



Gandhi Institute of Engineering and Technology University, Odisha, Gunupur
(GIET UNIVERSITY)

M.Sc. (Second Semester - Regular) Examinations, May - 2026
24MLSPC12001- Bioinformatics and Biostatistics
Life science

Time: 3 hrs

Maximum: 60 Marks

Answer ALL questions
(The figures in the right hand margin indicate marks)

PART – A

(2 x 5 = 10 Marks)

Q.1. Answer ALL questions	CO #	Blooms Level
a. Draw the flow chart of storing method of EMBL	CO1	K1
b. What is LIBRA ? Write its use	CO2	K1
c. Find the Hamming distance between two sequence ADVANCED and NEWDELHI	CO3	K1
d. Find the mean for the following data. 6, 15, 4, 10, 12, 11, 5, 3, 16	CO4	K1
e. Write the formula to calculate expected frequency and degrees of freedom	CO5	K1

PART – B

(10 x 5 = 50 Marks)

Answer ALL the questions		Marks	CO #	Blooms Level												
2. a.	Discuss the storing and retrieving method of NCBI database	5	CO1	K1												
b.	Explain briefly on Genbank database	5	CO1	K1												
(OR)																
c.	Explain the layer of PIR database	5	CO1	K1												
d.	Explain the Menu and sub menu of PDB database	5	CO1	K1												
3.a.	Find the alignment score and optimal alignment of two sequence GCATGCG and GATTACA (assume match = 1, mismatch = -1 and gap = -1) by using Needleman Wunch algorithm	5	CO2	K2												
b.	Make a PAM matrix of all amino acid of the given MSA ACGCTAFKI GCGCTAFKI ACGCTAFKL GCGCTGFKI GCGCTLFKI ASGCTAFKL ACACTAFKL	5	CO2	K2												
(OR)																
c.	Find the optimal alignment and alignment score between two sequence GGATCGA and GAATTCAGTTA (Assumeing match = 5, Mismatch = -3 and gap = -4) bu using smith-Waterman algorithm	5	CO2	K2												
d.	Suppose there are 20,000 aminoacid in the database of which 2000 are Serin and there are 5000 amino acids in helical conformation of which 500 are serine. Calculate the type of information	5	CO3	K2												
4.a.	Design a HMM of the given MSA VGA- -H V - - - -N VEA- - D VKG - - - VYS - -T FNA - - N IAGADN	5	CO3	K3												
b.	Design a phylogenetic tree by NJ method	5	CO3	K3												
<table border="1"><tr><td></td><td>A</td><td>B</td><td>C</td><td>D</td><td>E</td></tr><tr><td>A</td><td></td><td>22</td><td>39</td><td>39</td><td>41</td></tr></table>						A	B	C	D	E	A		22	39	39	41
	A	B	C	D	E											
A		22	39	39	41											

B			41	41	43
C				18	20
D					10
E					

(OR)

- c.

Convert the given molecular marker into MSA
TCYGIFVL, TCGIFVL, SCYGIFVLSG, ACGIFVLSG

5

CO4

K2
- d.

How many rooted phylogenetic tree is possible to construe of the given distance tableby using UPGMA method.

5

CO4

K3

	A	B	C	D
A				
B	8			
C	7	9		
D	12	14	11	

- 5.a.

Calculate the mean from the following data of presence of urea in the blood samples of 520 patients in a hospital

5

CO4

K2

Range of (mg/dl)	20-25	25-30	30-35	35-40	40-45	45-50	50-55	55-60
No. of patients	20	44	60	101	109	84	66	10

- b.

Find the standard deviation of intelligence quotient of 68 students of the following data

5

CO5

K2

Intelligence quotient	10-20	20-30	30-40	40-50	50-60	60-70	70-80
No.of students	5	12	15	20	10	4	2

(OR)

- c.

Calculate the mean deviation about the mean for the following data of Tryglyceride presents in the blood sample of 50 patients tested in pathology Laboratory on a certain day

5

CO5

K2

Tryglycerides	5	15	25	35	45	55	65
No. of patients	8	12	10	8	3	2	7

- d.

A shoe shop in Delhi had sold 100 pairs of shoes of a particular brand on a certain day with the following distribution .

5

CO5

K2

Size of shoe	4	5	6	7	8	9	10
No of pairs	10	15	20	35	16	3	1

Find the mode of the distribution

- 6.a.

Find the coefficient of correlation between the heights of father and sons from the following data

5

CO6

K2

Heights of fathers in inches	65	66	67	68	69	70	71
Heights of sons in inches	67	68	66	69	72	72	69

- b.

The pulse rate of a man due to the effect of Amtas AT 25 mg on different days in a month were found to be 66,65,69,70,69,71,70,63,64, and 68. Discuss whether the mean pulse rate of the man in the month is 65

5

CO6

K2

(OR)

- c.

From the given data calculate regration equation taking devation of item from the mean of X and Y series

5

CO1

K1

X	1	3	5	7	9	4	6
Y	2	4	3	7	5	6	1

- d.

Ten objects are chosen at random from a large population and their weight are found to be in gms 36,63,64,65,66,69,69,70,70,71. In the light of the above data, discuss the suggestion that the mean in 65

5

CO1

K1

--- End of Paper ---