ID 201: Core Principles of Biostatistics and Epidemiology for Public Health Practice

Fall 2022

# Instructor Information

## Faculty

Heather Mattie, PhD

Lecturer on Biostatistics, Harvard T.H. Chan School of Public Health

Co-Director, Health Data Science Master's Program

Office: Building I, 4th floor Room 421A

Email: [hemattie@hsph.harvard.edu](mailto:hemattie@hsph.harvard.edu)

Office Hours: Tuesdays, 3:30-4:30 ET, Building 2, Room 426

Murray A. Mittleman, MD, DrPH

Professor of Epidemiology, Harvard T.H. Chan School of Public Health

Director and Chair, MPH Program

Office: Kresge 505D

Email: [mmittlem@hsph.harvard.edu](mailto:mmittlem@hsph.harvard.edu)

Office Hours: Tuesdays, 3:30-4:30 ET, Building 2, Room 426

|  |  |  |
| --- | --- | --- |
|  | **Biostatistics** | **Epidemiology** |
| **Head Teaching Fellows** | Lauren Hsu [laurenhsu@hsph.harvard.](mailto:laurenhsu@hsph.harvard.edu)edu | Justin Farmer [jfarmer@hsph.harvard.edu](mailto:jfarmer@hsph.harvard.edu) |
|  |  |  |
| **Teaching Fellows** | Futu Chen  [fchen@g.harvard.edu](mailto:fchen@g.harvard.edu) | Isa Berzansky  [iberzansky@hsph.harvard.edu](mailto:iberzansky@hsph.harvard.edu) |
|  | Zach Clement  [zclement@hsph.harvard.edu](mailto:zclement@hsph.harvard.edu) | Sarah Christensen  [sarahchristensen@hsph.harvard.edu](mailto:sarahchristensen@hsph.harvard.edu) |
|  | Rebecca Danning  [rebecca\_danning@harvard.edu](mailto:rebecca_danning@harvard.edu) | Eric Fung  [efung@hsph.harvard.edu](mailto:efung@hsph.harvard.edu) |
|  | Ravi Dhawan  [ravidhawan@hsph.harvard.edu](mailto:ravidhawan@hsph.harvard.edu) | Kelly Fung  [klfung@hsph.harvard.edu](mailto:klfung@hsph.harvard.edu) |
|  | Anuraag Gopaluni  [anuraag\_gopaluni@hms.harvard.edu](mailto:anuraag_gopaluni@hms.harvard.edu) | Colleen McGrath  [cmcgrath@hsph.harvard.edu](mailto:cmcgrath@hsph.harvard.edu) |
|  | Fuyu Guo  [fuyuguo@hsph.harvard.edu](mailto:fuyuguo@hsph.harvard.edu) | Susan Rattigan  [susanrattigan@g.harvard.edu](mailto:susanrattigan@g.harvard.edu) |
|  | Omar Mansour  [omansour@g.harvard.edu](mailto:omansour@g.harvard.edu) | Emily Reichert  [ereichert@hsph.harvard.edu](mailto:ereichert@hsph.harvard.edu) |
|  | Avik Ray  [avikray@hsph.harvard.edu](mailto:avikray@hsph.harvard.edu) | Eva Rumpler  [erumpler@hsph.harvard.edu](mailto:erumpler@hsph.harvard.edu) |
|  | Arpan Sarkar  [arpan\_sarkar@g.harvard.edu](mailto:arpan_sarkar@g.harvard.edu) | Madikay Senghore  [msenghore@hsph.harvard.edu](mailto:msenghore@hsph.harvard.edu) |
|  | Denys Shay  [denys\_shay@g.harvard.edu](mailto:denys_shay@g.harvard.edu) | Zihan Sun  [zihansun@hsph.harvard.edu](mailto:zihansun@hsph.harvard.edu) |
|  | Mai-Han Trinh  [maihantrinh@hsph.harvard.edu](mailto:maihantrinh@hsph.harvard.edu) | Yingzhe Zhang  [yzhang1@hsph.harvard.edu](mailto:yzhang1@hsph.harvard.edu) |

**Credits**

7.5 credits

# Course Purpose and Description

This course introduces the methods of biostatistics and epidemiology in the context of public health and clinical research. ID 201 is designed to be a self-contained terminal course in biostatistics and epidemiology and is intended for students who do not anticipate taking further methodological courses in either domain. The main objectives of this course are to introduce the basic biostatistical techniques and epidemiological principles and methods used in public health research. Statistical and epidemiological reasoning will be emphasized through applications.

