

Total No. of Questions : 5]

SEAT No. :

PE-13461

[Total No. of Pages : 2

[6517]-1001
F.Y. M.Sc.
BIOINFORMATICS
BIN 101 (T) : Biological Data and Databases
(2024 Pattern) (Semester - I)

Time : 2 Hours]

[Max. Marks : 35

Instructions to the candidates :

- 1) *All questions are compulsory*
- 2) *The numbers in the right indicates full marks.*

Q1) Attempt Any 4 (Out of 6) :

[8]

- a) Define primary biological database.
- b) Write full forms of EBI, ExpASy, UniParc and miRBase.
- c) What is interPro?
- d) List any four protein sequence databases.
- e) What is NCBI Taxonomy database?
- f) Write a note on ProSite as a constituent database.

Q2) Attempt Any 2 (Out of 4) :

[8]

- a) Explain the organization and structure of EMBL nucleotide database.
- b) Describe the nature and classification of biological data.
- c) Write a detailed note on NCBI resources and services.
- d) Explain the importance of biological literature databases with reference to PubMed.

Q3) Attempt Any 2 (Out of 4) :

[8]

- a) Explain the UniProtKB database and its annotation system.
- b) Describe the UniRef, UniParc and Proteomes databases.
- c) Write a note on RefSeq database and its significance.
- d) Explain the ExpASy bioinformatics server and its utilities.

P.T.O.

Q4) Attempt Any 2 (Out of 4) :

[8]

- a) Describe the interPro database and its role in protein family classification.
- b) Write a note on RNACentral database.
- c) Explain the MIT/ICBP siRNA database.
- d) Write a short note on GBIF biodiversity database.

Q5) Attempt Any 1 (Out of 2) :

[3]

- a) Differentiate between GenBank and RefSeq databases.
- b) Differentiate between UniRef and UniParc databases.



Total No. of Questions : 5]

SEAT No. :

PE-13462

[Total No. of Pages : 2

[6517]-1002

F.Y. M.Sc.

BIOINFORMATICS

BIN102 : Biological Sequence Analysis

(2024 Pattern) (Semester - I)

Time : 2 Hours]

[Max. Marks : 35

Instructions to the candidates:

- 1) All questions are compulsory.
- 2) Figures to the right indicate full marks.

Q1) Attempt Any 4 (Out of 6) :

[8]

- a) Define biomolecular sequence analysis.
- b) What is meant by pairwise sequence alignment?
- c) Write any two differences between global and local alignment.
- d) What is a scoring matrix in sequence alignment?
- e) What is the purpose of BLAST search?
- f) Define sequence logo.

Q2) Attempt Any 2 (Out of 4) :

[8]

- a) Give an overview of biomolecular sequence analysis and its applications in bioinformatics.
- b) Explain methods used for analysis of a single nucleotide and protein sequence.
- c) Describe the Needleman-Wunsch algorithm for global sequence alignment.
- d) Explain the Smith-Waterman algorithm for local sequence alignment.

P.T.O.

Q3) Attempt Any 2 (Out of 4) :

[8]

- a) Explain the PAM and BLOSUM scoring matrices and their importance in sequence comparison.
- b) Describe the FASTA algorithm for database similarity searching.
- c) Explain the BLAST algorithm and its variants.
- d) Write a note on the CLUSTALW multiple sequence alignment algorithm.

Q4) Attempt Any 2 (Out of 4) :

[8]

- a) Describe the MUSCLE algorithm for multiple sequence alignment.
- b) Explain T-Coffee algorithm and its advantages over other MSA methods.
- c) Write a note on sequence logos, consensus sequences and sequence patterns.
- d) Explain the concept of sequence profiles and their applications.

Q5) Attempt Any 1 (Out of 2) :

[3]

- a) Differentiate between FASTA and BLAST algorithms.
- b) Differentiate between global alignment and local alignment.



Total No. of Questions : 5]

SEAT No. :

PE-13465

[Total No. of Pages : 2

[6517]-1003
F.Y. M.Sc.
BIOINFORMATICS
BIN 103 : Biomolecular Structure & Organization
(2024 Pattern) (Semester - I)

Time : 2 Hours]

[Max. Marks : 35

Instructions to the candidates :

- 1) *All questions are compulsory.*
- 2) *Numbers on the right indicates full marks.*

Q1) Attempt Any 4 (Out of 6) :

[8]

- a) Define :
 - i) Covalent bond,
 - ii) Hydrogen bond.
- b) What is a peptide unit?
- c) What are supersecondary structures in proteins?
- d) What is cruciform DNA structure?
- e) What are membrane microdomains?
- f) Write a note on membrane protein structure.

Q2) Attempt Any 2 (Out of 4) :

[8]

- a) Explain rectangular coordinate systems used in structural biology.
- b) Discuss Ramachandran plot and its significance.
- c) Describe the mechanism of protein folding.
- d) Explain the principle of NMR spectroscopy.

