

Your Roll No.....

Sr. No. of Question Paper : 5023
Unique Paper Code : 2164001206
Name of the Paper : Informatics and statistics for biology and allied sciences,
Name of Course : Botany (GE)
Semester : Sem VIII

Duration : 2 Hours

Maximum Marks : 60

Instructions for Candidates

- Write your Roll No. on the top immediately on receipt of this question paper.
- Attempt Part A and Part B on **separate answer sheets**.
- Question no. 1 is compulsory** of both the sections
- Attempt any **three** questions from Part A and **three** questions from Part B including compulsory Q. No.1 of both the sections.
- Attempt **All** parts of a question together
- Use of simple calculator allowed

Part A

1. (a). Expand the following (**any five**)

5×1=5

- BLAST
- KEGG
- HGP
- OTU
- MEGA
- PDB
- INSDC

(b). Fill in the blanks (**any five**)

5×1=5

- The mother and father of bioinformatics is _____.
- A _____ in a sequence alignment stands for an insertion or deletion of letters.
- Metabolomics is the scientific study of _____, which are the small molecules left behind by cellular processes.
- In the UniProt database, the _____ section contains high-quality entries, reviewed by experts.
- The first bioinformatics database was _____.
- _____ provides search and analysis tools for its extensive and growing collection of chemical data.
- The first bacterial genome fully sequenced was _____.

2. Write the differences between the following (**any two**)

2×5=10

- i. Primary and secondary databases
- ii. Transcriptomics and proteomics
- iii. Maximum parsimony and maximum likelihood

3. Write short notes on following (**any two**)

2×5=10

- i. GenBank
- ii. System biology
- iii. Sequence alignment

4. What is molecular phylogeny? Write its applications and limitations in biological research. Construct a hypothetical phylogram and label different important components in it.

2+5+3=10

Part B

1. a. Fill in the blanks (**any five**)

5×1=5

- i. Arrangement of statistical data into systematic rows and columns is called _____.
- ii. Systematic, stratified, and cluster sampling are all examples of _____ sampling methods.
- iii. The value that occurs most frequently in a data distribution is known as the _____.
- iv. Graphical representation of a continuous frequency distribution is best done using a _____.
- v. The square of the correlation coefficient is technically termed as the coefficient of _____.
- vi. The specific measure of central tendency that is most sensitive to, and heavily influenced by, extreme outlier values in a dataset is the _____.
- vii. Data collected by questionnaires and surveys are a type of _____.
- viii. _____ is an example of Inferential Statistics.

b. State True or False (**any five**)

5×1=5

- i. A sample must always be a true representative of the target population to avoid sampling bias.
- ii. A Type I error occurs when a statistical test incorrectly rejects a true null hypothesis.
- iii. Range is a highly stable measure of dispersion because it is unaffected by extreme outlier values.
- iv. When data is heavily concentrated at one extreme end of a distribution scale, it results in a perfectly normal distribution.
- v. Mean deviation is an absolute measure of dispersion calculated relative to a central value.
- vi. In a perfectly symmetrical, bell-shaped normal distribution, the values for the mean, median, and mode are all exactly equal to each other.
- vii. Mode always lies in between median and mean

2. Differentiate between (**any two**) 2×5=10

- i. Null Hypothesis vs. Alternative Hypothesis
- ii. Correlation Analysis vs. Regression Analysis
- iii. Histogram and bar diagram

3. Write short notes on the following with suitable examples (**any two**) 2×5=10

- i. Spearman's Rank Correlation Coefficient
- ii. Quartile Deviation and its merits over Range
- iii. Standard Error
- iv. Applications of biostatistics

4. Define the Chi-square test as used in biological experiments. In Maize (*Zea mays*), purple kernel color is dominant over yellow kernel color, and smooth/starchy texture is dominant over wrinkled/sweet texture. A dihybrid cross yielded an F₂ generation with the following observed counts:

Kernel Phenotype	Observed Number
Purple, Smooth	442
Purple, Wrinkled	158
Yellow, Smooth	143
Yellow, Wrinkled	57

Calculate chi square value and interpret the results. 2+6+2= 10