95-796, Statistics for IT Managers & Intermediate Statistics (90-777) (Fall 2024)

Course Description:

This is an introductory course in data analysis and statistical inference. Its objective is to provide individuals who aspire to enter management or policy analysis positions with the basic statistical tools for analyzing and interpreting data. For students intending to take more advanced courses in statistics and machine learning its purpose is also to ensure that you are well grounded in the fundamentals of statistics. The course is divided into three distinct modules: descriptive statistics, statistical inference, and regression analysis. The emphasis of the lectures on descriptive statistics is the calculation and interpretation of summary statistical measures and graphical methods for describing raw data. The sessions on statistical inference are designed to provide you with the background for executing and interpreting hypothesis tests and confidence intervals. The final component of the course focuses on regression analysis, a widely used statistical methodology. Throughout the course you will regularly analyze data relevant to management and policy analysis using a statistical platform developed here at CMU called ISLE.

Instructors

Eli Ben-Michael (ebenmichael@cmu.edu)
Office hours: 10am – 11am Monday and 2pm – 3pm Tuesday. HbH 2118D

Andy Garin (agarin@andrew.cmu.edu)
Office hours: 2pm–3pm Monday and 4–5PM Thursday. HbH 2224

Daniel Nagin (dn03@andrew.cmu.edu)
Office hours: 3pm-4pm Monday and 11am-12pm Friday HbH 2213

Schedule 95-796:

A: Mon-Wed 3:30pm-4:50pm, HbH 1202
    Fri 9:30am-10:50 am, HbH 1206
B: Mon-Wed 2pm-3:20pm HbH 1206
    Fri 9:30am-10:50am-, A301
C: Mon-Wed 9:30am-10:50am HbH 1206
    Fri 3:30pm-4:50pm, A301
D: Tue-Thu 3:30pm-4:50pm HbH 1206
    Fri 9:30am-10:50pm, A301
E: Tue-Thu 2:00pm-3:20pm HbH 1204
    Fri 9:30am-10:50am A301
F: Mon-Wed 11:00am-12:20pm HbH 1206
    Fri 3:30pm-4:50pm-11:30am, A301

Schedule 90-777:
A: Mon-Wed 9:30am - 10:50am HbH 1202
    Fri 3:30pm - 4:50pm HbH A301

B: Mon-Wed 5:00pm - 6:20pm HbH 1202
    Fri 3:30pm - 4:50pm HbH 1206

Teaching Assistants:

Name	Email	Student Last Name for Assignment Clarification Questions
Andre Ribeiro Cardoso	acardoso@andrew.cmu.edu	**A-H**
Lindsay Graff	lgraff@andrew.cmu.edu	**I-M**
Logan Crowl	lcrowl@andrew.cmu.edu	**N-U**
Johnna Sundberg	jsundber@andrew.cmu.edu	**V-Z**

TA Remote Office Hours begin on Tuesday, Sept. 3

TA Hours (All Remote)
- Logan: Monday 10:30am-noon
- Lindsay: Monday 7pm-8:30pm
- Johnna: Tuesday 10:30am-noon
- Andre: Tuesday 5pm-7pm

Course Materials

- Statistics for Business and Economics, 11th Edition, by McClave, Benson, and Sincich.
  *** *You should be able to buy a used copy cheaply on Amazon, etc. It's fine to use the newer 12th Edition, or the older 9th or 10th Edition instead, but the section numbers may be slightly different! ****
- Lecture slides, assignments, datasets, and any other class materials are posted on Canvas and ISLE.
- **ISLE** You will receive instructions on accessing ISLE. We highly recommend that you use either the Chrome or Firefox browsers to access ISLE. Do **not** use Internet Explorer. The platform developer Philipp Burckhardt (pburckhardt@cmu.edu) has graciously agreed to allow you to contact him directly to resolve any technical malfunctioning with ISLE that you may encounter. Do not, however, direct inquiries about course material or ISLE functionality (e.g., how to do a hypothesis test) to Philipp. Such questions should be directed to a TA.

