

Stat597D, Spring 2001
F. Chiaromonte, D. Hunter
INTRODUCTION

Who are we?

Francesca Chiaromonte, Statistics Department
chiaro@stat.psu.edu, Thomas 411, 5-7075
Off. Hours: Tue, Thur 1.00—2.00pm

Research interests relevant to this course: multivariate analysis, dimension reduction methods, analysis of global gene expression data (microarrays).

David Hunter, Statistics Department
dhunter@stat.psu.edu, Thomas 310, 3-0979
Off. Hours: Tue, Thur 11:00-12:00

Research interests relevant to this course: likelihood methods, computational techniques, application to statistical genetics.

Who are you?

... a question sheet to gather some information on the people attending this class.

Who's a statistician?

Do you have any training in computing?
In biology/genetics?

Who's a computer scientist/engineer?

Do you have any training in statistics?
In biology/genetics?

Who's in the life sciences?

Do you have any training in statistics?
In computing?

Working together:

Approximately biweekly homework assignments, and a larger final project (all open-ended, involving data analysis and/or computing).

Assignments will not have posted solutions, and will often be discussed in class.

Final projects will be presented and discussed in 2 or 3 meetings at the end of the semester.

Divide enrolled students and stable auditors and guests into working groups containing people with different backgrounds, to work together on assignments and projects.

We will bring a grouping proposal back to you in a week.
Comments, questions?

Computing and software:

Who's familiar with

??Fortran, Pascal, C?

??S+, Matlab or similar?

??Minitab or similar statistics packages?

What's available on the Statistics Department PC and Unix networks:

S+, SAS, Minitab, Mathematica (PC and Unix)

Gauss and SPSS (PC)

Matlab (Unix if demand warrants)

(also non-statistics students can access our networks if needed, with a special permit to open an account)

What are we planning to do in this course?

Present modeling approaches, statistical methods and computational techniques, focusing on their applications to problems in genetic and bioinformatics.

A methodological toolkit relevant to a certain class of applications.

We will adopt a hands-on approach, but also insist on an accurate understanding of the quantitative background.

Interaction/feedback among you, and between you and us is crucial: this is a course in the making... and we are not biologists!

What are you expecting to gain from this course?

Concerns, comments, questions?

Reading material and references:

For Statistical Genetics:

??Mathematical and Statistical Methods for Genetic Analysis (K. Lange, Springer-Verlag)

For global gene expression:

??Theory and methods from various books; handouts of specific chapters.

??Recent papers from genetics, bioinformatics and statistics journals. Web-available manuscripts.

Website for the course:

<http://www.stat.psu.edu/~dhunter/genetics/>:

??Announcements and communications on the course

??Class notes (hopefully, but not surely, beforehand)

??Assignments (but not solutions)

??Links to relevant websites

An “open spot” on our calendars:

Tue 4.00—5.00pm, Thomas 327
... on a NON-REGULAR basis

??For extra meetings

??For guest lectures

??For informal presentations and seminars

Keeping the tradition we have started last spring with our reading group meetings.

Announcements on the course website and through a mailing list.

A (very tentative) schedule for the regular lectures:

Jan 9 Introductory lecture

Jan 11, 16, 18, 23, 26

Dave (basic population genetics, gene counting, maximum likelihood, EM)

Jan 30, Feb 1, 6, 8, 13

Francesca (intro to microarrays, sources of error, data normalization, dimension reduction)

Feb 15, 20, 22, 27, Mar 1

Dave (categorical data hypothesis testing, permutation and resampling schemes, pedigree analysis)

Mar 6, 8 SPRING BREAK

Mar 13, 15, 20, 22

Francesca (working with a response)

Mar 26, 29, Apr 3, 5

Dave (evolutionary trees and more pedigree analysis, MCMC)

Apr 10, 12, 17, 19

Francesca (clustering methods)

Apr 24, 27

Final projects presentations

