

Stats M254 Statistical Methods in Computational Biology

Course site: <http://www.stat.ucla.edu/~zhou/courses/Stats254>

Lecture: TR 1–2:20pm, GEOLOGY 3656 .

Instructor: Qing Zhou (zhou@stat.ucla.edu), OH: Thursday 4:30-6pm, MS 8979.

Prerequisite: Stats 100A (Probability) or Stats 200A (Probability) or Bioinfo 260A (Bioinformatics).

Grading

Your final grade of this course will be composed of two parts:

1. Three homework assignments (60%). Some problems need computer programming.
2. Oral presentation (40%): A 40-minute presentation on a recent research paper.

Topics

Introduction to statistical and computational methods in computational biology and bioinformatics. The topics are grouped into four chapters:

1. Introduction and Data: molecular biology of gene regulation, typical data.
2. Multivariate Methods: gene expression data, multiple tests, principal component analysis, clustering methods, liquid association.
3. Statistical Sequence Analysis: Bayesian inference, Markov chain, hidden Markov model, missing data, Monte Carlo, motif discovery.
4. Predictive modeling: Bayes classifier, discriminant analysis, logistic regression, support vector machine, boosting, sparse modeling, and their applications in gene regulation.

References

- Ewens, W.J. and Grant, G.R. Statistical methods in bioinformatics: An introduction.
- Hastie, T. et al. The elements of statistical learning.
- Watson, J.D. et al. Molecular biology of the gene.
- Other papers posted on the course webpage.