# Pre-Requisites

No prior knowledge of biostatistics or epidemiology is assumed. However, students are expected to have a working knowledge of computing and algebra, including logarithms and exponentials. Calculus is not a prerequisite.

# Course Learning Objectives

Upon successful completion of this course, you should be able to:

* Develop comprehensive knowledge of the study design principles of epidemiology to provide a quantitative approach for addressing health problems
* Calculate and interpret epidemiologic measures of frequency and association
* Implement and interpret basic statistical techniques, including methods of estimation and hypothesis testing, either by hand (if appropriate) or using the Stata statistical software package
* Develop judgment about which statistical technique to use in a given situation
* Evaluate published clinical and public health articles, including identification of sources of bias, confounding, and random error and how they may impact the results
* Explain quantitative methods and evaluation methods to address health issues at multiple (individual, group, organization, community and population) levels
* Explain the use and limitations of surveillance systems and national surveys in assessing, monitoring and evaluating policies and programs and to address a population’s health

The course introduces evidence-based approaches to public health, enabling students to:

* Apply epidemiological methods to the breadth of settings and situations in public health practice
* Select quantitative and qualitative data collection methods appropriate for a given public health context
* Analyze quantitative and qualitative data using biostatistics, informatics, computer-based programming and software, as appropriate
* Interpret results of data analysis for public health research, policy, or practice

# Course Readings

Readings for each lecture are listed on the course website. They include readings from the textbooks listed below, supplementary resources, and articles published in peer-reviewed journals. Textbooks are on reserve at Countway Library. Pagano, Gauvreau, and Mattie (2022) is available through the Canvas Library Reserves tool (three copies are available at a time). Rothman (2012) is available electronically at the link below (there is no limit to the number of simultaneous copies). Both textbooks are also available for rent or purchase in paper or electronic format from a variety of booksellers.

* *Principles of Biostatistics,* 3rd Edition (2022), Pagano, Gauvreau, and Mattie; CRC Press
* *Epidemiology: An Introduction,* 2nd Edition (2012), Rothman; Oxford University Press. http://ezp-prod1.hul.harvard.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=2114235&site=ehost-live&scope=site

# Course Structure

This course satisfies the school-wide requirements for biostatistics and epidemiology. All components of the course listed below are integral to understanding the material presented.

**Canvas Course Website**

Canvas is an essential learning tool for this course. This is where you will access the official syllabus, required readings and videos, discussion boards, and course and school policies. It is also where you will submit course assignments and share other resources with the class. Course announcements will be posted on the Canvas site, and individual and group messages also may be sent through the Canvas message Inbox (Conversations tool). Since the Canvas site is the official source of course information, you are strongly encouraged to check the course website daily and/or to make sure that Canvas notifications about announcements and Canvas Conversations are being sent to your primary email address. You are expected to stay informed regarding course communications. Ignorance of posted information or course changes will not be considered a valid excuse for late or missing work.

**Lectures**

Course lectures will be prerecorded and posted on the course website. This will be the primary presentation of the course material. Lecture handouts will be available on the Canvas course website.

**Interactive Class Sessions**

Faculty will offer weekly in-person, interactive class sessions in which students will participate in a variety of concept- and skill-related activities, such as answering poll questions, problem-solving, and conceptual discussions. Students are required to attend the weekly interactive class session they signed up for during registration (Tuesday or Thursday at 8:00 am). Interactive class sessions will be recorded; however, since these sessions may include some small group work, the recordings may be of limited value.

**Lab Sessions**

Students must register for the required in-person weekly lab session, with time slots available on Thursdays and Fridays. Each lab session will be led by two teaching fellows, one from biostatistics and the other from epidemiology. The labs will include applications of material covered in the prerecorded modules for that week. A handout of the lab practice problems and links to any lab readings will be available on the course website. Lab sessions provide the opportunity for group discussion. They are designed to reinforce the principles covered in the prerecorded lectures by working through practice problems and by reviewing Stata commands.

**Notice of Sensitive Content**

Throughout this course, materials draw on a wide range of important topics that reflect the breadth of impact of public health issues on society. Students should be aware that some may find these issues disturbing and difficult to read.

**Office Hours**

Each week there will be office hours led by teaching fellows and the faculty. Students may attend as many as they find helpful to discuss course material and Stata syntax with the teaching team. Office hour times are posted on the course website.

