Unique Paper Code : **32375301**

Name of the Paper : **Basics of Statistical Inference**

Name of the Course : **Statistics: Generic Elective for Honours (GE-III) under CBCS**

Semester : **III**

Duration : **3 Hours**

Maximum Marks : **75**

Instructions for Candidates

Attempt any ***four*** questions. All questions carry equal marks.

Use of simple calculator is allowed.

1. What do you understand by point estimation? When would you say that estimator of a parameter is 'good'? In particular, discuss the requirements of consistency and unbiasedness of an estimator.

   In a sample of 600 men from a certain large city, 400 are found to be smokers. In another sample of 900 from another large city, 450 are smokers. Do the data indicate that the cities are significantly different with respect to prevalence of smoking among men?

2. Explain the method of constructing a 95% confidence interval for mean of a population following normal distribution with mean μ and variance σ^2, if (i) σ^2 is known, (ii) σ^2 is unknown but sample size is large (i.e. $n \geq 30$) and (iii) σ^2 is unknown but sample size is small (i.e. $n < 30$).

   Independent random samples of sizes $n_1 = 16$ and $n_2 = 25$ from normal populations with $\sigma_1 = 4.8$ and $\sigma_2 = 3.5$ have the means $\bar{x} = 18.2$ and $\bar{y} = 23.4$, find 95% confidence interval for $\mu_1 - \mu_2$.

3. Explain, with the help of examples, where would you use parametric tests and where would you use non-parametric tests. Describe the sign test to test the null hypothesis H_0: $\mu = \mu_0$ against the alternative hypothesis H_1: $\mu > \mu_0$, stating clearly the assumptions made.

   The following are the amounts of time, in minutes, that it took a random sample of 16 technicians to perform a certain task: 18.1, 20.3, 18.3, 15.6, 22.5, 16.8, 17.6, 16.9, 18.2, 17.0, 19.3, 16.5, 19.5, 18.6, 20.0 and 18.8. Assuming that this sample came from a symmetrical continuous population, use the signed-rank test at the 0.05 level of significance to test the null hypothesis that the mean of this population is 19.4 minutes against the alternative hypothesis that it is not 19.4 minutes. ($Z_{0.025} = 1.96$)

4. Discuss the chi-square test of goodness of fit of a theoretical distribution to an observed frequency distribution.

   830 college students were classified according to their intelligence and economic conditions. Test whether there is any association between intelligence and economic conditions. ($\chi^2_{0.05,\,3} = 7.815$, $\chi^2_{0.05,\,2} = 5.991$, $\chi^2_{0.05,\,1} = 3.841$)

		Intelligence			Total
		Excellent	good	Mediocre	
Economic conditions	Good	50	180	170	400
	Bad	80	170	180	430
Total		130	350	350	830

5. What is the difference between 'variability within classes' and 'variability between classes'? Explain with a suitable example.

An experiment was carried out for testing the variety effect. Given the following information:

	A	B	C
No. of observations (*n*)	5	5	5
Sum of observations	90	125	85
Sum of squares of observations	1646	3235	1469

Use analysis of variance technique at the 5% level of significance to test whether the three varieties are significantly different in their mean yields, ~~(~~showing all the steps in the general test procedure~~)~~. ($F_{0.05}$ (2, 12) = 3.89, $F_{0.05}$ (5, 12) = 3.11)

6. A randomized block experiment has been carried out in 4 blocks with 5 treatments A, B, C, D and E. Derive the expected value of mean sum of squares due to errors in the above design.

An experiment was carried out on wheat with three treatments in four randomised blocks. Complete the following table for the analysis of variance of a fixed effects randomised block design:

Source of variation	Degrees of freedom	Sum of squares	Mean sum of squares	Variance ratio
Blocks	3	6.67	-	-
Treatments	2	-	-	-
Error	-	-	2.22	
Total	-	28		

Test the hypothesis that the treatment effects are equal to zero, showing all the steps in the general test procedure. ($F_{0.05}$ (2, 6) = 5.14, $F_{0.05}$ (3, 6) = 4.76)