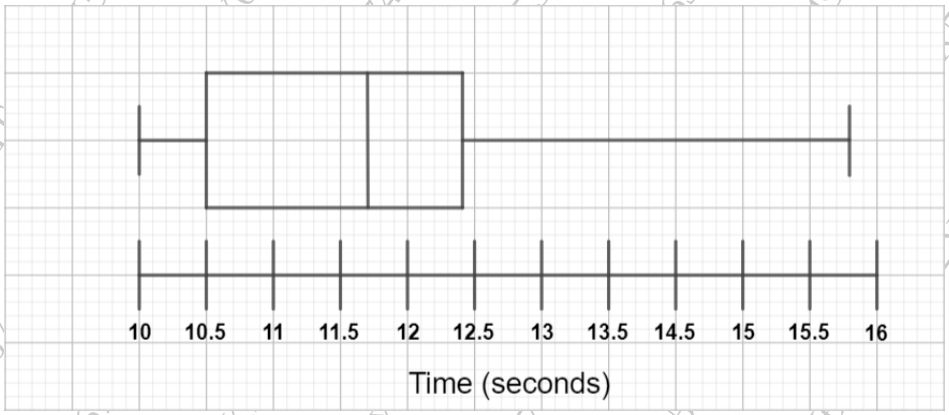
(3 Hours) **Total Marks: 80**

N.B. 1. Question No. 1 is compulsory.
2. Attempt any three questions out of remaining five.
3. All questions carry equal marks
4. Assume Suitable data, if required and state it clearly.

1 Attempt any four: **20**

(a) The box plot below was constructed from a collection of times taken to run a 100 m sprint. Using the box plot, determine the range and interquartile range. **5**



(b) Define Type I and Type II Errors **5**

(c) In 800 families with 4 children each. Classify according to given criteria, how many families would you expect to have? **5**
2 boys and 2 girls

(d) A coin is tossed three times. Calculate the probability of obtaining more heads than tails. **5**

(e) Explain the various decomposition models used in time series data. Also, state which decomposition model will be appropriate for the following condition: **5**
a) When the seasonal variation is relatively constant over time.

2 (a) You have just taken ownership of a pizza shop. The previous owner told you that you would save money if you bought the mozzarella cheese in a 4.5 pound slab. Each time you purchase a slab of cheese, you weigh it to ensure that you are receiving 72 ounces of cheese. The results of 7 random measurements are 70, 69, 73, 68, 71, 69 and 71 ounces. Are these differences due to chance or is the distributor giving you less cheese than you deserve? **10**
a. State the hypotheses.
b. Calculate the test statistic.
c. Would the null hypothesis be rejected at the 10% level? The 5% level? The 1% level?

(b) Elaborate moving average and exponential smoothing techniques? **10**

3 (a) Define sampling and central limit theorem ? Elaborate stratified sampling, judgment sampling , systematic sampling and cluster sampling **10**

(b) Use multiple regression derive equation for y given x1 and x2. **10**

y	x1	x2
-3.7	3	8
3.5	4	5
2.5	5	7
11.5	6	3
5.7	2	1

4 a) In a manufacturing unit, four teams of operators were randomly selected and sent to four different facilities for machining techniques training. After the training, the supervisor conducted the exam and recorded the test scores. At 95% confidence level does the scores are same in all four facilities? (Kruskal–Wallis and chi-square table) **10**

Facility 1	Facility 2	Facility 3	Facility 4
88	77	71	52
82	76	56	65
86	84	64	68
87	59	51	81

b) In the context of Multiple linear regression explain what is Over fitting & multicollinearity? **10**

5 a) Some vehicles pass through a junction on a busy road at an average rate of 300 per hour. **10**

a. Find out the probability that none passes in a given minute.
b. What is the expected number of passing in two minutes?
c. Find the probability that this expected number found above actually pass through in a given two-minute period.

b) Find the simple linear regression equation for the data given below: **10**

X	Y
2	21
4	27
6	29
8	64
10	86
12	92

6 a) Explain any 3 numerical measures for : **10**

a. Measures of variability
b. Measures of location
c. Measures of distribution shape

b) Difference between **10**

a. Parametric and non-parametric test
b. Discrete and Continuous probability distribution.