ESTIMATION OF IMPACT DUE TO MILK SUPPLY SCHEMES IN A DYNAMIC POPULATION

By
H.P. Singh, J.P. Jain and B.C. Saxena
Indian Agricultural Statistics Research Institute, New Delhi-110 012

(Received : September, 1982)

Summary

Jain and Rajagopalan [1] suggested earlier a simple heuristic approach for estimating the change in the total of a character in dynamic populations. However, when psu's change partially over time its use may mean sacrificing part of the data. To deal with this situation a suitable methodology has been delineated. The formulae developed have been illustrated with the data collected from the rural areas covered under the Madhavaram Milk Supply Scheme, Chingleput (T.N) during 1975-76 and 1979-80.

1. Introduction

For studying the changes that have taken place due to the introduction of new developmental programme, it is necessary to conduct a bench mark survey at the commencement of the programmes and repeat the enquiry after a suitable interval of time. In order to correctly assess the changes due to the several specific measures of the developmental programme as distinct from changes due to the general development, it is necessary to carry out concurrently the study of the various response indicators in the 'Control area', which is economically similar to the development area but where the programme has not been introduced. A comparison of the changes over the years in the development area with that in the corresponding control area provides a measure of impact of the new programme. The methodology is quite simple when the size and structure of the population remains the same over time. Since these conditions are no longer tenable for dynamic populations, the available methodology needs refinement. For instance, in the study of impact of milk supply schemes on rural economy in milk collection areas, the population of clusters of villages categorised as supplying milk and those not supplying milk to the organised agency, may undergo a change in their status partially or wholly over time. Srivastava [2] has given a some what general approach to study the changes in dynamic populations. Subsequently, Jain and Rajagopalan [1] gave a simple heuristic approach for studying the changes for the specific situation, when there are no partial changes in the denomination of the clusters. But when partial changes are also there, its use would mean sacrificing a substantial part of the data. To overcome this difficulty a suitable methodology has been discussed in this paper. The methodology developed has been illustrated with the data collected under the Madhavaram Milk Supply Scheme, Chingleput (Tamil Nadu) during 1975-76 and 1979-80.

2. Sampling Design

The design employed in assessment surveys is generally one of stratified two stage random sampling. The Community Development Blocks or Talukas or groups of Talukas in each of the supplying and non-supplying areas, constitute the strata, with clusters of villages within each stratum as the first stage sampling units and households within a cluster as the second stage sampling units. The same sample units selected on the first occasion are convassed on the second occasion. The reason for retaining the same units on the second occasion is that we are mainly interested in studying the changes over time.

3. The Situation Considered

The situation considered here is the one actually realized in one of the assessment surveys carried out by Indian Agricultural Statistics Research Institute in the milk shed areas of Madhavaram Milk Supply Scheme, Chingleput (Tamil Nadu). The initial classification of the psu's (clusters of three villages each) as supplying milk to the organised agencies and those non-supplying was found to have undergone a denominational change at the time of repeat survey in the sense that some of the earlier non-supplying villages became supplying and vice-versa This resulted in complete change in the composition of some clusters and partial change in others. That is to say, the selected clusters of villages got split up into eight distinct sub-populations consisting of clusters of villages in which:

1. all the 3 villages remained supplying on both the occasions ;
2. all the 3 villages were supplying on the first occasion but one village was found non-supplying on the second occasion ;
3. all the 3 villages were supplying on the first occasion but 2 villages were non-supplying on the second occasion ;
4. all the 3 villages were supplying on the first occasion but were non-supplying on the second occasion ;
5. all the 3 villages were non-supplying on both the occasions ;
6. all the 3 villages were non-supplying on the first occasion but one village was supplying on the second occasion ;
7. all the 3 villages were non-supplying on the first occasion but two villages were supplying on the second occasion and ;
8. all the 3 villages were non-supplying on the first occasion but were supplying on the second occasion.

