

A PRACTICAL INTRODUCTION TO THE UNDERLYING STATISTICAL CONCEPTS AND TECHNIQUES FOR SUCCESSFUL USE OF BIOINFORMATICS TOOLS

Effective use of the tools and methods of bioinformatics requires a careful understanding of not only the relevant biology and computational problems, but also critical statistical principles. *Statistical Bioinformatics* provides an essential understanding of the novel statistical concepts necessary for the analysis of genomic and proteomic data using various bioinformatics and computational techniques. Dr. Jae Lee and the authors present both basic and advanced topics, focusing on those that are relevant to the efficient and rigorous analysis of large data sets in biology.

The book starts with an introduction to probability and statistics for genome-wide data, and moves into topics such as clustering, classification, multidimensional visualization, experimental design, statistical resampling, and statistical network analysis. Chapters begin with a description of a statistical concept and practical examples from biomedical research, followed by more detailed presentation, discussion of limitations, and problems.

- Clearly explains the use of bioinformatics tools in life sciences research without requiring an advanced background in math/statistics
- Enables biomedical and life sciences researchers to successfully evaluate the validity of their results and make inferences
- Enables statistical and quantitative researchers to rapidly learn novel statistical concepts and techniques appropriate for large biological data analysis
- Carefully revisits frequently used statistical approaches and highlights their limitations in large biological data analysis
- Offers programming examples and datasets
- Includes chapter problem sets, a glossary, a list of statistical notations, and appendices with references to background mathematical and technical material
- Features supplementary materials, including datasets, links, and a statistical package available online

Statistical Bioinformatics is an ideal textbook for students in medicine, life sciences, and bioengineering, aimed at researchers who utilize computational tools for the analysis of genomic, proteomic, and many other emerging high-throughput molecular data. It may also serve as a rapid introduction to the bioinformatics science for statistical and computational students and audiences who have not experienced such analysis tasks before.