Stat 140: DESIGN OF EXPERIMENTS
Department of Statistics, Faculty of Arts and Sciences, Harvard University

Instructor: Marie-Abele Bind (ma.bind@mail.harvard.edu)
Office hours: 1-2pm on Mondays in MD129 or by appointment
Course website: https://canvas.harvard.edu/courses/165553
Teaching fellow: Chloe Yu (chloeyu@fas.harvard.edu)

Class meetings: Monday and Wednesday 10:30-11:45am in MD G135

Sections: Monday 4:30pm in SC 228

Course objectives: The course covers topics on design of randomized experiments and causal inference at an undergraduate level. The objective is to equip students with classical and modern methodologies for designing and analyzing large and complex experiments in different scientific disciplines.

Intended audience: Stat 140 is intended for undergraduate and graduate students interested in designing and analyzing randomized experiments to estimate causal effects. Students from life sciences, biostatistics, epidemiology, economics, social science, political science, education, and any other field are warmly encouraged to attend.

Prerequisites: Both introductory courses in probability and inference (i.e., Stat 110 and Stat 111) are required. An introductory background in programming and linear models (e.g., Stat 139) is also preferred.

Assignments: There will be four assignments that give students an opportunity to master the methods discussed in class. There will be two in class midterms. These will generally be conceptual in nature. There will be one final project.

Project and class presentation: The class projects will be completed individually. Students should design and conduct a randomized experiment and analyze the collected data using the statistical methods from the course. Grading of projects will be based on (i) a short written report and (ii) an in-class project presentation.

Grading: Assignments (40%, two days of grace period), class participation (10%), midterms (30%), and final project (20%).

References:

- On the Application of Probability Theory to Agricultural Experiments. Essay on Principles. Section 9, Neyman J., *Statistical Science*, 1923, 1990
- On the Two Different Aspects of the Representative Method: The Method of Stratified Sampling and the Method of Purposive Selection, Neyman J., *Journal of the Royal Statistical Society*, 1934
- Design of Experiments, Fisher R., *Olivier and Boyd*, 1935
- The Design and Analysis of Experiments, Kempthorne O., *Robert Krieger Publishing Company*, 1952
- Experimental Design, Cochran W., *Wiley Classics Library*, 1957
- The Theory of the Design of Experiments, Cox D. and Reid N., *Chapman & Hall/CRC*, 2000
- A Modern Theory of Factorial Designs, Mukerjee, R. and Wu J., *Springer Series in Statistics*, 2006
- Experimental Design for Biologists, Glass D. and Glass D., *Cold Spring Harbor Laboratory Press*, 2007
- Rerandomization to improve covariate balance in experiments, Morgan K. and Rubin D., *Annals of Statistics*, 2012
- Causal Inference for Statistics, Social, and Biomedical Sciences, Imbens G. and Rubin D., *Cambridge University Press*, 2015

Topics:

1. History and basic principles of design of experiments
2. History of "big ideas" of causal inference (e.g., missing data problem and potential outcomes framework)
3. Assignment mechanism and essential assumptions
4. One-factor with two levels experiment
    - Fisherian inference
    - Neymanian inference
    - Bayesian inference
5. Rerandomization
6. Matched-pair experiment
7. One-factor with more than two levels experiment
8. Complex randomization structure (e.g., crossover)
9. Randomized block
10. Two-factor experiments
11. Full and fractional factorial experiment
12. Complex settings and applications (e.g., observational studies, non-compliance, high-dimensional)

Tentative schedule:

Date	Modules	Assignments
Jan 26	Course overview and Module 1 (History and basic principles of design of experiments)	
Jan 28	Module 2 (History of "big ideas" in causal inference)	
Feb 2	Module 3 (Assignment mechanism)	Assignment 1
Feb 4	Module 4 (Fisherian inference)	
Feb 9	Guest lecture (Controlled experiments): Florian Engert	**Assignment 1 due** and Assignment 2 posted
Feb 11	Module 5 (Neymanian inference)	
Feb 16	Holiday	
Feb 18	Module 6 (Bayesian inference)	**Assignment 2 due**
Feb 23	**Midterm 1**	
Feb 25	Module 7 (Rerandomization)	Assignment 3 posted
March 2	Module 8 (Matched-pair experiments)	

March 4	Module 9 (>2 levels)	**Assignment 3 due**
March 9	Guest lecture (Essential concepts): Donald Rubin	
March 11	Guest lecture (Essential concepts): Donald Rubin	**Abstract of project due**
March 16	Spring break	
March 18	Spring break	
March 23	Module 10 (Crossover experiment)	
March 25	Module 11 (Block randomized experiments)	
March 30	Module 12 (Two-factor)	Assignment 4
April 1	Guest lecture (Randomized Controlled Trials): Beow Yeap	
April 6	Guest lecture (p-value and statistical significance): Ronald Wasserstein	
April 8	Module 13 (Full factorial)	**Assignment 4 due**
April 13	Guest lecture (Fractional factorial): Tirthankar Dasgupta	
April 15	Guest lecture (Fractional factorial): Tirthankar Dasgupta	
April 20	**Midterm 2**	
April 22	Module 14 (Non-compliance)	
April 27	Module 15 (High-dimensional)	
April 29	Module 16 (Observational studies)	
May 4	Reading period	
May 6	Reading period	
May 11	**Presentations**	**Full project due**
May 13	**Presentations**	