Goa Vidyaprasarak Mandal's

GOPAL GOVIND POY RAITURCAR COLLEGE OF COMMERCE AND ECONOMICS, PONDA-GOA

B.C.A. (SEMESTER –IV) EXAMINATION, APRIL 2018

DATA ANALYSIS AND STATISTICAL TECHNIQUES

Duration: 2 Hours **Marks: 50**

Instructions: Standard normal distribution table and t-distribution table will be given.

Answer the following questions. Justify your answer.

Q.1.(a) (7)

(i) From a bag containing 3 red and 2 black objects, 2 objects are drawn at random.

Find the probability that they are of the same color.

(ii) When A and B are two mutually exclusive events such that P(A)=1/2 and P(B)=1/3,find $P(A \cup B)$.

(iii) If P(A)=1/3,P(B)=3/4 and $P(A \cup B)$= 11/12,find P(A/B).

(iv) If P(A)=.65, P(B)=.4 and $P(A \cap B)$=.24, can A and B be independent events?

(v) Using Poisson distribution formula, find P(x=2) for λ= 2.5

(vi) Find the value of t for the t distribution for the following data,

Area in the right tail=0.5 and df = 9

(vii) Find the area from the t distribution table for the data given as t=2.467 and df=28

1. (b) If $A \subset B$, P(A)=1/4 and P(B)= 1/3, find P(A/B) and P(B/A). (3)

Q.2.

(i) A factory production line is manufacturing bolts using three machines, A, B and C. Of the total output, machine A is responsible for 25%, machine B for 35% and machine C for the rest. It is known from previous experience with the machines that 5% of the output from machine A is defective, 4% from machine B and 2% from machine C. A bolt is chosen at random from the production line and found to be defective. What is the probability that it came from

(a) machine A (b) machine B (c) machine C? (5)

(ii) Fit a straight line of the form x= a+by to the following data. Here y is independent variable and x is dependent variable.

x	-16	-11	-6	-1	14	19	23·5	29·5
y	-4	-3	-2	-1	2	3	4	5

Calculate linear correlation coefficient for the above data. (7)

Q.3.

(i) Find the value of z so that the area under the standard normal curve,

(a)From 0 to z is ·4772 and z is positive (b) from 0 to z is approximately ·4785 and z is negative. (4)

(ii) For the data given below

15,9,8,7,6,9,14,3,6,9

Calculate median and mode. Determine the values of three quartiles

and the inter quartile range. (5)

Q.4.

(i) (a) For a data set obtained from a sample, n=20 and $\bar{x}$=20.4. It is known that σ=2.3. Make a 99% confidence interval for μ. (2)

(b) Suppose for a sample, $\bar{x}$=68.5 and s=8.9. Construct a 95% confidence interval for μ, assuming that n=16. (3)

(ii) Given the following points,

2, 3, 4,10,11,12,20,25,30,

assume k=2 and the initial two clusters are (2, 3) and (4,10,11,12,20,25,30). Determine the clusters obtained using k-means algorithm after one iteration. (5)

Q.5.

(i) For the data, 7,8,14,7,20,12, list all possible samples of size five that can be selected from this population. Calculate the mean for each of those samples, write the sampling distribution of $\bar{x}$. (5)

(ii) The following transaction table is given

TID	Items bought
1	a,pe,pi
2	o,a,m,t
3	a,pe,t,c
4	a,t,pe,o

Find one item set of size 2 and 3 having minimum support 50% . Using the large item set of size 3, determine two association rules with minimum confidence 50%. (4)

xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx