| RST-003 |

Ph.D. IN STATISTICS
(PHDSTAT)

Term-End Examination

□□□□55   **June, 2018**

RST-003 : ADVANCED SAMPLE SURVEYS

Time : 3 hours   *Maximum Marks : 100*

Note : '*Question no. 1 is compulsory. Attempt any four questions from questions no. 2 to 7. Only non-programmable scientific calculator is allowed. Symbols have their usual meanings.*

1. (a) State whether the following statements are *true* or *false*. Give reasons in support of your answers.   4×2=8

    (i) If a population contains 4 items and 2 items are selected using SRSWR, then all possible samples will be 16.

    (ii) If a surveyor misunderstood the directions and included children up to 10 years of age instead of only children 5 – 54 months of age, this is a case of sampling error.

    (iii) If the non-circular intra-class correlation coefficients are all negative, the estimator based on stratified sampling will be superior to the estimator based on systematic sampling.

(iv) The ratio estimator of population mean will be more efficient than the sample mean ($\bar{y}$) using SRSWOR if $\rho > \dfrac{1}{2}\dfrac{C_y}{C_x}$.

(b) Differentiate between the following and give examples to illustrate the differences : $4 \times 3 = 12$

(i) Ordered and Un-ordered estimators

(ii) Response and Non-response errors

(iii) Stratified and Cluster sampling

(iv) Ratio and Regression method of estimation

2. (a) Explain the procedure of post-stratification. Also determine the expected value and variance of the estimator of the population mean under this scheme.

(b) 450 farms of wheat in a region were stratified according to farm size (in acres) into 2 strata. The population values of strata means ($\bar{y}_i$) and standard deviation (σ_i) for the area under wheat are given as follows :

Strata No.	Farm Size (in acres)	No. of Farms	Strata Means	Standard Deviation
1	0 – 100	300	45	15
2	100 – 200	150	90	60

For a sample of size 45, determine the sample size from each stratum using (i) Proportional allocation; and (ii) Neyman allocation. $12+8=20$

3. (a) A sample of 10 villages is selected from a population of 100 villages in a tehsil, using SRSWOR. The study variable Y is the number of cows and the corresponding numbers in census data (X) in the previous year are as given below :

Villages	Y	X
1	25	23
2	15	14
3	22	20
4	24	25
5	13	12
6	18	18
7	35	30
8	30	27
9	10	8
10	29	31

Given that the total number of cows in the tehsil based on the census, is 19,700, estimate the number of cows in the current year using the ratio method of estimation.

Also compare the efficiency of this estimator with the estimator based on sample mean without using auxiliary information.

(b) Describe the regression method of estimation under double sampling scheme along with an example. *15+5=20*

4. For studying the yield of sheep wool, a two-stage sampling design, with the tehsil as the first-stage units and villages in the tehsil as

second-stage units, was adopted. The data given below are the sheep population in the selected villages in each of the 4 tehsils selected from 12 tehsils.

Tehsil No.	Number of Villages in Tehsil (M_i)	Sheep Population
1	20	26, 31, 17, 18, 10, 12, 22
2	18	12, 11, 16, 27, 20, 18
3	15	22, 27, 23, 26, 20
4	10	28, 26, 22, 18

Estimate the average sheep population, along with its standard error, when $\overline{M} = 16$, using any two estimators under this scheme. *20*

5. (a) Describe the estimators of population proportion under SRSWR and SRSWOR. Also obtain expressions for their variances.

(b) A list of 3000 voters of a ward in a city was examined for measuring the accuracy of age of individuals. A random sample of 300 names was taken, which revealed that 51 citizens were shown with wrong ages. Estimate the total number of voters having a wrong description of age in the list, and estimate the standard error assuming that sample was taken using SRSWOR. *12+8*

6. Derive the expected value and MSE of the estimator of the population mean using the regression method of estimation, when the regression coefficient is computed from the sample. *20*

7. (a) Describe the cumulative total method for selecting probability proportional to size sample.

   (b) The result of a sample survey on the number of lime trees and the area reported under lime, in each of the 15 villages growing lime in one of the tehsils of a state, are given below :

S. No. of Village	Area under Lime (in acres)	No. of Lime Trees
1	32	1328
2	8	618
3	1	105
4	16	900
5	43	2000
6	40	1600
7	10	840
8	6	311
9	5	12
10	95	3500
11	54	2400
12	1	0
13	2	5
14	123	8400
15	5	20

From this population,

(i) Select five villages with probability proportional to area under lime with replacement by Lahiri's method.

(ii) From this selected sample, estimate the total number of lime trees along with its standard error. *5+15=20*