Welcome to DSCI 552: Statistical Inference and Computation I

Print to PDF

Contents

This course reviews classical and simulation-based techniques for estimation and hypothesis testing, including inference for means and proportions. We make particular emphasis on case studies and real data sets, as well as reproducible and transparent workflows when writing computer scripts for analysis and reports.

High-Level Goals

By the end of the course, students are expected to:

- Build a solid foundational understanding of frequentist Statistical Inference (computational and classical!).
- Become competent using `R` to perform computation for frequentist Statistical Inference.

Back to top

Learning Objectives

- Describe real-world examples of questions that can be answered with the statistical inference methods presented in this course (e.g., estimation, hypothesis testing) and apply inference skills and concepts to answer such questions.
- Explain what random and representative samples are and how they can influence estimation.
- Write computer scripts to perform estimation and hypothesis testing via simulation-based inference approaches, as well as by applying results from exact and approximate distributional theory.
- Interpret and explain results from confidence intervals and hypothesis tests.
- Compare the application of simulation-based inference approaches with the application of exact and approximate distributional results.
- Effectively visualize point estimates and different measures of uncertainty (e.g., confidence intervals, standard errors) by writing computer scripts.
- Discuss the impact of type I & II errors as well as responsible use and reporting of p-values on hypothesis tests.
- Explain estimator bias and uncertainty, and write a computer script to calculate it.
- Discuss how an estimator’s bias arises (e.g., sample bias, study design), and its implications in statistical inference.
- Perform all aspects of a statistical analysis (from data consumption to reporting) using reproducible and transparent computer scripts.

Teaching Team

Position	Name	Slack Handle	GHE Handle	Section
Lecture/Lab Instructor	Katie Burak	@Katie	@katieburak	1
Lecture/Lab Instructor	Alexi Rodríguez-Arelis	@Alexi	@alexrod6	2
Teaching Assistant	Ramin Rezaeianzadeh	@Ramin Rezaeianzadeh (TA)		1
Teaching Assistant	Anne-Sophie Fratzscher	@Anne-Sophie (TA)	@afratz	1
Teaching Assistant	Sky Sheng	@Sky Sheng		1
Teaching Assistant	Maria Stephenson	@Maria	@msteph11	1
Teaching Assistant	Haley Oleynik	@Haley Oleynik (TA)	@holeynik	2
Teaching Assistant	Sidney Saint	@Sidney (TA)		2
Teaching Assistant	Jeremy Thomas	@Jeremy Thomas		2

Lecture Topics

This course occurs during **Block 2** in the 2024/25 school year. The course notes can be accessed here. Typically, there will be assigned pre-lecture videos and readings that you should review before each lecture. The videos for this course can be found here and the recommended viewing schedule is available on Canvas.

Lecture	Topic	Optional Readings
1	Populations and Sampling	Modern Dive: Chapter 7
2	Bootstrapping and its Relationship to the Sampling Distribution	Modern Dive: Chapter 8, sections 8 - 8.2 inclusive
3	Confidence Intervals via Bootstrapping	Modern Dive: Chapter 8, sections 8.3 - 8.7.1 inclusive
4	Hypothesis Testing via Simulation/Randomization	• Modern Dive: Chapter 9, sections 9 - 9.5 inclusive (omit 9.4.2) • Fun interactive visualization of permutation hypothesis test
5	Confidence Intervals Based on the Assumption of Normality or the Central Limit Theorem	Modern Dive: • Appendix, section A.2 Normal distribution • Chapter 7, section 7.5.2 Central Limit Theorem • Modern Dive: Chapter 8, section 8.7.2
6	Classical Tests Based on Normal and t-Distributions	• Modern Dive: Chapter 9, section 9.6.1 • Degrees of freedom video
7	Tests for Multiple Group Comparisons	• Open Intro Statistics • Stats and R: ANOVA in R • Learning Statistics with R
8	Errors in Inference / There is only one test!	

Lecture	Topic	Optional Readings
		• There is still only one test • Modern Dive: Chapter 9, section 9.4 • Interactive Simulation • Video: Introduction to power in hypothesis test

Deliverables

This is an **assignment-based course**. The following deliverables will determine your course grade:

Assessment	Weight
Lab 1	10%
Lab 2	10%
Lab 3	10%
Lab 4	10%
Worksheet 1	1%
Worksheet 2	1%
Worksheet 3	1%
Worksheet 4	1%
Worksheet 5	1%
Worksheet 6	1%
Worksheet 7	1%
Worksheet 8	1%
iClicker	2%
Quiz 1	25%
Quiz 2	25%

Class Schedule & office hours

See calendar.

Textbook

We are using an open source textbook: *ModernDive: Statistical Inference via Data Science* developed by Chester Ismay and Albert Y. Kim. This book is available as:

- an HTML free on the web.
- a book for purchase on Amazon.

Policies

See the general MDS policies.

Use of LLMs

LLMs, such as ChatGPT, can be helpful tools if we use them responsibly. In this course, students are permitted to use these tools to gather more information, review concepts, or brainstorm, and students must cite these tools if they use them for assignment. Having said all this, it is **not** permitted to write any given assignment via copying and pasting AI-generated responses.

Attribution

The course is built upon previous years' materials developed by previous instructors.

License