**Computing**

This course will use the Stata statistical package available through the HSPH VDI system:

[*https://www.hsph.harvard.edu/information-technology/frequently-asked-questions/vdi-faq/*](https://www.hsph.harvard.edu/information-technology/frequently-asked-questions/vdi-faq/)

Information on how to use Stata has been posted on the course website in the Stata Resources folder; additional details will be provided in the weekly lab handouts and reviewed during lab sessions. For some homework problems that do not require a computer and for the in-class examinations, a hand calculator will be useful. The only required features for the hand calculator are basic arithmetic operations, logarithms, and exponentiation.

**Grading, Progress, and Assessment**

The final grade for this course will be based on:

* Exams (54%)
* Homework (36%)
* Watch prerecorded material and complete weekly self-assessment questions before each Thursday at 1:00 pm ET (10%)

**Exams (54%)**

There will be three examinations administered on Canvas. Exams will be cumulative, but they will focus on the most recently covered material. Exams will be held on either Tuesday or Thursday depending on the day of the interactive class session a student registered for. Those dates are Tuesday, October 4 or Thursday, October 6 (18%); Tuesday, November 8 or Thursday, November 10 (18%); and Tuesday, December 13 or Thursday, December 15 (18%). Together they are worth 54% of the final course grade. See the detailed course calendar at the end of the syllabus or course website for further information regarding the timing of the examinations. Students must take all three exams to pass the course.

**Homework (36%)**

Beginning in the third week of the course, there will be weekly homework assignments that will include applications of the material covered in class and reviewed in the lab sessions. Homework assignments will cover material up to and including material presented in the prior week with a strong emphasis on more recent material. Assignments will be posted on Canvas and will be due on Thursdays and must be submitted online on the due date prior to 1:00 pm ET. The teaching fellows’ feedback on homework responses should be reviewed carefully to ensure that the concepts are clear.

Students may discuss the problems, but each student must complete the homework independently and should turn in their own written assignment. Assignments must be submitted as Word or PDF documents by the posted deadline. Canvas will remain open until 10:00 pm ET on Friday. Late homeworks will be penalized with a 20% deduction. Homework will not be accepted after 10:00 pm ET on Friday. Students must speak with a faculty member (not a teaching fellows) in advance of the deadline if they are facing extreme life circumstances affecting their ability to meet upcoming deadlines. No makeup homework assignments will be provided.

There will be 10 homework assignments during the semester. Point values for each homework will vary and will be listed on the homework cover sheet. At the end of the semester, the lowest homework grade (based on percentage of points earned) will be dropped, and the final homework grade will be calculated as the number of points earned summed over the other 9 assignments divided by the total number of possible points for those assignments. Homework is worth 36% of the final course grade.

**Weekly Prerecorded Lectures and Self-Assessment Questions (10%)**

During each week, students are expected to watch all prerecorded lectures and answer self-assessment questions. These knowledge checks will help students ensure understanding of the most important concepts in those lectures. Students may take the self-assessments as many times as they wish. Full points will be awarded for completion by Thursday at 1:00 pm ET.

# Technical Information

## Assistance

Canvas

If the issue is Canvas-related (e.g., you can’t figure out how to use something or a feature seems broken), first try the documentation located under the Help menu found on the left-hand side of each Canvas page. If the issue is not covered there, contact Instructure directly, also via the Help menu. You can e-mail, text, or speak live with them at any time day or night. If you cannot access Canvas to view the Help menu, you can reach Instructure by phone at +1 (844) 326-4466.

Zoom

For help with Zoom video conferencing, first check the variety of video tutorials and online help at <https://support.zoom.us>. In addition, you may contact the Helpdesk by emailing [helpdesk@hsph.harvard.edu](mailto:helpdesk@hsph.harvard.edu) or calling +1 (617) 432-HELP (4357).

Harvard-Specific Issues

If the issue seems Harvard-specific (e.g., HUID or myHarvardChan authentication, email not working, etc.), contact the Helpdesk at [helpdesk@hsph.harvard.edu](mailto:helpdesk@hsph.harvard.edu) or +1 (617) 432-HELP (4357).

