



# STATISTICS 890

## SELECTED TOPICS: STATISTICAL METHODS IN COMPUTATIONAL BIOLOGY

---

Spring 2003  
DAY COURSE

---

Instructor: DR. BRAD MCNENEY

---

### Prerequisites:

Permission of instructor. The Equivalent of Stat 380 recommended.

---

### Audience:

Statistics graduate students and other interested graduate students with some background in statistics.

### Assessment:

There will be several (probably three) small projects throughout the term and a final project at the end of the term. Group work on all projects will be encouraged.

### Materials:

We will use the text "Statistical Methods in Bioinformatics" by Ewens and Grant.

### Computing:

Students should have access to the statistics package R version 1.5 or higher. R is freely available for both Windows and Unix at:

<http://cran.r-project.org/>

---

### Course Description:

We will introduce a limited set of problems in computational molecular biology and discuss associated statistical considerations. No background in molecular biology will be assumed.

0. Basic notions in biology
1. Applications to be covered:
  - sequence alignment
  - gene finding
  - phylogenetic tree estimation
2. Review of basic probability:
  - univariate and multivariate probability distributions
  - conditional probability and independence
3. Review of statistical inference:
  - point estimation, confidence intervals, hypothesis testing
  - maximum likelihood, the EM algorithm, likelihood ratio tests

4. Stochastic processes:
  - Poisson processes and Markov chains
  - continuous-time Markov chains
  - hidden Markov models
  - random walks
  
5. Computer intensive approaches to statistical inference:
  - permutation and randomization tests
  - Monte Carlo tests
  - bootstrap confidence intervals
  - Markov chain Monte Carlo

---

*Students should be aware that they have certain rights to confidentiality concerning the return of course papers and the posting of marks. Please pay careful attention to the options discussed in class at the beginning of the semester.*

---

Revised October 2002