Q3) Attempt Any 2 (Out of 4) :

[8]

- a) Describe different types of DNA double helices.
- b) Explain energetics of RNA structure.
- c) Discuss base pairing in DNA and RNA.
- d) Describe triple helical DNA structures.

P.T.O.

Q4) Attempt Any 2 (Out of 4) :

[8]

- a) Explain carbohydrate conformations and ring structures.
- b) Discuss glycoproteins and glycolipids with, structural aspects.
- c) Discuss membrane organization and lipid bilayer structure.
- d) Explain methods used to study protein folding.

Q5) Attempt Any 1 (Out of 2) :

[3]

- a) Differentiate between tertiary and quaternary protein structure.
- b) Differentiate between RNA secondary structure and DNA secondary structure.



Total No. of Questions : 3]

SEAT No. :

PE13466

[Total No. of Pages : 2

[6517]-1004

F.Y.M.Sc. (Bioinformatics)

BIE-106: MATHEMATICS & STATISTICS

(2024 Pattern) (Semester - I)

Time : 2 Hours]

[Max. Marks : 35

Instructions to the candidates:

- 1) *All questions are compulsory.*
- 2) *Figures to the right indicate full marks.*

Q1) Attempt Any 6 (Out of 8)

[12]

- a) Solve the equation $4y'' + 12y' + 9y = 0$
- b) Find A^{-1} of matrix A

$$A = \begin{bmatrix} 3 & -3 & 4 \\ 2 & -3 & 4 \\ 0 & -1 & 1 \end{bmatrix}$$

- c) Define difference of set, complement of set, infinite set and subset.
- d) Explain first-order and second-order ordinary differential equations.
- e) Define variance and coefficient of variation.
- f) What is meant by skewness of a distribution?
- g) In a DNA sequencing experiment, the probability of detecting a mutation in a sample is 0.2. If 5 samples are tested, calculate the probability that exactly 2 samples show mutation.
- h) Explain the concept of hypothesis testing.

Q2) Attempt Any 5 (Out of 7)

[15]

- a) Solve $3yy' = \frac{1+3x^2}{y-2}$, $y(0)=1$ by separating the variables.
- b) Evaluate the following integral $\int x\sqrt{2x+1}dx$.
- c) Explain the concepts of limit and continuity of a function with suitable examples.
- d) Explain the concepts of correlation and regression. Describe the method of fitting a regression line.

P.T.O.

- e) In a population of protein lengths, the mean length is 300 amino acids with a standard deviation of 20. Calculate the Z-score for a protein with length 340 amino acids.
- f) In a genetic study, the probability of a person having a certain gene mutation is 0.3 and the probability that the mutation causes disease is 0.4. Find the probability that a person both has the mutation and develops the disease assuming independence.
- g) Write short note on Chi-Square test.

Q3) Attempt Any 2 (Out of 4).

[8]

- a) Solve the following first-order ordinary differential equation using the method of separation of variables: $\frac{dy}{dx} = \frac{x^2}{y}$, given the initial condition $y(0)=2$.
- b) Find the Laplace transform of the function $f(t)=t^2 e^{-3t}$.
- c) An experiment yields the following data points for drug dosage (X) and patient response (Y):
 X: [1, 2, 3, 4, 5]
 Y: [3, 5, 7, 8, 11]
 - i) Calculate the Pearson correlation coefficient (r).
 - ii) Fit a linear regression line $Y = a + bX$ to this data.
- d) A researcher tests a new drug meant to reduce cholesterol. A sample of 30 patients shows an average reduction of 18 mg/dL with a standard deviation of 6 mg/dL.
 - i) Formulate the null (H_0) and alternative (H_a) hypotheses.
 - ii) Conduct a Z-test (or t-test) at a 5% level of significance ($\alpha = 0.05$) to determine if the drug is effective. (Assume the critical Z-value for a one-tailed test is 1.645).



Total No. of Questions : 5]

SEAT No. :

PE-13467

[Total No. of Pages : 2

[6517]-1005
F.Y. M.Sc. (Bioinformatics)
BIE 108: MOLECULAR BIOLOGY
(2024 Pattern) (Semester - I)

Time : 2 Hours]

[Max. Marks : 35

Instructions to the candidates :

- 1) *All questions are compulsory.*
- 2) *Numbers on the right indicate full marks.*

Q1) Attempt Any 4 (Out of 6) :

[8]

- 1) What is circular DNA?
- 2) What are exons and introns?
- 3) What is recombination repair?
- 4) What is genome editing?
- 5) Define transposon and retroposon.
- 6) Define operon with example.

Q2) Attempt Any 2 (Out of 4) :

[8]

- 1) Explain globin gene family and its significance.
- 2) Describe nucleoid organization in bacteria.
- 3) Explain organization of chloroplast genome.
- 4) Discuss gene clusters and pseudogenes.

Q3) Attempt Any 2 (Out of 4) :

[8]

- 1) Explain DNA replication in eukaryotes.
- 2) Describe mismatch repair mechanism.
- 3) Explain non-homologous recombination.
- 4) Discuss chemical agents causing DNA damage.