## Technical Requirements

* Reliable, high-speed internet connection
* Your laptop must meet the minimum technical requirements found on the [Student Guide page](https://www.hsph.harvard.edu/information-technology/student-guide/#laptop-requirements).
* Modern and updated web browser (e.g., a recent version of Firefox or Chrome)
* Web camera and microphone (integrated into computer or USB peripheral)
* Throughout this course, you will be using VDI to access certain applications (e.g., Stata); in turn, your computer must meet the minimum hardware and software requirements displayed on the [VDI page](https://www.hsph.harvard.edu/information-technology/frequently-asked-questions/vdi-faq/).
* Please contact [helpdesk@hsph.harvard.edu](mailto:helpdesk@hsph.harvard.edu) with questions.

Please note that while it is possible to access most of the course materials via mobile and wireless devices, video conferencing and other bandwidth-intensive sessions will have the greatest reliability on a wired high-speed connection.

# Harvard Chan Policies and Expectations

## Inclusivity Statement

Diversity and inclusiveness are fundamental to public health education and practice. Students are encouraged to have an open mind and respect differences of all kinds. I share responsibility with you for creating a learning climate that is hospitable to all perspectives and cultures; please contact me if you have any concerns or suggestions.

## Bias Related Incident Reporting

The Harvard Chan School believes all members of our community should be able to study and work in an environment where they feel safe and respected. As a mechanism to promote an inclusive community, we have created an anonymous bias-related incident reporting system. If you have experienced bias, please submit a report [here](https://reportinghotline.harvard.edu/) so that the administration can track and address concerns as they arise and to better support members of the Harvard Chan community.

## Title IX Sexual Harassment and Other Sexual Misconduct

For information on Harvard University policies and procedures and Title IX Resource Coordinators at Harvard Chan, please see:

* Harvard University Interim Title IX Sexual Harassment and Interim Other Sexual Misconduct policies and procedures: <https://titleix.harvard.edu/policies-procedures>
* Title IX Resource Coordinators: <https://titleix.harvard.edu/coordinators>
* Title IX Sexual Harassment and Other Sexual Misconduct resource guide: <https://titleix.harvard.edu/resource-guide>

## Academic Integrity

Each student in this course is expected to abide by the Harvard University and the Harvard T.H. Chan School of Public Health School’s standards of Academic Integrity. All work submitted to meet course requirements is expected to be a student’s own work. In the preparation of work submitted to meet course requirements, students should always take great care to distinguish their own ideas and knowledge from information derived from sources.

Students must assume that collaboration in the completion of assignments is prohibited unless explicitly specified. Students must acknowledge any collaboration and its extent in all submitted work. This requirement applies to collaboration on editing as well as collaboration on substance.

Should academic misconduct occur, the student(s) may be subject to disciplinary action as outlined in the Student Handbook. See the [Student Handbook](https://www.hsph.harvard.edu/student-handbook/) for additional policies related to academic integrity and disciplinary actions.

## Accommodations for Students with Disabilities

Harvard University provides academic accommodations to students with disabilities. Any requests for academic accommodations should ideally be made before the first week of the semester, except for unusual circumstances, so arrangements can be made. Students must register with the Local Disability Coordinator in the Office for Student Affairs to verify their eligibility for appropriate accommodations. Contact Colleen Cronin [ccronin@hsph.harvard.edu](mailto:ccronin@hsph.harvard.edu) in all cases, including temporary disabilities.

## Religious Holidays, Absence Due to

According to Chapter 151c, Section 2B, of the General Laws of Massachusetts, any student in an educational or vocational training institution, other than a religious or denominational training institution, who is unable, because of his or her religious beliefs, to attend classes or to participate in any examination, study, or work requirement on a particular day shall be excused from any such examination or requirement which he or she may have missed because of such absence on any particular day, provided that such makeup examination or work shall not create an unreasonable burden upon the School. See the [student handbook](https://www.hsph.harvard.edu/student-handbook/) for more information.

Grade of Absence from Examination  
A student who cannot attend a regularly scheduled examination must request permission for an alternate examination from the instructor in advance of the examination. See the [student handbook](https://www.hsph.harvard.edu/student-handbook/) for more information.

## Final Examination Policy

No student should be required to take more than two examinations during any one day of finals week. Students who have more than two examinations scheduled during a particular day during the final examination period may take their class schedules to the director for student affairs for assistance in arranging for an alternate time for all exams in excess of two. Please refer to the [student handbook](https://www.hsph.harvard.edu/student-handbook/) for the policy.

