MULTI-PURPOSE SURVEYS ON SUCCESSIVE OCCASIONS IN A TWO-STAGE DESIGN

C. L. AGARWAL AND P. C. GUPTA
Rajasthan University, Jaipur
(Received : October, 1975)

INTRODUCTION

In repeated sampling enquiries the application of successive sampling technique, with partial replacement of sampling units on subsequent occasion, has many advantages. It may be convenient, cheaper and some times necessary to study a number of correlated characters instead of a single one. The several characters on successive occasions was first discussed by Tikkiwal [6], [8] for unistage sampling design. He, however, assumed the following correlation pattern :—

The correlation between the i^{th} and j^{th} characters studied on r^{th} and s^{th} $(s \geqslant r)$ occasions respectively, is given by

$$(i) \quad \rho_{rs}^{ij} = \rho_{rs} = \prod_{t=r}^{s-1} \rho_{t\, t+1} \text{ for } i=j,\ r \neq s$$

where $\rho_{t\, t+1}$ is the correlation between t^{th} and $(t+1)^{th}$ occasions for the same character.

$$(ii) \quad \rho_{rs}^{ij} = \rho'_{ij} = \prod_{t'=1}^{j-1} \rho'_{t'\, t'+1} \text{ for } i \neq j,\ r=s$$

where $\rho'_{t'\, t'+1}$ is the correlation between t'^{th} and $(t'+1)^{th}$ characters for the same occasion.

$$(iii) \quad \rho_{rs}^{ij} = \rho'_{ij}\, \rho_{rs} \text{ for } i \neq j \text{ and } r \neq s$$

Singh and Singh [3] has developed estimators for various characters for $h (h \geqslant 2)$ occasions without restricting the correlation pattern in a unistage design. Further, they have extended these results for a two-stage sampling design for a specific replacement pattern which allows replacement of first stage units (fsu's) only.

Abraham *et al* [1] have considered a general replacement pattern for two-stage sampling design which allows the replacement of fsu's as well as of second stage units (ssu's). They have obtained the results for two occasions only. Srivastava and Shivtar Singh [4] have obtained the results for more than two occasions under this general replacement pattern. In both these papers the authors have assumed that the two-stage population has equal first stage units. Mostly in multi-purpose surveys the population has unequal first stage units. In the present investigation both the cases are considered and the two different sampling schemes have been used for these two cases such that the unbiased estimators of the population mean belonging to the same class of estimators *i e.*, belong to T_{11}-class given by Koop [2].

2. The Replacement Pattern and Sampling Schemes

Let the population consists of N fsu's and within i^{th} fsu there are M_i ssu's. Further let n and m be the sample sizes for fsu's and ssu's respectively. The most general replacement pattern for a two-stage sampling design for two successive occasions with partial replacement of units can be defined as follows :

Retain a fraction 'p' of fsu's and from each retained fsu retain a fraction r of ssu's from the earlier occasion and select a fraction s $(r+s=1)$ of ssu's afresh. Further select a fraction q of fsu's afresh $(p+q=1)$. We shall consider the following two cases :

Case I : Two-Stage Sampling : Equal First Stage Units

In this case $M_i=M$ for all $i=1, 2,..., N$ and we adopt simple random sample without replacement (SRSWOR) at both the stages.

Case II : Two-Stage Sampling : Unequal First Stage Units

In this case $M_i \neq M$ for all $i=1, 2,..., N$ and we select fsu's with probabilities proportional to their sizes and with replacement (PPSWR), size being the number of ssu's with fsu, and the ssu's are selected with SRSWOR within each selected fsu. If the i^{th} fsu happens to be selected δ_i times then δ_i samples of ssu's each of size m are drawn independently.

3. Notations :

In this section we shall define all the statistics and notations which are used in subsequent sections for both the cases discussed above.

$x_{h(k)ij}$ = observation on j^{th} ssu of i^{th} fsu for the k^{th} character studied at h^{th} occasion, $i=1, 2,...N$; $j=1, 2,...M_i$; $k, h=1, 2.$

$$\bar{X}_{h(k)i}=\frac{1}{M}\sum_{j=1}^{M_i} x_{h(k)ij}$$

$$\bar{X}_{h(k)}=\frac{1}{M_o}\sum_{i=1}^{N}\sum_{j=1}^{M_i} x_{h(k)ij}$$

$$M_o=\sum_{i=1}^{N} M_i$$

$$\bar{x}_{h(k)i}=\frac{1}{m}\sum_{j=1}^{m} x_{h(k)ij}$$

$$\bar{x}_{h(k)}=\frac{1}{n}\sum_{i=1}^{n} \bar{x}_{h(k)i}$$

$$\bar{x}'_{h(k)}=\frac{1}{npmr}\sum_{i=1}^{np}\sum_{j=1}^{mr} x_{h(k)ij},$$ mean per ssu for the k^{th} character on the h^{th} occasion based on $npmr$ ssu's common to preceding occasion.

$$\bar{x}''_{h(k)}=\frac{1}{npms}\sum_{i=1}^{np}\sum_{j=1}^{ms} x_{h(k)ij},$$ mean per ssu, for the k^{th} character, on h^{th} occasion based on np fsu's which are common to preceding occasion but within fsu no ssu is common to preceding occasion.

$$\bar{x}'''_{h(k)}=\frac{1}{nqm}\sum_{i=1}^{nq}\sum_{j=1}^{m} x_{h(k)ij},$$ mean per ssu for the k^{th} character, on the h^{th} occasion, based on nqm ssu's which are not common to preceding occasion.

