Intro to Experimental Analysis

Dmitry Arkhangelsky
darkhangel@cemfi.es

Fall, 2021

Logistics

Lectures: Thursday, December 2, 11:30-13:30; Friday, December 3, 10:00-11:30, 12:00 - 13:30
Location: Sala Europa, Banco de España

Plan

Social scientists, government agencies, and tech companies embrace experiments as the best tool for answering causal questions and improving policies/products. In their simplest form – randomized control trials, or A/B tests – experiments are simple to analyze. However, in practice, we need both to analyze and design experiments most efficiently. In this short (3 lectures) mini-course, we will discuss various econometric aspects of these problems.

We start with static experiments and ask a question: what is the most efficient way of splitting the observed units into treatment and control groups. This question is of vital interest whenever the number of observed units is modest, and there is a considerable degree of observed heterogeneity among them (think markets, classrooms, firms). We will show that this question can be approached as an optimization problem that modern tools can solve efficiently.

As a next step, we consider adaptive experiments (AKA multi-armed bandits), i.e., situations in which units are assigned to treatments sequentially based on the knowledge obtained so far. This approach is ubiquitous in the industry, where companies want to learn quickly which product design works best. It is also natural in social and medical experiments where we do not want our subjects to suffer from inferior treatments. We will discuss various algorithms for this problem, their advantages, and problems.

Finally, we will look at dynamic experiments where we treat the same unit multiple times, and our goal is to achieve some long-term outcomes. This approach is most developed in medical experiments, where one looks for the best dynamic treatment regime. However, various social experiments naturally fit into this framework. Here we will discuss critical foundational concepts and analogs of simple estimators from static experiments.

Prerequisites

I expect basic knowledge of the standard causal framework (Neyman-Rubin model), familiarity with ordinary least squares and basic data analysis concepts (e.g., cross-validation, sample splitting, etc.). There is a (incomplete) list of references below, I do not expect you to read them, but the material will be loosely based on them.

References

Dimitris Bertsimas, Mac Johnson, and Nathan Kallus. The power of optimization over randomization in designing experiments involving small samples. *Operations Research*, 63(4):868–876, 2015.

Iavor Bojinov, Ashesh Rambachan, and Neil Shephard. Panel experiments and dynamic causal effects: A finite population perspective. *arXiv preprint arXiv:2003.09915*, 2020.

Maria Dimakopoulou, Zhengyuan Zhou, Susan Athey, and Guido Imbens. Estimation considerations in contextual bandits. *arXiv preprint arXiv:1711.07077*, 2017.

Vitor Hadad, David A Hirshberg, Ruohan Zhan, Stefan Wager, and Susan Athey. Confidence intervals for policy evaluation in adaptive experiments. *arXiv preprint arXiv:1911.02768*, 2019.

Miguel A Hernán and James M Robins. Causal inference: What if, 2020.

Nathan Kallus. Optimal a priori balance in the design of controlled experiments. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 80(1):85–112, 2018.

Maximilian Kasy and Anja Sautmann. Adaptive treatment assignment in experiments for policy choice. *Econometrica*, 89(1):113–132, 2021.

Tze Leung Lai and Herbert Robbins. Asymptotically efficient adaptive allocation rules. *Advances in applied mathematics*, 6(1):4–22, 1985.

Susan A Murphy. Optimal dynamic treatment regimes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 65(2):331–355, 2003.

James M Robins. Optimal structural nested models for optimal sequential decisions. In *Proceedings of the second seattle Symposium in Biostatistics*, pages 189–326. Springer, 2004.

William R Thompson. On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. *Biometrika*, 25(3/4):285–294, 1933.