Statistics 2: Introduction to Statistics

UC Berkeley, Fall 2024

> **Note**
>
> Please note that this syllabus is subject to change and may be updated during the semester.

Course Details

- Course format: online
- Four (4) semester credits

Course Description

This course will introduce you to statistical reasoning, which is a framework for learning from data. We'll start with descriptive statistics that allow us to summarize important features of data we've collected without requiring much in the way of mathematical theory or assumptions. Then we'll move on to probability and inference.

Probability gives us a language for talking about how random variability arises in the data we collect. Inference uses the language of probability to help us determine what we can learn from our data despite the presence of that random variability.

Prerequisites

There are no prior course requirements.

Course Learning Objectives

After successfully completing this course, you will be able to:

- Identify and create appropriate visual and numerical summaries for different data types
- Describe where an observation lies relative to a distribution (e.g. using standard units and/or percentile ranks)
- Apply the normal approximation to data, and identify situations when this is or isn't appropriate
- Calculate chances of events using basic laws of probability
- Understand what the sampling distribution (probability histogram) of a statistic is
- Carry out and interpret standard statistical inference techniques including correlation, regression, plug-in estimation, confidence intervals, and hypothesis testing
- Explain the relationship between the sampling distribution and the precision of standard statistical inference techniques
- Evaluate and critique reported statistical results in the real world

Instructor Information and Communication

Course Instructor

- Adam Lucas

Graduate Student Instructors (GSIs)

- Isabel Moreno: (labs 105,106)
- Frederik Stihler: (labs 101,103)
- Lauren Murai: (labs 102,104)
- Jasmine Khamuani (labs 107,110)
- Yimeng Yin (labs 108,109)
- Austin Tao (labs 111,117)
- Cancan Jin (labs 112, 113)
- Minh Le (labs 114, 116)
- Yichen Xu (labs 115, 118)

While the instructor will interact with the whole class and will oversee all activities and grading, as well as being available to resolve any issues that may arise, the GSIs will be your main point of contact. Your GSIs are responsible for assisting you directly with questions about assignments and course requirements. The GSIs will also facilitate ongoing discussion and interaction with you on major topics in each module.

GSI Office Hours

The GSIs will offer office hours in person or via Zoom. These office hours allow for synchronous interaction with the GSIs and are a good opportunity to discuss your questions relevant to the course. You may attend any office hours, not only those for your assigned GSI and instructor. Times and links to the appropriate Zoom meeting will be available in bCourses. See bCourses for the time and place.

Question & Answer Sessions

The instructor will hold in person Q&A sessions to address any questions you have on the homework and other course related questions. See bCourses for the time and place.

Course Mail

Please contact your instructor and GSI using your Berkeley email (not through bCourses).

Course Help

You're not alone in this course; the instructor and GSIs are here to support you as you learn the material. It's expected that some aspects of this course will take time to grasp, and the best way to grasp challenging material is to ask questions.

We will be using Ed Discussion as our discussion forum. A link to ED is on bCourses. You can also reach out to the course staff in office hours, Q&A sessions, lab sections, and/or via email.

If you need to contact the instructor through email, please ensure to also include your assigned GSI in the thread, unless you need to discuss a personal matter.

Students with Disabilities

If you require course accommodations due to a physical, emotional, or learning disability, contact UC Berkeley's Disabled Students' Program (DSP). Notify the instructor and GSI through course email of the accommodations you would like to use. You must have a Letter of Accommodation on file with UC Berkeley to have accommodations made in the course.

UC Berkeley is committed to providing robust educational experiences for all learners. With this goal in mind, we have enabled the ALLY tool for this course. You can now automatically generate "Alternative Formats" for course files and bCourses Pages. Depending on the context, these formats can include Tagged PDF, HTML, BeeLine Reader, Electronic Braille, ePub, Immersive Reader, MP3, and translated versions. For more information, watch the video entitled, "Ally Tutorial".

Course Materials and Technical Requirements

Textbook / Required Materials

The required textbook is Statistics: Fourth Edition by Freedman, Pisani, and Purves. The paperback version of the fourth edition is fine and usually much cheaper, as is the international version. You will find the assigned readings posted on bCourses.

