M.Sc. BIOTECHNOLOGY
Third Semester (Repeat)
BIOSTATISTICS, BIOINFORMATICS & IPR
(MBT - 301)

Duration: 3Hrs.

Full Marks: 70

Part-A (Objective) =20
Part-B (Descriptive) =50

(PART-B: Descriptive)

Duration: 2 hrs. 40 mins.

Marks: 50

Answer any *four* from *Question no.* 2 to 8
Question no. 1 is compulsory.

1. What are the different measures of central tendency? What are the desirable properties for arithmetic mean? A clerk calculated arithmetic mean of 50 values as 39.2. However; it was found that instead of taking two values as 25 and 32, he took them as 52 and 23.Find the corrected arithmetic mean. $(3+3+4=10)$

2. Answer the following questions: $(2+4+4=10)$

   (A) Write conditional probability definition.

   (B) If A and B are events with $P(A) = \frac{1}{3}$, $P(B) = \frac{1}{4}$ and $P(A \cup B) = \frac{1}{2}$, find

      (i) $P(A/B)$

      (ii) $P(B/A)$

      (iii) $P(A \cap B)$

      (iv) $P(A/B')$

   (C) A bag contains 10 black and 5 white balls. Two balls are drawn from the bag one after the other without replacement. What is the probability that one black and one white ball will be selected?

3. Briefly describe the different types of sampling techniques with examples. (10)

4. Define multiple sequence alignment? Explain match, mismatch, gap, insertion in a sequence alignment event. $(2+8=10)$

5. What do you mean by biological databases? Explain different nucleic acid sequence database. (2+8=10)

6. What do you mean by molecular phylogeny? Explain the steps of phylogenetic tree construction? Mention the names of different methods of constructing such trees.

(2+5+3=10)

7. What do you understand by patent? How a patent is registered and what benefits do the scientific community or the society as a whole receives from patent protection?

(3+4+3=10)

8. Write short notes on: (*any two*) (5×2=10)

a) Trade secret

b) Traditional knowledge

c) Turmeric case study

*****

M.Sc. BIOTECHNOLOGY
Third Semester (Repeat)
BIOSTATISTICS, BIOINFORMATICS & IPR
(MBT - 301)

Duration: 20 minutes

Marks – 20

(PART A - Objective Type)

I. Choose the correct answer:

$1 \times 20 = 20$

1. Which of the following is a relative measure of dispersion?
   a) Standard deviation
   b) Variance
   c) Coefficient of variation
   d) All of the above

2. Two events A and B are statistically independent when,
   a) $P(A \cap B) = P(A) \times P(B)$
   b) $P(A/B) = P(A)$
   c) $P(A \cup B) = P(A) + P(B)$
   d) Both (a) and (b)

3. A binomial distribution may be approximated by a Poisson distribution provided
   a) n is small and p is large
   b) n is large and p is small
   c) n is large and p is large
   d) n is small and p is small

4. The standard deviation of the binomial distribution is:
   a) np
   b) $\sqrt{np}$
   c) npq
   d) $\sqrt{npq}$

5. All normal distributions are:
   a) Bell – shaped
   b) Symmetrical
   c) Defined by its parameters μ and σ
   d) All of the above

6. Which of the following is non-probability sampling?
   a) Purposive sampling
   b) Random sampling
   c) Cluster sampling
   d) Stratified sampling

7. The sum of squares of deviations from mean is:
   a) Maximum
   b) Minimum
   c) Zero
   d) None of the above

8. Which of the following relationship is true in a multimodal distribution?
   a) Mean-mode = 3(mean-median)
   b) Mode= 3median – 2mean
   c) 3median =(2mean + mode)
   d) All of the above

9. Coefficient of quartile deviation is calculated by the formula:
   a) $\frac{(Q_3 + Q_1)}{4}$
   b) $\frac{(Q_3 + Q_1)}{2}$
   c) $\frac{(Q_3 - Q_1)}{(Q_3 + Q_1)}$
   d) None of the above

10. BLASTx program is used for:
    a) Translate protein sequence
    b) Translate DNA database
    c) Translate input sequence
    d) None of these

11. Which is data retrieving tool?
    a) ENTREZ
    b) EMBL
    c) PHD
    d) All

12. PDB is a
    a) Primary database for macromolecules.
    b) Type of gel electrophoresis.
    c) Composite database.
    d) Database for three dimensional structure of biological macromolecule.

13. Which of the following is correct in respect to global sequence alignment?
    a) Compares sequences and gives best overall alignment.
    b) Will return only the best matching segment.
    c) May fail to find the best local region of similarity.
    d) All.

14. Multiple Sequence Alignment (MSA) is very often used for the following purposes:
    a) Detecting similarities between sequences.
    b) Detecting conserved regions.
    c) Detection of structural homologies.
    d) All.

15. A lead compound is a drug originally discovered by:
    a) Natural products
    b) High-throughput screening
    c) Serendipity
    d) All

16. Which of the following will fall under GI protection of intellectual property right?
    a) Darjeeling tea
    b) Madhubani painting
    c) Tezpur lichi
    d) All of the above

17. Following is an example of trademark.
    a) Logo of adidas
    b) Joha rice of Assam
    c) Secret formula of Coca cola
    d) All of the above

18. Which of the following regarding a patent is true?
    a) Patent protection can be given to both product and process
    b) The facilities provided under patent protection is for the lifetime of the patent holder and also for their successors up to certain time period.
    c) A generic drug is one which has become off patent.
    d) All of the above.

19. Copyrights can be given to:
    a) Software
    b) Scientific invention
    c) Novel
    d) Both a and c

20. Which of the following can be included in the sign of a trademark?
    a) Colour
    b) Design of letter/words
    c) Three dimensional structure
    d) All of the above

*****