2023/08

B.Sc. BIOTECHNOLOGY
FIFTH SEMESTER (SPECIAL REPEAT)
BIOINFORMATICS & BIOSTATISTICS
BBT-504
[USE OMR SHEET FOR OBJECTIVE PART]

SET

A

Duration: 3 hrs.

Full Marks: 70

Time: 30 mins.

Marks: 20

Choose the correct answer from the following:

1×20=20

1. GC hairpin region is not found in:
   a. Promoters
   b. Introns
   c. Poly A site
   d. All

2. Maximal segment pairs of blast are added in?
   a. Linear manner
   b. Exponential manner
   c. Both
   d. None

3. Which of the following average is not well defined?
   a. Mean
   b. Median
   c. Mode
   d. Both b and c

4. In a certain distribution, CV = 60%, mean = 5, then the value of the variance is:
   a. 1.5
   b. 1.73
   c. 3
   d. 9

5. In a Poisson distribution, mean = 4, the standard deviation is:
   a. 4
   b. 2
   c. 0.5
   d. 0.25

6. If X be a normal variate with mean 30 and variance 25, what is the value of the standard normal variate Z for X = 22?
   a. 1.6
   b. 0.32
   c. -1.6
   d. -0.10

7. In a certain test, the null hypothesis $H_0{:}\mu = 23$, find the value of the test statistic. It is given that $\bar{x} = 27, s^2 = 16$ and $n = 15$
   a. Z = 3.87
   b. Z = 0.97
   c. t = 3.87
   d. t = 0.97

8. In analysis of variance, if total SS (TSS) = 387, column SS (SSC) = 299, the value of SSE is:
   a. 88
   b. 686
   c. 1.29
   d. None of the above

9. If the value of one of the regression coefficients is – 2.87, the possible value of the other regression coefficient, is:
   a. 0.65
   b. -0.31
   c. 0.31
   d. -1.23

1

0. If the correlation coefficient between two variables is 0.87 and one of the regression coefficients is 1.13, then the value of the other regression coefficient is:
   a. -0.75
   b. 0.75
   c. -0.67
   d. 0.67

1. Which is the best method to identify an unknown sequence?
   a. blast
   b. fasta
   c. MSA
   d. All

2. Which of the following is a criteria for constructing composite database?
   a. Non redundancy
   b. Non Duplication of data
   c. Up to dated
   d. All

3. Which of the following is sequence retrieval system can be easily customised?
   a. Entrez
   b. SRS
   c. Both
   d. None

4. Which of the following protein databases helps to identify protein families and domains?
   a. Prosite
   b. Pfam
   c. iproclass
   d. All

5. In MSA, adding of homologous sequences are done in:
   a. Progressive way
   b. Linear way
   c. Exponential way
   d. All

6. Which of the following programs is used for predicting genes in unknown DNA?
   a. GENSCAN
   b. Glimmer
   c. Genemark
   d. HMM Gene

7. Which softwares are used for drug designing?
   a. CADD
   b. QSAR
   c. SAR
   d. All

8. Which type of alignment is allowed in BLAST?
   a. Local
   b. Global
   c. Both
   d. None

9. Which of the following branch of bioinformatics is used to study human disease?
   a. Pharmacogenomics
   b. Functional genomics
   c. Both
   d. None

10. Which is not a heuristic method of alignment?
   a. blast
   b. fasta
   c. MSA
   d. All

-- --- --

(Descriptive)

Time : 2 hr. 30 mins.   Marks : 50

[Answer question no.1 & any four (4) from the rest]

1. a) Discuss the relative advantages and disadvantages of the various   6
measures of dispersion.
   b) A bag contains three white and four black balls. Four balls are drawn   4
at random without replacement. What is the probability that the first
ball is white, second ball is black, third ball is white and fourth ball is
black?

2. a) Align the following sequences using BLOSUM   5
Sequence 1: WYPRS
Sequence 2: LVIDK
   b) How bioinformatics evolved as a subject?   5

3. a) Write in brief the importance of normal distribution in Biological   5
Science.
   b) Students of a class were given a mechanical aptitude test. These   5
marks were found to be normally distributed with mean 60 and
standard deviation 5. What percent of students scored between 45
and 65? [Given, Z = 3, Area = 0.9987, Z = 1, Area = 0.8413]

4. If a person is suffering from a type of infection caused by a microbe,   10
what bioinformatics tools and techniques will be used to identify the
microbe at gene and protein level. Explain with examples.

5. a) State the difference between Type I error and Type II error.   5
Which one is more harmful?
   b) In a bivariate data, the mean of X = 32, the mean of Y = 12, the   5
variance of X = 3.8, variance of Y = 4.3 and the correlation coefficient
between X and Y = 0.45. Estimate Y, when X = 24.

6. a) What does the BLAST e value 6e-0.01 indicate? Explain the   5
significance of different colors in graphical representation of BLAST.
   b) How bootstrapping can test the reliability of a result? If you perform   5
MSA and construct phylogenetic tree of 5 sequences A, B, C, D and E,
and get result that A and B are homologous, C and D form a separate
group and E is homologous to C, then draw the phylogenetic tree.

7. a) "Primary data are raw data, whereas, the secondary data are the   6
finished product of raw data"- explain.
   b) In a certain distribution, mean = 80, mode = 76 and coefficient of   4
skewness = -0.45, determine the coefficient of variation.

8. a) State the differences between Genbank, EMBL, DDBJ.   5
   b) Explain a heuristic method of similarity search with the help of a   5
matrix.

= = *** = =