D 41693   (Pages : 2)   Name..............................

                                                             Reg. No...........................

THIRD SEMESTER M.Sc. DEGREE EXAMINATION, FEBRUARY 2008

M.Sc General Biotechnology

GBT-211 —BIOSTATISTICS AND BIOINFORMATICS

Tine:: Three Hours   Maximum : 80 Marks

Section A

Answer any **two** *questions.*
Each question carries 10 *marks.*

1. The following are the number of **genomes** of 21 groups of microbes, plants and animals which are available in **genome** database. Compute the mean, standard deviation and variance of these data and infer the result :

    26 28 34 26 25 26 26 30 34 28 25 26 31 25 25 25 25 28 25 25 25.

2. Describe in detail about the MS-Office software used for word processing and spreadsheet.

3. How does the **FASTA** algorithm work on **pairwise** sequence alignment—Explain.

                                                             $(2 \times 10 = 20$ marks)

Section B

Answer any **ten** *questions.*
Each question carries 5 *marks.*

4. Explain the different steps involved in a statistical investigation.

5. Define range and **interquartile** range. Find out the range and **interquartile** range for the following theoretical PI values of 13 protein sequences :

    2.10 1.74 1.68 1.83 1.57 1.71 1.73 1.65 1.74 1.57 2.67 1.90 1.77.

6. From the following data, conclude that there is a relationship between **HPV** status and stage of HIV infection using CM-square test :—

    HPV status and stage of HIV infection among 96 human beings

HPV	Seropositive symptomatic	Seropositive asymptomatic	Seronegative	Total
Positive	23	4	10	37
Negative ...	10	14	35	59
Total	33	18	45	96

7. Describe the relational data model to construct a database.

8. Write a C program to find standard deviation of any 10 observations.

9. Explain the differences between measures of central tendency and measures of dispersion with suitable example.

10. Describe the facilities available in **powerpoint.**

                                                             Turn over

11. Write the significance of Corel Draw.
12. Explain bubble sort with suitable example.
13. Write an algorithm to find the biggest of 3 given numbers.
14. What is Internet ? How does it is important for Bioinformatics study ?
15. What is pairwise alignment ? Explain the matrices used for BLAST algorithm.

(10 x 5 = 50 marks)

Section C

Answer **all** *questions.*
Each question carries 2 *marks.*

16. Mention the *two* types of errors while setting the hypothesis.
17 What is flowchart ? Write its importance in programming development.
18. Explain the input and output statements of QBASIC.
19. What are the uses of regression equation ?
20. For the following data, construct a cumulative frequency distribution :—

Age	20-29	30-39	40-49	50-59	60-69	70-79	80-89	90-99
Number	55	93	113	90	85	73	29	5

(5 x 2 = 10 marks)