Question Paper

Exam Date & Time: 19-Apr-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)
DEGREE EXAMINATION-APRIL 2018
Thursday, April 19, 2018
Time: 10:00am to to 1:00pm

Probability and Statistical Inferences [BDA 605]

Marks: 100 **Duration: 180 mins.**

Answer all the questions.

1) A) Describe the Von Mises's Statistical definition of probability with its advantages and limitations (5)

B) State the Bayes theorem with a detailed explanation on terms involved in the theorem. Provide an example of the theorem. (5)

2) A) Define the following concepts with the help of an example (6)
(i) Independent events
(ii) Conditional probability
(iii) Mutually exclusive events (3x2=6 marks)

B) Following table shows the relationship between gender and eye color for a group of 167 German men. (4)

		Eye Colour					
		Black	Brown	Blue	Green	Gray	Total
Gender	Female	20	30	10	15	10	85
	Male	25	15	12	20	10	82
	Total	45	45	22	35	20	167

(i) What is the probability that a randomly selected person will have green eyes?
(ii) What is the probability that a randomly selected female will have brown eyes?
(iii) What is the probability that a randomly selected person will have brown eyes OR will be male?
(1+1+2=marks)

3) (i) Define a univariate random variable. What are the two (5)

A) major classifications of a univariate random variable? Give two examples for the each classification.

(2marks)

(ii) For an univariate random variable X, Define the following terms.

a) probability mass function

b) probability density function

c) distribution function (1+1+1=3 marks)

B) (i) Let X be a random variable which denotes "the number of heads observed" defined for the random experiment of throwing a coin three times and observing the sequence of heads and tail. Obtain the functional form of the probability mass function (p.m.f) and Cumulative Distribution Function (CDF) for the random variable X . (5)

(3 marks)

(ii) Let the distribution function of a random variable X is defined as follows,

$$F(x)=\begin{cases} 0 & if \quad x<0 \\ x & if \quad 0 \leq x \leq 1 \\ 1 & if \quad x>1 \end{cases}$$

Find $P\left(\dfrac{2X+6}{3}>2.2\right)$? (2 marks)

4) Write a short note on Poisson distribution (5)

A)

B) Calculate the values of mean and variance of the binomial distribution specified by the parameters n = 5 and 1-p = 0.4? (5)

5) Define the following terms (5)

A) (i) Parameter

(ii) Estimate

(iii) Statistic

(iv) Sampling distribution

(v) Standard error

B) What you mean by an unbiased estimator? Let X be a random variable following a normal distribution with mean μ and variance σ^2. Propose an estimator for the population mean μ. Check whether the proposed estimator is an unbiased estimator for the population mean μ. (5)

6) Let X be a random variable following a normal distribution (10)

with mean μ and variance σ^2. Then,

(i) Give the sampling distribution of the sample mean.
(ii) State the central limit theorem
(iii) When a batch of a chemical product is prepared, the amount of an impurity (in grams) in the batch is a random variable X with: μ = 4.0g and $\sigma^2 = (1.5g)^2$. Suppose that n = 50 batches are prepared (independently). What is the probability that the sample mean impurity amount will be greater than 4.2grams? (Hint: $\phi(4.2) = 0.8272$)

(3+2+5=10 marks)

7) A) Derive the expression for $100(1-\alpha)\%$ confidence interval for the mean of a normal distribution $N(\mu,\sigma^2)$ when σ is unknown and the sample size is small. (5)

B) (i) Give the expression for $100(1-\alpha)\%$ confidence interval for the population proportion p. (5)
(ii) A medical study showed 57 of 300 persons failed to recover from a particular disease. Obtain the 95 % C.I for the mortality rate of the disease.

(2.5+2.5=5marks)

8) (i) Discuss about the different type of errors in testing of hypothesis and how to manage these errors. (10)
(ii) What do you mean by critical region and discuss the decision criteria based critical region
(iii) Mention the steps involved in hypothesis testing (5+3+2=10 Marks)

9) Discuss the following tests in detail (10)
(i) ANOVA
(ii) Mann-Whitney U test

10) Suppose we measure the thickness of plaque (mm) in the carotid artery of 10 randomly selected patients with mild atherosclerotic disease. Two measurements are taken, thickness of plaque before treatment with Vitamin E (baseline) and after two years of taking Vitamin E daily. Do the data provide enough evidence to indicate that there is a difference in plaque before and after treatment with vitamin E for two years? (10)
[α = 0.01, $t_{1-\alpha/2}$ (9) = 3.25]

Before	0.66	0.72	0.85	0.62	0.59	0.63	0.64	0.70	0.73	0.68

After	0.61	0.65	0.79	0.63	0.51	0.55	0.62	0.67	0.65	0.64

-----End-----