## Course Evaluations

Constructive feedback from students is a valuable resource for improving teaching. The feedback should be specific, focused and respectful. It should also address aspects of the course and teaching that are positive as well as those which need improvement.

Completion of the evaluation is a requirement for each course. Your grade will not be available until you submit the evaluation. In addition, registration for future terms will be blocked until you have completed evaluations for courses in prior terms.

# Course Schedule

|  |  |  |  |
| --- | --- | --- | --- |
| **Weekly topics** | **Objectives** | **Readings** | **Activities/**  **Assignments** |
| **Week 1 (8/29/2022–9/4/2022)** | | | |
| **Course Introduction**  **Descriptive Statistics** | **Course Introduction**  Describe course requirements, topics, and primary course objectives  Describe the role of quantitative methods in addressing public health questions  Differentiate between descriptive and analytic aspects of epidemiology and biostatistics  Describe the uses and limitations of surveillance systems, including national surveys, to monitor and assess the health of populations  **Descriptive Statistics**  Identify the different types of quantitative data and understand how they differ  Identify which methods of data presentation are most appropriate and effective for each type of quantitative data  Identify which numerical summary measures are most appropriate for each type of quantitative data  Distinguish between measures of central tendency and measures of variation, and explain how they should be reported | **Course Introduction**  Lash, T. L., VanderWeele, T. J., Haneuse, S., & Rothamn, K. J. (2021). *Modern Epidemiology.* (4th ed.), p 185-194; 200-208  **Descriptive Statistics**  P,G,&M: 2.1-2.4 | **Watch prerecorded material and complete all self-assessment questions by 1:00 pm EDT on Thursday, 9/1/2022.**  **Lab Session Practice Problems 1**  **Complete in Lab**  Descriptive Statistics  Introduction to Stata |
| **Week 2 (9/5/2022–9/11/2022)** | | | |
| **Probability**  **Measures of Occurrence**  **Measures of Association** | **Probability**  Describe situations in terms of operations on events  Calculate and interpret probabilities and conditional probabilities  Apply the addition and multiplication rules of probability, and the rule of total probability  **Measures of Occurrence**  Distinguish among proportions, rates, and ratios  Describe the utility of person-time analysis and rates  Calculate prevalence, cumulative incidence, and incidence rate and describe how they are related  **Measures of Association**  Calculate and interpret ratio and difference measures of effect and association  Distinguish between the appropriate applications for each measure of disease frequency and association | **Probability**  P,G,&M: 5.1-5.3  **Measures of Occurrence**  Rothman: p.38-56  **Measures of Association**  Rothman: p.57-68 | **Watch prerecorded material and complete all self-assessment questions by 1:00 pm EDT on Thursday, 9/8/2022.**  **Lab Session Practice Problems 2**  **Complete in Lab**  Probability  Measures of Association  Measures of Occurrence |
| **Week 3 (9/12/2022–9/18/2022)** | | | |
| **Overview of Study Designs**  **Cohort Studies**  **Case-Control Studies** | **Overview of Study Designs**  Describe key features, strengths, and limitations of common study designs  Identify the study design from a description in peer-reviewed literature  **Cohort Studies**  Describe key features, strengths, and limitations of cohort studies  Explain the measures of association that can be calculated in cohort studies and when they are appropriate  **Case-Control Studies**  Describe the conceptualization of a case-control study as an efficient cohort study  Evaluate the validity of different ways to select controls  Explain the measure of association used in case-control studies and why it is appropriate | **Overview of Study Designs**  **Cohort Studies**  Rothman: p.69-87  **Case-Control Studies**  Rothman: p.87-96  p.99-102, p.104-107 | **Watch prerecorded material and complete all self-assessment questions by 1:00 pm EDT on Thursday, 9/15/2022.**  **Homework 1:** Due 1:00 pm EDT on Thursday, 9/15/2022  **Lab Session Practice Problems 3**  **Complete in Lab**  Overview of study designs  Cohort  Case-control  Matching |
| **Week 4 (9/19/2022–9/25/2022)** | | | |
