Odd Semester, 2020

(Held in March, 2021)

STATISTICS

(Elective/Honours)

[STEH–3 (TH)]

(Categorical Data, Survey Sampling and Design of Experiments)

Marks : 56

Time : 3 hours

The figures in the margin indicate full marks for the questions

Answer **five** questions, taking **one** from each Unit

UNIT—I

1. *(a)* What do you understand by consistency of data? State the conditions for consistency of data of three attributes A, B and C. Consequently show that

$$(AB)+(BC)+(AC) \geq (A)+(B)+(C)-N$$

2+3+1=6

(b) If a collection contains N items, each of which is characterised by one or more of the attributes A, B, C and D, show that with the usual notations

(i) $(ABCD) \geq (A)+(B)+(C)+(D)-3N$

(ii) $(ABCD)=(ABD)+(ACD)-(AD)+(AD\beta\gamma)$

where β and γ represent the characteristics of the absence of B and C respectively. 3+3=6

2. *(a)* What are the various methods of finding whether two attributes, A and B, are associated, disassociated or independent? Deduce that

$$-1 \leq Q \leq 1$$

where Q is the coefficient of association. 3+3=6

(b) If the attributes A and B are independent, then show that the proportion of AB's in the population is equal to the product of the proportions of A's and B's in the population. 6

UNIT—II

3. *(a)* What is meant by sampling method? In what situations, sampling is inevitable? 2+3=5

(b) Show that in SRSWOR, the probability of selecting any specified unit of the population in any draw is equal to the probability of its being selected in the first draw. 3

(c) With the help of random numbers and using any method, explain how you will select a random sample of size 20 from a population of size 200. 3

4. *(a)* Prove that in SRSWOR, the sample mean square is an unbiased estimator of the population mean square. 6

(b) Explain the method of systematic sampling. Discuss its merits and demerits. 2+3=5

UNIT—III

5. *(a)* Under what conditions, stratified random sampling is preferred over simple random sampling and why? 5

(b) Prove that in stratified random sampling using SRSWOR in each stratum

$$\bar{y}_{st} = \sum_{i=1}^{k} P_i \bar{y}_{n_i}$$

is an unbiased estimator of the population mean and

$$V(\bar{y}_{st}) = \sum_{i=1}^{k} P_i^2 \left(\frac{1}{n_i} - \frac{1}{N_i} \right) S_i^2 \qquad 2+4=6$$

6. *(a)* Briefly explain the proportional allocation and optimum allocation methods of allocating a sample to different strata. Obtain the variance of the estimate of the population mean under proportional allocation. 4+2=6

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

"The larger the size of the stratum, the larger should be the size of the sample to be selected from that stratum." 5

UNIT—IV

7. *(a)* Explain what is meant by ANOVA. Discuss the layout of one-way ANOVA under fixed effects model. 2+3=5

(b) Starting from a linear model, show the breakup of the total sum of squares into the different components sum of squares. Also write down the ANOVA table. 5+1=6

8. *(a)* What do you understand by design of experiments? Explain with example how the principle of replication broadens the scope of an experiment. 2+3=5

(b) Define the following terms which occur in the design of experiments : 2+2+2=6

(i) Treatment

(ii) Experimental unit

(iii) Experimental error

UNIT—V

9. *(a)* What is meant by CRD? Give a complete statistical analysis of CRD. 2+4=6

(b) Explain the situation in which RBD is considered an improvement over CRD. What are the disadvantages of RBD? 3+2=5

10. *(a)* Give the layout and analysis of a Latin square design. 5

(b) Define the terms 'main effects' and 'interaction effects' in relation to a 2^2-factorial experiment. Show that in a 2^2-factorial experiment the main effects are mutually orthogonal contrast. 3+3=6

★ ★ ★