[This question paper contains 4 printed pages.]

Sr. No. of Question Paper : 6704 **D** **Your Roll No.....**

Unique Paper Code : 237303

Name of the Course : **B.Sc. (H) Statistics**

Name of the Paper : STHT–304 : Survey Sampling

Semester : III

Duration : 3 Hours Maximum Marks : 75

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. Attempt **five** questions in all, selecting **three** from **Section A** and **two** from **Section B.**

SECTION A

1. (a) Define the following :

   (i) simple random sampling and pps sampling

   (ii) sampling errors and non-sampling errors

   (iii) sampling units and frame

   (b) A sample of size n is drawn from a population having N units by simple random sampling without replacement. A sub-sample of n_1 units is drawn from the n units by simple random sampling without replacement. Let $\bar{y}_1$ denote the mean based on n_1 units and $\bar{y}_2$ be the mean based on $(n - n_1)$ units. Consider the estimator of the population mean $\bar{Y}_N$ given by :

$$\bar{y}_w = w\bar{y}_1 + (1 - w)\bar{y}_2 .$$

   (i) Show that $E(\bar{y}_w) = \bar{Y}_N$, and obtain its variance.

P.T.O.

(ii) Find the optimal value of w for which $V(\bar{y}_w)$ is minimum.

(iii) Find the optimal estimator and its variance.

(c) For srswor, prove that

$$\text{cov}(x_i, \bar{y}_n) = \frac{N-n}{Nn} \cdot \frac{1}{N-1} \sum_{i=1}^{N} (X_i - \bar{X}_N)(Y_i - \bar{Y}_N) = \frac{N-n}{n(N-1)} \text{cov}(X,Y).$$

Also evaluate $E(\bar{x}_n \bar{y}_n)$. (4½,5½,5)

2. (a) Discuss the method given by Dalenius and Hodges for obtaining the optimal points of stratification.

(b) What is meant by non-response in sample surveys ? Explain the Hansen and Hurwitz technique for removing the bias arising from non-response in mail surveys. Obtain the variance of the estimator of population mean.

(c) In a population of N units, the number of units possessing a certain characteristic is A and in a simple random sample of size n from it, the number of units possessing that characteristic is '**a**'. Obtain the estimate of A and its variance. Also obtain an unbiased estimate of var $(\hat{A})$. (5,6,4)

3. (a) Define linear systematic and circular systematic sampling. Prove that systematic sampling is more precise than srswor if the variation within the systematic samples is larger than population variation as a whole.

(b) Estimate the gain in efficiency due to stratification for arbitrary allocation over simple random sampling.

(c) A simple random sample of size 3 is drawn from a population of size N with replacement. Show the probabilities that the sample contains 1, 2 and 3 (e.g aaa, aab, abc) different units respectively are :

$$P_1 = \frac{1}{N^2}, \quad P_2 = \frac{3(N-1)}{N^2}, \quad P_3 = \frac{(N-1)(N-2)}{N^2}$$ (5,6,4)

4. (a) Derive the variance of the estimate of population mean under systematic sampling in terms of intra-class correlation coefficient ρ. Prove that reduction in this variance over srswor will be 100% if ρ assumes the minimum possible value. If ρ assumes the maximum value, what is the relative efficiency of systematic sampling over simple random sampling ?

(b) Establish the result which justifies the following statements :

(i) The smaller the size of stratum, the smaller should be the size of sample to be selected from that stratum.

(ii) The smaller the variability within a stratum, the smaller should be the size of sample selected from that stratum.

(iii) The cheaper the cost per unit in a stratum, the larger should be the size of sample selected from that stratum.

Hence obtain minimum size required for estimating population mean with fixed variance under optimum allocation. (7,8)

SECTION B

5. (a) Derive to the first approximation the bias for ratio estimator and find the condition under which this bias vanishes altogether. Also obtain the variance of the ratio estimator to the first approximation.

(b) Discuss, what is meant by 'model unbiased' ? Stating clearly the underlying assumptions, show that for a super population model the ratio estimator is BLUE. Also derive the expression for the minimum variance. (7,8)

6. (a) Prove that if clusters are formed at random, cluster sampling is as efficient as simple random sampling without replacement.

(b) In two-stage sampling with equal size of first stage units, prove that sample mean is an unbiased estimator of population mean and obtain the estimated variance of sample mean.

(c) Compare two-stage sampling with cluster sampling. (3,8,4)

P.T.O.

7. (a) Prove that the mean of cluster means $\bar{\bar{y}}$ is an unbiased estimator of population mean with variance given by :

$$V(\bar{\bar{y}}) = \frac{N-n}{N-1} \cdot \frac{\sigma^2}{nM}\left[1+(M-1)\rho\right].$$

(b) For two-stage sampling, if the total cost of survey is proportional to the size of the sample, discuss the problem of determining the optimal values of n and m to estimate the population mean with maximum precision for given cost or has desired precision for minimum cost. (10,5)

(600)