| **Probability Distributions** | **Probability Distributions**  Explain the utility of probability distributions  Describe the characteristics and applications of the binomial distribution  Explain the difference between discrete and continuous probability distributions  Define the characteristics and applications of the normal and standard normal distributions  Identify the properties of the sampling distribution of the mean | **Probability Distributions**  P,G,&M: 7.1-7.2  P,G,&M: 7.4, 8.1-8.3 | **Watch prerecorded material and complete all self-assessment questions by 1:00 pm EDT on Thursday, 9/22/2022.**  **Homework 2:** Due 1:00 pm EDT on Thursday, 9/22/22  **Lab Session Practice Problems 4**  **Complete in Lab**  Probability Distributions |
| **Week 5 (9/26/2022–10/2/2022)** | | | |
| **Confidence Intervals**  **Causation** | **Confidence Intervals**  Estimate a population mean using a confidence interval  Interpret the interval and the level of confidence  Apply the t distribution when the underlying population standard deviation is not known  Estimate a population proportion using a confidence interval  Calculate and interpret confidence intervals for measures of association and differences in means  Distinguish between sources of systematic and random error  Identify the difficulties in concluding that a given exposure causes an outcome of interest  **Causation**  Differentiate between selection bias, information bias, and confounding  Define component, necessary and sufficient causes  Present criteria for judging whether an association is causal | **Confidence Intervals**  P,G,&M: 9.1, 9.3  P,G,&M: 14.1-14.3  **Causation**  Rothman: p.23-29, 32-35 | **Watch prerecorded material and complete all self-assessment questions by 1:00 pm EDT on Thursday, 09/29/2022.**  **Homework 3:** Due 1:00 pm EDT on Thursday, 9/29/2022  **Lab Session Practice Problems 5**  **Complete in Lab**  Confidence Intervals  Causation |
| **Week 6 (10/3/2022–10/9/2022)** | | | |
| **Hypothesis Tests for One Mean**  **Hypothesis Tests for Two Means** | **Hypothesis Tests for One Mean**  Implement the steps for conducting a one-sample hypothesis test for a population mean  Explain and interpret a p value  Draw conclusions about a population mean based on the results of the hypothesis test  Explain the relationship between hypothesis tests and confidence intervals  **Hypothesis Tests for Two Means**  Distinguish between paired and independent samples  Conduct and interpret two-sample hypothesis tests comparing means for both paired and independent samples | **Hypothesis Tests for One Mean**  P,G,&M: 10.1-10.3  **Hypothesis Tests for Two Means**  P,G,&M: 11.1-11.2 | **Watch prerecorded material and complete all self-assessment questions by 1:00 pm EDT on Thursday, 10/6/2022.**  **Examination 1**  **Tuesday, 10/4/2022 *or***  **Thursday, 10/6/2022**  The exam will be in the classroom from 8:00 am until 9:30 am.  **Lab Session Practice Problems 6:**  **Complete in Lab**  Hypothesis Tests for One or Two Means |
| **Examination 1: 10/4/2022 or 10/6/2022 – See details above regarding timing.** | | | |
|  |  |  |  |
| **Week 7 (10/10/2022–10/16/2022)** | | | |
| **One-Way Analysis of Variance**  **Hypothesis Tests for Proportions** | **One-Way Analysis of Variance**  Conduct and interpret one-way ANOVA  Interpret the results of post hoc tests which use a Bonferroni correction  **Hypothesis Tests for Proportions**  Conduct and interpret a one-sample test for a proportion  Conduct and interpret two-sample hypothesis tests comparing proportions  Interpret the results of hypothesis tests for r x c contingency tables | **One-Way Analysis of Variance**  P,G,&M: 12.1-12.2  **Hypothesis Tests for Proportions**  P,G,&M: 14.4, 14.6, 15.1-15.2 | **Watch prerecorded material and complete all self-assessment questions by 1:00 pm EDT on Thursday, 10/13/2022.**  **Homework 4:** Due 1:00 pm EDT on Thursday, 10/13/22  **Lab Session Practice Problems 7**  **Complete in Lab**  ANOVA  Hypothesis Tests for Proportions |
| **Week 8 (10/17/2022–10/23/2022)** | | | |
