

PAPER IV: BIOINFORMATICS AND BIOSTATISTICS

Unit I: Computer Architecture, Internal and External devices, computer software, operating system windows, Unix, Application software like word processor, spread sheet, Database, RDBMS. Computer Network- Advantages of network, types of network (LAN, WAN & MAN), Network protocols, Internet protocol (TCP, IP), File transfer protocol (FTP), WWW, HTTP, HTML, VRL.

Unit II: Computer words coding (ASCII and EBCDIC), Numeric data. Introduction to programming languages, C⁺⁺ Perl. Databases: Introduction to databases- Relational databases- Oracle, SQL, Database generation, Sequence databases- Resources-Human Genome project (HGP), Microbial genomes, Structural databases- protein data bank (PDB).

Unit III: Principles behind computational analysis, Sequence analysis, sequence alignment, scoring matrices for sequence alignment, Restriction mapping, Similarity searching (FASTA and BLAST), Pair wise comparison of sequences, Multiple alignment of sequences.

Unit IV: Brief description of tabulation of data and their graphical representation, measures of central tendency and dispersion: Mean, median, mode, range, standard deviation, variance. Simple linear regression and correlation. Brief idea of statistical softwares and their applications.

Unit V: Elementary idea of probability, definition and properties of binomial, poisson and normal distributions. Elementary idea of random sampling, selection of simple random samples from a finite population, definition of sampling distribution, Randomized block design, sampling variance and standard error. Analysis of Variance (ANOVA), Idea of two types of errors and level of significance, test of significance, chi-square test of independence and homogeneity test based on Z and T statistics.