

PH 240D – Biostatistical Methods: Applications of Statistics to Genetics and Molecular Biology

Spring 2003

Time	TuTh 2:00-3:30pm (no lab)
Location	2030 VLSB
CCN	76077
Units	4

Instructor.

Sandrine Dudoit

E-mail: sandrine@stat.berkeley.edu

Website: www.stat.berkeley.edu/~sandrine

Office hours: Tu 3:30-4:30pm and Th 5:30-6:30pm in 109 Haviland Hall.

Website.

www.stat.berkeley.edu/~sandrine/PH240D.S03.

Please check the website regularly. All relevant class information (e.g., lecture notes, assignments, references, announcements) will be posted on this website. In addition, I will use e-mail for important announcements. Please make sure you are on the class mailing list.

Summary.

This course surveys applications of probability and statistics to genetics and molecular biology, from the early Mendelian experiments to modern day genomic research. Three main topics will be covered: genetic mapping; DNA microarray experiments; biological sequence analysis. In addition to discussing specific statistical methods, the course will provide an introduction to basic notions in genetics and molecular biology and to the main software packages for the analysis of biological data, with emphasis on the R language

and environment. The course will also involve the critical reading of articles related to statistical analyses in the biological and medical sciences.

Prerequisites.

Statistics 200 A and B (may be taken concurrently) or consent of instructor.

Some familiarity with R or S-Plus. Tutorials are available on the R (www.r-project.org) and Bioconductor (www.bioconductor.org) websites.

No formal training in biology is required; basic notions will be presented in class and references will be provided for further reading.

Grading policy.

50% homework and 50% final project.

There will be at least one assignment on each of the three sections of the course. Assignments will involve both theory and biological data analysis using R.

The final project will consist of a written report and poster presentation on a topic that involves the application of statistical and computational methods to address a particular biological question. I will provide a list of suggested topics.

References.

There is no required text. I will provide lecture notes and references on the website.