Technical Requirements

This course is built on a Learning Management System (LMS) called Canvas (UC Berkeley's instance of Canvas is called bCourses). You'll need to meet these computer specifications to participate within this online platform.

You will also need a calculator or calculator app that can raise numbers to a power and do square roots. Statistical or graphing calculators are not necessary.

Technical Support

If you're having technical difficulties, please alert one of the GSIs immediately. However, understand that neither the GSIs, nor the instructor can assist you with technical problems. You must call or email tech support to resolve any technical issues.

To contact tech support, click on the "Help" button on the bottom left of the global navigation menu in bCourses. Be sure to document all interactions (save emails and transaction numbers).

Learning Activities and Assignments

You're expected to fully participate in all the course activities described below. A weekly course schedule is provided on the Syllabus page in bCourses. Check bCourses for specific assignment due dates.

Asynchronous/Synchronous Format

Some of your learning activities will be asynchronous (i.e., lectures, readings, assignments), which you can complete at any time before the due date. Other activities will be synchronous (i.e., lab sections, office hours, Q&A sessions, final exam, etc.), which you will attend at a specific time.

All times listed are Pacific Time—please adjust for your time zone. If you prefer, you can set your own time zone to display throughout bCourses.

Modules and Parts

This course is organized into Modules, each of which is further divided into Parts. You can think of a Part as being equivalent to the material covered in a traditional in-person lecture. Each Part has other learning activities associated with it. As you move through the course materials online, you are expected to fully participate in all the course activities described below.

Reading Assignments and Suggested Practice Problems

Most Parts have a reading assignment and suggested practice problems from the textbook. You may prefer to do these either before or after watching the lecture videos. The answers to the suggested practice problems are in the back of the textbook. They are for your practice only and are not graded.

Lecture Videos and Discussion Forums

Each Part has lecture videos for you to watch. All lectures have been pre-recorded, for you to watch on your own: *there are no live lecture meetings for this course.* Recorded lectures support your readings and assignments but also contain additional material that may be included in the exams. You're expected to take notes while viewing the lectures as you would in a regular classroom. **Please watch the lectures for each Part before the following day's lab section**, as indicated in the schedule on the Syllabus page in bCourses.

The slides for the lecture videos are also posted. I suggest you download the slides and use them to follow along and add your own notes while watching the videos. Post your comments and questions on ED and the instructor/GSIs will respond.

Comprehension Quizzes (10%)

Each Part has a short quiz associated with it, to be completed after watching the lecture videos. Quizzes are due by 11:59 pm Pacific Time on the days indicated in the schedule. **It's important to keep up with the lecture videos and quizzes so that you're ready to participate in labs on the scheduled dates.**

Labs/Section Participation (20%)

Labs allow you to practice what you are learning by working on problems in a collaborative environment. There are two ways to participate in labs. You may either attend synchronously in person, or you may complete the lab problems on your own and upload a copy of your work and solutions. If you choose to complete the lab on your own, it is strongly recommended that you check your answers against the solutions (posted to bCourses the following day) and bring any questions to office hours.

Labs are graded based on participation, with 0-2 points awarded for each lab. If you attend your lab and participate according to the guidelines your GSI describes, you will automatically get full credit. If you do the lab problems on your own, you will get credit based on the percentage of lab problems you have made a good-faith effort to complete, with 2 points for attempting everything, 1 point for attempting at least half the problems, and 0 points otherwise. You should upload your work by **11:59PM PST** on the day of the lab.

Your lab section number will be the one in which you enrolled. You can view this section number by exploring the "Section" column within the "People" page.

Homework Assignments (30%)

There are 5 assignments. We use the Gradescope platform for submitting and grading assignments. You can log in directly from the Gradescope link in bCourses left navigation. Further details and instructions about uploading your assignments are included on bCourses. Assignments are due at **11:59PM PST** on the dates indicated in the course schedule (see the Syllabus page in bCourses). Because of the size of the class we cannot accept late assignments. Instead we allow all students to drop one assignment for emergencies.

