Jour. Ind. Soc. Ag. Statistics
47(2), 1995 : 192–196

On Systematic Sampling Allowing Estimation of Variance of Mean

S. Ray and M.N. Das*
Central Statistical Organisation, New Delhi
(Received, August, 1994)

SUMMARY

A slightly modified circular systematic sampling scheme is presented. It is equally simple and provides estimate of variance of mean. As joint probabilities of inclusion of pairs of units, are unequal, Horvitz-Thompson method of estimation has been adopted. These probabilities can be obtained quite easily.

Key words : Circular Systematic Sampling, Horvitz-Thompson Method of estimation, Estimation of Variance.

Introduction

Systematic sampling provides a very convenient scheme of sampling as compared to other sampling schemes. It is used widely in different situations. It is being used extensively in different surveys conducted by the National Sample Survey Organisation, Government of India. This technique has, however, the drawback that though it provides unbiased estimate of population mean, it cannot provide an estimate of the variance of the estimate of mean. As much the user cannot get any idea of precision of the estimate. The usual systematic sampling scheme requires that population size N is an exact multiple of sample size. Lahiri (1954) suggested a modification of systematic sampling which does not have the above drawback and called it circular systematic sampling. This scheme also has the drawback that it cannot provide estimate of variance. In this paper a modified technique of circular systematic sampling is provided. This technique ensures unbiased estimates of variance of mean and at the same time does not affect the simplicity of the existing scheme. Though the scheme works for both linear and circular systematic technique, we are emphasising on circular systematic sampling as this scheme works for any population and sample sizes. The linear systematic sampling is also covered by the present method.

2. *A Modified Circular Systematic Sampling Technique.*

Let N denote the size of population and n, the sample size. The units are provided serial numbers in any order.

* Formerly at, IASRI, New Delhi.

5. *Conclusions*

In general the LER can be analysed using any of the standardisation without much loss in the precision of comparisons. The fear that different divisors used for computation of PLER and LER would violate the basic structure of additive model and usher in fresh problems about the assumptions of the non-normality appears to be little unfounded. And agronomists may use different sole crop treatments in their experiment, for meaningful interpretation. Considering all the three aspects of additivity, normality and precision, the standardisation treatmentwise over all the blocks appears to be the best.

REFERENCES

[1] Jagannath, M.K. and Sundararaj, N., 1987. Productivity Equivalent Ratio and Statistical Testing of its advantages in intercropping. *Jour. Ind. Soci. Ag. Statistics.* **30** : 289-300

[2] Mead, R and Wiley, R.W., 1980. The concept of 'Land Equivalent Ratio' and advantages in yield from intercropping. *Experimental Agriculture* **16** : 217-228.

[3] Oyejola, B.A. and Mead, R., 1982. Statistical assessment of different ways of calculating Land Equivalent Ratios (LER). *Experimental Agriculture*. **18** : 125-138.

[4] Kendall McDonald and S.K. Katti., 1975. Test for normality using a characterisation in '*Statistical Distribution of Scientific Work*'. **2**. 91-104. Eds. G.P. Patil, S. Kortz and J.K. Ord. D. Reidel Publishing Co. Boston. U.S.A.

[5] Ramanath Chetty and Narayana Reddy. M., 1984. Analysis of Intercrop Experiments in dryland agriculture. *Experimental Agriculture*. **20** : 31-40.

[6] Scheffe, 1959. *Analysis of Variance*. 345-358. John Wiley Publication, Canada.

[7] Snedecor and Cochran, 1969. *Statistical Methods. Fifth Edition. 321. The Iowa State University Press, Ames, Iawa, USA.*

Step : I. A circular systematic sample of size (n–1) is first drawn by following the existing method by taking $K = \frac{N}{n-1}$, that is, the integer part smaller than $\frac{N}{n-1}$. When N is divisible by n–1 which is the case of linear systematic sampling, K has to be so taken that it does not divide N. Normal choice of K in this case is $K = \frac{N}{n-1} - 1$ or $K = \frac{N}{n-1} + 1$.

Step : II. Next, one more unit is drawn at random from among the $N - (n-1) = N - n + 1$ units left after the sample in Step : I is drawn.

Step : III. The n units obtained from Step–I and Step–2 form the required sample.

3. *Estimation*

The proposed estimation procedure is Horvitz-Thompson method of estimation. Accordingly, we require the inclusion probability of the i-th unit, πi which, of course, is constant and thus independent of i and the inclusion probability of the i-th and j-th units together, π_{ij} $(i, j = 1, 2, ..., N)$

3 (a). Values of π_i's

The i–th unit is included in a sample if the systematic sample drawn contains the i–th unit or the systematic sample drawn does not contain the i–th unit but the extra unit drawn subsequently is the i–th unit. The probability of the i–th unit being included in a sample is then

$$\pi_i = P_r \text{ (i–th unit in systematic sample)}$$

$$+ P_r' \text{ (i–th unit not in systematic sample and the extra unit is the i–th unit)}$$

$$= \frac{n-1}{N} + \left(1 - \frac{n-1}{N}\right) \frac{1}{N-(n-1)}$$

$$= \frac{n-1}{N} + \frac{(N-n+1)}{N} \cdot \frac{1}{(N-n+1)}$$

$$= \frac{n-1}{N} + \frac{1}{N}$$

$$= \frac{n}{N} = \text{constant.}$$

Hence, every unit being included in a sample has the same probability.

