

Syllabus: Stats UB-103

Professor Joshua Loftus

Fall 2019

E-mail: loftus@nyu.edu

Office hours: email for appointment

Office: KMC 8-60

TF: Vinu Abeywickrama

TF E-mail: va774@stern.nyu.edu

TF Office hours: TBD

TF Office: KMC 8-174

Web: <http://joshualoftus.com/page/introstats/>

Class hours: Tu/Th/F 3:30-4:45pm

Classroom: Tisch 201

Course Description

This course examines modern statistical methods as a basis for decision making in the face of uncertainty. Topics include probability theory, discrete and continuous distributions, hypothesis testing, estimation, and statistical quality control. With the aid of computers, these statistical methods are used to analyze data. Also presented are an introduction to statistical models and their application to decision making. Topics include the simple linear regression model, inference in regression analysis, sensitivity analysis, and multiple regression analysis.

Exam schedule

Tentative dates for in-class midterms: TBD.

Required Materials

- We will make use of a number of free resources linked on the course page.
- On certain days students will be asked to bring a computer to class.

Tentative Outline

Items listed under “topics” and the last section, on machine learning, should be considered outside of the core concepts of the class. They are included to indicate the context and importance of the core concepts in the “modern” era of data science and machine learning. Course assessments will not require any advanced knowledge of these topics.

1. Measurement and experiments

- Reliability and validity
- Randomized controlled trials
- Observational studies, confounding

2. Description and visualization

- Summarizing data
- Histograms
- Other visualizations

3. Probability basics

- History, games, sets, counting
- Random variables, expectation, variance
- Theorems: Chebyshev, Bayes

4. Probability in statistics

- Estimation, bias and variance
- Sampling distributions, central limit theorem
- Topics: Stein's paradox, error propagation

5. Confidence intervals and tests

- Intervals, error bars, t -distribution
- Interpretation
- Hypotheses, falsification
- Bootstrap and resampling methods
- Topics: multiple testing

6. Regression

- Covariance, correlation, simple regression
- Prediction, intervals, tests, and interpretation
- Non-linearity, transformations, discrete predictors
- Multiple regression
- Topics: logistic regression, mixed effects models, causal inference

7. Machine learning

- Model selection, complexity, validation
- Classification

The order of this outline is not intended to be chronological. We may revisit certain topics at various points throughout the course.

Course Policies

General Rules

- Be on time to class. Attendance is mandatory and part of the grade.
- During lectures, everyone is encouraged to ask and answer questions and actively participate, but please raise your hand first. After you have already spoken once try to be mindful of giving other students time to speak as well.
- During lectures, electronic devices may be used only for taking notes unless otherwise specified. If you are distracting your classmates or not paying attention I will stop the lecture and ask you to put the device away.
- Late submission of assignments may be permitted with prior notice on a case-by-case basis, but is generally highly discouraged. There is no guarantee of acceptance of late submissions even for partial credit.
- Students are encouraged to form study groups review course material together, but unless otherwise specified every assignment must be completed alone. Receiving help or copying answers from other students, the internet, or any other source may be considered plagiarism, a very serious violation that may be investigated by the university.

These rules and the Code of Conduct mentioned below exist for your benefit and that of your classmates by helping create a positive and constructive environment which is a safe space for everyone to learn.

Grading Policy

Stern policies regarding grades for undergraduate core courses apply to this class.

To avoid grades which are heavily biased toward specific skills like test-taking, we will use a variety of assessments including in-class tests, problem sets (homework), in-class participation, online participation, and data analysis projects. Before determining final grade weights for these different assessments we will have a survey so everyone can have some input into the decision. Attendance is included under in-class participation.

Code of Conduct

All students are expected to follow the Stern Code of Conduct <http://www.stern.nyu.edu/uc/codeofconduct>. A student's responsibilities include, but are not limited to, the following:

- A duty to acknowledge the work and efforts of others when submitting work as one's own. Ideas, data, direct quotations, paraphrasing, creative expression, or any other incorporation of the work of others must be clearly referenced.
- A duty to exercise the utmost integrity when preparing for and completing examinations, including an obligation to report any observed violations.
- To minimize the temptation for copying or sharing during an exam, there will be multiple versions of every exam, and the seating order will be randomly assigned.

Accommodations for Different Abilities

If you are qualified for academic accommodation of any kind during this course, you must notify me at the beginning of the course and provide a letter from the Moses Center for Students with Disabilities (CSD, 998-4980, www.nyu.edu/csd) verifying your registration and outlining the accommodations they recommend. If you will need to take an exam at the CSD, you must submit a completed Exam Accommodations Form to them at least one week prior to the scheduled exam time to be guaranteed accommodation.

Health and Wellness

To access the University's health and mental health resources, contact the NYU Wellness Exchange. You can call its private hotline (212-443-9999), available 24 hours a day, seven days a week, to reach out to a professional who can help to address day-to-day challenges as well as other health-related concerns.