On the first occasion, the first four sub-populations together constitute the supplying area and the last four sub-populations the non-supplying area as conceived at the first occasion. As such, the respective four sub-populations could be treated as one population for building up the estimates of the characters for the first occasion. As regards the estimates on the second occasion, the estimates of sub-populations 1 and 8 could be combined to obtain the estimate of the total of a character for the supplying area. The corresponding estimate for the non-supplying area would be obtained likewise from sub-populations 4 and 5. Since these estimates pertain to overlapping populations on the two occasions, they will not provide valid estimate of change in the parameter between the two occasions. Moreover, this approach sacrifices data on the four sub-populations viz. 2,3,6 and 7. An appropriate procedure would be to first estimate the impact per cluster, say total of a character for each of the sub-populations separately and then pooled appropriately to obtain the impact for the entire milk shed area. The procedure of estimation is presented in the following section.

4. Estimation Procedure

Assume, for any given stratum

M_i : total number of clusters of villages in the ith sub-population (i=1,2,................8);

m_i : number of clusters sampled out of M_i,

N_{ijk} : total number of households on the jth occasion in the kth selected cluster of the ith sub-population (j=1,2).

N_{i*k} : number of common households between N_{i1k} and N_{i2k}.

n_{ijk} : number of households selected out of N_{ijk} from the ith sub-population on the jth occasion in the kth selected cluster.

n_{i*k} : number of common units between n_{i1k} and n_{i2k}.

X_{ijkl} : the value of a character recorded in the lth selected household of the kth cluster on the jth occasion in the ith sub-population.

We now define the estimate of mean per cluster and total of a character for the ith sub-population on the jth occasion as

$$\hat{\bar{X}}_{ij..}=\frac{1}{m_i}\sum_{k=1}^{m_i}\hat{X}_{ijk}.$$

where,

$$\hat{X}_{ijk.}=\frac{N_{ijk}}{n_{ijk}}\sum_{l=1}^{n_{ijk}}X_{ijkl}=N_{ijk}\ \hat{\bar{X}}_{ijk}.$$

The corresponding estimate of the total is

$$\hat{X}_{ij..}=M_i\hat{\bar{X}}_{ij..}$$

Further, let

$\bar{X}_S$ = true cluster mean for supplying clusters on the first occasion ;

$\bar{X}_N$ = the corresponding cluster mean for the non-supplying clusters ;

I_i = the impact or effect of milk supply scheme on the cluster mean of the character in the *i*th sub-populations (i=1,2..8).

I = overall impact on a character in the entire area and

E = the effect of general development in the area.

Then, ignoring sampling fluctuations, the estimated cluster means of the first four sub-populations on the first occasion viz. $\hat{\bar{X}}_{11..}$, $\hat{\bar{X}}_{21..}$, $\hat{\bar{X}}_{31..}$ and $\hat{\bar{X}}_{41..}$ are all equal to $\hat{\bar{X}}_S$ and of the last four sub-populations viz. $\hat{\bar{X}}_{51..}$, $\hat{\bar{X}}_{61..}$, $\hat{\bar{X}}_{71..}$ and $\hat{\bar{X}}_{81..}$ to $\hat{\bar{X}}_N$. The corresponding estimates for the second occasion can be represented as :

$$\hat{\bar{X}}_{i2..}=\begin{cases} \bar{X}_S+I_i+E, \text{ for } i=1,2,3,4 \\ \bar{X}_N+I_i+E, \text{ for } i=5,6,7,8 \end{cases}$$

Clearly, I_4 and I_5 are zero. Hence from sub-populations 4 and 5, two estimates of 'E' can be obtained as (*i*) $\hat{\bar{X}}_{42..}-\hat{\bar{X}}_{41..}$, and (*ii*) $\hat{\bar{X}}_{52..}-\hat{\bar{X}}_{51..}$ which are then combined to obtain a single estimate as :

$$\hat{E}=\frac{M_4(\hat{\bar{X}}_{42..}-\hat{\bar{X}}_{41..})+M_5(\hat{\bar{X}}_{52..}-\hat{\bar{X}}_{51..})}{(M_4+M_5)}$$

$$=\frac{(\hat{X}_{42..}-\hat{X}_{41..})+(\hat{X}_{52..}-\hat{X}_{51..})}{(M_4+M_5)}$$

