Course Syllabus

BIOE 505 – Computational Bioengineering

Instructor:
Sergei Maslov
Professor of Bioengineering and Physics,
Faculty member at Carl R. Woese Institute for Genomic Biology,
and National Center for Supercomputing Applications

Office: 3406 IGB (by appointment)
E-mail: maslov@illinois.edu (start subject line with [BIOE 505])
Phone: (217) 265-5705

Suggested Textbooks:
Montgomery, DC and Runger, GC: Applied Statistics and Probability for Engineers, 4th (2007) to 6th edition (2014) John Wiley & Sons, Inc, Hoboken, NJ, USA

Montgomery, DC and Runger, GC: Student Solutions Manual Applied Statistics and Probability for Engineers, 4th (2007) to 6th edition (2014) John Wiley & Sons, Inc, Hoboken, NJ, USA

Vidakovic, B: Statistics for Bioengineering Sciences with MATLAB and WinBUGS Support (2011) Springer, New York.

Pevsner, J: Bioinformatics and functional genomics, 3rd edition (2015), Wiley-Blackwell,

Ewens, WJ and Grant, GR: Statistical Methods in Bioinformatics: An Introduction, 2nd ed, (2005). Springer,

Course website: https://courses.engr.illinois.edu/bioe505
Grades will be on https://my.bioen.illinois.edu/gradebook

Meeting Schedule: two 75-minute lectures per week

Topical Outline:

Probability and Statistics in Bioengineering 15

1. Big Data in Bioengineering
2. Fundamentals of Probability and Statistics
    a. Probability space, events, axioms of probability
    b. Foundations of combinatorics
    c. Conditional probabilities and Bayes theorem
    d. Random variables and Independence
3. Discrete Random Variables
    a. Uniform
    b. Binomial
    c. Poisson
    d. Geometric

    e. Negative binomial
4. Continuous Random Variables and their Distributions
    a. Uniform
    b. Normal
    c. Exponential
    d. Erlang
    e. Gamma
    f. Lognormal
    g. Power law
5. Multi-Variable Probability Distributions
    a. Joint, marginal and conditional probability distributions
    b. Independence of two or more random variables
    c. Covariance and correlation
    d. Pearson product-moment and Spearman's rank correlation coefficients
6. Sampling Distributions
    a. Sample mean and variance
    b. Central Limit Theorem
    c. Box plots and histograms, quantiles
    d. Probability plots, fitting distribution to a sample
7. Parameter Estimation
    a. Method of moments estimation
    b. Maximum Likelihood Estimation
    c. Least squares and multiple regression
8. Confidence Intervals
9. Hypothesis Testing
    a. Type-I and Type-II errors
    b. P-value
    c. One- and two-sample t-tests
    d. Goodness-of-fit tests
    e. Regression

Computational Genomic and Systems Biology Methods

10. Statistical Methods in Sequence Analysis
    a. Genome assembly
    b. Evolutionary trees

11. Expression data analysis
    a. Finding differentially expressed genes
    b. Clustering: hierarchical, K-means, Principal Component Analysis (PCA)
12. Biomolecular network analysis
    a. Types of biomolecular networks
    b. Degree distributions and correlations
    c. Network Modularity
    d. Network visualization using Gephi

Grading:

Projects 20%, Midterm 30%, and a final 50%. Homework (ungraded) will build on topics covered in lectures and will consist of problem sets related to topics covered in lecture