ACCEPTABLE SEQUENTIAL ESTIMATORS OF POPULATION MEAN*

FAURAN SINGH, AND D. SINGH
H.A.U , Hissar *I.A.S.R.I., New Delhi*
(Received : April, 1979)

SUMMARY

The sampling structure D (Ω, t, d) has been discussed to bring a class of acceptable estimators for population mean. To devolve sequential estimation the probability samples have been used with a class of acceptable estimators. Ultimately sampling terminates if and when the risk function of D touches a fixed or a minimum value.

INTRODUCTION

Wald (1947) used concept of admissible decisions in decision theory. Godambe (1960), Roy and Chakravorty (1960) and Hanurav (1968) applied the admissibility concept to sampling theory. Murthy and Singh (1969), Joshi (1968) Prabhu Ajgaonkar (1969) and others contributed on best and admissible estimators in sampling for finite population. Singh (1977) has applied the concept to sequential sampling with some modification and a class of acceptable estimators has been discussed. The probability samples with variable sample size have been used and an attempt has been made to devolve sequential estimation with a class of acceptable estimators.

Definition 1.

Let a finite population u consist N distinguishable units U_i associated with a real variate y_i, $i=1, 2......, N$. A parameter $\theta=[\underset{\sim}{\theta}=(y_1, y_2,..., y_N)]$ is a point in the Euclidean space or class of point sets (for bravity class R).

Usually the problem is to estimate θ on the basis of individuals i sampled from the population u and the values y_i associated them, *i.e.*, on the basis of s (s, $x_i \in s$) where s is a subset of u drawn with a given sampling design p.

*The work was supported partially by a grant from the University Grants Commission (Grant No. SR I-25-3/7585/77).

Sampling Design : is any function p on A, the set of all possible subsets of s of u such that $p(s) \geqslant 0$, $\Sigma\, p(s)=1$, $s \in A$.

Probability Field : consider a non-negative function P defined for every combination $(y_{j1}, y_{j2}, .. y_{jn})$ of s_j. A probability measure may be constructed in which the combination $(y_{j1}, y_{j2}, ... y_{jn})$ will be sampled with probability proportionate to $P\,(y_{j1}, y_{j2}, ..., y_{jn})$ over the combination such that $\Sigma P=1$ and it will be referred as a probability field (Ω).

Definition 2.

Parametric Functions : Let us consider the parametric functions defined by Chaudhary and Singh (1979) say, $\theta\,[=\underset{\sim}{\theta}\,(y)]$ that can be expressed over the class A, *i.e.*, a more general method to express a parametric function may be

$$\theta = \sum_{ai \in A} \lambda_i\, \pi_{ai}\, f(ai) \qquad ...(1.1)$$

where $f(ai)$ is a single-valued set function defined over the class A, is the summation over all sets 'ai' belonging to class A, π_{ai} is a probability measure defined over ai in the class A, and λ_i is some adjustment constant.

Definition 3.

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

$$t = \sum_{ai \in A} f(ai)\ \phi\,(ai,\ s) \Big/ \sum_{ai \in s} \phi\,(ai, s) \qquad ...(1.2)$$

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

In case $\sum_{ai \in s} \phi\,(ai, s)=1$, then the estimator t is called an unbiased estimators.

The estimator (1.2) may also be as

$$t = \sum_{ai \in s} f(a)\, p\,(ai, s/E_{ai}) \Big/ \sum_{ai \in s} p\,(ai, s/E_{ai}) \qquad ...(1.3)$$

where E_{ai} is an event depending on the occurence of the set 'ai' in sample s and $p\,(ai, s/E_{ai})$ is a probability measure for ai when E_{ai} has occured.

Definition 4.

Sampling System : It may be considered as the specification of all possible samples alongwith their probability fields over the combination of units in the sample with reference to θ, *i.e.*, it is a combination of estimators of ordered sequence of samples s from u with probability field (Ω) symbolically, $F \equiv F(t, \Omega)$.

