Efficient Sampling Design in High Density Apple Orchards

Tabasum Mushtaq[1], S.A Mir[2], Nageena Nazir[3]

[1,2,3]Division of Agricultural Statistics, Sher-e-Kashmir University of Agricultural Sciences and Technology, Shalimar Campus, Srinagar, Jammu & Kashmir, (India)

ABSTRACT

In horticultural surveys determination of appropriate sampling design is of foremost importance. Inappropriate, Inadequate sampling strategy can lead to wrong results and can influence the quality of estimates of population parameters. The paper describes the appropriate method of sampling strategy choosing to gain maximum efficiency of estimates in survey of high density apple orchards. The data used for this research was primary data collected over two years from high density apple block of SKUAST-Kashmir (HDP, plate-1). The study was carried on the basis of trunk cross sectional area of high density apple trees of two varieties viz. Gala Red Lum and Fuji Zehn Aztec. Various sampling schemes were employed for obtaining the sample to estimate the population parameters. It is observed that stratified sampling scheme with proportional allocation procedure gives most efficient estimates for estimating population parameters of both the HDA varieties under study.

Keywords*: *Efficiency, Proportional allocation, Simple Random Sampling, Stratified Random Sampling,

I.INTRODUCTION

In any scientifically based decision making, the sampling design is a fundamental part of data collection. In ensuring that the data are sufficient to draw the conclusions needed, a well- developed sampling design plays a critical role. Accurate information is important for sound science-based decision. Developing a sampling design is a crucial step in collecting appropriate and defensible data that accurately represent the problem being investigated. Sampling plays an important role in research design involving human population and commands increasing attention from the social scientists, chemists, engineers, accounts, biologists and medical practitioners engaged in marketing, commerce, industry, public health, biostatistics, education, public administration, sociology, accountants, economics, anthropology, political science, and even social workers [1].

Sampling designs are given as a measure to an end originating in essential research problems especially in the social sciences and their applications [2]. For making inference from sample value to the population value the sampling design covers the tasks of selection and estimation. There are the problems of making inferences from one survey population to another and generally broader population, with measurements free from error. Various sampling designs would result in unlike standard errors, and selecting the design with the least error is the principle aim of sampling design.

Within a population an effective sampling technique represents an appropriate extraction of useful data which provides meaningful knowledge of the important aspects of the population [3]. Probability samples are usually designed to be measurable. They are designed in such a way that statistical inference to population values can be based on measures of variability, usually standard errors, computed from the sample data. So there is demand to devise a sampling strategy which is economical and easy to function, that yields unbiased estimates, and minimizes the effects of variation due to sampling.

Normally in sample surveys more than one population parameters are estimated and these parameters may be of contradictory nature. Stratified sampling has been designed to ascertain that all important views are represented in samples. Stratification is a method of sampling design by which the entire heterogeneous population of interest is divided in to a number of homogeneous groups, usually known as strata, each of these groups is homogeneous within itself, and then units are sampled at random from each of these stratums. The sample size in each stratum varies according to the relative importance of the stratum in the population. The technique of the drawing this stratified sample is known as Stratified Sampling. In other words, stratification is the technique by which the population is divided into subgroup/strata. Sampling will then be conducted separately in each stratum. Strata or Subgroup are chosen because evidence is available that they are related to outcome. The selection of strata will vary by area and local conditions. After stratification, sampling is conducted separately in each stratum. In stratified sample, the sampling error depends on the population variance within stratum but not between the strata. Stratified random sampling also defined as where the population embraces a number of distinct categories, the frame can be organized by these categories into separate "strata." Each stratum is then sampled as an independent sub-population, out of which individual elements can be randomly selected. Stratified sample designs are put to use for various reasons. These include: to increase the precision of estimates for the whole population for one or more key data items being collected in the survey, to obtain more precise estimates for interesting domains, to allow the use of different sampling, non-response adjustment, editing, or estimation methods for domains with differing characteristics affecting the choice of method, and to facilitate administration of the survey.

The purpose of the study is to find the efficient sampling design in order to gain the maximum efficiency in surveys of high density apple orchards.

II.DATA USED

The data used for this research was primary data on various tree/ fruit characteristics collected over two years from high density apple block of SKUAST-Kashmir (HDP, plate-1). The study was carried on the basis of trunk cross sectional area of high density apple trees of two varieties viz. Gala Red Lum and Fuji Zehn Aztec.

III.METHODOLOGY

In this study, for selection of samples, both the procedures simple random sampling (wor) and stratified random sampling techniques have been employed. Trunk cross sectional area (TCA) was stratified into three strata in each variety of high density apple plants. The stratification was done following the principles that –

IJARSE
ISSN: 2319-8354

i) The strata (i.e. range of TCA) are non-overlapping and together comprise the whole population.

ii) The strata (i.e. range of TCA) are homogeneous within themselves with respect to the characteristics under study.

