Question Paper

MANIPAL UNIVERSITY

SCHOOL OF INFORMATION SCIENCES
FIRST SEMESTER Master of Engineering - ME (BIG DATA AND ANALYTICS)
DEGREE EXAMINATION - NOVEMBER 2017
DATE : Tuesday, November 18, 2017
TIME : 10:00AM - 1:00PM

Probability and Statistical Inferences [BDA 605]

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

Answer all the questions.

1) (A) Describe the mathematical definition of probability with its advantages and its limitations (10)

(b) 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) A. Define the following concepts with the help of an example (10)
(i) Random experiment
(ii) Mutually exclusive events
(iii) Equally likely events

B. Explain conditional probability with the help of an example

((3X2)+4=10 Marks)

3) Given the following probability distribution, (10)

X	0	1	2	3	4	5	6	7
P(X=x)	0	c	2c	2c	3c	c^2	$2c^2$	$7c^2+c$

Find the value of
(i) c
(ii) P(X=7)
(iii) $P(X < 3)$
(iv) $P(X \geq 5)$

(4+1+2+3=10 marks)

4) A. Write a short note on Binomial distribution (10)

B. What is the mean and variance of the bionimal distribution, with binomial parameters n = 3 and $1-p = 0.6$? **(5+5=10 marks)**

5) A. Define and differentiate in detail between point estimation and interval estimation. (10)

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. Let $(X_1, X_2, \ldots\ldots, X_n)$ be a random sample. Verify whether the sample mean $\frac{1}{n}\sum_{i=1}^{n} X_i$ is an unbiased estimator for the population mean μ.

(5+5=10 Marks)

6) Let X be a random variable following a normal distribution with mean μ and variance σ^2. (10)

Then,

(i) Give the 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. (10)

B. A random sample of 16 values from a normal population showed a mean of 41.5 inches and the sum of squares of deviation from this mean equal to 135 square inches. Obtain the 95% and 99% confidence interval for the population mean.

(Hint: $t_{15,\frac{0.05}{2}} = 2.131$, $t_{15,\frac{0.01}{2}} = 2.947$)

8) A. Define the term hypothesis? Explain the null and alternative hypothesis with the help of suitable examples. (10)
B. Define P value and discuss the decision criteria based on P value.
C. Mention the steps involved in hypothesis testing

(5+3+2=10 Marks)

9) Discuss the following tests in detail (10)
A. Paired t test
B. Kruskal-Wallis test
(5+5=10 Marks)

10) A. Given the Birth weight (in kg) of the child, 7 born to a non-smoker and 6 to a smoker mothers. Investigate whether there is statistically significant difference in the birth weights of children born to smoker at 5% level of significance and those born to a nonsmoker mother. (Tabulated value= 2.201) (10)

	N	Mean	Standard. Deviation
Non smoker	7	3.3	0.34
Smoker	6	3.8	0.2

B. Mention the assumptions related to the test used
(7+3 =10 Marks)

-----End-----