Evaluation Method

Grades will be based on:
Completed problem worksheets for problem solving sessions in weeks 2, 3, 5, and 6: 20% (5% each)
Four homework assignments (**completed without the assistance of others**): 20% (5% each)
Mid-term quiz: 25%
Final exam: 35%

Important Note on problem worksheets: You will be required to submit your solutions by Midnight EST on the Thursday _prior_ to the Friday problem solving session. *The worksheet will _not _be evaluated on whether your solutions is correct.* You will receive full credit for demonstrating a sincere attempt at solving the problem. If you work in a group, you must individually submit the group's solution.

Reminder: All assignments must be submitted through Gradescope no later than the designated time. Late homework or worksheet assignments **will not be accepted**. **Please refer to the Gradescope guide located on Canvas for instructions on how to submit your work properly. Failure to follow directions can lead to mis-graded assignments!**

Type	Due date
Worksheet Week 2	Midnight EST, Thursday, Sept. 5
Homework 1	Midnight EST Tuesday Sept. 10
Worksheet Week 3	Midnight EST Thursday, Sept. 12
Homework 2	Midnight EST Tuesday Sept. 17
Mid-term	Scheduled review session on Friday, Sept 20
Worksheet week 5	Midnight EST Thursday, Sept. 26
Homework 3	Midnight EST Tuesday, Oct. 1
Worksheet week 6	Midnight EST Thursday, Oct. 3
Homework 4	Midnight EST Tuesday Oct 8
Final Exam	Friday, Oct. 11 Time TBD

Frequently asked questions:

Do you curve our grades? Answer: Students' numeric grades for homework, quizzes and the final exam will not be curved. However, the instructor will apply a curve when deciding on the final assignment of letter grades in order to meet Heinz College standards for mean student GPA in core courses.

Cheating and Plagiarism Notice

Completed homework assignments must be your work alone. Also, your answers on the final exam and mid-term quiz must reflect your work alone. Sanctions for cheating include lowering your grade which may include failing the course. In egregious instances, the instructors may recommend the termination of your enrollment at CMU. Don't cheat! It is not worth it.

Use of generative AI tools

To best support your own learning, you should complete all graded assignments in this course without the use of generative artificial intelligence (AI). Refrain from using AI tools to generate any content (text, video, audio, images, code, etc.) for worksheets, homework, or exams. Passing off any AI generated content as your own (e.g., cutting and pasting content into written assignments, or paraphrasing AI content) constitutes a violation of CMU's academic integrity policy. If you have any questions about using generative AI in this course please discuss with your assigned instructor—Prof. Ben-Michael, Garin or Nagin.

Late Work Policy: You are expected to turn in all work on time. **Late assignments WILL NOT be accepted.**

Re-grade Policy: Any requests for a re-grade must be submitted *in writing* via Gradescope within seven (7) days from when the graded assignment is <u>returned</u>. **You must include a clear written explanation of why you think your answer was not properly graded, within the comment section of Gradescope**. **Regrade requests without such an explanation will be automatically denied.** If you are dissatisfied with the decision, you should then redirect your appeal to your assigned TA. If you are still dissatisfied you can contact your assigned instructor— Prof. Ben-Michael, Garin or Nagin. If a re-grade request is submitted, we may re-grade your entire assignment, which may either raise or lower your score. In general, we will only raise scores in cases where we have made an error in grading, and all decisions on re-grade requests are final.

E-mail Questions Policy: To balance the workload fairly among the teaching assistants, and to ensure a reasonable response time for questions received via e-mail, each student has been assigned one TA as a "first contact". This is the person that you should e-mail first with any questions that you might have regarding the course material, policies, etc. We will do our best to answer questions within 24 hours, or 48 hours on weekends. **First contacts are assigned based on your last name. For lasts names beginning with A-H email** Andre Ribeiro Cardoso acardoso@andrew.cmu.edu**, for I-M email** Lindsay Graff lgraff@andrew.cmu.edu**, N-U email** Logan Crowl (lcrowl@andrew.cmu.edu), **for V-Z email and Johnna Sundberg (jsundbe@andrew.cmu.edu) for V-Z**

Accommodations for Students with Disabilities:

If you have a disability and have an accommodations letter from the Disability Resources office, we encourage you to discuss your accommodations and needs with your assigned instructor as early in the semester as possible. We will work with you to ensure that accommodations are provided as appropriate. If you suspect that you may have a disability and would benefit from accommodations but are not yet registered with the Office of Disability Resources, we encourage you to contact them at access@andrew.cmu.edu.

