Bioinformatics Course Announcement
Statistics/Math 547 and 548

547: Probability Models in Bioinformatics (3 credits)
548: Computing with Probability Models in Bioinformatics (1 credit)

Instructors: Kerby Shedden (547), Carlos Santos (548).

The focus of this course will be computational and algorithmic aspects of the analysis of gene and protein sequence data. A special emphasis will be placed on the use of probability models for the motivation and derivation of algorithms, but heuristically-motivated algorithms will be considered as well. The mathematical and statistical topics will include: review of elementary probability and statistical estimation theory, multinomial and Dirichlet distributions, simulation, Markov chains, dynamic programming, hidden Markov models, change point detection, Bayesian methods, classification, EM algorithm, and Gibbs sampling. The biological topics will include a review of the relevant molecular biology, pairwise and multiple alignment, identification of coding regions, splice prediction, motif searches, evolutionary inference, BLAST, structure prediction, TF binding site identification, and shotgun sequencing.

The mathematical/statistical and biological material will be developed from first principles. However a certain level of quantitative maturity is expected. Strongly motivated students with a background in only one of the two areas are encouraged to enroll.

The laboratory course 548 will focus on using PERL as a tool for computational biological sequence analysis. The required work for both courses 547 and 548 will include programming exercises. Students planning to enroll in 547 who have little or no programming background are strongly encouraged to enroll in 548 concurrently. Students with a moderate or advanced programming background (not necessarily in PERL) may choose not to enroll in 548.


547 meets Monday, Wednesday, Friday 9-10, in 271 Dennison
548 meets Tuesday 9-10, 271 Dennison

Contact Kerby Shedden (kshedden@umich.edu) for more information.