

BACHELOR WITH BIOTECHNOLOGY AS MAJOR

5TH SEMESTER

BTG 522J3: Biostatistics and Bioinformatics

CREDITS: THEORY – 4, INTERNSHIP – 2

- **Course Learning Objective:** Through this course, students will learn about the applications of statistics in life sciences and about essentials of bioinformatics.
- **Expected Learning Outcomes:** At the end of the course students should be able to;
 - ☐ collect, represent and compute different parameters of data.
 - ☐ establish the relationship between different data variables.
 - ☐ use basic bioinformatics tools.
 - ☐ predict protein structure and make phylogenetic trees.

Unit – 1 (15 HOURS)

Sample, Population, sampling techniques (random and non-random). Data collection and its representation (Histogram, Bar Chart, Pie chart, Frequency curve). Mean, Median, Mode and their comparison. Dispersion (mean deviation, standard deviation and variance, coefficient of variance).

Unit – 2 (15 HOURS)

Sampling distribution. Confidence intervals. Test of significance: standard error of mean, p-value & statistical significance, students' 't' test, paired and unpaired t test. Chi square test. F-test and analysis of variance (null hypothesis, analysis of variance 'ANOVA').

Unit – 3 (15 HOURS)

Introduction to bioinformatics and its applications. Introduction to biological databases (types: sequence, structure, pathway & disease database). Nucleic acid databases (NCBI, GenBank, EMBL). Protein databases (PIR, Swiss-Prot, PDB). Sequence similarity, identity and homology. Alignment: local and global alignment, pairwise and multiple sequence alignments, BLAST, FASTA and CLUSTALW. Introduction of PubMed.

Unit - 4 (15 hours)

Reading frames and identification of reading frames. Basics of genome annotation and gene identification tools (NCBI genome). Basic idea of phylogenetic trees. Protein structure analysis: importance of protein structure determination/prediction. primary structure analysis (protparam), secondary structure prediction (JPred, PROSECSC). Introduction to protein

motifs and domains (prosite). Tertiary structure prediction methods: homology, threading and ab initio. Steps of homology modeling (swiss-model and BHAGEERATH).

<http://www.scfbio-iitd.res.in/bioinformatics/bioinformaticssoftware.htm>

INTERNSHIP (2 CREDITS)

BOOKS RECOMMENDED

1. W.W Daniel, Biostatistics: A foundation for analysis in the Health Sciences, JohnWiley and Sons.
2. VB Rastogi, Biostatistics, Medtech
3. Kulkarni, A.P., Basics of Biostatistics, CBS Publishers & Distributors
4. Rastogi, Bioinformatics: Methods and Applications: Genomics, Proteomics and Drug Discovery, Prentice Hall India Learning Private Limited.
5. Jin Xiong, Essential Bioinformatics, Cambridge University Press.
6. Ghosh, Z. and Mallick, B, Bioinformatics – Principles and Applications, Oxford University Press.