Basic Mental Health:

As a student, you may experience a range of challenges that can interfere with learning, such as strained relationships, increased anxiety, substance use, feeling down, difficulty concentrating and/or lack of motivation. These mental health concerns or stressful events may diminish your academic performance and/or reduce your ability to participate in daily activities. CMU services are available, and treatment does work. You can learn more about confidential mental health services available on campus at: http://www.cmu.edu/counseling/. Support is always available (24/7) from Counseling and Psychological Services: 412-268-2922.

Course Outline

Module I: Descriptive Statistics and Probability (Chapters 1-4)

Week 1: Course Overview and Descriptive statistics (Aug. 26-Aug. 30)
Lectures:
Module 1 Course Introduction
Module 2a Numerical Summaries
Module 2b Graphical Methods
Module 3a Probability Basics
Module 3b Conditional Probability and Independence
Note: Friday review session will be split between lecturing on modules 3a and 3b as needed and a tutorial on use of the ISLE platform
Readings:
What is statistics? (1.1-1.3)
Types of data (1.5)
Random sampling (1.6)
Histograms (2.2)-
Measures of central tendency (2.4)
Measures of variability (2.5-2.6)
Box plots (2.8)

Week 2: Probability (Sept 4-6)

NO CLASS SEPT. 2 OR 3

Lectures:
Module 3c Law of Total Probability and Bayes Rule
Module 3d Discrete Random Variables
Reading:
Bayes' Theorem (3.8)
Discrete Random Variables (4,1-4.3)

==Week 2 worksheet due by Midnight EST, Thursday, Sept. 5==

Week 3: Random Variables and Confidence Intervals (Sept. 9-Sept. 13)

Lectures:
Module 3e Binomial Random Variables
Module 3f Continuous Random Variables
Module 3g Central Limit Theorem
Module 4a Confidence Intervals for μ

Readings:
Binomial random variables (4.3)
Continuous random variables (4.5)
The Uniform distribution (4.9)
The Normal distribution (4.6)
Central Limit Theorem (4.11)

Large-sample confidence intervals for the mean (5.2)

Hypothesis Testing and Inference (Chapters 5-7)

Week 4: Confidence intervals and Hypothesis Testing (Sept. 16-Sept. 20)

Lectures:
Module 4b Confidence Intervals for p
Module 4c Specifying the Null and Alternative Hypothesis
Readings:
Small-sample confidence intervals for the mean (5.3)
Large-sample confidence intervals for the population proportion (5.4)
Determining the sample size (5.5)
Introduction to hypothesis testing (6.1-6.2)

Week 5: Hypothesis testing (Cont.) & Regression Sept. 23-Sept. 27)

Lectures
Module 4d Conducting a Hypothesis Test
Module 4e Type I and II Errors, p-values
Module 4f Two Population Hypothesis Tests
Module 5a Simple Regression

Readings
Large-sample hypothesis tests for the mean (6.3)
p-values (6.4)
Small-sample hypothesis tests for the mean (6.5)
Large-sample hypothesis tests for the population proportion (6.6)
Comparing two population means (7.2)
Comparing two population proportions (7.4)
Linear models (10.1)
Least squares linear regression (10.2)

Module III: Regression (Chapters 10-11)

Week 6: Simple Regression and Multiple Regression (Sept. 30-Oct. 4)

Lectures:
Module 5b Simple Regression Continued
Module 5c Multiple Regression

Readings:
Estimating the error of the model (10.3)
Making inferences using the model (10.4)
Coefficients of correlation and determination (10.5)

Homework 3 due by midnight EST Tuesday, Oct. 1

Week 6 worksheet due by Midnight pm EST Thursday, Oct. 3

Week 7: Multiple regression (Oct. 7-Oct. 11)
Lectures:
Module 5d Multiple Regression Continued

Reading:
Multivariate linear models (11.1)
Least squares linear regression (11.2)
Estimating the error of the model (11.2)
Multiple coefficient of determination (11.3)

Homework 4 due by midnight EST Tuesday Oct. 8

Final exam Friday, Oct. 11th (Time TBD)