Roll No. .... Total Printed Pages - 8

F- 567

M.Sc. (Second Semester)
EXAMINATION, May - June. 2022
BIOTECHNOLOGY
Paper Fifth
Biostatistics, Bioinformatics and Computers in Biotechnology

Time : Three Hours] *[Maximum. Marks:80*

Note: Attempt all sections as directed.

Section - A
(Objective/Multiple Type Questions)

(1 mark each)

Note : Attempt all questions. Choose the correct/ most appropriate answer and write it in your answer book :

1. The sum of all the squared deviations is divided by the total number of observations to calculate

   (A) Population deviation

   (B) Population variance

   (C) Sample deviation

   (D) Sample variance

2. Standard deviation of the following sample data. 2,4,6,8,10, and 12 will be

   (A) 4.42

   (B) 2.42

   (C) 3.42

   (D) 5.42

3. While computing the arithmetic mean of a frequency distribution, each value of a class is considered equal to?

   (A) Lower limit

   (B) Upper limit

   (C) Lower class boundary

   (D) Class mark

4. Cumulative frequency is the frequency........than the upper class boundary of a class.

   (A) Distribution

   (B) Curve

   (C) Greater

   (D) Less

5. ....are used to represent data classified on qualitative characters.

   (A) Vertical bars

   (B) Simple bars

   (C) Horizontal bars

   (D) Pie charts

6. Which Chi square distribution looks the most like a normal distribution?
   - (A) A Chi square distribution with 4 degrees of freedom
   - (B) A Chi square distribution with 5 degrees of freedom
   - (C) A Chi square distribution with 6 degrees of freedom
   - (D) A Chi square distribution with 16 degrees of freedom

7. A hypothesis which defines the population distribution is called:
   - (A) Null hypothesis
   - (B) Simple hypothesis
   - (C) Statistical hypothesis
   - (D) Composite hypothesis

8. In regression analysis, total of the squared residuals is called
   - (A) Coefficient of determination
   - (B) Sum of squares of error
   - (C) Standard error of the estimate
   - (D) r - squared

9. If Y = 2-0.2X, then the value of Y intercept is equal to
   - (A) -0.2
   - (B) 0.2X
   - (C) 2
   - (D) None of these

10. When the k population means are truly different from each other, it is likely that the average error deviation.
    - (A) Is relatively small compared to the average treatment deviations
    - (B) Is about equal to the average treatment deviation
    - (C) Is relatively large compared to the average treatment deviations
    - (D) Is not equal to the average treatment deviation

11. The..........in a document and the order in which it is presented is referred to as the documents logical structure.
    - (A) Data
    - (B) Information
    - (C) Features
    - (D) Functions

12. The symbol for “Insert Worksheet” is located at
    - (A) Top left
    - (B) Top right
    - (C) Bottom left
    - (D) Bottom right

13. What is an external sorting algorithm?
    - (A) Algorithm that uses tape or disk during the sort
    - (B) Algorithm that uses main memory during the sort
    - (C) Algorithm that involves swapping
    - (D) Algorithm that are considered ‘in place’

14. Find the variance of data set: 4.2, 4.3, 4, and 4.1
    - (A) 0.30
    - (B) 0.31
    - (C) 0.27
    - (D) 0.28

15. The change in which of following terms does not affect the standard deviation?
    - (A) Origin
    - (B) Scale
    - (C) Origin and scale
    - (D) Neither origin nor scale

16. Which of the following is not a software for dot plot analysis?
    - (A) DOTMATCHER
    - (B) LALIGN
    - (C) DOTLET
    - (D) SIMMI

17. Margaret Dayhoff developed the first protein sequence database called:
    - (A) SWISS PROT
    - (B) PDB
    - (C) Atlas of protein sequence and structure
    - (D) Protein sequence databank

18. The alignment method suitable for finding out conserved patterns in protein or DNA sequences is:
    - (A) Multiple sequence alignment
    - (B) Local alignment
    - (C) Pair wise alignment
    - (D) Global alignment

19. BankIt and Sequin are sequence submission tools in
    - (A) DDBJ
    - (B) PDB
    - (C) GenBank
    - (D) EMBL

20. The human genome contains approximately:
    - (A) 6 billion base pairs
    - (B) 5 billion base pairs
    - (C) 4 billion base pairs
    - (D) 3 billion base pairs

Section - B

(Very Short Answer Type Questions)

(2 marks each)

Note: Attempt all questions. Answer using 2-3 sentences.

1. What do you mean by multiple sequence alignment?
2. What is SWISS PROT?
3. What are application software?

4. What is spreadsheet?
5. What we do in matched pair t-test?
6. What is degree of freedom?
7. Why most of the statistical analyses are done at 95% confidence limit?
8. What do you understand by positively skewed data?

Section - C

(Short Answer Type Questions)

(3 marks each)

Note : Attempt all questions. Answer precisely < 75 words.

1. How BTIS is helpful in conducting and planning of biotechnological research?
2. What do you mean by database mining? Discuss uses of it in medical biotechnology.
3. What do you mean by word processing? Mention applications of it.
4. Below data provides age in years in case of child deaths. Find out the average age of child.

Age in Years	0	1	2	3	4	5
Number of deaths	42	55	32	22	15	06

5. An IQ test was conducted on 5 MRs before and after they were trained. The results are given below.

MRs	1	2	3	4	5
Before training	110	120	123	132	125
After training	120	118	125	136	121

Test whether there is any change in IQ after the training program?

6. What we do in regression analysis? Discuss giving a suitable example.
7. Write a note on idea of two types of errors in statistical analysis.
8. Discuss different modes popularly used for presentation of research data.

Section - D

(Long Answer Type Questions)

(each 5 marks)

Note : Attempt all questions. Answer precisely using 150 words.

1. Discuss entire process of searching biological databases for similar/matching sequence.
2. Write down an exhaustive note on bubble sorting and its applications.
3. What do you understand by correlation analysis and what are different types of it? Explain procedures followed for its calculation.
4. Calculate standard deviation of the below given data.

Age in Years	7	8	9	10	11	12
Number of Children	13	13	18	17	15	14