Mastery Quizzes (25%)

Mastery quizzes are designed to evaluate and reinforce mastery of key skills and are held in place of midterm exams. They are held on bCourses, have a default time limit of 50 minutes, and consist of multiple-choice questions. These quizzes are open book, open notes, but working with other students is prohibited. Quiz problems are also randomized to deter cheating. You should finish the Mastery quizzes by **11:59PM PST** on the dates indicated in the course schedule (see the Syllabus page in bCourses).

Final Exam (15%)

The final exam will be held online (on bCourses) on **December 19, 2024, 11:30–2:30pm**. If you have a conflict with this time, please inform your instructors **at least one month in advance** to arrange an alternative time to take the exam. The final will have a similar format to the Mastery Quizzes, but it will be cumulative—covering material from the entire course—and with more questions.

Grading

Your final course grade will be calculated as follows:

Table 1: Table 1: Final Grade Percentages

Category	Percentage of Grade	Notes
Comprehension Quizzes (24 total)	10%	No extensions. The lowest 5 quiz scores will be dropped.
Labs (22 total)	20%	No extensions. The lowest 5 lab scores will be dropped.
Assignments (5 total)	30%	No extensions. The lowest 1 assignment scores will be dropped.
Mastery Quizzes (5 total)	25%	You can ask for 1 extension if necessary.
Final Exam	15%	N/A

Course letter grades will be assigned based on the following intervals:

Table 2: Table 2: Grade Intervals

Letter Grade	Percentage Interval
A+/A/A-	100-90%
B+/B/B-	90-80%
C+/C/C-	80-60%
F	Below 60%

If you are taking the class pass-fail, the cutoff for passing is 60%. The bCourses grade book uses the grading scheme above, so you can always check your current course grade there. **We do not round up when calculating letter grades, even when the score is very close.** This is because grades tend to have a continuous distribution, so that no matter where we set the threshold, there are always going to be students who just miss it.

Strategies for Successful Learning

Take Care of Yourself

Do your best to maintain a healthy lifestyle this semester by eating well, exercising, getting enough sleep, and taking time to recharge your mental health. Taking time to care for yourself, and avoiding burnout, will help you achieve your academic, professional, and personal goals.

If you start to feel overwhelmed, be kind to yourself and reach out for support. Remember that seeking help is a courageous thing to do—for yourself and for those who care about you.

Support Resources include emotional, physical, safety, social, and other basic wellbeing resources for students. Academic resources can be found at the Student Learning Center and English Language Resource sites. Berkeley's Office of Emergency Management has resources to prepare for emergencies.

Course Policies

Late Work Policy

Comprehension quizzes and labs and assignments are not accepted for credit late, but we automatically drop the **lowest 5** comprehension quizzes and labs and **lowest 1** assignment when calculating your course grade. This is meant to cover a variety of issues that can arise, including technical issues, illness, work conflicts, family concerns, etc. This adjustment is already built into the bCourses gradebook. If you add the course late and miss the first assignment deadline you will need to use your 1 assignment drop on that assignment.

DSP students whose accommodations include extensions on assignments will automatically receive a 48-hour extension on assignments on Gradescope. If you don't see this set up for you please contact the instructor.

Scheduling Conflicts

Please notify the instructor in writing by the second week of the term about any known or potential academic or extracurricular conflicts. I will try my best to help you with making accommodations but cannot promise them in all cases.

Academic Integrity

Any assignment or exam submitted by you is presumed to be your own original work. While you are encouraged to study together and to work together on the lab problems and assignments, there is no value to you in copying someone else's work. The standard for what constitutes your "own original work" is that if asked, you should be able to convincingly explain out loud what you have written to the instructor or a GSI. As you work on assignments, you can talk to each other about the problems, but you need to get to the point where the answer is something you have internalized as something you, yourself, understand well enough to convincingly explain to someone else. In cases where cheating is suspected, the instructor or GSI may ask you to make this verbal explanation.

