STRATIFICATION AND ESTIMATION PROCEDURE IN CULTIVATED FODDER SURVEYS

By

ANAND PRAKASH AND S.D. BOKIL

I.A.S.R.I., New Delhi-110 012.

(Received : October, 1978)

SUMMARY

Different methods of construction of strata on the basis of auxiliary character have been studied to determine number of strata and strata boundaries. Efficiency of different methods of estimation of area under fodders as compared to the method of simple random sampling has also been studied. The method 'Equal interval of cumulative $\sqrt[3]{f}$, in general, gave less variance as compared with other methods. The regression estimate provided an estimate of area under cultivated fodders with the smallest variance.

The development of livestock industry depends to a considerable extent on the production of nutritious feeds. At present, a major portion of these are obtained as by-products in the production of various crops. In the progressive areas, however, appreciable areas are put under fodder crops. Unless efforts are made to ensure increasing production of nutritious fodders, substantial increase in the production of various livestock products, particularly milk, may not be possible. This will involve a regular annual review of production of fodder in various parts of the country on the lines being done for food-grains at present. The estimation of production in case of fodder crops is complicated by the fact that the fodder is sometimes harvested in instalments and in case of some fodder crops, repeated cuttings are taken. An essential step therefore is to develop the necessary crop cutting survey technique for the fodder crops. The Indian Agricultural Statistics Research Institute had made in this direction a beginning in the year 1972-73 by taking up a pilot sample survey in Meerut district to evolve a suitable sampling technique for estimation of yield of cultivated fodder crops. A similar pilot survey has been subsequently undertaken in Karnal

district of Haryana. In these surveys as under foodgrain surveys, stratification had been done on the basis of geographical contiguity under the assumption that units near one another may not show wide variation. When the primary objective is not to obtain estimates of production parameters for each geographical region there is scope for looking for better methods of stratification. One possible alternative is to construct strata on the basis of a variable highly correlated with the one under study, in such a way as to give maximum precision for the character under study. In this paper an attempt is made to examine critically the various aspects involved in stratification for estimation of area under cultivated fodders. In brief, the problems studied are :

1. Alternative methods of construction of strata on the basis of auxiliary character. This study will include determining number of strata and strata boundaries.

2. Efficiency of different methods of estimation as compared to the method of simple random sampling.

The various methods of construction of strata that will be examined in the present investigation are as follows :—

(*i*) Equal aggregate output method due to Mahalanobis (1952).

(*ii*) Equal interval of cumulative $\sqrt{f}$ proposed by Dalenius and Hodges (1957) where f is the frequency function,

(*iii*) Equal interval on cumulative $\sqrt[3]{f}$ as proposed by Singh (1967).

(*iv*) $W_h=(X_h-X_{h-1})$=constant where W_h is the total frequency in stratum h and $X_h e$ & X_{h-1} are the upper and lower boundary points of stratum h.

(*v*) Durbin's method : Equal interval on the cumulative of

$$[r(x)+f(x)]\times h$$

where r is the frequency function for a rectangular distribution.

For the construction of strata by these methods, the auxiliary variate (x) chosen was the area under cultivated fodders in the year 1970-71. The coefficient of correlation between X_a and the area under fodder crops in 1972-73 (Y) was of the order of 0.89. Using the information of area under cultivated fodders in 1970-71 for Hapur tehsil in Meerut district of Uttar Pradesh, the optimum points of stratification were determined using the above five methods for number of strata varying from 2 to 4. For determining the strata boundaries, the points of stratification will be approximated

to the value of the class interval nearest to it and as these boundaries are changing by steps of 25, the estimated boundaries are multiples of the number.

Construction of Strata

(*a*) *Equal aggregate output method* :

Using this method, the strata boundaries have been worked out from the data and are shown in table No. 1.