Case I : For $M_i=M$; $i=1, 2,\ldots, N$.

$$S_{h(k)h'(k')b}=\frac{1}{N-1}\sum_{i=1}^{N}(\bar{X}_{h(k)i}-\bar{X}_{h(k)})(\bar{X}_{h'(k')i}-\bar{X}_{h'(k')})$$

$$\bar{S}_{h(k)h'(k')w}=\frac{1}{N(M-1)}\sum_{i=1}^{N}\sum_{j=1}^{M}(x_{h(k)ij}-\bar{X}_{h(k)i})(x_{h'(k')ij}-\bar{X}_{h'(k')i})$$

Case II : For $M_i\neq M$; $i=1, 2,\ldots, N$.

$$\sigma_{h(k)h'(k')b}=\sum_{i=1}^{N}\frac{M_i}{M_o}(\bar{X}_{h(k)i}-\bar{X}_{h(k)})(\bar{X}_{h'(k')i}-\bar{X}_{h'(k')})$$

$$S_{h(k)h'(k')i}=\frac{1}{M_i-1}\sum_{i=1}^{M_i}(\bar{X}_{h(k)ij}-\bar{X}_{h(k)i})(\bar{X}_{h'(k')ij}-\bar{X}_{h'(k')i})$$

4. Estimation of Population Mean and its Variance

We shall first consider the case of two characters studied on two successive occasions. The unbiased estimator for the population mean, for second character on second occasion, $\hat{\bar{X}}_{2(2)}$ may be obtained as

$$\hat{\bar{X}}_{2(2)}=AX' \qquad \ldots(4.1)$$

where

$$X=C(\bar{x}'_{1(1)}-\bar{x}''_{1(1)}),\ldots,(\bar{x}'_{1(2)}-\bar{x}'''_{1(2)}),\ \bar{x}'_{2(2)},\ \bar{x}''_{2(2)},\ \bar{x}'''_{2(2)}) \ldots(4.2)$$

and $A=(a_1, a_2,\ldots, a_9)$ is a 1×9 row vector which is so chosen that (4.1) provides the unbiased estimate,

i.e., $AE'=1$

where $E=(0, 0,\ldots, 1, 1, 1)$, a 1×9 row vector.

The variance of $\hat{\bar{X}}_{2(2)}$ defined in equation (4.1) is given by

$$V(\hat{\bar{X}}_{2(2)})=AVA' \qquad \ldots(4.3)$$

where V is the variance-covariance matrix (V_{ij}) of X given in (4.2).

To minimise the variance given in (4.3) consider a function

$$\phi=AVA'-2\lambda(AE'-1)$$

where λ is the Lagrange's multiplier.

The minimum value of (4.3) can be obtained by solving the equation

$$\frac{\partial \phi}{\partial A}=0$$

or $$VA'=\lambda E' \qquad ..(4.4)$$

If V is a non-singular matrix then from equation (4.4) we have

$$A'=\lambda V^{-1}E' \qquad ..(4.5)$$

Substituting the value of A' from equation (4.5) in (4.3) we have

$$\text{Var}\,(\hat{\bar{X}}_{2(2)})=\lambda=\frac{1}{EV^{-1}E'} \qquad ..(4.6)$$

In order to obtain the value of V^{-1} we shall have to evaluate V_{ij}'s in both the cases, which can be obtained easily by the following two lemmas. The proofs of the lemmas are simple hence omitted.

LEMMA 4.1.

Let $\bar{x}_{h(k)}$ and $\bar{x}_{h'(k')}$ denote the sample means for variates $X_{h(k)}$ and $X_{h'(k')}$ based on the observations for the two variates on each of the m ssu's selected out of M ssu's of each of the n fsu's which in turn are selected out of N fsu's. If the method of selection at both the stages is one that of SRSWOR, then

(*i*) If n and m both are common for $X_{h(k)}$ and $X_{h'(k')}$

$$\text{Cov}\,(\bar{x}_{h(k)},\bar{x}_{h'(k')})=\left(\frac{1}{n}-\frac{1}{N}\right)S_{h(k)h'(k')b}+\frac{1}{n}\left(\frac{1}{m}-\frac{1}{M}\right)\bar{S}_{h(k)h'(k')w}$$

(*ii*) If n fsu's are common for $X_{h(k)}$ and $X_{h'(k')}$ but in each fsu the m ssu's are different

$$\text{Cov}\,(\bar{x}_{h(k)},\bar{x}_{h'(k')})=\left(\frac{1}{n}-\frac{1}{N}\right)S_{h(k)h'(k')b}-\frac{1}{nM}\bar{S}_{h(k)h'(k')w}$$

(*iii*) If the n fsu's for variate $X_{h(k)}$ are different from the $X_{h'(k')}$

$$\text{Cov}\,(\bar{x}_{h(k)},\bar{x}_{h'(k')})=-\frac{1}{N}S_{h(k)h'(k')b}.$$

This lemma is the direct generalisation of Tikkiwal's lemma 2.4 (1964).

