BMI206: Statistical Methods for Bioinformatics

Course Overview

Course Objectives

- **Goal:** understand statistical methods that are used in bioinformatics research
- You will learn both ideas/concepts and how to implement methods in code.
- By the end of the course, you should be able to make good decisions about how to statistically analyze bioinformatics data.

Course Approach

Learn statistical methods in the context of biomedical bioinformatics research.

- Lectures for key concepts
- Board work to highlight details
- Paper discussions to see methods in action
- Coding to implement methods on data
- Projects bring it all together in your hands

Weekly Structure

Asynchronously watch recorded short lectures

- Monday board work, lecture Q&A
- Tuesday coding examples, introduce lab
- Wednesday paper discussion
- Thursday work on lab

  Project assignments due Tuesdays

  Paper preparation due Wednesdays

  Lab homework due Fridays

Journal Articles

Several recent papers will be assigned as background reading and to demonstrate the complexities around how the concepts from lecture are used in practice.

- Papers will be discussed on Wednesdays.
- Lectures and labs may refer to the papers.
- Papers will form the basis for projects.

If you want additional reading on any topic, we can provide a longer list of references.

Labs

Most Thursdays will be computer lab sessions.

- Collaborate and ask each other questions

- We will discuss our findings

- The final portion of each lab will be an open problem to solve on your own.

You will need to install software packages on your laptop. Labs will be in R. TAs can help as needed.

- To start, please install R (r-project.org)

A summary of your solution plus your code will be due Friday midnight by email to the TAs & me.

Project Objectives

Goals of the Project:

- Learn to critically read bioinformatics papers from a statistical perspective
- Obtain primary data from a publication
- Practice selecting appropriate analysis methods
- Practice making figures and other data displays
- Compare and evaluate different bioinformatics and statistical approaches to answering a biological question
- Evaluate sensitivity of results to analysis choices

Wynton HPC accounts

Pre-Requisites

BMI206 is intended for students with a strong undergraduate background in statistics. Students should

1. Know statistics background topics (slides).
2. Complete David Quigley's biostatistics course:

   https://courses.ucsf.edu/course/view.php?id=3223

3. Complete MBoC tutorial
4. Know how to program in at least one language. Lab will be R, project may use Python.

Grading

- Participation, paper discussions, and attendance = 10%
- Lab assignments (homework) = 50%
- Project = 40%

Participation and attendance are critical

* absences should be communicated

* plan for Thanksgiving week and finals week

Communication

- Website

  http://docpollard.org/teaching/bmi206/

- Slack (please monitor!)
- Email: Katie Pollard: kpollard@gladstone.ucsf.edu
- Office hours