Saurabh Srivastava

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

TEST -3 EXAMINATION- May-2018

B.Tech. IV Semester

COURSE CODE: **10B12MA421** MAX. MARKS: **35**

COURSE NAME: **BIOSTATISTICS**

COURSE CREDITS: **5** MAX. TIME: **Two Hours**

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

1(i) A study to assess the capability of subsurface flow wetland systems to remove biochemical oxygen demand (BOD) and various other chemical constituents resulted in the accompanying data on $x =$ BOD mass loading (Kg/ha/d) and $y =$ BOD mass removal (Kg/ha/d). Values of relevant summary quantities are, $n=14, \sum x_i = 517, \sum y_i = 346, \sum x_i^2 = 39095, \sum y_i^2 = 17454, \sum x_i y_i = 25825$. Determine the equation of the estimated regression line. **[3][CO-1]**

(ii) Write the formula for simple correlation coefficient and define t-statistic to test the hypothesis that $\rho = 0$. **[1+1][CO-2]**

2 (i) To find out whether a new serum will arrest leukemia, 9 patients who have all reached on advanced stage of the disease are selected. Five patients receive the treatment and 4 do not. The survival times in years from the time, the experiment commenced are,

Treatment	**2.1**	**5.3**	**1.4**	**4.6**	**0.9**
No Treatment	**1.9**	**0.5**	**2.8**	**3.1**	

Use the ***Wilcoxon rank sum test*** at the 0.05 level of significance, to determine if the serum is effective. *(Given that u(4,5)=1 for 0.05 level of significance)* **[3][CO-3]**

(ii) For the stochastic process $X(t) = \zeta \cos \omega t, -\infty < t < \infty, \omega$ is constant, the amplitude ζ is a random variable with Uniform distribution in the interval $(0,1)$, check whether the process is WSS or not? **[2][CO-4]**

3. A student's study habits are as follows: if he studies one night, he is 70% sure not to study the next night. On the other hand if he does not study one night, he is 60% sure not to study the next night as well. In the long run, how often does he study? Also draw the transition state diagram. **[4+1][CO-5]**

[P.T.O.]

4. Suppose that the blood samples of patients arriving at a pathology unit according to a Poisson's process with a mean rate of 2 samples per minute. Find the probability that in an interval of 5 minutes, the number of blood samples arriving is **(i)** exactly 3 **(ii)** less than 3. **[2.5+2.5][CO-5]**

5. Apply ***Agglomerative Hierarchical Clustering algorithm*** to find the clusters, if the proximity matrix is as follows;

Subjects	P	Q	R	S	T
P	0	1	3	5	4
Q	1	0	2	3	5
R	3	2	0	1	6
S	5	3	1	0	3
T	4	5	6	3	0

Also draw the ***Dendrogram*** and find its height. **[3+1][CO-6]**

6. Apply ***K-means algorithm*** up to only two iterations for the following data;

Items	A	B	C	D	E	F	G	H	I	J
X	5	5	3	0	2	4	2	2	1	5
Y	0	2	1	4	1	2	2	3	3	4

[3+3][CO-6]

7. ***CLARA*** and ***CLARANS*** are algorithms employed for clustering of large database. Compare their

(i) Characteristics

(ii) Advantages

(iii) Disadvantages. **[3+1+1][CO-6]**