Jour. Ind. Soc. Ag. Statistics
Vol. XL, No. 2 (1988), *pp*. 105-110

ON OPTIMUM SAMPLE SIZE WITH CONSTRAINTS ON PRECISIONS OF STRATUM ESTIMATES

SHAMBHU DAYAL* and RAJESH KUMAR
Indian Institute of Sugarcane Research, Lucknow-226002

(Received : June, 1986)

SUMMARY

The procedures for determining optimum sample size have been given when stratum level estimates are required with specified precisions and the precision of the overall mean is also fixed.

Keywords : Optimum sample size, Constraints, Optimum allocation.

Introduction

The procedure for allocation of sample size when estimates of stratum level means are required with specified precisions was given by Dayal [2]. There are, however, associated problems of determining the optimum sample size when estimates of stratum level means are required with specified precisions and the precision of the over-all mean is also fixed; for which solutions are not available at present. The problems are acutely faced by practising statisticians because of the greater emphasis being laid these days on reasonably precise estimates at lower (stratum) level; whereas the method given by Neyman [3] aims at determining the optimum sample size subject to given precision of the over-all mean only. In this paper, solutions to these problems are given following Dayal [2].

* Central Water Commission, New Delhi.

2. Procedure for Optimum Sample Size

Let us first minimise

$$n = n_1 + n_2 + \ldots + n_L$$

subject to

$$\sum_{h=1}^{L} W_h^2 S_h^2 / n_h - \sum_{h=1}^{L} W_h^2 S_h^2 / N_h$$

$$= V, \text{ a constant,} \tag{2.1}$$

where $N_h/N = W_h$. The minimum value of n will be obtained when (Cochran, [1])

$$n_h = \left(W_h S_h \sum_{h=1}^{L} S_h \right) \Big/ \left(V + \sum_{h=1}^{L} W_h^2 S_h^2 / N_h \right) \tag{2.2}$$

and will be given by

$$n = \left(\sum_{h=1}^{L} W_h S_h \right)^2 \Big/ \left(V + \sum_{h=1}^{L} W_h^2 S_h^2 \Big/ N_h \right) \tag{2.3}$$

Suppose the minimum precision in terms of relative variance of the mean for the hth stratum is given by

$$\left\{ V(\bar{y}_h) / \bar{Y}_h^2 \right\} \leqslant b_h, \quad h = 1, 2, \ldots, L.$$

where b_h's are given constants. This gives

$$n_h \geqslant a_h, \quad h = 1, 2, \ldots, L$$

or

$$a_h - n_h \leqslant 0, \quad h = 1, 2, \ldots, L \tag{2.4}$$

where a_h are constants depending upon b_h, N_n, and S_h. New, the problem is to minimise 'n' subject to constraints (2.1) and (2.4). To solve the problem, we first minimise n subject to (2.1) only, i.e. put the value of n_h as given at (2.2) for each stratum. The values of n_h, thus found, may not satisfy constraints (2.4) for all the strata. To such strata where (2.4) is not satisfied, we allocate a_h and find the variances of means of these strata and add them up. That is, we find

$$\Sigma\, W_h^2\, S_h^2/a_h \;-\; \Sigma\, S_h^2\, W_h^2/N_h = V', \text{ say} \tag{2.5}$$

for such strata for which (2.4) has not been satisfied by values of n_h as found above.

Next, we minimise

$$n' = \Sigma\, n_h$$

for the remaining strata subject to

$$\Sigma\, W_h\, S_h^2/n_h \;-\; \Sigma\, W_h^2\, S_h^2/N_h = V - V' \tag{2.6}$$

for the remaining strata in the same manner as done earlier. Again, if the values of n_h, thus found, do not satisfy constraints at (2.4) for all the remaining strata, we allocate a_h to such strata where (2.4) has not been satisfied and further proceed as above. The process will continue till we reach a stage when constraints at (2.4) are satisfied for all the strata. The constraints at (2.1) will obviously be satisfied because of constraints like (2.5) being put at each stage of minimisation. The allocation, thus found, will be the optimum allocation and the value of n, obtained by adding the sample allocated to each stratum, will give the overall optimum sample size.

3. Procedure for Minimum Cost

Let the cost function be

$$C = C_0 + \sum_{h=1}^{L} C_h\, n_h \tag{3.1}$$

where C_0 is the overhead cost, C_h is the cost of surveying one unit in hth stratum and C is the total cost. The problem is to minimise (3.1) subject to (2.1) and (2.4).

