Question Paper

Exam Date & Time: 27-Nov-2018 (10:00 AM - 01:00 PM)



MANIPAL ACADEMY OF HIGHER EDUCATION

SCHOOL OF INFORMATION SCIENCES

FIRST SEMESTER MASTER OF ENGINEERING - ME (BIG DATA AND DATA ANALYTICS) Probability and Statistical Inferences [BDA 605]

s.

Maulas 100	Trobability and Statistical Interences [DDA 005]					
Marks: 100		30 min				
Answer	all the questions					
1)	1A) Discuss in detail the axiomatic definition of probability.	(10)				
	1B) State the Bayes theorem with a detailed explanation on terms involved in the theorem. Provide an example of the theorem. (5+5=10 Marks)					
2)	 2A) Define the following concepts with the help of an example a) Sample space b) Exhaustive events c) Equally likely events. 	(10)				
	2B) State the multiplication theorem in probability and discuss the change in the theorem when the events are statistically independent? $(6+4=10 \text{ Marks})$					
3)	3A) Define a univariate random variable. What are the two major (10) classifications of a univariate random variable? Give two examples for the each classification.					
	(2+1+2=5 marks)					
	 3B) Define the following terms for an univariate random variable X, a) Probability mass function b) Probability density function c) Distribution function. (2+2+1=5 marks) 					
4)	 4) A coin is thrown three times, and the sequence of heads and tail is observed. Let X be a random variable which denotes "the number of heads observed". Then, A) Define the sample space of the random experiment. B) Identify the functional form for the probability mass function (p.m.f) of the random variable X. C) Identify the distribution function for the random variable X. Explain its properties using a neat diagram. (1+4+5= 10 marks) 	(10)				

5) 5A) Write a short note on normal distribution with the help of a neat (10) diagram and by clearly specifying the properties.

5B) Suppose that an average of 5 traffic accidents per month occur at Manipal. Let X be a random variable which denotes the number of traffic accidents that occur according to a Poisson distribution. Assuming the average number of accidents in a month is constant throughout an year, calculate the probability of observing more than 5 accidents in a given month. (5+5=10 marks)

6A) State the Central Limit Theorem (CLT)?

6B) The population is the weight of six pumpkins (in pounds) displayed in a carnival "guess the weight" game booth. You are asked to guess the average weight of the six pumpkins by taking a random sample without replacement from the population.

Pumpkin	А	В	С	D	Е	F
Weight (in pounds)	19	14	15	9	10	17

a) Calculate the population mean?

b) Obtain the sampling distribution of the sample mean for a sample size of 2 when one samples without replacement?

c) Obtain the sampling distribution for the sample mean when the sample size is 5. $(2.5 \times 4 = 10 \text{ marks})$

7) 7A) Draw the table of statistic versus standard error of sample mean, (10) observed sample proportion, difference of two sample mean and difference of two sample proportion?

7B) A random sample of 500 apples was taken from a large consignment and 60 were found to be bad. Obtain the 95% confidence limits for the percentage number of bad apples in the consignment.

7C) In a certain factory there are two independent processes manufacturing the same item. The average weight in a sample of 250 items produced from one process is found to be 120 ozs. with a standard deviation of 12 ozs. while the corresponding figures in a sample of 400 items from the other process are 124 and 14. Obtain the standard error of difference between the two sample means. Also find the 99% confidence limits for the difference in the average weights of items produced by the two processes respectively. (4+3+3=10 marks).

8A) Explain hypothesis testing, Type I and Type II errors for hypothesis ⁽¹⁰⁾ testing, power of test, critical value, significant value, and p-value interpretation? Use figures and tables.

8B) Suppose we take a random sample of 500 Penn State students and find that 278 are from Pennsylvania. Can we conclude that the proportion is larger than 0.5 at a 5% level of significance? Also solve using the P-value Approach? Given: (tabulated $Z_{0.05} = 1.645$, P(Z > 2.50) = 0.0062)

(5+5=10)

marks)

(10)

8)

9A) Explain the statistical testing procedure for comparing two independent population means.

9B) In a packing plant, a machine packs cartons with jars. It is supposed that a new machine will pack faster on the average than the machine currently used. To test that hypothesis, the times it takes each machine to pack ten cartons are recorded. The results, in seconds, are shown in the following table.

New machineOld machine42.141.342.443.241.842.743.842.543.144.04141.842.842.342.743.643.343.541.744.1Do the data provide sufficient evidence to conclude that, on the average,
the new machine packs faster? Perform the required hypothesis test at
the 5% level of significance. Given: $-t_{0.005} = -1.734$ for 18 d.f.

(5+5 =10 marks)

¹⁰⁾ 10 A) The data represent the number of algebra problems solved in 3 minutes time ⁽¹⁰⁾ by three different high school students.

Sample 1	Sample 2	Sample 3		
5	5	6		
3	3	4		
1	3	4		
4	1	2		
2	4	3		

Do the problem solving ability of the students of 3 different schools varies with each other? Given: $F_{2,12} = 3.88$.

10 B) Write a short note on non-parametric statistical testing methods? State the advantages and disadvantages of non-parametric methods?

(5+5=10 marks)

9)

-----End-----