The estimates of impact per cluster on the total of a character for each of remaining six sub-populations are obtained as

$$\hat{I}_i=(\hat{\bar{X}}_{i2..}-\hat{\bar{X}}_{i1..})-\hat{E},\ i=1,\ 2,\ 3,\ 6,\ 7,\ 8$$

then the impact (I) on the total of a character for the entire area is estimated as :

$$\hat{I}=M_1\hat{I}_1+M_2\hat{I}_2+M_3\hat{I}_3+M_6\hat{I}_6+M_7\hat{I}_7+M_8\hat{I}_8$$

$$=\sum_i(\hat{X}_{i2..}-\hat{X}_{i1..})-\frac{(\sum_i M_i)}{(M_4+M_5)}[(\hat{X}_{42..}-\hat{X}_{41..})+$$

$$+(\hat{X}_{52..}-\hat{X}_{51..})]$$

$(i=1,\ 2,\ 3,\ 6,\ 7,\ 8)$

The estimate of variance of $\hat{I}$ is given by

$$\hat{V}(\hat{I})=\sum_{i}[\hat{V}(\hat{X}_{i2..})+\hat{V}(\hat{X}_{i1..})-2\ \hat{\text{Cov}}(\hat{X}_{i2..},\hat{X}_{i1..})]$$

$$+\frac{(\sum M_i)^2}{(M_4+M_5)^2}[\hat{V}(\hat{X}_{42..})+\hat{V}(\hat{X}_{41..})+\hat{V}(\hat{X}_{52..})+\hat{V}(\hat{X}_{51..})$$

$$-2\ \hat{\text{Cov}}\ (\hat{X}_{42..}\ ;\ \hat{X}_{41..})-2\ \hat{\text{Cov}}\ (\hat{X}_{52..},\hat{X}_{51..})]$$

$$(i=1,\ 2,\ 3,\ 6,\ 7,\ 8)$$

Where for any $i=1,\ 2\ldots 8$ and $j=1,2$

$$\hat{V}(\hat{X}_{ij..})=M_i^2\left[\left(\frac{1}{m_i}-\frac{1}{M_i}\right)s_{ij}^2\right.$$

$$\left.+\frac{1}{m_iM_i}\sum_{k=1}^{m_i}N_i\left(\frac{1}{n_{i*k}}-\frac{1}{N_{i*k}}\right)s_{ijk}^2\right]$$

$$\hat{\text{Cov}}(\hat{X}_{i2..},\hat{X}_{i1..})=M_i^2\left[\left(\frac{1}{m_i}-\frac{1}{M_i}\right)s_{i(b)12}^2\right.$$

$$\left.+\frac{1}{m_iM_i}\sum_{k=1}^{m_i}N_{i*k}^2\left(\frac{1}{n_{i*k}}-\frac{1}{N_{i*k}}\right)s_{i(w)12}\right]$$

where $$s_{ij}^2=\frac{1}{(m_i-1)}\sum_{k=1}^{m_i}(\hat{X}_{ijk.}-\hat{\bar{X}}_{ij..})^2$$

$$s_{ijk}^2=\frac{1}{(n_{i*k}-1)}\sum_{l=1}^{n_{ijk}}(X_{ijkl}-\hat{\bar{X}}_{ijk.})^2$$

$$s_{i(w)12}=\frac{1}{(n_{i*k}-1)}\sum_{l=1}^{n_{i*k}}(X_{i1kl}-\hat{\bar{X}}_{i1k.})(\hat{X}_{i2kl}-\hat{\bar{X}}_{i2k.})$$

and

$$s_{i(b)12}=\frac{1}{(m_i-1)}\sum_{k=1}^{m_i}(\hat{X}_{i1k.}-\hat{\bar{X}}_{i1..})(\hat{X}_{i2k.}-\hat{\bar{X}}_{i2..})$$