If we have only constraint (2.1), the minimum value of (3.1) will be given when (Cochran, [1]):

$$n_h = \frac{(W_h S_h/\sqrt{C_h}) \sum_{h=1}^{L} \sqrt{C_h}\; W_h S_h}{V + \sum_{h=1}^{L} \frac{W_h^2 S_h^2}{N_h}} \tag{3.2}$$

If the constraint (2.4) is also imposed, we first minimise (3.1) subject to (2.1) only, i.e. put the value of n_h as given at (3.2) for each stratum. The values of n_h, thus found, may not satisfy constraints at (2.4) for all the strata. To such strata where (2.4) is not satisfied, we allocate a_h and find the variances of means of these strata and add them up. That is, we find

$$\sum_{h=1}^{L} W_h S_h^2/a_h - \sum_{h=1}^{L} S_h^2 W_h^2/N_h = V_1, \text{ say} \tag{3.3}$$

for such strata for which (2.4) has not been satisfied by values of n_h as found above.

Next, we minimise

$$C' = C_0 + \sum_{h=1}^{L} C_h n_h$$

for the remaining strata subject to

$$\sum_{h=1}^{L} W_h S_h^2/n_h - \sum_{h=1}^{L} W_h^2 S_h^2/N_h = V - V_1 \tag{3.4}$$

for the remaining strata in the same manner as done earlier. The process will continue till we reach a stage when constraints at (2.4) are satisfied for all the strata. The allocation, thus found, will be the optimum allocation and the cost with this allocation will be the minimum cost.

Note. Situations may arise that in order to achieve the specified precision at stratum level, all the a_h values are more than n_h values due to Neyman allocation. In such cases we have to choose a_h values only and the overall constraint becomes superfluous.

4. Illustration

The procedure given in Section 2 is illustrated with the help of an example, Table 1 gives the values of $W_h S_h$ and a_h for 16 strata. It is also given that the variance of the overall mean, V, ignoring finite population correction (f.p.c.) should be 220. The problem is to determine optimum sample size along with its allocation to various strata. In column 4, the allocation is done following (2.2) ignoring f.p.c. The sample size comes as 4210 but it is found that for stratum numbers 1, 3, 5, 8, 9 and 16, constraints (2.4) are not satisfied. For these strata, we

TABLE 1 : OPTIMUM SAMPLE SIZE AND ITS ALLOCATION

Stratum No.	$W_h S_h$	a_h	n_h	n_h	n_h
1	*2*	*3*	*4*	*5*	*6*
1.	27.30	213	110	213	213
2.	267.22	756	1169	1106	1099
3.	6.71	34	29	34	34
4.	2.25	10	10	9	10
5.	2.09	21	9	21	21
6.	76.77	176	327	309	307
7.	142.80	524	625	591	587
8.	12.77	94	56	94	94
9.	18.68	132	82	132	132
10.	46.46	52	203	192	191
11.	111.96	207	490	463	460
12.	47 60	200	208	196	200
13.	36.03	34	158	149	148
14.	71.69	310	314	270	310
15.	51.49	78	226	213	211
16.	42.54	388	186	388	388
TOTAL :	964.36	3229	4202	4380	4405
VARIANCE	—	382	220	220	220

therefore put $n_h = a_h$ in column 5, and then, V, the sum of the variances of means for these six strata, comes to 14.06. $V - V'$ is therefore equal to 205.94. Putting $V - V'$ for V in (2.2) ignoring f.p.c., n_h are determined for the remaining 10 strata and put in column 5. The total sample size comes to 4380. It is seen that for stratum numbers 4, 12 and 14, constraints (2.4) are not satisfied. For these strata also, we therefore put $n_h = a_h$ in column 6 and repeat the process, as given above. It is found that now the constraints (2.4) are satisfied for all the strata. The optimum sample size is found by adding all the values of n_h given in column 6 and is 4405. The optimum allocation of the sample size is also given in column 6.

ACKNOWLEDGEMENT

The authors are grateful to the referees for their valuable comments.

REFERENCES

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

[2] Dayal, S. (1982) : Sample size and its optimum allocation with constraints on precision of stratum estimates — An extension of Neymen's method, *Jour. Ind. Soc. Agri. Stat.* XXXIV (3).

[3] Neyman, J. (1934) : On the two different aspects of the representative method. The method of stratified sampling and the method of purposive selection, *Jour. Roy, Stat. Soc.* 97 : 558-606.