To copy text or ideas from another source without appropriate reference is plagiarism and will result in a failing grade for your assignment and usually further disciplinary action. If you're unclear about the expectations for completing an assignment or taking a test or examination, be sure to seek clarification from your instructor or GSI beforehand. For additional information, read these UC Berkeley Library guides on How to Avoid Plagiarism and Statistics: Writing + Citing (the latter includes some recommended citation management tools).

As a member of the campus community, you're expected to demonstrate integrity in all of your academic endeavors and will be evaluated on your own merits. The consequences of cheating and academic dishonesty—including a formal discipline file, possible loss of future internship, scholarship, or employment opportunities, and denial of admission to graduate school—are simply not worth it. Read more about Berkeley's Honor Code.

Incomplete Course Grade

Students who have substantially completed the course but for serious extenuating circumstances, are unable to complete the remaining course activities, may request an Incomplete grade. This request must be submitted in writing to the GSI and instructor. You must provide verifiable documentation for the seriousness of the extenuating circumstances.

Refer to the Office of the Registrar's website for more information on the university's policy on Incomplete Grades.

End of Course Evaluation

UC Berkeley is committed to improving our online courses and instruction. Before your course ends, please take a few minutes to participate in the course evaluation. We are interested in your online learning experience, and your feedback will help us plan for the future and make improvements. The evaluation does not request any personal information, and your responses will remain strictly confidential. Information about the course evaluation will be made available in bCourses.

Course Outline

Note: All dates are included in the bCourses calendar and in the Schedule of Activities on the Syllabus page in bCourses.

Module 1: Data collection and variables (Weeks 1-2)

Topics: Data collection; experimental design; limitations of observational studies; randomized clinical trials; confounding variables; variable types

Lecture video – How this course is taught Lecture video - Statistics: The big picture

Comprehension Quiz 1 Lecture video - Data collection Lecture video - Variable types Comprehension Quiz 2

Lab 1: Data collection and variable types

Module 2: Describing one quantitative variable (Week 3)

Topics: Histograms, standard measures of center and spread, summarizing shapes of distributions; percentiles; standard units (z- scores)

Lecture video - Interpreting histograms Lecture video - Crowding and the density scale Lecture video - Drawing a histogram Comprehension Quiz 3

Lab 2: Histograms

Lecture video - Describing the center Lecture video - The interquartile range Lecture video - The standard deviation Comprehension Quiz 4

Lab 3: Summary statistics Assignment 1 (covers Modules 1 & 2)

Module 3: The normal approximation (Week 4)

Topics: Normal curve calculations using a table; normal approximation to data; percentiles and percentile ranks using the normal approximation; linear transformations

Lecture video - Standard units

Lecture video - The normal curve

Lecture video - The normal approximation to data Lab 4: Standard units and the normal curve Comprehension Quiz 5

Lecture video - Normal approximation caveats and percentiles Lecture video - Three ways to represent an observation Lecture video - Linear transformations

Comprehension Quiz 6

Lab 5: Normal approximation and percentiles Mastery Quiz 1 (covers Module 3)

Module 4: Describing two quantitative variables (Weeks 5-7)

Topics: Scatterplots; correlation coefficient; regression line; errors in regression; normal approximations in regression; regression effect

Lecture video - The correlation coefficient Lecture video - Advanced use of scatterplots

Lecture video - Use and misuse of scatterplots and correlation Comprehension Quiz 7

Lab 6: Correlation

Lecture video - The regression line

Lecture video - The ”two variable roadmap” Lecture video - The regression effect Comprehension Quiz 8

Lab 7: Regression - prediction and estimation

Mastery Quiz 2 (covers Module 4, Part 2 -- material from Lecture 8.1 through Lab 7)

Lecture video - Errors in regression Lecture video - RMS error

Lecture video - Normal approximations in regression Comprehension Quiz 9

Lab 8: Regression - Roadmap, regression effect, and RMS error Lecture video - Intercept and slope of the regression line Lecture video - Regression review

Comprehension Quiz 10

Lab 9: Regression - Residual plots, “slice problems,” and review Assignment 2 (covers Module 4)

Module 5: The language of probability (Weeks 8-9)