In the derivation of the expression for the estimate of the variance, $\hat{I}$ it has been assumed that M_i's are known. In fact these have to be estimated as follows :

$$M_i=\begin{cases}\dfrac{m_i}{m_S}M_S \text{ for } i=1,\ 2,\ 3,\ 4\\[2ex] \dfrac{m_i}{m_N}M_N \text{ for } i=5,\ 6,\ 7,\ 8\end{cases}$$

where M_S and M_N are respectively the number of clusters of supplying and non-supplying villages in the population on the first occasion and m_S and m_N are the number of clusters sampled out of M_S and M_N respectively. Treating the estimated sub-population sizes as true values will not materially affect the variance estimates because of large sample size.

5. Illustration

The foregoing methodology was used to obtain the impact of Madhavaram Milk Supply Scheme, Chingleput (T.N.) in respect of total cow milk production per day in commercial milk producer households. The outline of the bench-mark and repeat surveys carried out for the purpose are as indicated in Sections 2 and 3.

Table 1 gives the estimates of daily cow milk production in commercial households on the two occasions and as also the estimate of impact of the milk supply scheme. The distribution of clusters of villages sampled and in the population is also shown in the table for each sub-population

It is to be noted that in the present context, sub-populations 1, 2, 3, 6, 7 & 8 together constitute the milk shed area. If however, the partial changes are ignored, sub-population-1 alone will constitute milk shed area.

After making allowance of the effect of general development in the area, it will be seen that the production increased from the first to the second occasion except in sub-population 3 and 8. This would perhaps be ascribed to small sample size particularly from sub-population 8. The overall impact of the scheme in terms of increased milk production per day is of the order of 220 kg., with standard error of about 3%. This works out to about 9% increase in production over that of the first occasion. With the earlier approach (Jain and Rajagopalan, 1978), the estimated increase in milk production in the milk shed area comprising sub-population-1 works out to be about 17%. Thus the earlier approach gives an over estimate of the impact as expected because with the inclusion of sub-populations with partial changes the impact will be sufficiently lowered down.

TABLE 1

Estimates of cow milk production per day in commercial households on the two occasions along with the estimate of impact of the scheme in the area

Sub-Population	No. of clusters in the population (M_i)	No. of clusters in the sample (m_i)	Occasion	Average milk production per cluster. $\hat{\bar{X}}_i$ (kg.)	Sub-population total milk production $M_i \hat{\bar{X}}_{ij}$ (tonne)	Estimate of impact for cluster, $\hat{I}_i$ (kg.)
1	39	7	BS	18.97	0.74	2.62
			RS	23.59	0.92	
2	13	5	BS	12.46	0.16	4.69
			RS	19.15	0.25	
3	21	11	BS	17.00	0.36	0.86
			RS	18.14	0.38	
4	23	8	BS	32.17	0.74	0
			RS	40.00	0.92	
5	403	15	BS	21.81	8.79	0
			RS	23.83	9.46	
6	34	2	BS	20.00	0.68	5.65
			RS	27.65	0.94	
7	12	2	BS	5.00	0.06	3.83
			RS	10.83	0.13	
8	16	2	BS	18.12	0.29	—4.62
			RS	15.00	0.24	

Effect of general development ($\hat{E}$)=2.00 kg.
Impact of the scheme ($\hat{I}$) : 217.26±6.64 kg.

REFERENCES

[1] Jain, J.P. and Rajagopalan, M. (1978) : An approach for estimating changes in dynamic populations. *Jour. Ind. Soc. Agri. Stat.* 30, 81-90.

[2] Srivastava, A.K. (1977) : Sampling from dynamic populations. Thesis submitted for the degree of Doctor of Pholosophy, University of Delhi, Delhi.