TABLE 1

Strata boundaries for varying number of Strata

No. of Strata	*Stratum boundaries*			
	1	2	3	4
$L=2$	0—200	>200	—	—
$L=3$	0—150	151—275	>275	—
$L=4$	0—125	126—200	201—300	>300

(*b*) *Equal interval on cumulative* $\sqrt{f}$ *method.*

The strata boundaries obtained by this procedure are indicated in table No. 2.

TABLE 2

Strata boundaries using $\sqrt[3]{f}$ **for varying number of strata**

No. of strata	*Stratum boundaries*			
	1	2	3	4
$L=2$	0—175	>175	—	—
$L=3$	0—125	126—250	>250	—
$L=4$	0—100	100—175	176—300	>300

(*c*) *Equal interval on cumulative* $\sqrt[3]{f}$

The strata boundaries for $L=2$ $L=3$ and $L=4$ are determined and presented in table No. 3.

TABLE 3

Strata boundaries using $\sqrt[3]{f}$ method for varying number of strata

No of strata	*Stratum boundaries*			
	1	2	3	4
$L=2$	0 - 200	>200	—	—
$L=3$	0 – 125	126 – 300	>300	—
$L=4$	0 – 100	101 – 200	201 – 350	>350

(*d*) *Ekman's Rule*

In this method, the product of the cumulated frequency within the stratum and width of the stratum are equalised. This method is somewhat complicated and difficult to apply because the value of $\Sigma f(X_h - X_{h-1})$ is not constant but depends on both L and the position of the boundaries. Hence for given L, it is not obvious as to what figure should be tried to equate $\Sigma W_h(X_h - X_{h-1})$ in the individual strata. A rough guide suggested by Ekman (1959) is to compute the product $Q = WX$ range x. The constant value per stratum is approximately Q/L^2 when the number of strata is L. This relation is exact for a rectangular distribution but tends to give high results in case of highly skewed distributions. The strata boundaries obtained by this procedure are given in table No. 4.

TABLE 4

Strata boundaries for varying number of strata

No. of strata	*Stratum boundaries*			
	1	2	3	4
$L=2$	4 – 175	>175	—	—
$L=3$	0 – 125	126 – 300	>300	—
$L=4$	0 – 100	101 – 200	201 – 375	>375

(*e*) *Durbin's Rule*

Durbin proposed a method in review of Dalenius' Doctoral thesis. The rule suggested amounts to making the strata by taking equal areas under a frequency distribution with density half-way between the original distribution and a rectangular distribution.

Table No. 5 provides the strata boundaries for varying number of strata.

TABLE 5

Strata boundaries for varying number of strata

No. of strata	*Stratum boundaries*			
	1	2	3	4
$L=2$	0—200	>200	—	—
$L=3$	0—125	126—275	>275	—
$L=4$	0—100	101—200	201—350	>350

Comparison of variances of estimates under different methods of stratification.

Table No. 6 givess the varianee of Total area under fodder crops for various sample sizes distributed in proportion to the total area under fodder crops in each stratum and for varying number of strata.

It may be seen from the above table that as the number of strata increases from 2 to 4 the variances in case of all methods considered decreases as is to be expected. Also there is conspicuous decreasing tendency of variance as the sample size is increased from 15 to 25.

Examining the different methods of construction of strata it will be noted that $\sqrt[3]{f}$ method of construction has the smallest variance as compared with all the other methods considered here and as such it may be concluded that $\sqrt[3]{f}$ method of stratifying the population in respects of cultivated fodder crops is most suitable for the type of data utilized in this study, whatever be the number of strata to be chosen.