P.T.O.

Q4) Attempt Any 2 (Out of 4) :

[8]

- 1) Explain transcription initiation and termination in eukaryotes.
- 2) Describe RNA splicing and editing.
- 3) Explain elongation and termination in translation.
- 4) Discuss post-translational modifications of proteins.

Q5) Attempt Any 1 (Out of 2) :

[3]

- 1) Differentiate between homologous and non-homologous recombination.
- 2) Differentiate between snRNA and snoRNA.



Total No. of Questions : 7]

SEAT No. :

PE-13468

[Total No. of Pages : 4

[6517]-1006

F.Y. M.Sc. (Bioinformatics)
BIN-105 : Research Methodology
(2024 Pattern) (Semester - I)

Time : 3 Hours]

[Max. Marks : 70

Instructions to the candidates:

- 1) *All questions are compulsory.*
- 2) *The numbers in the right indicates full marks.*

Q1) Multiple Choice Questions :

[10]

- i) The scientific method generally begins with:
 - a) Data analysis
 - b) Observation of a phenomenon
 - c) Publication
 - d) Hypothesis rejection
- ii) Skepticism in research encourages:
 - a) Blind acceptance
 - b) Questioning of evidence
 - c) Avoiding experimentation
 - d) Ignoring data
- iii) Deductive reasoning moves from:
 - a) Specific to general
 - b) General to specific
 - c) Data to hypothesis
 - d) Observation to assumption
- iv) Replication in research refers to:
 - a) Repeating experiments
 - b) Writing reports
 - c) Literature review
 - d) Data collection
- v) Correlation analysis is used to:
 - a) Summarize qualitative data
 - b) Measure strength of association between variables
 - c) Test cause-effect relationship
 - d) Select sample randomly

P.T.O.

- vi) MS Office tools are primarily used for:
 - a) Writing and formatting reports
 - b) Data analysis
 - c) Sampling
 - d) Statistical modeling
- vii) GEO database mainly stores:
 - a) Gene expression datasets
 - b) Protein structures
 - c) Genomic DNA sequences
 - d) Chemical structures
- viii) Bibliography in research includes:
 - a) List of references
 - b) Abstract
 - c) Methods
 - d) Results
- ix) Self-plagiarism refers to:
 - a) Copying others work
 - b) Copying own previously published work without citation
 - c) Referencing properly
 - d) Referencing with software tool
- x) Ethical committees review:
 - a) Research proposals for ethical compliance
 - b) Sampling methods
 - c) Data analysis
 - d) Literature review

Q2) Match the Pairs :

[5]

<u>Column A</u>	<u>Column B</u>
a) Empiricism	i) manipulated by the researcher
b) Independent	ii) knowledge derived from sensory experience and variable observation
c) Nominal scale	iii) divides the data into homogeneous subgroups
d) PDB	iv) categories without any order
e) Stratified sampling	v) protein structural data

Q3) Fill in the blanks/Answer in one word : [5]

- a) A _____ refers to entire group under study.
- b) A _____ is a comprehensive document submitted by a student detailing research findings.
- c) Normal distribution is _____ distribution.
- d) IPR stands for _____.
- e) _____ committees review research proposals for ethical compliance.

Q4) Answer in short (Attempt any 5) : [10]

- a) Discuss rationalism in research.
- b) What is hypothesis? Give one example.
- c) Explain inductive and deductive approaches in research.
- d) Explain the concepts of replication in research.
- e) Discuss the importance of paraphrasing in scientific writing.
- f) Differentiate between primary and secondary literature.
- g) Discuss the significance of correlation analysis.
- h) What are the different types of publications?

Q5) Write short note on (Attempt any 4) : [20]

- a) Discuss exploratory, descriptive and experimental research designs.
- b) Explain the concepts of measurement, causality, generalization and replication in research.
- c) Describe probability distributions used in biological research, including binomial, Poisson, normal and exponential distributions.
- d) Discuss the process of writing a research proposal for funding.

Q6) Attempt the following questions : [10]

- a) Discuss the criteria for selecting a suitable journal for publication, including impact factor, scope, indexing and audience.

OR

- a) Explain the concept of Intellectual Property Rights (IPR). Discuss royalties and commercialization of research.
- b) Explain the concept of sampling in research. Discuss statistical population, sample, sampling frame, sampling error and sample size determination.

OR

- b) Discuss the levels of measurement in research (nominal, ordinal, interval and ratio scales) and explain their importance in scientific data analysis.

Q7) Explain the following (Attempt any 2) :

[10]

- a) Describe the philosophical foundations of research, including Rationalism, Empiricism, Skepticism, Reductionism and Pseudoscience. Explain their relevance in modern research.
- b) Explain different methods of data analysis and interpretation, including graphical representation, univariate analysis and bivariate analysis.
- c) Discuss research ethics and explain issues related to plagiarism, self-plagiarism, citation practices, reproducibility, and accountability in research.

