ON GENERALIZED SEQUENTIAL ESTIMATORS

By

FAURAN SINGH[1]

AND

DAROGA SINGH[2]

(Received : April, 1979)

1. Introduction

The concept of linear estimators has been generalized by Godambe [1]. Murthy [4] gave admissible linear estimators while Koop [2], Prabhu-Ajgeonkar [3] and others generalized the concept of admissible linear estimators and sampling designs. Singh [5] extended the idea by discussing acceptable sequential estimators. Chaudhary and Singh [6] gave generalized sequential estimators. Let us consider only such parametric functions, $\theta(=\theta(Y))$ that can be expressed as a sum of single-valued set functions defined over a class A (for details, Op. Cit. 1977), *i.e.*, θ is such a function then

$$\theta = \sum_{ai \in A} \lambda_i f(ai) \tag{1.1}$$

where f (ai) is a single valued set function defined over the class A,

$\sum_{ai \in A}$ is the summation over all sets 'ai' belonging

to class A, and λ_i is some adjustment constant.

Elaborating the idea further a more general method to express a parametric function (1.1) may be written as

$$\theta = \sum_{ai \in A} \lambda_i \quad \pi_{ai} \quad f(ai) \tag{1.2}$$

where π_{ai} is a probability measure defined over a_i in the class A.

1. Directorate of Research, Haryana Agricultural University, Hissar-125004,

2. I.A.S.R.I. New Delhi.

For instance-the population mean $\bar{y}$ for a characteristic y can be expressed as θ in (1.2) with 'ai' as a point set $\{U_i\}$, $f(ai)$ as y_i along with probability measure πa_i as $\frac{1}{N}$ and λ_i as 1. Similarly the population variance σ^2 can be expressed as (1.2) a point set of two units $\{U_i, U_j\}$ with $f(ai)$ and πai as $(y_i - y_j)^2$ and $\frac{1}{N^2}$ respectively with λ_i as 1.

Thus it is one of the main reasons that the probability samples are preferred over non-probability samples, whatever the sampling scheme of design may be, there is always search of a method which provides a technique to evaluate it (the probability measure) in a form acceptable to all occasions. In the present paper an attempt has been made to derive some generalized sequential estimators to the parameters expressed in (1.2) in terms of the above lines.

2. Sequential Estimators

A 'Statistic' t defined over the probability field is a function over the samples. A statistic used to estimate a parametric function θ is called an estimator to θ and the most general form of a linear estimator may be

$$t = \sum_{ai \in s} \phi(ai)\, f(ai, s) \Big/ \sum_{ai \in s} f(ai, s) \qquad \ldots(2.1)$$

where $f(ai, s)$ is a probability measure defined over s and $\phi(ai)$ is a function of point set ai in s.

In case

$$\sum_{ai \in s} f(ai, s) = 1,$$

then t is called an unbiased estimator. The estimator (2.1) may also be written as

$$t = \sum_{ai \in s} \phi(ai)\, p(s/E\,ai) \Big/ \sum_{ai \in s} p(s/E\,ai) \qquad \ldots(2.2)$$

Where E_{ai} is an event depending on the occurrence of the set. 'ai' in the sample and $p(s/E_{ai})$ is a probability measure for a given E_{ai}. In other words, the event E_{ai} may be termed as a 'terminal event', since it decides the number of observations in the sample and sampling terminates accordingly *i.e.*, the stopping rules are defined in terms of this event. In non-sequential process, the minimum value of the loss in estimating the mean value with known variance is 2 $C^{1/2}\sigma$

where C is the cost per observation. So in a sequential process, sampling will terminate as soon as it touches this value. Similarly there may be a number of ways to define the terminal event.

In a sequential sampling scheme the requirement

$$\sum_{ai \in s} P\ (s/E\ ai)=1$$

is always satisfied at every step of sampling. Thus a class of sequential estimators may be obtained by defining the event E_{ai} in a number of ways and the estimator given by the relation (2.2) may be termed as a 'generating estimator' for sequential sampling schemes.

3. Illustrations

Let the units selected in the sample be $(U_1, U_2, \ldots U_n \ldots)$ with variate $(y_1, y_2 \ldots y_n \ldots)$ respectively. The sampling procedure is sequential, *i.e.*, the units are selected one-by-one, the estimate of the populations mean shall be given sequentially as below:

After first observation $\quad \bar{y}_1=y_1$

After second observation $\quad \bar{y}_2=\dfrac{y_1+y_2}{2}$

After third observation $\quad \bar{y}_3=\dfrac{y_1+y_2+y_3}{3}$

.
.
.
.