Comparison of variances of estimates under different methods of estimation

With the above findings. an attempt is made to estimate the area under cultivated fodders during 1972-73. From the list of 322 villages in Hapur tehsil a sample of 25 villages was selected and allocated under proportional allocation to 4 size strata constructed using $\sqrt[3]{f}$ method on the basis of the area figures relating to the

TABLE 6

Variance of total area under cultivated fodder crops for varying sample size according to different methods of stratification

Variance in $(Acre)^2$

Sample size	*No. of Strate*	*Method of Stratification*				
		Equal aggregate output method	*Equal interval on cumulative* $\sqrt{f}$	*Equal interval on cumulative* $\sqrt[3]{f}$	*Ekman's rules*	*Durbin's method*
15	2	213.90×10^5	112.25×10^5	87.90×10^5	117.25×10^5	139.10×10^5
	3	78.21×10^5	62.85×10^5	60.27×14^5	70.52×10^5	70.09×10^5
	4	33.16×10^5	29.95×10^5	21.72×10^5	46.21×10^5	39.46×10^5
20	2	120.04×10^5	118.21×10^5	82.27×10^5	115.15×10^5	150.21×10^5
	3	58.29×10^5	70.45×10^5	52.54×10^5	56.63×10^5	67.71×10^5
	4	32.15×10^5	34.17×10^5	13.74×10^5	24.95×10^5	30.70×10^5
25	2	57.04×10^5	58.30×10^5	41.15×10^5	65.31×10^5	82.55×10^5
	3	46.22×10^5	35.27×10^5	25.23×10^5	43.41×10^5	60.01×10^5
	4	31.13×10^5	25.60×10^5	8.65×10^5	18.81×10^5	20.85×10^5

year 1970-71, the method of selection of villages was S.R.S. and independent in each stratum. The estimates built up using different methods of estimation are given in table No. 7.

TABLE 7

Relative efficiency of different methods of estimation

Method of estimation	*Variance (Acre)*^{2}	*Estimated area (in acres)*
Simple random sampling	224×10^5	47,327
Ratio method of estimation	178×10^5	46,856
Regression method of estimation	154×10^5	45,235

Examining the results of different methods of estimation it is seen that the estimate of variance is least when regression method of estimation is used. Hence $\sqrt[3]{f}$ method together with regression method of estimation could be used with advantage.

Further, the bias involved in the ratio and regression method of estimation has also been worked out and is found to be of small order. Consequently the efficiency of either of the above methods remains unchanged.

Discussion and Conclusions

Before a sample survey is initiated the first step is to specify the number of strata into which the population units could be grouped and the strata boundaries. The problem of optimum stratification, was first discussed by Dalenius (1950) who gave sets of equations giving optimum strata boundaries for optimum and proportional methods of allocating the sample to different strata. Subsequently various other research workers gave methods of optimum stratification of population.

In the present work we have tried to consider the problem of optimum stratification on the basis of the area under cultivated fodder crops during 1970-71 in Hapur tehsil of Meerut district. The different methods for demarcation of stratification points such as :—

(*i*) Equal aggregate output method,
(*ii*) Equal interval on cumulative $\sqrt{f}$,
(*iii*) Equal interval on cumulative $\sqrt[3]{f}$,
(*iv*) Ekman's rule,
(*v*) Durbin's rule,

were examined and the efficiencies for estimating area under fodders were worked out. The results are indicative of the fact that $\sqrt[3]{f}$ method, in general, gave less variance as compared with the other methods.

In this work we have also compared stratified simple random sampling estimate, ratio estimate and regression estimate using $\sqrt[3]{f}$ rule of construction of strata which gave smallest variance as compared with other methods of stratification. It is found that regression estimate provides an estimate of area under cultivated fodders with smallest variance.

ACKNOWLEDGEMENT

The authors are grateful to the referees for their valuable suggestions for improving the paper.

REFERENCES

Dalenius, T. (1950) : The problem of optimum stratification skandinavik, Vol. 33, pp. 203-213.

Dalenius, T. and Hodges, J.L. (1957) : The choice of stratification points, *Standinavik*, Vol. 40, pp. 198-203.

Ekman, G. (1959) : An approximation useful in univariate stratification. *The Annals of Mathematical Statistics*, Vol. 30, pp. 219-229.

Singh, R. (1967) : Some contribution to the theory of construction of strata. Unpublished thesis for the award of Ph. D. degree.

Mahalanobis, P.C, (1952) : Some aspects of the designs of sample surveys, *Sankhya*, pp. 1-7.