| **Nonparametric Tests**  **Sample Size and Power**  **Screening** | **Nonparametric Tests**  Interpret the results of two-sample nonparametric tests for paired and independent samples  Explain the advantages and disadvantages of nonparametric tests of hypothesis  **Sample Size and Power**  Explain the types of errors which can occur when conducting a test of hypothesis  Explain the roles of power and sample size in hypothesis testing  Identify the factors needed to compute power and sample size for one- and two-sample tests of means and proportions  **Screening**  Identify characteristics of disease and screening tests which would be suitable for a screening program  Calculate sensitivity, specificity, predictive value positive, and predictive value negative  Define lead time bias and length time bias  Describe how the choice of a cutoff influences the sensitivity and specificity of the test | **Nonparametric Tests**  P,G,&M: 13.1-13.5  **Sample Size and Power**  P,G,&M: 10.4-10.6  **Screening**  Rothman: p.235-242 | **Watch prerecorded material and complete all self-assessment questions by 1:00 pm EDT on Thursday, 10/20/2022.**  **Homework 5:** Due 1:00 pm EDT on Thursday, 10/20/22  **Lab Session Practice Problems 8**  **Complete in Lab**  Nonparametric Tests  Sample Size and Power  Screening |
| **Week 9 (10/24/2022–10/30/2022)** | | | |
| **Selection Bias and Information Bias**  **Confounding**  **Directed Acyclic Graphs** | **Selection Bias and Information Bias**  Identify and differentiate between different sources of bias  Provide examples of selection and information bias  Describe methods to avoid selection bias and information bias  **Confounding**  Describe how confounding arises and how it impacts study results  Evaluate the properties of confounding using data and subject matter expertise  Anticipate the potential direction of confounding  Describe approaches to control for confounding in the design and analysis phase and the strengths and limitations of each approach  **Directed Acyclic Graphs**  Construct causal diagrams that depict the relationship between variables for examining a research question  Explain concepts presented in the archived edX course by Dr. Miguel Hernan: "Causal Diagrams: Draw Your Assumptions Before Your Conclusions" | **Selection Bias and Information Bias**  Rothman: p.124-135  **Confounding**  Rothman: p.136-145, p.188-192  **Directed Acyclic Graphs**  Lesson 1: Theory of casual diagrams:  <https://www.youtube.com/playlist?list=PLPHYAoFw7OElMOt-PHVOAjdNte_Ir331G>  Lesson 2: Applications of causal diagrams:  <https://www.youtube.com/playlist?list=PLPHYAoFw7OEmSfp7NtBq_Xcj-wDVl0IcG> | **Watch prerecorded material and complete all self-assessment questions by 1:00 pm EDT on Thursday, 10/27/2022.**  **Homework 6:** Due 1:00 pm EDT on Thursday, 10/27/22  **Lab Session Practice Problems 9**  **Complete in Lab**  Selection and Information Bias  Confounding  DAGs |
| **Week 10 (11/2/2022–11/6/2022)** | | | |
| **Effect Measure Modification**  **Stratified Data**  **Standardization** | **Effect Measure Modification**  Describe the concept of effect measure modification (EMM)  Evaluate the presence of EMM on the difference and ratio scale  Interpret results in the presence of EMM  **Stratified Data**  Construct, calculate and interpret point estimates, confidence intervals and p-values for stratified data  Construct, calculate and interpret statistical tests of heterogeneity  **Standardization**  Describe the purpose and assumptions of standardization  Define crude, stratum-specific, and adjusted rates  Describe how differences in population composition confound a comparison between crude rates | **Effect Measure Modification**  Rothman: p.198-209  **Stratified Data**  Rothman: p.176-188, P,G,&M: 15.2, 15.4  **Standardization**  P,G,&M: 3.1-3.2  Rothman: p.188-192 | **Watch prerecorded material and complete all self-assessment questions by 1:00 pm EDT on Thursday, 11/3/2022.**  **Homework 7:** Due 1:00 pm EDT on Thursday 11/3/2022  **Lab Session Practice Problems 10**  **Complete in Lab**  Effect Measure Modification  Stratified Data  Standardization |
| **Week 11 (11/7/2021–11/13/2022)** | | | |
| **Correlation**  **Introduction to Linear Regression** | **Correlation**  Describe the uses and limitations of correlation coefficients  Identify differences between Pearson and Spearman correlation coefficients  **Introduction to Linear Regression**  Explain the concept of a linear regression model | **Correlation**  P,G,&M: 16.1-16.3  **Introduction to Linear Regression**  P,G,&M: 17.1 | **Watch prerecorded material and complete all self-assessment questions by 1:00 pm EST on Thursday, 11/10/2022.**  **Examination 2**  **Tuesday, 11/8/2022 *or***  **Thursday, 11/10/2022**  The exam will be in the classroom from 8:00 am until 9:30 am.  **Tuesday, 11/8/2022**  **Lab Session Practice Problems 11:**  **Complete in Lab**  Correlation  Introduction to Linear Regression |
| **Examination 2: 11/8/2022 or 11/10/2022 – See details above regarding timing.** | | | |
|  | | | |
| **Week 12 (11/14/2022–11/20/2022)** | | | |
