

BACHELOR OF SCIENCE

6th SEMESTER

DISCIPLINE SPECIFIC ELECTIVES (DSEs)

BT620D1: BIO-TECHNOLOGY: ELEMENTARY BIOSTATISTICS AND BIOINFORMATICS

CREDITS: THEORY – 4, PRACTICAL – 2(4+2)

OPTION-I

THEORY (4 CREDITS: 60 HOURS)

MAXIMUM MARKS: 60, MINIMUM MARKS: 24

Objective: This course introduces students to basic statistical concepts involved in biology and illustrates the power of computing in modern biology.

Unit-1 (15 HOURS)

Introduction to statistics; Understanding of data & variables (with their types and categories); Data production – experiments vs sample surveys, principles & types of experimental design, idea of randomization, detailed account of sampling designs; Graphical representation of data (bar graph, pie chart, stemplot, histogram).

Unit – 2 (15 Hours)

Measures of central tendency (mean, median, mode) & dispersion (quartiles, standard deviation) with their properties and comparison; Understanding of correlation, least-squares regression & scatterplots; Overview of probability & probability rules, statistical inference with emphasis on confidence intervals and p-values.

Unit – 3 (15 Hours)

Introduction to bioinformatics; Scope and application of bioinformatics; Introduction to biological databases (types-sequence, structure & pathway), Nucleic acid databases (NCBI, GenBank, EMBL), Protein databases (PIR, Swiss-Prot, PDB); Introduction to PubMed.

Unit – 4 (15 Hours)

Sequence similarity and alignment – local & global alignment, pairwise & multiple sequence alignments, BLAST, FASTA & CLUSTALW; Basic idea of phylogenetic tree; Protein structure analysis - levels of protein structure, primary structure analysis (protparam), secondary structure predictions (ExPASy, JPred), tertiary structure prediction methods (homology, threading).

PRACTICALS (2 CREDITS: 60 HOURS) MAXIMUM MARKS: 30, MINIMUM MARKS: 12

1. Use of excel for calculating: Mean, Mode, Median.
2. Use of excel for drawing, histogram, bar-chart & piechart.
3. Use of NCBI, GenBank, EMBL, SwissProt, PDB, TREMBL.
4. Pairwise and multiple sequence alignment (BLAST and ClustalW)
5. Use of protparam, Expasy and JPred.

BOOKS RECOMMENDED

1. *Basic Biostatistics*: Bert Gurtzman
2. *Bioinformatics: Methods and Applications: Genomics, Proteomics and Drug Discovery*, -Rastogi, Prentice Hall India Learning Private Limited.
3. *Essential Bioinformatics*, Jin Xiong, - Cambridge University Press.
4. *Bioinformatics – Principles and Applications*, Ghosh, Z. and Mallick, B., -Oxford University Press (India).

Expected Learning Outcomes:

1. Understanding of basic statistical methods as applied to biological sciences.
2. Concept of Bioinformatics, types of data and databases.
3. Understanding of tools used for data analysis and prediction of different levels of protein structure.