The allocation of the samples to the different ranges of TCA was carried out through proportional allocation method. In this method, the sampling fraction, n/N is same in all strata. This allocation was used to obtain a sample that can estimate size of the sample with greater speed and a higher degree of precision. The allocation of a given sample of size n to different stratum was done in proportion to their sizes i.e. in the i^{th} stratum,

$$i \qquad (1)$$

where n represents sample size, N_i represents population size of the i^{th} strata and N represents the population size. In our study, N = 270; n = 41

The variance of the estimate of population mean for stratified random sampling with proportional allocation and simple random sampling (wor) is given by

$$Var\,(\bar{y}_{st})_P = \left(\frac{1}{n} - \frac{1}{N}\right)\Sigma \qquad (2)$$

$$Var\,(\bar{y}_n)_R = (\frac{1}{n} \qquad (3)$$

where $$S^2 = \frac{1}{N-1}\sum_{i=1}^{k}\sum_{j=1}^{N_i}(Y_i$$

$$S^2 = \sum_{i=1}^{k}(N_i - 1)S_i^2 + \sum_{i=1}^{k} N_i(\bar{Y}_N \qquad (4)$$

IV. SUMMARY OF RESULTS

In the selection of sample in each stratum, for proportional allocation, sample size varies depending upon the total number of plants in each range of trunk cross sectional area for both the varieties. The results obtained are shown in Table 1, Table 2 and Table 3.

Table 1: Empirical study on Stratification of Trunk Cross Sectional Area for Gala Red Lum

Strata	N_i	n_i	$\bar{Y}_{N_i}$	$\bar{Y}_{N_i}^2$	$N_i\bar{Y}_{N_i}$	$N_i\bar{Y}_{N_i}^2$	S_i	S_i^2	N_iS_i	$N_iS_i^2$	$(N_i-1)S_i^2$
1	77	12	6.02	36.24	463.54	2790.51	0.85	0.72	65.45	55.63	54.91
2	150	23	8.87	78.68	1330.50	11801.54	1.31	1.72	196.50	257.42	255.70
3	43	06	12.79	163.58	549.97	7034.12	1.47	2.16	63.21	92.92	90.76
Total	270	41	27.68	278.50	2344.01	21626.16	3.63	4.60	325.16	405.97	401.37

Table 2: Empirical study on Stratification of Trunk Cross Sectional Area for Fuji Zehn Aztec

Strata	N_i	n_i	$\bar{Y}_{N_i}$	$\bar{Y}_{N_i}^2$	$N_i\bar{Y}_{N_i}$	$N_i\bar{Y}_{N_i}^2$	S_i	S_i^2	N_iS_i	$N_iS_i^2$	$(N_i-1)S_i^2$
1	15	08	3.19	10.18	47.85	152.64	0.99	0.98	14.85	14.70	13.72
2	186	25	6.42	41.22	1194.12	7666.25	1.80	3.24	334.80	602.64	599.40
3	69	08	10.75	115.56	741.75	7973.81	2.26	5.11	155.94	352.42	347.32
Total	270	41	20.36	166.96	1983.72	15792.70	5.05	9.33	505.59	969.77	960.44

IJARSE
ISSN: 2319-8354

Table 3: Variance Based on Simple Random Sampling (wor) and Stratified Random sampling

HDP Variety	$V(\bar{y}_n)_R$	$V(\bar{y}_{st})_P$
Gala Red Lum	0.13	0.03
Fuji Zehn Aztec	0.16	0.07

V.GAIN IN EFFICIENCY (GE) IN STRATIFIED RANDOM SAMPLING OVER SIMPLE RANDOM SAMPLING WITHOUT REPLACEMENT

In order to observe how the sample size gets affected due to different types of allocation, an analysis on gain in efficiency (GE) due to different types of allocations is utmost required. Hence gain in efficiency due to stratified random sampling (proportional allocation procedure) over simple random sampling *(wor)* is given by

$$(E-1) = \frac{V(\bar{y}_n)_R}{V'} \qquad (5)$$

Table 4: Gain in Efficiency Due To Proportional Allocation

Variety	Gala Red Lum	Fuji Zehn Aztec
GE	3.3333	1.2548
% GE	333.33	125.48

VI.DISCUSSION OF RESULTS

In the selection of sample in each stratum for proportional allocation, sample size varies depending on the total number of plants in each range of trunk cross sectional area. The variance for Gala Red Lum and Fuji Zehn Aztec under simple random sampling is 0.13 and 0.16 respectively and stratified random sampling with proportional allocation is 0.03 and 0.07 respectively. It is observed that the gain in efficiency (GE) due to stratification with proportional allocation of sample size to each range of trunk cross sectional area for Gala Red Lum and Fuji Zehn Aztec is 3.3333 and 1.2548 respectively.

VII. CONCLUSION

Finally, after employing various sampling schemes for obtaining the sample to estimate the population parameters it is concluded that stratified sampling scheme with proportional allocation procedure gives most efficient estimates for estimating population parameters of both the high density apple varieties under study.

REFERENCES

[1] L. Kish, *Survey sampling* (New York, Wiley, 1965).

[2] N.Hunt and S.Tyrell, *Stratified sampling* (Coventry University Press, 2004) http://www.coventry.ac.uk/ec/-nhunt/meths/strati.html

[3] J.A. Diaz-Garcia, and L.U. Cortez, Multi-objective optimization for optimum allocation in multivariate stratified sampling. *Survey Methodology*, *34(2)*, 2008, 215-222.