3 (b) Value of π_{ij}'s

The inclusion probability of the i–th and the j–th units as per original numbering of the units depends on i and j. If the i–th and j–th units occur together in λ_{ij} samples in the original systematic sampling scheme at *step-I*, then the total number of samples in which both the i–th and j–th units occur together finally comes out as

$$\begin{aligned}\Delta_{ij} &= \lambda_{ij}(N-n+1)+2(n-1-\lambda_{ij})\\ &= \lambda_{ij}(N-n+1-2)+2(n-1)\\ &= \lambda_{ij}(N-n-1)+2(n-1)\end{aligned} \tag{3.1}$$

As the total number of samples is N (N – n + 1)

$$\pi_{ij} = \frac{\Delta_{ij}}{N(N-n+1)} \tag{3.2}$$

The values of λ_{ij} can be obtained conveniently as below when the sample units have the same numbering as used initially for drawing the systematic sample in *step-I*.

Let p = abs (i – j).

If $\frac{p}{K}$ or $\frac{N-p}{K}$ is an integer, s then

$$\left.\begin{array}{l}\lambda_{ij} = n-s-1 \text{ when } (n-s-1)>0\\ \text{Otherwise } \lambda_{ij}=0\end{array}\right| \tag{3.3}$$

This shows that Δ_{ij} is non-zero for all possible choices of i and j. Hence, when Horvitz-Thompson method of estimation is adopted variance of the mean is estimable. As the inclusion probability of each unit is the same, usual arithmatic mean of the sample observations gives an unbiased estimate of the population mean.

4. *Estimate of variance*

In almost all sampling schemes π_i and π_{ij} are both constants or both unequal. The present scheme has π_i's constant but π_{ij}'s are unequal. As such the formula for variance estimatae can be expressed in an alternative simpler form. But putting

$\pi_i = \pi_{(i=1,2,\ldots N)}$, and after some reduction the estimated variance becomes

$$\hat{V} = \frac{1}{N^2}\left[N\,\bar{y}^2 - \frac{N}{n}\left\{\sum_{i=1}^{n} y_i^2 + \frac{2n}{n}\sum_{i=1}^{n-1}\sum_{j>i}^{n}\frac{y_i\,y_j}{\pi_{ij}}\right\}\right]$$

For linear systematic sample if adopted π_{ij} takes two values viz. $\frac{1}{K}$ for two units which occur in a systematic sample $K = N/(n-1)$ and $\frac{2(n-1)}{N(N-n+1)}$ for two units one of which is a unit drawn subsequently.

5. *An Example : $N = 25, n = 6$*

The size of the systematic sample is $6 - 1 = 5$. As (n–1) divides N, we have to take the interval $K = 5 - 1 = 4$ required for systematic sample. For actually drawng the sample take a non-zero random number less than or equal to 25. If this number be R, then the sample is R, R+4, R+8, R+12, R+16. If any of these numbers be greater than N (= 25), then 25 is to be substracted from each of them. If R = 10, the samle is 10, 14, 18, 22, 1. If any number happends to be 0, the it is to be replaced by 25. Next, from the units other than those in the above sample, one more units is selected at random. Let it be denoted by Ro = 17, say. Then the final sample is 10, 14, 18, 22, 1, 17. The mean of these observations from these units gives an unbiased estimate of the population mean.

For obtaining an unbiased estimate of variance of the mean, all the π_{ij}'s for the units in the sample are required. These can be obtained easily from (3.1), (3.2) and (3.3).

Here, $\Delta_{ij} = 18\ \lambda_{ij} + 10$

where λ_{ij} is as given in (3.3).

Case I. To get. $\Delta_{10,14}$ we need $\lambda_{10,14}$.

As $p = \text{abs}\ (10 - 14) = 4$ and $k = 4$, we have $s = \frac{p}{k} = \frac{4}{4} = 1$.

Hence,

$$\begin{aligned}\lambda_{10,14} &= n - s - 1\\ &= 6 - 1 - 1\\ &= 4.\end{aligned}$$

$$\begin{aligned}\Delta_{10,14} &= \lambda_{10,14\,(N-n-1)+2\,(n-1)}\\ &= 18 \times 4 + 10\\ &= 72 + 10\end{aligned}$$

$$\pi_{10,14} = \frac{82}{25 \times 20} = \frac{82}{250}$$

Case 2. Again $i = 17$ and $j = 18$.

$$p = \frac{1}{4} \text{ or } \frac{25-1}{4} = 6$$

So, $\lambda_{17,18} = 6 - 6 - 1 = -1 < 0$

As $\lambda_{17,18} < 0$, $\lambda_{17,18} = 0$

$\because$ $\Delta_{17,18} = 10$

and $$\pi_{17,18} = \frac{10}{250} = \frac{1}{25}$$

Similarly, other Δ_{ij}'s can be obtained easily.

REFERENCE

[1] Lahiri, D.B., 1954. On the question of bias of systematic sampling; Proceedings of World Population Conference, 6, 349-362.