PRELIMINARIES (SEQUENTIAL ESTIMATES)

Let the units selected in the sample be $U_1, U_2, \ldots, U_n$ with variate values $y_1, y_2, \ldots, y_n$, respectively. The sampling procedure is sequential with replacement, *i.e.*, the units are selected 1— by —1, with replacement, a sequential mean may be defined as

$$\bar{y}_{n_{seq}} = \sum_{i=1}^{n} y_i/n \qquad \ldots(2.1)$$

where n is also a variate.

Which can also be expressed in terms of the previous estimators.

$$\bar{y}_{n_{seq}} = (n-1)/n \; \bar{y}_{n-1} + 1/n \; y_n \qquad \ldots(2.2)$$

where $\bar{y}_{n-1}$ is the previous estimate, and y_n is the observation under consideration with $\frac{1}{n}$ as its probability of inclusion in the sample.

Thus a sampling system in the present procedure is as follows :

(*i*) after the first observation, the estimate is $\bar{y}_1 = y_1$ with probability field $\Omega_1 \equiv \{1\}$ so as to form a sample of size 1.

(*ii*) after the second observation, the estimate is $\bar{y}_2 = \frac{1}{2} \bar{y}_1 + \frac{1}{2} y_2$, with probability field $\Omega_2 = \{\frac{1}{2}, \frac{1}{2}\}$ so as to form a sample of size 2 such that

$$\sum_{i \in s} p_i = 1$$

(*iii*) after the third observation, the estimate is $\bar{y}_3 = \frac{2}{3} \bar{y}_2 + \frac{1}{3} y_3$ with probablitity field $\Omega_3 = \{\frac{2}{3}, \frac{2}{3}\}$ so as to form a sample of size 3 such that

$$\sum_{i \in s} p_i = 1$$

and so on Thus sampling will contiuue,

Another illustration of sequential estimate is

$$\bar{y}_{n_{seq}} = (\bar{y}_{n-1} + y_n)/2 \qquad \text{...(2.3)}$$

which gives equal importance to the value in hand (nth observation) and the estimate on previous occasion, *i.e.*, $\bar{y}_{n-1}$.

Thus the sampling system in this case will be as follows :

(*i*) after the first observation, the estimate is $\bar{y}_1 = y_1$ with probability field $\{1\}$ so as to form a sample of size 1.

(*ii*) after the second observation, the estimate is $\bar{y}_2 = \frac{1}{2}\bar{y}_1 + \frac{1}{2} y_2$ with probability field $\Omega_2 = \{\frac{1}{2}. \frac{1}{2})$ so as to form a sample of size 2 such that

$$\sum_{i \in s} p_i = 1.$$

(*iii*) after the third observation, the estimate is $\bar{y}_3 = \frac{1}{2}\bar{y}_2 + \frac{1}{2} y_3$ with probability field $\Omega_3 = \{\frac{1}{2}, \frac{1}{2}\}$ so as to form a sample of size 3 such that

$$\sum_{i \in s} p_i = 1.$$

Thus sampling will continue.

If we expand (2.3), it can be seen that

$$\bar{y}_{n_{seq}} = y_1/2^{n-1} + y_2/2^{n-1} + y_3/2^{n-2} + \ldots + y_{n-1}/2^2 + y_n/2 \qquad \text{...(2.4)}$$

which gives rise to a sequential sampling with unequal probability to different units.

The relations (2.2), (2.3) and (2.4) give the idea that these simple estimators may be extended with generalized probablity measures and can be written in a generalized form as

$$\left.\begin{aligned} \bar{y}_{n_{seq}} &= \pi_{n-1}\ \bar{y}_{n-1} + \pi_n\ y_n \\ &= (1-\pi_n)\, \bar{y}_{n-1} + \pi_n\ y_n \end{aligned}\right\} \qquad \text{...(2.5)}$$

where π_n is the probability of introducing the nth unit in the sample, and π_{n-1} is the probability of getting $\bar{y}_{n-1}$ when introducing the nth unit in the sample such that $\pi_{n-1} + \pi_n = 1$.