Topics: Definitions of probability; probability rules; two-way tables; conditional probability; independent and mutually exclusive events; multiplication and addition rules

Lecture video - Why study probability? Lecture video - Defining probability

Lecture video - Basic rules of probability and conditional probability Comprehension Quiz 11

Lab 10: Basic probability rules Lecture video - The multiplication rule Lecture video - Independent events

Lecture video - Checking for independence Comprehension Quiz 12

Mastery Quiz 3 (Covers Module 5, Parts 1 & 2 -- material from

Lectures 11 and 12)

Lab 11: Independence

Lecture video - Mutually exclusive events Lecture video - The addition rule

Lecture video - The complement rule, revisited Comprehension Quiz 13

Lab 12: Independence, mutual exclusivity, and review Assignment 2 (Covers Module 5)

Module 6: Chance processes and the box model (Weeks 10- 11)

Topics: Box model for repeated experiments (probability distributions); expected value and standard error calculations; Central Limit Theorem

Lecture video - Chance processes and repeated experiments Lecture video - The box model

Lecture video - Box model practice: Roulette Comprehension Quiz 14

Lab 13: Box models for chance processes Lecture video - Probability histograms

Lecture video - Expected value and standard error of the sum of draws

Lecture video - The Central Limit Theorem for the sum of draws Lecture video - Applying the CLT

Comprehension Quiz 15

Lab 14: Expected value and standard error Lecture video - Boxes for classifying and counting

Lecture video - Probability histograms of 0-1 boxes Lecture video - Simulations for box models

Comprehension Quiz 16

Lab 15: Classifying and counting Mastery Quiz 4 (Covers Module 6)

Module 7: Sampling from a population (Weeks 11-12)

Topics: Simple random sample; critiquing design of sample surveys; accuracy of the sample percentage

Lecture video - Sample surveys and biases in surveys

Lecture video - Case study: 2016 and 2020 U.S. Presidential polls Lecture video - Other survey methods

Comprehension Quiz 17 Lab 16: Sample surveys

Lecture video - Probability histograms of statistics Lecture video - The sample percentage

Lecture video - Accuracy of the sample percentage Comprehension Quiz 18

Lab 17: The sample percentage Assignment 4: Covers Modules 6 & 7

Module 8: Estimation and confidence intervals (Weeks 12- 13)

Topics: Plug-in method of estimation; calculating and interpreting confidence intervals

Lecture video - Estimation and the plug-in method

Lecture video - Confidence intervals for the population percentage Lecture video - Interpreting confidence intervals

Comprehension Quiz 19

Lab 18: Estimation and confidence intervals for the population percentage

Lecture video - Probability histogram of the sample average Lecture video - Confidence intervals for the population average Lecture video - Estimation and confidence interval review

Comprehension Quiz 20

Lab 19: Estimation and confidence intervals for the population average

Mastery Quiz 5 (Covers Module 8)

Module 9: Tests of significance (Weeks 13-15)

Topics: Hypothesis testing framework; one and two sample z tests; p-values; one and two tailed tests; limitations of hypothesis testing

Lecture video - Tests of significance

Lecture video - The z test and computing the p-value Lecture video - One-tailed and two-tailed tests Comprehension Quiz 21

Lab 20: Tests of significance and the z test statistic Lecture video - Generalizing the z test

Lecture video - Two-sample z tests for independent samples Lecture video - Two-sample z tests for experiments Comprehension Quiz 22

Lab 21: Tests for comparing groups Lecture video - Interpreting p-values

Lecture video - The wide world of test statistics Lecture video - Hypothesis testing caveats

Comprehension Quiz 23

Lab 22: Testing caveats and testing review

Module 10: Statistics: The big picture (Week 15)

Topics: Identifying appropriate statistical techniques for a given applied problem; knowing when NOT to use a statistical technique; using probability to describe the limitations of statistical inference

Lecture video - Topic review and synthesizing questions, part 1 Lecture video - Synthesizing questions, part 2

Lecture video - Main ideas of the course Comprehension Quiz 24

Assignment 5 (Covers Modules 8-10)

RRR Week activities (Week 15)

Activities TBD

Final Exam