Epi & Biostats 101: An Intro to Public Health Research Tools & Techniques (PBHLTH 198)

Course Syllabus

Facilitators
Andrew O'Connor (andrewoc@berkeley.edu)

Faculty Advisor
Andrew Mertens, Ph.D., M.S., M.A.

Course Email
epibiostats101@gmail.com

Time and Location
Dwinelle 104

Units: 1 P/NP

Course Introduction

Mastery of the principles and methods in biostatistics is fundamental for students seeking to contribute meaningfully to the contemporary landscape of public health research. We aim to expose undergraduates from any discipline to the foundational concepts and applications of biostatistics. This course, for individuals who already have a foundational understanding of statistics, is designed to provide a survey of advanced statistical techniques, emphasizing their practical applications in the field of public health.

Course Prerequisites

PBHLTH 142, DATA 8 or Exposure to basic statistics

Course Materials

All the course materials (lecture powerpoints, Jupyter Notebooks, assignment instructions, etc.) will be provided via bCourses and the course website. Supplemental materials will also be available on the course website or linked in assignments.

Learning Outcomes

- Describe the role of epidemiology and biostatistics in addressing public health problems.
- Describe the role of epidemiology and biostatistics in addressing public health problems
- Describe the role of data analysis in the biostatistics, epidemiology, data science, and related fields
- Describe common epidemiological study designs
- Utilize Python for statistical computing and data cleaning
- Understand the role of regression and predictive modeling in applications to health

Method of Instruction

Facilitators and students will meet weekly on ***Mondays 6-8pm*** to learn theoretical concepts. The first hour will be lecture and the second hour will be reserved for in-class collaboration on assignments. Students are encouraged to discuss worksheets and projects but should submit individual copies of their own work.

Weekly Assignments: Each week, a worksheet will be assigned after the lecture is given. Time in class will be reserved to complete the worksheets. During the rest of the week, students will complete any unfinished problems and read/watch optional material to reinforce concepts. All worksheets are due by the beginning of the next class (5:59pm on the Monday following the class the worksheet was assigned).

Group Project: In small groups, students will complete a comprehensive project requiring knowledge of Biostatistics and Epidemiology. Instructions, grading and guidelines will be clearly noted on the assignment, formatted as a Jupyter Notebook. No independent research question is required. More details TBA.

Final Presentation: In your small groups, you will present the findings of your group project.

Expectations and Communication

Check bCourses regularly for any announcements regarding assignments, lecture material, etc. The e-mails of the facilitators are provided above, so do not hesitate if you have any questions or anything you would like to tell us (including feedback!). **Weekly topics and assignments are subject to change based on student interests/previous experience.**

Grading

In order to receive a passing grade in this class, you must complete the following:

1. Attend 10 out of the 12 in-person DeCal meetings, unless excused for a previously communicated situation.
    a. We will be taking attendance at the beginning of each meeting.
    b. Week 12 attendance is mandatory in light of the project presentations.
    c. Please notify us in advance of any midterm or exam conflicts and we will accommodate you.
2. Complete 10 out of the 12 of the weekly worksheets, unless excused for a previously communicated situation.
3. Participate in the group project and final presentation.

Attendance	25%
Assignments	30%
Group Project	25%
Final Presentation	20%
Extra Credit	5% (Opportunities will be dispersed throughout the assignments)
Total	**105%** Possible

Lecture Schedule

Module	Week	Topics
Module 1: Foundations	1	Lecture 1 • Introduction to Public Health • Role of Epidemiology and Biostatistics • Fundamental tools and skills for pursuing career in Biostatistics/Epidemiology
	2	Lecture 2 • Python review • Fundamental computing concepts ◦ Functions, Conditions, Variables, Scope • Dependency management: Anaconda, Virtual Envs
	3	Lecture 3 • Descriptive Statistics: Measures of Center/Spread • Set notation • Introduction to Probability • Random Variables • Distributions
	4	Lecture 4 • Distributions review (Binomial, Poisson, Normal) • Data cleaning with Pandas
	5	Lecture 5 • Introduction to Predictive Modeling • Relating Models to Biostatistics and Epidemiology • Types of Models: Regression, Clustering, Classification • Training Models: train-test split • Conceptual overview of over/underfitting, feature engineering/selection
	6	Lecture 6 • Modeling Process Overview • Evaluation Metrics • Regression Methods: Simple Linear, Multiple Linear • Clustering Methods: KMeans • Classification Methods: Decision Trees, Random Forests
	7	Lecture 7 • Review difference between Epidemiology and Biostatistics • Types of Prevention; Natural History of Disease • Mortality vs Morbidity; Incidence vs. Prevalence • Study Designs Overview • Brief intro to Risk and Association

	8	Lecture 8 • Introduction to Causal Inference in Public Health • Terminology: Cause, Bias, DAGs, Mediators, Colliders, Confounding, Effect Measure Modification • Video: Causal Structures • Causal Relationships: Necessity and Sufficiency • Rothman's Sufficient Cause Model • 1964 Surgeon General's Report
	9	Lecture: 9 • Introduction to Infectious Disease Epidemiology • Chain Model • Characteristics of Infectious Agents • SIR Model, R0, RE, SAR, Herd Immunity • Example: COVID-19
Module 2: Final Project	10	Group Project Work Time
	11	Group Project Work Time
	12	Group Project Presentations

Recommended Reading: Additional Supplementary articles and videos will be posted on bCourses and updated weekly. Epidemiology supplementary text is Gordis Epidemiology, 6th edition by Celentano, D. and Moyses Szklo. Biostatistics supplementary text is The Practice of Statistics in the Life Sciences by Brigitte Baldi and David S. Moore.

Plagiarism and Cheating
Plagiarism of any kind will not be tolerated and will result in a “No Pass” grade in the class. In addition, the department will be notified.

Community and Student Support
Disability-related accommodations or personal/emergency accommodations will be available to all students.
We acknowledge that the field of public health has been used for justifications of racist, eugenicist policies in the past, and we openly denounce such approaches to health and education. We, as facilitators, will make active efforts to create a safe, non-discriminatory learning environment for students of all backgrounds.
We welcome all students and hope to create an environment in which you always feel included. If any parts of the course create barriers to your inclusion, please let us know.

If you, a classmate, or a friend are in need of mental health support, UC Berkeley has several services and resources.

If at any point you would like support or have feedback about the class, please feel free to talk to us after class or email us at the addresses listed above.