Here it should be noted that at every step of sampling the relation between probabilities holds good. To make the idea more clear for defining a sequential sample system with a probability field generated, let us take an example to illustrate the concept by drawing samples of size 2, 3, . with some known variate values and with some defined probability measures. After selection of a first unit U_1 if a second unit U_2 with its variate value y_2 will be included in the sample with probability of inclusion $(1+p)/2$ then U_1 with its variate value y_1 will be included in sample $(0 \leqslant p \leqslant 1)$ with probability $(1-p)/2$. So the generated probability field is a set $\Omega_2=\{(1-p)/2, (1+p)/2\}$ and the estimator t_{seq} may be written as

$$\bar{y}_2=y_1\ (1-p)/2+y_2\ (1+p)/2$$

Further if a third unit U_3 with variate value y_3 is included in the sample with probability of inclusion $\pi_3=(1+p)/3$ then the changed generated probability field is a set $\Omega_3=\{(1-p)/3, (1+p)/3\}$ and hence the estimator t_{seq} may be written as

$$\bar{y}_3=\bar{y}_2\ (1-p)/3+y_3\times(1+p)/3.$$

Succeeding in a similar manner and assuming y_4,...the variate values at 4th,...draws, the system of sequential sampling may be continued.

It can be shown that the estimators defined by (2.2), (2.3), (2.4), and (2.5) are unbiased at every step of sampling *i.e.*, the sequential mean $\bar{y}_{seq}$ is an unbiased estimator of $\bar{Y}$.

SEQUENTIAL DECISION RULE

Before deciding how the sampling process will terminate, some terms used within the text may be defined :

A sequential decision rule is a pair (ϕ, δ) in which ϕ is a stopping rule and δ is a terminal rule.

Stopping Rule : A stopping rule is a sequence of function

$$\phi\ (y)=(\phi_0,\ \phi_1\ (y_1),\ \phi_2\ (y_1,\ y_2),\ldots) \qquad \ldots(3.1)$$

with $\phi_j\ (y_1, y_2,\ldots y_j)$ such that $0 \leqslant \phi_j \leqslant 1$ for all j.

Where ϕ stands for conditional probability that the experimenter will cease sampling, given that he has taken j observations.

Terminal Rule : A terminal rule is a sequence of functions

$$\delta\ (y)=(\delta_0,\ \delta_1,\ (y_1),\ \delta_2\ (y_1,\ y_2),\ldots) \qquad \ldots(3.2)$$

for all j.

δ_j is a sequential terminal rule for a statistical decision problem in the probability distribution σ-field for which expected loss $E\ [\theta\ (\phi,\ \delta)]$ is finite.

Risk Function : The risk function of a sequential decision rule (ϕ, δ) is the expected value of the risk when θ is the true value of the parameter and will be denoted as $d\equiv E\ [\theta,\ (\phi,\ \delta)]$, ...(3.3)

Sampling Structure : A sampling system F alongwith its risk function d defined over the probability field Ω is called a sampling structure for estimation of θ, symbolically,

$$D \equiv D\ (F, d) \equiv (t, \Omega, d) \qquad \ldots(3.4)$$

A sampling structure D is said to be unbiased if t is an unbiased estimator of θ.

A sampling structure is said to be ultimate acceptable if t is consistent and a minimum risk unbiased estimator (*MRUE*). If the risk of D_1 is smaller than that of D_2, *i.e.*, $d_1 < d_2$, then t_1 is said be an acceptable estimator.

ACCEPTABLE ESTIMATORS

Let us attempt to develop a most acceptable form of a sampling structure D. If or when the risk function of D touches a fixed or a minimum value, the process will be terminated.

Given by a sequential scheme the sample units, arranged in ascending order of their unit indices form an 'order statistic' which may be written as

$$t = [y_{(1)},\ y_{(2)},\ \ldots,\ y_{(n)}]$$

Another order statistic if all units are distinct, may be formed

$$t' = [y_{(1)},\ y_{(2)},\ \ldots, y_{(v)}]$$

Such statistics are sufficient to form a class of estimators which belong to some defined probability fields. Let us consider sequential sampling with replacement method in which two types of estimators of population mean are defined as

$$\bar{y}_n = y/n \equiv \text{ave of all sample units, and}$$

$$\bar{y}_v = y/v = \text{ave of all } v \text{ distinct units in the sample.}$$

In a sequential sample with n as variate size the number of distinct units v is also a random variate. So a generalized form of sequential estimators of the population mean may be written as

$$\bar{y}_{n_{seq}} = \phi_1(v) + \phi_2(v)\bar{y}_v \qquad \ldots(4.1)$$

where $\phi_1(v)$ and $\phi_2(v)$ are two functions.

