



THAPAR INSTITUTE OF ENGINEERING AND TECHNOLOGY

End-Semester Examination – December 2006

**BT-009 BIOSTATISTICS AND BIOINFORMATICS**

Instructor – Dr. N.Tejo Prakash

Time: 3 Hours

Maximum Marks: 36

The question paper contains two sections (A & B)

Follow serial order in answering questions

**Section A (Bioinformatics)**

**ANSWER TWO OUT OF THREE QUESTIONS (2 X 6 MARKS = 12 MARKS)**

1. What are the major tools and extensions of BLAST. Discuss the areas of application of these programs. (6)
2. Generate Needleman-Wunsch matrix for the following pair of sequences and show the alignment (use score of -2 for gap; +1 for match; and -1 for mismatch). (6)  
GGCTTGACCGG  
GGATTGACCCG  
Differentiate between NW and SW algorithms in terms of their applications.
3. Describe the application of multiple sequence alignment and differentiate between progressive and iterative methods. (6)

**Section B (Biostatistics)**

**ANSWER THREE OUT OF FOUR QUESTIONS (3 X 8 MARKS = 24 MARKS)**

4. Using Tukey Test on the following data giving selenium (Se) concentrations (mg/ml) in 5 different groundwater wells, determine if mean Se concentrations are same in all 5 groundwater wells. ( $F_{table} = 2.75$ ;  $q_{table} = 4.16$ ) (8)

Well 1	Well 2	Well 3	Well 4	Well 5
28.2	39.6	46.3	41.0	56.3
33.2	40.8	42.1	44.1	54.1
36.4	37.9	43.5	46.4	59.4
34.6	37.1	48.8	40.2	62.7
29.1	43.6	43.7	38.6	60.0
31.0	42.4	40.1	36.3	57.3

5. Protein content of 5 varieties of leafy vegetables estimated in the five regions is given below. Test the variation due to Regions and Varieties. ( $F_{table} = 2.642$ ) (8)

Varieties	Region1	Region2	Region3	Region4	Region5
1.	1.82	1.36	0.9	1.20	1.82
2.	1.37	1.18	1.42	2.22	2.08
3.	1.67	1.26	1.34	1.54	1.35
4.	1.64	1.00	1.04	2.62	1.22
5.	1.18	1.72	1.60	0.74	1.03

6. The data of wing length of thirteen sparrows of various ages are given below. Determine whether using lengths are associated with age. Define the regression equation and coefficient of determination. Test the significance of regression using one-way ANOVA ( $F_{table} = 4.84$ ) (8)

Age (days)	3	4	5	6	8	9	10	11	12	14	15	16	17
Wing length (cm)	1.4	1.5	2.2	2.4	3.1	3.2	3.2	3.9	4.1	4.7	4.5	5.2	5.0

7. The concentrations of two pollutants were determined in a certain urban area ( $\mu\text{g}/\text{m}^3$ ). (8)
- (a) Test the hypothesis that both classes of pollutants were present in the same concentration, using paired sample hypothesis ( $t_{table} = 2.228$ )
- (b) Also test the hypothesis with Wilcoxon paired sample test ( $T_{table} = 10$ )

Day	1	2	3	4	5	6	7	8	9	10	11
Pollutant 1	104	116	84	77	61	84	81	72	61	97	84
Pollutant 2	108	118	89	71	66	83	88	76	68	96	81