum_{h=1}^{L} c_h n_h^3 \quad (\alpha = 3)$

Using Lagrange function we get

$$n_h = n\left(\frac{W_h^2 S_h^2}{c_h}\right)^{1/4} \sum_{h=1}^{L}\left(\frac{c_h}{W_h^2 S_h^2}\right)^{1/4} \quad (5.1)$$

$$\lambda = \frac{1}{3n^4}\left[\sum_{h=1}^{L}\left(\frac{W_h^2 S_h^2}{c_h}\right)^{1/4}\right]^{1/4} \quad (5.2)$$

$$V\left(\bar{y}_{st}\right) = \frac{1}{n}\left\{\sum_{h=1}^{L}\left(W_h^6 S_h^6 c_h\right)^{1/4} \sum_{h=1}^{L}\left(\frac{W_h^2 S_h^2}{c_h}\right)^{1/4}\right\} \quad (5.3)$$

6. NUMERICAL EXAMPLE

Consider a population divided in two strata with single characteristic under study for which the values of W_h, S_h, c_h are given in the following table:

Table 1.1

Stratum h	W_i	S_h	C_h
1	0.40	4	1
2	0.30	5	2

Let us fix the budget at 100 units

Solving variance for ($\alpha = 1/2$, $\alpha = 1, \alpha = 2, \alpha = 3$) , taking the value of $n = 1000$, we get

$$V\left(\bar{y}_{st,\alpha=1/2}\right) = 0.010182$$

$$V\left(\bar{y}_{st,\alpha=1}\right) = 0.00939$$

$$V\left(\bar{y}_{st,\alpha=2}\right) = 0.013947$$

$$V\left(\bar{y}_{st,\alpha=3}\right) = 0.004745$$

7. CONCLUSSION

So for optimum allocation in multivariate sampling we will minimize the cost by minimizing the variance. In future we will develop a computer program for above formulation

8. REFERENCES

[1] Cochran, W.G., Sampling Techniques. John Wiley & Sons, New York.(1977)

[2] Kokan, A.R and Khan, S.U.,Optimum Allocation in Multivariate Surveys-An analytical Solution. Jour. Roy. Stat. Soc. Ser. B.29, 115-125 (1967).

[3] Khan, M.G.M.,Chand,Munish A. and Ahmad Nesar., Optimum Allocation in Two-stage and Stratified Two-stage Sampling for Multivariate Surveys. ASA Section on Survey Research Methods (2006).

[4] Singh, S., Advanced Sampling Theory with Applications:How Michael “selected” Amy. Springer(2003)

[5] Chatterjee, S., Multivariate Stratified Surveys. Jour. Amer. Stat. Assoc. 63, 530-534 (1968).

[6] Neyman, J., On the Two Different Aspects of Representative Method: The Method of Stratified Sampling and the Method of Purposive Selection. Jour Roy. Stas. Soc., 97, 558-606 (1934)

[7] Yates, F., Sampling Methods for Censuses and Surveys (2nd ed). Charles Griffin and Co. Ltd. London (1960)

[8] Aoyama, H., Stratified Random Sampling with Optimum Allocation for Multivariate Populations. Ann. Inst. Stat. Math., 14, 251-258 (1963)

[9] Ghosh, S. P., A Note on Stratified Random Sampling with Multiple Characters. Cal. Stat. Assoc. Bull., 8, 81-89 (1958).

[10] Bethel, J., An Optimum Allocation Algorithm for Multivariate Survey. Proceedings of the Survey Research Section, ASA, 204-212 (1985)