If the number of distinct units v in the sample is taken as given, we have,

$$E(\bar{y}_{n_{seq}}/v) = E\{[\phi_1(v) + \phi_2(v)y_v]/v\}$$

$$= \phi_1(v) + \phi_2(v)\bar{Y}$$

Therefore $E(\bar{y}_{n_{seq}}) = E(\phi_1(v)) + E(\phi_2(v))\bar{Y} \qquad \ldots(4.2)$

Obviously the necessary and sufficient conditions of $\bar{y}_{n_{seq}}$ being an unbiased estimator of $\bar{Y}$ are

$$E(\phi_1(v))=0 \text{ and } E(\phi_2(v))=1 \qquad \ldots(4.3)$$

In order to choose an acceptable estimator from the field, the risk function of sampling structure D should be minimized.[5] Thus the problem is to choose unbiased estimators in Ω, which have a minimum variance and satisfy the conditions in (4.3). Hence we can minimize the variance

$$V(\bar{y}_{n_{seq}})=V(\phi_1(v)+(v)\bar{y}_v)$$

Firstly, we find the variance of $\bar{y}_v$ when v is given, *i.e.*,

$$V(\bar{y}_{v_{seq}}/v)=(1/v-1/N)S^2$$

where $\quad S^2=N\sigma^2/(N-1)$

Also we know that the variance of a variate is the sum of the expected value of the conditional variance and the variance of the conditional expected value or symbolically,

$$V(x)=E_1(V_2(x))+V_1(E_2(x))$$

Thus we get,

$$V(\bar{y}_{n_{seq}})=E(\phi^2_2(v)).(1/v-1/N)S^2$$
$$+E[\phi_1(v)+\phi_2(v)\bar{y}_v-E\{(\phi_1(v)+\phi_2(v)\bar{y}\}]^2 \quad \ldots(4.4)$$

To choose a sampling structure which gives unbiased estimators having, uniformly minimum risk, a proper choice of $\phi_1(v)$ and $\phi_2(v)$ should be made in conformity with conditions in (4.3). Minimizing the first term in (4.4) which involves $\phi_2(v)$ only, we have by using Swartz's inequality,

$$\phi_2(v)=(1/v-1/N)^{-1}/E(1/v-1/N)^{-1}$$
$$=vN/(N-v)/E(vN/(N-v)) \qquad \ldots(4.5)$$

Minimizing the sccond term in (4.4) in terms of $\phi_2(v)$, we have,

$$\phi_1(v)=\bar{Y}(1-\phi_2(v)) \qquad \ldots(4.6)$$

Since (4.6) expresses $\phi_1(v)$ in terms of unknown population parameter $\bar{Y}$, the value of $\phi_1(v)$ cannot be determined except the case where $\phi_2(v)=1$. Thus $\bar{y}_{n_{seq}}$ reduces to $\bar{y}_v$ only and it is one of the explanations that the sample size was fixed in advance in random samples. If some prior information about $\bar{Y}$ is avaliable say $\bar{X}$ then (4.6) may be written as

$$\phi_1(v)=\bar{X}(1-\phi_2(v)) \qquad \ldots(4.7)$$

Hence the estimator (4.1) becomes

$$\bar{y}_{n_{seq}}=\bar{X}[1-vN/(N-v)/E(vN/(N-v))] + vN/(N-v)/E(vN/(N-v))\,\bar{y}_v \qquad ...(4.8)$$

On the other hand if no information about $\bar{X}$ is given then taking $\bar{X}=0$, we write (4.8) as

$$\bar{y}''_{n_{seq}}=vN/(N-v)/E(vN/(N-v))\,\bar{y}_v \qquad ...(4.9)$$

Further if v/N can be ignored then (4.8) and (4.9) become

$$\bar{y}^{*}_{n_{seq}}=\bar{X}(1-v/E(v))+v/E(v)\bar{y}_v \qquad ...(4.10)$$

and

$$\bar{y}^{**}_{n_{seq}}=v/E(v)\bar{y}_v \qquad ...(4.11)$$

respectively.

