

School of Computing Science and Engineering

Master of Computer Applications
Semester End Examination - May 2024

Duration : 180 Minutes

Max Marks : 100

Sem IV - MCAN2412 - Data Analytics with Python

General Instructions

Answer to the specific question asked

Draw neat, labelled diagrams wherever necessary

Approved data hand books are allowed subject to verification by the Invigilator

- 1) A letter of the English alphabet is chosen at random. Calculate the probability that the letter so chosen precedes m and is a vowel. K1 (2)
- 2) Illustrate How do you create dummy variables in a dataset K2 (4)
- 3) In a large number of parts manufactured by a machine , the mean number of defectives in a sample of 20 is 2. Out of 1000 such samples , how many would be expected to contain atleast 3 defective parts. K2 (6)
- 4) Suppose that during any hour in a large department store, the average number of shoppers is 448, with a standard deviation of 21 shoppers. K3 (9)

 What is the probability that a random sample of 49 different shopping hours will yield a sample mean between 441 and 446 shoppers?
- 5) Apply ROC analysis to evaluate the performance of a binary classifier. Given a dataset with 100 instances, where the classifier correctly predicts 80 positive instances and 10 negative instances, calculate the True Positive Rate (TPR) and False Positive Rate (FPR). K3 (9)
- 6) Calculate the F-statistic for two samples with variances of 20 and 30, respectively. K5 (10)
- 7) Compare and contrast the least squares method with the method of moments in estimating parameters of a statistical model. K4 (12)
- 8) It is known that 5% of the books bound at a certain bindery have defective bindings. Find the probability that 2 of 100 books bound by this bindery will have defective bindings using K5 (15)

 (a)The formula for the binomial distribution;
 (b)The Poisson approximation to the binomial distribution
- 9) Given the data points (1, 3), (2, 5), (3, 7), (4, 9), calculate the least squares regression line and use it to predict the value of y for x = 5. K5 (15)
- 10) Four technicians regularly make repairs when breakdowns occur on an automated production line. K6 (18)
 Janet, who services 20% of the breakdowns, makes an incomplete repair 1 time in 20;
 Tom, who services 60% of the breakdowns, makes an incomplete repair 1 time in 10;
 Georgia, who services 15% of the breakdowns, makes an incomplete repair 1 time in 10;
 and Peter, who services 5% of the breakdowns, makes an incomplete repair 1 time in 20.
 For the next problem with the production line diagnosed as being due to an initial repair that was incomplete, what is the probability that this initial repair was made by Janet?