After n^{th} observation $\bar{y}_n \quad =\dfrac{y_1+y_2+y_3+\ldots\ldots y_n}{n}$

Hence we may define a sequential mean as

$$\bar{y}_{seq} \quad =\sum_{i=1}^{n} y_i/n$$

Where n is not fixed in advance.

The sequence of these means in terms of previous one can be expressed as below:

$$\bar{y}_2 \quad =\tfrac{1}{2}\,\bar{y}_1+\tfrac{1}{2}\,y_2$$

$$\bar{y}_3 \quad =\tfrac{2}{3}\,\bar{y}_2+\tfrac{1}{3}\,y_3$$

$$\bar{y}_4 \quad =\tfrac{3}{4}\,\bar{y}_3+\tfrac{1}{4}\,yn$$

.
.
.

$$\bar{y}_{seq} \quad =\bar{y}_n=\frac{n-1}{n}\,\bar{y}_{n-1}+\frac{1}{n}\,y_n$$

Another interesting example of sequentinal mean is

$$\bar{y}_n = \frac{\bar{y}_{n-1} + y_n}{2}$$

$$= \frac{y_1}{2^{n-1}} + \frac{y_2}{2^{n-1}} + \frac{y_3}{2^{n-2}} + \ldots + \frac{y_{n-1}}{2^2} + \frac{y_n}{2}$$

which gives rise to a random sample mean with unequal probability for different items.

Hence we can write sequential estimators in a generalized form as

$$\bar{Y}_{seq} = \pi_{n-1} \bar{y}_{n-1} + \pi_n y_n$$

where $\pi_{n-1} + \pi_n = 1$

π_n is the probability of introducing the n^{th} unit in the sample.

Here it should also be noted that at every stage this relation holds good.

Now let us consider a finite population with N units, and let a sample of size 2, 3, ... be drawn with known variate values and with some defined probability field. The sequential estimates of the mean are sketched as below:

Sample size (n)	*Probability eld/probability* (Ω) (π_n)	*Variate values* (y)	*Estimate* ($\bar{y}_{seq}$)
2	$\Omega_2 = (\frac{1}{2}, \frac{1}{2})$	1, 2	$\bar{y}_2 = \frac{3}{2}$
3	$\pi_3 = \frac{1}{3}$	1	$\bar{y}_3 = \frac{4}{3}$
4	$\pi_4 = \frac{1}{4}$	0	$\bar{y}_4 = 1$
5	$\pi_5 = \frac{1}{5}$	1	$\bar{y}_5 = 1$

We can define terminal event in terms of cost/loss or risk functions. As soon as it touches the given limit d, it is decided to terminate sampling accordingly.

SUMMARY

Starting with generalized linear estimators for simple random sampling we proceed to define generalized sequential estimators for a parametric function by taking the basic concept of probability measure into consideration. Recently some authors have given an approach todefining both parameter and estimator in the light of this concept. This paper is a modest attempt to extend the idea further and contributes that sequential estimators can be used in restricted situations also. To establish practical utility, some illustrations have been made.

ACKNOWLEDGEMENT

The authors acknowledge the valuable comments given by the referees of this paper.

REFERENCES

[1] Godambe, V.P. (1955) : A unified theory of sampling from finite populations, *J.R. Statist. Soc.* B, **17**, 268-78.

[2] Koop. J.C. (1963) : On the axioms of sample formation and their bearing on the construction of linear estimators in sampling theory from finite univers, Part I and Part II. *Matrika*, **7**, 81-114 and 165-204.

[3] Prabhu-Ajgaonkor, S.G. (1969) : A best estimator for the entire class of linear estimators. *Sankhyā*, A, **31**, 455-62.

[4] Murthy, M.N. (1963) : Generalized unbiased estimation in sampling from finite populations. *Sankhyā*, A, **25**, 245-62.

[5] Singh, F. (1977) : Sequential approach to sample surveys. Ph.D. thesis, Meerut University.

[6] Chaudhary, F.S. and Singh, J.B. (1979) : On the generalized sequential estimators, *J. Ind. Soc. Agri. Statist.*, **31**, 72.