LEMMA 4.2.

Let $\bar{x}_{h(k)}$ and $\bar{x}_{h'(k')}$ denote the sample means for the variates $X_{h(k)}$ and $X_{h'(k')}$ based on the observations for the two variates on each of the m_i ssu's selected out of M_i ssu's ($i=1, 2,..., N$) of each of the n fsu's which in turn are selected out of N fsu's. If the method of selection at first stage is PPSWR, size being M_i's, and at the second stage is SRSWOR, then :

(*i*) If n and m_i both are common for $X_{h(k)}$ and $X_{h'(k')}$

$$\text{Cov}\ (\bar{x}_{h(k)}, \bar{x}_{h'(k')}) = \frac{1}{n}\left[\sigma_{h(k)h'(k')b} + \sum_{i=1}^{N} \frac{M_i - m_i}{m_i M_o} S_{h(k)h'(k')i}\right]$$

(*ii*) If n fsu's are common for $X_{h(k)}$, $X_{h'(k')}$ but in each of the n fsu's the m_i ssu's are different

$$\text{Cov}\ (\bar{x}_{h(k)}, \bar{x}_{h'(k')}) = \frac{1}{n}\left[\sigma_{h(k)h'(k')b} - \frac{1}{M_o} \sum_{i=1}^{N} S_{h(k)h'(k')i}\right]$$

(*iii*) If the n fsu's are not common for $X_{h(k)}$, $X_{h'(k')}$

$$\text{Cov}\ (\bar{x}_{h(k)}, \bar{x}_{h'(k')}) = 0$$

A consistent estimate of the variance of $\hat{\bar{x}}_{2(2)}$ for both the cases can be obtained by substituting the unbiased estimates of the population values involved, which can be obtained as discussed by Sukhatme and Sukhatme [5].

REMARKS

1. The entire results of section 4 can, however, be generalised for any number of characters with suitable modification in the vectors X and A.

2. If the study is to be continued for more than two occasions the sample can be drawn in many ways, each of them gives a separate estimate for the character under study.

3. The use of SRSWOR in Case II has not been considered as the unbiased estimate of the population mean no more belongs to T_{11} class but to T_{21} class of linear estimators.

SUAMMARY

The paper gives the estimator of population mean for each of the two characters in a multipurpose sampling enquiry conducted on two occasions under a more general replacement pattern than one discussed by Singh and Singh [3]. Further, the results have been obtained without the assumption of infinite population. As these

estimators utilize all the information of preceding occasions, obviously the efficiency of the estimator is increased. The estimators considered in both the cases are MVLUE in T_{11} class of unbiased linear estimator given by Koop [2].

ACKNOWLEDGEMENT

The authors are thankful to Dr. B. D. Tikkiwal, Professor of Statistics, University of Rajasthan, Jaipur for his valuable suggestions while going through the manuscript. We also wish to thank the referee for his suggestions which lead to improvement in this paper.

REFERENCES

[1] Abraham, T.P., Khosla, R.K. and Kathuria, O.P. (1969) : Some investigations on the use of successive sampling in pest and disease surveys, *Jour. Ind. Soc. Agril. Stat.*, 21, 43-57.

[2] Koop, J.C. (1963) : On the axioms of sample formation and their bearing on the construction of linear estimators in sampling theory for finite population, *Metrika*, 7 (2 and 3), 8–114 ; 165–204.

[3] Singh, D. and Singh, R. (1973) Multipurpose surveys on successive occasions, *Jour. Ind. Soc. Agril. Stat.*, 25, 81-90.

[4] Srivastava, A.K. and Shivtar Singh (1974) : A note on two-stage successive sampling, *Jour. Ind. Soc. Agril. Stat.*, 26, 57-64.

[5] Sukhatme, P.V. and Sukhatme, B.V. (1976) : *Sampling Theory of Surveys with Applications* Indian Society of Agricultural Statistics, New Delhi.

[6] Tikkiwal, B.D. (1955) : *Multiphase Sampling on Successive Occasions*, Ph.D. Thesis, North Carolina State College, USA.

[7] ——— (1964) : A Note on Two-stage Sampling on Successive Occasions, *Sankhya*, A, 26, 97-100.

[8] ——— (1967) : Theory of Multiphase Sampling from a finite or an infinite population on successive occasions, *Rev. Inter. Stat. Instt.*, 35, 247-263.