

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

TEST -2 EXAMINATION- 2016

B.Tech IV Semester

COURSE CODE: 10B12MA421

MAX. MARKS: 25

COURSE NAME: BIOSTATISTICS

COURSE CREDITS: 4

MAX. TIME: 1Hr 30 Min

Note: All questions are compulsory. Carrying of mobile phone during examinations will be treated as case of unfair means.

1. The number of failures $N(t)$, which occur in a computer network over the time interval $[0, t)$, can be described by a homogeneous Poisson process $\{N(t), t \geq 0\}$. On an average, there is a failure after every 4 hours, i.e. the intensity of the process is equal to $\lambda = 0.25$ per 30 minutes
 - a) What is the probability of 1 failure in $[0, 8)$, at least 2 failures in $[8, 16)$, and at most 1 failure in $[16, 24)$ (time unit: hour)?
 - b) What is the probability that the third failure occurs after 8 hours? Marks 6
2. A system has three possible states, 0, 1 and 2. Every hour it makes a transition to a different state, which is determined by a coin flip. For example, from state 0, it makes a transition to state 1 or state 2 with probabilities 0.5 and 0.5.
 - a) Make flow chart of one step transition probabilities and then construct the transition probability matrix.
 - b) Find the three-step transition probability matrix.
 - c) Find the steady-state distribution of the Markov chain. Marks 6
3. Consider the following data set consisting of the scores of two variables on each of ten individuals:

Subject	1	2	3	4	5	6	7	8	9	10
A	1	1.5	3	5	3.5	4.5	3.5	2	4	2.5
B	1	2	4	7	5	5	4.5	2.5	5.5	4.5

As a first step in finding a sensible initial partition, let the A & B values of the two individuals furthest apart (using the Euclidean distance measure) define the initial cluster means, giving:

	Individual	Mean Vector (Centroid)
Group 1	1	(1.0, 1.0)
Group 2	4	(5.0, 7.0)

Apply **k-Means** clustering and group the data set into two clusters. Marks 6

4. Given data concerning the dichotomous data of 6 species in 7 samples.

	S1	S2	S3	S4	S5	S6
A	1	1	1	0	1	0
B	1	1	0	1	1	1
C	0	1	1	0	1	1
D	0	0	0	1	0	0
E	1	1	1	0	1	0
F	0	1	0	1	1	1
G	0	1	1	0	1	0

On this presence-absence data;

- a) Find the dissimilarities, based on the Jaccard index, between all pairs of seven samples.
- b) Apply hierarchical clustering Using the 'maximum' (or 'complete linkage') method to and make the Dendrogram.
- c) Write the solution which provides all clusters with respective cut point. Marks 7