| **Linear Regression continued** | **Linear Regression continued**  Interpret and make inference on the coefficients of a linear regression model with a single continuous explanatory variable  Explain how a regression model is used for prediction  Assess the strength of the linear relationship using the coefficient of determination  Identify assumptions of the linear regression model, and use graphical techniques to evaluate these assumptions  Describe the utility of variable transformations  Interpret and make inference on coefficients of a multivariable linear regression model  Describe how indicator variables are used to model the effects of categorical predictors  Describe how linear regression models can be used to control for confounding  Test for interaction / effect measure modification using a regression model | **Linear Regression continued**  P,G,&M: 17.2-17.3  P,G,&M: 18.1-18.3 | **Watch prerecorded material and complete all self-assessment questions by 1:00 pm EST on Thursday, 11/17/2022.**  **Homework 8:** Due 1:00 pm EST on Thursday, 11/17/22  **Lab Session Practice Problems 12**  **Complete in Lab**  Linear Regression |
| **Week 13 (11/21/2022–11/27/2022) Thanksgiving Week** | | | |
| **Introduction to Logistic Regression** | **Introduction to Logistic Regression**  Explain the concept of a logistic regression model  Differentiate between linear and logistic regression models  Interpret and make inference on the coefficients of a logistic regression model  Explain how a regression model is used for prediction | **Introduction to Logistic Regression**  P,G,&M: 19.1 | **Watch prerecorded material and complete all self-assessment questions by 1:00 pm EST on Sunday, 11/27/2022.** |
| **Week 14 (11/28/2022–12/4/2022)** | | | |
| **Logistic Regression continued**  **Application of Regression Models** | **Logistic Regression continued**  Interpret and make inference on coefficients of a multivariable logistic regression model  Describe how logistic regression models can be used to control for confounding  Test for interaction / effect measure modification using a regression model  Evaluate discrimination of a model using the c statistic  **Application of Regression Models**  List the advantages and disadvantages of using regression models compared to tabular analyses of crude and stratified data  Using logistic regression as an example, calculate and interpret associations using different types of data  Explain how to examine the presence of effect measure modification with regression models | **Logistic Regression continued**  P,G,&M: 19.2-19.3, 19.5-19.6  **Application of Regression Models**  Rothman: p.211-234 | **Watch prerecorded material and complete all self-assessment questions by 1:00 pm EST on Thursday, 12/1/2022.**  **Homework 9:** Due 1:00 pm EST on Thursday, 12/1/2022  **Lab Session Practice Problems 13**  **Complete in Lab**  Logistic Regression |
| **Week 15 (12/5/2022–12/11/2022)** | | | |
| **Experimental Studies**  **Survival Analysis** | **Experimental Studies**  Differentiate between observational and interventional studies with respect to validity, generalizability, ethics, and feasibility  Explain the goal of randomization with regards to causal inference  Analyze and interpret data from experimental studies  **Survival Analysis**  Describe applications of a life table  Explain the concept of years of potential life lost  Explain how time to event data differs from other types of data  Define the concept of censoring  Construct and interpret a Kaplan-Meier estimate of a survival curve  Interpret the results of a log-rank test for comparing survival curves  Explain the objective of a Cox proportional hazards model  Interpret a hazard ratio estimated using a Cox proportional hazards model | **Experimental Studies**  Rothman: p.71-75, p.242-249  **Survival Analysis**  P,G,&M: 4.2-4.3, 20.1-20.4 | **Watch prerecorded material and complete all self-assessment questions by 1:00 pm EST on Thursday, 12/8/2022.**  **Homework 10:** Due 1:00 pm EST on Thursday, 12/8/2022  **Lab Session Practice Problems 14**  **Complete in Lab**  Experimental Studies  Survival Analysis |
| **Week 16 (12/12/2022–12/16/2022)** | | | |
| **Review** | Review course content |  | **Examination 3**  **Tuesday, 12/13/2022 *or***  **Thursday, 12/15/2022**  The exam will be in the classroom from 8:00 am until 9:30 am. |
| **Examination 3: 12/13/2022 or 12/15/2022 – See details above regarding timing.** | | | |

Please note, session topics and activities may be subject to change during the course.