ZOS555- STATISTICS AND BIOINFORMATICS
Teaching Hours 10 /Unit

COURSE OUTCOME

1. It is an important paper and students are trained to use different statistical tools applicable to biological research and field work.
2. Terminologies used, data collection, tabulation, graphical representation of data are dealt.
3. Data analysis, types of data analysis and errors, accuracy, methods for large sample analysis are discussed.
4. Some of the recent tools and various soft wares used for statistical analysis are discussed.
5. Information on databases for protein and genome analysis, networking and bioinformatics tools are also introduced.
6. Skill based course that grooms students in application of statistics and bioinformatics in data analysis. On successful completion they can take up data analysis assignments

They learn data handling skills such as recording, collating and analysing data using appropriate techniques and equipment.

UNIT I
Basic Concepts & Descriptive Statistics: Biostatistics terminology, variables in biology, Levels and measurements of biological data, Classification, tabulation and frequency distribution of the data, graphical representation of data by histogram, Polygon, Ogive curve and pia diagram. Measures of central tendency, measures of dispersion, Comparison of two CVs; Skewness- Kurtosis.

UNIT II
Correlation & Regression: Positive and Negative correlation and calculation of Karl-Peasrson's Co-efficient of correlation, Spearman's rank correlation, Partial and multiple correlation, regression analysis; Sample linear and non-linear regression; multiple regression, regression equation, Calculation of an unknown variable using regression equation, Confidence interval level of confidence.

UNIT III
Errors in measurements & Statistical Analysis: Errors, Accuracy, Precision, general theory of Errors, Classification, standard errors. Ways of expression of precision, Accuracy detection of determinates errors, Statistic analysis of biochemical data with spread sheet applications, Use of statistical packages, Data management with computer Basic idea of significance test – Hypothesis testing. , Null and alternative hypothesis; Large sample tests (z-test); Test of significance of single and two sample means; Testing of single and two proportions - Small sample tests: F-test — testing of single mean; Testing of two sample means using

independent t test, paired t test; ANOVA and Chi-Square Tests: One-way and two-way ANOVA – Latin Square tests for association and goodness of fit; testing linkage; segregation ratio.

UNIT IV

Information theory and Bioinformatics Network: Biological data exploration through internet Resources– EMB net, NCBI, BTIS network, Bioinformatics landscape intrinsic & extrinsic view, Cheminformatics and medical informatics. Biological databases sequence databases, Protein sequence databases, Structural databases, PDBs, Motif databases, Protein motif database, Genome databases, Proteome databases etc.

UNIT V

Bioinformatics tools: Pair wise Alignment, Alignment algorithms, sequence analysis tools, BLAST (Basic Logical Alignment Search Tool) FASTA, Multiple Alignment, Sequence analysis, using EMBOSS, DNA micro array technique.

REFERENCES

1. Arora, P. W., Malhan P.K. (2002) Biostatistics, Himalaya's pub. House, Mumbai.
2. Bailey, (2000), Statistical Method in biology.
3. Bliss, C. I.K. (1967), Statistics in biology vol. 1 Mac-Graw Hill.
4. Campbell, R.C. (1974) Statistics for biologist, Cambridge University Press.
5. Daniel, W. W., Biostatistics (A foundations for analysis in health sciences).
6. Dixit, M. (2000) Internet an Introduction, Tata McGraw-Hill.
7. Khan, I. A and Khanum A. (1994) Fundamental of Biostatistics.Ukaaz publications.
8. Lachin, J.M. (2010) Biostatistical Method: the assessment of relative risks, 2nd edition, Wiley online publications, ISBN: 978-0-470-50822-0.
9. SurnderR. P. S. S. and Richard J. (1996) An introduction to Biostatistics, Prentice Hall of India.
10. Wardlaw, A.C (1985), Practical Statistics for Experimental biologist.