Department Name: Biotechnology

Semester – IV

Course Title: Bioinformatics and Biostatistics	Course Code: 24MJBTEC4S
Total Contact Hours: 30 Hrs	No. of Credits: 02
L:T:P: 1:0:2	
Internal Assessment Marks: 10	Duration of SEE: 1hr, 30 Min
Semester End Exam Marks: 40	

Course Outcomes (Cos):

At the end of the course, students will be able to:

1. Remember, understand and apply the skills gained, in basic concepts of bioinformatics including databases, database searches and role of internet in bioinformatics.
2. Understand, analyze and evaluate methods to characterize and manage the different types of biological data and gain an insight into the basics of sequence alignment and analysis.
3. Remember, understand and apply basic concepts in biostatistics exemplifying sampling methods, graphical representation of data, measuring central tendencies and dispersion
4. Understand, apply and evaluate hypothesis testing using statistical methods for analyzing one or two variables, interpret and explain a p-value, perform a two-sample t-test and interpret the results.

Unit	Description	Hours
1	**Bioinformatics definition, history, scope and applications** **Bioinformatics web portals:** NCBI, EBI, ExPASy **Biological databases:** Classification of databases – primary (Genbank), secondary (PIR) and tertiary or composite (KEGG) databases **Sequence databases** – DNA sequence databases (ENA,DDBJ) **Protein sequence databases** (Swissprot, PROSITE) **Basics of sequence alignment –** match, mismatch, gaps, gap penalties, scoring alignment **Types of sequence alignment –** pairwise and multiple alignment, local and global alignment **Dot matrix comparison of sequences** **Concepts of phylogeny –** Distance based (NJ method) and character based (ML method) tree construction methods	15Hrs
2	**Introduction to Biostatistics, kinds of data and variables –** based on nature (numerical discreetee and continuous, categorical-ordinal and nominal) – based on source (primary and secondary data), sample size, sampling methods and sampling errors	15Hrs

	Data tabulation and representation methods: graphical methods– stem and leaf plot, line diagram, bar graphs, histogram, frequency polygon, frequency curves; diagrammatic method- pie diagram **Measures of central tendency**- mean, median, mode; merits and demerits **Measures of dispersion**- range, variance, standard deviation, standard error and coefficient of variation; merits and demerits Correlation and regression analysis and their applications to biology **Hypothesis testing**- steps in testing for statistical hypothesis, null and alternative hypothesis, level of significance- type-1 and type-2 errors **Test of significance for large samples-** Z-test for means and proportions **Test of significance for small samples**- student's t-test(one sample and two samples) **Chi-square test and its applications**- goodness of fit (not based on distribution), test of independence **Analysis of variance (One-way ANOVA)** and their applications to biology	

References:

1. Khan &Khanum (2004), Fundamentals of Biostatistics, II Revised Edition, UkaazPublication
2. Bailey, N.T.J ,Statistical methods in Biology, Cambridge Univ.Press
3. Fundamentals of Biostatistics, P Hanmanth Rao and K.Janardhan
4. Danial, W. W ,Biostatistics, Wiley
5. Introduction to Bioinformatics By Aurther M lesk
6. Developing Bioinformatics Computer Skills By: Cynthia Gibas, Per Jambeck
7. Bioinformatics second edition By David Mmount
8. Essential Bioinformatics By JinXiong
9. Bioinformatics Computing By Bryan Bergeron
10. Bioinformatics: Concepts, Skills & Applications By R.S.Rastogi

*****