

EASTERN UNIVERSITY, SRI LANKA

THIRD EXAMINATION IN SCIENCE -(2002/2003)

(JUNE/JULY, 2003)

Repeat

FIRST SEMESTER

ST 302 - SAMPLING THEORY

Answer all questions

Time: Three hours

1. In sample survey to study the yield of mango trees, a simple random sample of 10 of 150 villages in a district was selected and the number of mango trees, Y , and the area under mango trees, X , were recorded for each.

Village(i)	1	2	3	4	5	6	7	8	9	10
Total number of mango trees(y_i)	50	99	81	130	190	66	40	70	88	80
Area(in hectares) under mango trees(x_i)	1.2	1.6	1.2	2.4	3.4	1.8	0.7	1.3	1.6	2.0

The total area under mango trees in this district is 88 hectares. Estimate the total number of mango trees in the district using

- (a) The simple random sample mean.
(b) The ratio estimator.

(c) The regression estimator.

Compute the relative efficiencies of your estimators (a) and (b) relative to (c).

2. (a) Prove that the sample mean and sample variance are unbiased estimators of the population mean and population variance respectively in Simple Random Sampling without replacement (*SRSWOR*).

(b) In a sample of 50 households drawn by *SRSWOR* from a village of 250 households, only 8 households were found to possess a bicycle. These houses had 3, 5, 3, 4, 7, 4, 4, and 5 members respectively. Estimate unbiasedly the total number of households in the village possessing a bicycle as well as the total number of persons in such households. Also estimate the variances of these estimates by using the unbiased estimators of their variances.

3. Two dentists *A* and *B* make a survey of the state of teeth of 200 children in a village. Dr. *A* selects a simple random sample of 24 children and counts the number of decayed teeth for each child, with the following results.

Number of decayed Teeth/child	0	1	2	3	4	5	6	7	8	9	10
Number of children	9	3	4	2	2	1	1	0	0	1	1

Dr. *B*, using the same dental techniques, examines all 200 children, recording merely those who have no decayed teeth. He finds 60 chil-

dren with no decayed teeth.

Estimate the total number of decayed teeth in the village children.

- (a) Using A 's results only.
 - (b) Using both A 's and B 's results.
 - (c) Are the estimates unbiased?
 - (d) which estimate do you expect to be more precise?
4. (a) Define an estimator for population mean in a Stratified Random Sampling Scheme and derive its variance. (You may use the results that you have obtained in part (a) of question (2)).
- (b) Distinguish between Proportional and Optimal allocation in stratified random sampling. Deduce the expression of the variance in proportional allocation using the results you obtained in part (a).
- (c) In a survey on the area under a crop, a total of 186 villages in a district were divided into 4 strata according to the area of the villages. From each stratum, simple random samples were selected by proportional method and the area under crop were noted. The following is the data obtained from the survey.

Stratum	Size(N_i)	Sample size(n_i)	Area under the crop
1	72	8	14,12,8,11,12,10,13,16
2	53	5	27,20,21,22,30
3	35	4	36,47,52,61
4	26	3	92,105,82

Obtain an estimate of the total area under the crop in the district and estimate the variance of this estimate.

5. An investigator proposes to take a stratified random sample with two strata. He estimates that the relevant quantities for the two strata are as follows.

Stratum	W_h	S_h	c_h (Rs)
1	0.4	10	4
2	0.6	20	9

where h = Stratum,

W_h = $\frac{N_h}{N}$ the stratum weight,

S_h^2 = Population variance of Y, the item of interest,

c_h = cost of sampling one unit from stratum h,

n_h = number of units randomly selected from stratum h.

He assumes that his total field cost will be of the form $c_1n_1 + c_2n_2$.

Finds the values of $\frac{n_1}{n}$ and $\frac{n_2}{n}$ that minimize the total field cost for a given value of $Var(\bar{y}_{st})$, where n is the total number of units selected and \bar{y}_{st} is the usual estimate of the population mean from a stratified sample. Find the total number of units n , required, under this optimum allocation, to make $Var(\bar{y}_{st}) = 1$. Assume that the finite population correction is negligible. How much will the total field cost be?

6. (a) State the advantages of adopting a "Ratio estimation" in simple random sampling. What are the assumptions required to perform ratio estimation for estimating population mean in *SRS*. Explain why a "Linear Regression Estimator" is proposed instead of "Ratio estimator" under certain circumstances.

(b) A sample survey was carried out to study the total area of corn planted in a region. Altogether ten farms were drawn by *SRSWOR* method from the total of 100 farms available in the region and the details are as follows. The average size of the farms in the region was found to be 200 acres.

Sample farm	Total area (Acres)	Areas in Corn (Acres)
1	100	20
2	140	60
3	160	80
4	180	90
5	200	100
6	110	60
7	190	90
8	220	110
9	240	100
10	230	120

i. Estimate the "Average area of corn per farm" and the "Total area of corn in the region". Calculate also the vari-

ances of these estimates.

- ii. Examine whether the above estimation is satisfactory or not. If not, obtain an improved estimator by ratio estimation. Calculate its variance and examine whether this estimation is more efficient or not.
- iii. Verify that the above ratio estimator satisfies the assumptions that you have described in part (a). State whether a "*Regression estimator*" is necessary for further improvement in the estimation.