[This question paper contains 2 printed pages.]

Your Roll No.....

Sr. No. of Question Paper : 237 **L**

Unique Paper Code : 2373012004

Name of the Paper : Advanced Techniques of Sample Survey, DSE-2B

Name of the Course : **B.Sc. (Hons.) Statistics (Under NEP-UGCF)**

Semester : IV

Duration : 3 Hours Maximum Marks : 90

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. Attempt **six** questions in all.
3. **All** questions carry equal marks.
4. Use of non-programmable scientific calculator is allowed.

1. (a) Explain the concept of Horvitz-Thompson estimator.

(b) If $\pi_i > 0, (i = 1, 2, 3, \ldots, N)$, then show that $\hat{Y}_{HT} = \sum_i^n \frac{y_i}{\pi_i}$ is an unbiased estimator of Y, with variance $V(\hat{Y}_{HT}) = \sum_{i=1}^N \frac{(1-\pi_i)}{\pi_i} y_i^2 + 2\sum_{i=1}^N \sum_{j>i}^N \frac{(\pi_{ij}-\pi_i\pi_j)}{\pi_i\pi_j} y_i y_j$,

where π_{ij} is the probability that units i and j both are in the sample. (5, 10)

2. (a) Explain the concept of post-stratification. Prove that the post-stratification, with large sample, is almost as efficient as stratified sampling with proportional allocation.

(b) For large N, show that the relative efficiency of cluster sampling compared with simple random sampling in terms of intra-class correlation (ρ) is $[1 + (M-1)\rho]^{-1}$, where M is the size of the cluster. (7, 8)

3. (a) Explain the concept of super-population model with suitable example.

(b) Consider the model $Y_i = \beta x_i + \epsilon_i [V(x_i)]^{1/2}$ and assume that $V(x_i) = x_i$. Show that under this assumption the model-based estimator of $T = \sum_{i=1}^N Y_i$ is the usual ratio-estimator. Derive its model variance and the condition under which it is minimum. (5,10)

4. (a) Explain the concept of optimal and balanced sampling with example.

P.T.O.

(b) Show that the estimator $\hat{T}$ of the population total T obtained under a balanced sampling scheme, which is BLUE under the model $Y_i = \beta x_i + \epsilon_i [x_i]^{1/2}$; $i = 1, 2, \ldots, N$. remains BLUE under the extended model $Y_i = \alpha + \beta x_i + \epsilon_i [x_i]^{1/2}$. $i = 1, 2, \ldots, N$. (5, 10)

5. (a) Explain the concept of two-stage sampling with the help of suitable example and diagram.

(b) Suppose that n first-stage units are selected by simple random sampling without replacement (SRSWOR), and from each selected first-stage unit, m second-stage units are further selected using SRSWOR. Show that the sample mean $\bar{y}$ is an unbiased estimator of the population mean $\bar{Y}$. Also, show that its sampling variance is given by

$$V(\bar{y}) = \frac{(N - n)}{N} \frac{S_b^2}{n} + \frac{(M - m)}{M} \frac{\bar{S}_w^2}{nm}$$

where notations have their usual meanings. (5, 10)

6. A population consists of N clusters each containing M elements. A simple random sample of n clusters is selected to estimate the population mean per element. Assume that $\bar{S}_w^2$, the variance within clusters, is given by aM^b, $b > 0$ and the cost function is $C = c_1 nM + c_2 \sqrt{n}$. Find the optimum size of the cluster that minimizes the variance of the estimator given that the cost of survey is fixed. (15)

7. In a survey, the expected response rate is 1/3 and $S_2^2 = \frac{3}{4} S^2$. It costs Rs 1.00 to include a unit in the sample, Rs 5.00 of getting information per unit in the response class, and Rs 9.00 of collecting information per unit in the non-response class. Obtain

(a) the optimal values of k and n so that the mean value of the population is estimated with $400/S^2$ precision.

(b) If the overhead cost is Rs 267, determine the expected cost of the survey. (7, 8)

8. Write a short note on the following.

(a) Hansen and Hurwitz - Call-back methods.

(b) Types of non-response.

(c) Circular systematic sampling. (5, 5, 5)

(500)