In relations (4.8), (4.9), (4.10) and (4.11) some sequential estimators of mean have been derived and they can generate commonly used estimators if some restrictions are imposed. In order to define a sampling structure D completely, the risk function should be derived or in other words the sampling variance of these estimators should be calculated. Since the estimators (4.8) and (4.9) involve the term $E(vN/(N-v))$ which requires a ready reference and hence it has been postponed to discuss them at present. Now we have considered the estimators given in (4.10) and (4.11) and their sampling variances have been derived.

If v is taken as given, we have,

$$V(y_v/v)=(1/v-1/N)S^2$$

therefore,

$$V(\bar{y}_v)=(E(1/v)-1/N)S^2=kS^2 \qquad ...(5.1)$$

Similarly we find that

$$V(\bar{y}^{**}_{n_{seq}}=[E(v)^{-1}-E(v^2)/N(E(v)^2]S^2 + \bar{Y}^2(V(v)/(E(v))^2) \qquad ...(5.2)$$

Comparing (5.1) and (5.2) we get,

$$V(\bar{y}_v)-V(\bar{y}^{**}_{n_{seq}}=k_1^2\;S^2-k_2^2\;\bar{Y}^2 \qquad ...(5.3)$$

where

$$k_1^2=[E(v)E(1/v)]/E(v)\;+V(v)/N(E(v))^2$$

and

$$k_2^2=V(v)/(E(v))^2$$

Since $E(v)E(1/v)>1$ and therefore co-efficient of both $\bar{Y}^2$ and S^2 are positive. Hence we conclude that $\bar{y}_v$ is better than

$$\bar{y}^{**}_{n_{seq}} \quad \text{if } cv \leqslant |\, k_2/k_1 | \qquad \text{...(5.4)}$$

worse otherwise.

Similarly the variance of the eastimator given by (4.10) may be obtained by

$$V(\bar{y}^{*}_{n_{seq}}) = V(v/E(v) \quad (\bar{y}_v - \bar{X})$$

So if we have some prior information of $\bar{Y}$, we find that there is only a change, *viz.*, $\bar{y}_v$ has been replaced by $(\bar{y}_v - \bar{X})$. Thus proceeding on similar lines as above it can be shown easily that $\bar{y}_v$ is better than $\bar{y}^*$

if

$$S^2/(\bar{Y} - \bar{X})^2 \leqslant k_2^2 \,/\, k_1^2 \qquad \text{...(5.5)}$$

worse otherwise.

This result also shows that if $\bar{X}$ provides a close approximation to $\bar{Y}$ then it is always better to use $\bar{y}^{*}_{n_{seq}}$ rather than $\bar{y}_v$.

REFERENCES

Chaudhary, F.S. & Singh (1979) J.B. : On generalized sequential estimators, *J. Ind. Soc. Agri. Statist.* **31**, 72.

Godambe, V.P. 1960 : An admissible estimate for any sampling design *Sankhya*, **22**, 285-88.

Hanurav, T.V. (1968) : Hyper-admissibility and optimum estimators for sampling finite populations. *Ann. Math. Statist.*, **39**, 621-42.

Joshi, V.M. (1968) : Admissibility of the sample mean as estimate of the mean of a finite population. *Ann. Math. Statist.*, **39**, 606-20.

Murthy, M.N. & Singh, M.P. (1969) : On the concepts of best admissible estimators in sampling theory. *Sankhya*, 31, 343-54.

Prabhu Ajgonkar, S.G. (1969) : On the non-existence of a best estimators for the entire class of linear estimators, *Sankhya*, **31**, 455-62.

Roy, J. & Chakravorty, I.M. (1960) : Estimating the mean of a finite population. *Ann. Math. Statist.*, 31, 392-98.

Singh, F. (1977) : A sequential approach to sample surveys. *Ph.D. Thesis, Meerut Uni.., India.*

Wald, A. (1947) : An essentially complete class of admissible decision functions. *Ann. Math. Statist.*, **18**, 549-55.