

Probability and Statistics Workshop Overview

Part I: Probability, Counting Techniques and Bayes' Theorem

Topics covered:

- Framework for probability
- Axioms and basic theorems for computing probabilities
- Basic counting techniques
- Combinations and permutations
- Independence and dependence of events
- Conditional probability
- Law of Total Probability and Bayes' Theorem
- Application of Bayes' Theorem in Statistics (prior and posterior probabilities)

Learning Objectives:

Students will gain an understanding of basic rules for computing probabilities and gain insight into the difference between independent and dependent events, and how Bayes' Theorem can be used to compute certain types of conditional probabilities. Finally, students will learn how Bayes' Theorem is used in statistics to dynamically adjust probabilities for certain events based on the observation of data.

Part II: Markov Chains

Topics covered:

- Random variable and stochastic processes
- Definition of Markov chains
- Representations of Markov chains (transition matrix and graph)
- Computation of path and state probabilities
- Stationary distribution

Learning Objectives:

Students will gain a basic understanding of Markov chains and be able to determine the transition matrix and the graph for a given system. Students will learn how to compute probabilities for certain events, such as occurrence of a particular sequence of states, or the distribution for states at a given point in time, utilizing a computer algebra system (e.g. Matlab).

Part III: Hidden Markov Models and Significance of Alignment

Topics covered:

- Definition of Hidden Markov Model (HMM)
- Sequence Alignment Algorithm based on HMM
- Population mean and variance for discrete and continuous random variables
- Significance of Alignment
- Z-score, P-Value and E-value (definition)
- Distributions from Data
- Sample Mean and Sample Variance
- Extreme value distribution
- Estimation of parameters for scoring distributions
- Difference between the measures of significance
- Visualization of P-value

Learning Objectives:

Students will gain a basic understanding of the Hidden Markov model, building on their knowledge of Markov Chains, and will learn about a Multiple Sequence Alignment algorithm based on the HMM. Students will learn about the different ways to measure whether an alignment has arisen by chance or not, and how the different measures of significance relate to each other.