UNIVERSITY OF PETROLEUM AND ENERGY STUDIES
End Semester Examination, April/May 2018

Course: Applied Statistical Analysis **Semester: IV**
Program: B.tech CSE+BAO
Time: 03 hrs. **Max. Marks: 100**

SECTION A (20 Marks)

S. No.		Marks	CO
Q 1	Explain the difference between following a. Primary and Secondary Data b. Descriptive Statistics Inferential Statistical	**4**	**CO1**
Q 2	Discuss various types of data and data transformation techniques.	**4**	**CO1**
Q 3	What type of null hypothesis is tested using analysis of variance? State basic assumptions of this analysis.	**4**	**CO3**
Q 4	Here are the ages in years of the cars worked on by the Village Autohaus last week. 5, 6, 3, 6, 11, 7, 9, 10, 2, 4,10, 6, 2, 1,5 Computer mode and means for this data set. Comment on which is better measure of the central tendency.	**4**	**CO2**
Q 5	Differentiate between parametric test and non-parametric test with example.	**4**	**CO4**

SECTION B (40 Marks)

Q 6	The customer accounts of a certain departmental store have an average balance of Rs.1200 and a standard deviation of Rs.400. Assuming that the account balances are normally distributed. a. What percentage of the accounts is over Rs.1500? b. What percentage of the accounts is between Rs.1000 and Rs.1500? c. What percentage of the accounts is below Rs.1500?	**10**	**CO2**
Q 7	The null hypothesis is that 20 per cent of the passengers go in first class, but management recognizes the possibility that this percentage could be more or less. A random sample of 400 passengers includes 70 passengers holding first class tickets. Can the null hypothesis be rejected at 10 per cent level of significance?	**10**	**CO3**
Q 8	Two independent samples of observations were collected. For the first sample of 60 elements the mean was 86 and the standard deviation 6. The second sample of 75 elements had a mean of 82 and standard deviation of 9. a) Compute the estimated standard error of the difference between two means. b) Using alpha=0.01, test whether the two samples can reasonably be considered to have come from population with the same mean.	**10**	**CO3**

Q 9	Explain following classification algorithm. a) Decision Tree b) Cluster Analysis OR Explain the following algorithm. a) Neural Network b) Factor Analysis	**10**	**CO5**
	SECTION-C(40 Marks)		
Q 10	Cost accountants often estimate overhead based on the level of production. They have collected information on the overhead expense and units of produced at the different plants and want to estimate a regression to predict future overhead. **Overhead:** 191 170 272 155 280 173 234 116 153 178 **Units:** 40 42 53 35 56 39 48 30 37 40 **a)** Develop the regression equation for cost accountants. **b)** Predict overhead when 50 units are produced. **c)** Calculate the standard error of estimate. OR Explain the following. a. Kruskal-Wallis Tests (H Test) b. Wilcoxon matched pairs Signed Ranks Test c. Sign Test d. Partial correlation.	**20**	**CO5**
Q 11	Why Chi-Square test is used? At the .10 level of significance, can we conclude that the following 400 observation follow a Poisson distribution with **λ=3?** No of arrivals per hours: 0 | 1 | **2** | **3** | **4** | **5 or more** No of hours: 20 | 57 | 98 | 85 | 78 | 62	**20**	**CO4**
Name: **Enrolment No:**		UPES	