

15.095 Project Proposal

Optimizing Drug Allocation to Prostrate Cancer patients using Policy Trees

Moritz Bartusch moritzba@mit.edu

Rachit Jain rachitj@mit.edu

1 Problem Motivation

Prostate cancer is the 2nd most commonly occurring cancer in men and the 4th most common cancer overall. There were more than 1.4 million new cases of prostate cancer in 2020. The American Cancer Society estimates 268,490 new cases and about 34,500 deaths in the US due to prostate cancer in 2022. This calls in for effective treatment drugs to be given to patients to elongate their average survival time.

2 Dataset

Diethylstilbestrol is a form of synthetic estrogen that has been historically used to treat late-stage prostate cancer. Byar and Green (1979) explored the effect of a placebo and three different levels, 0.1mg, 0.2mg and 0.5mg, of the drug involving estrogen, over 502 patients affected with stage 3 or 4 prostate cancer.

We shall utilize a publicly available dataset from the study by Byar and Green. It contains 502 patients with 18 different features. These include, cancer stage, months of follow-up, age in years, blood pressure levels, weight index, history of cardiovascular disease, etc. For each patient, we have data on of the 4 types of treatments.

3 Methodology

In lecture 16, we learned how to identify exceptional responders in randomized trials using the example of prostate cancer treatments. Instead of finding the subset of patients on which the effect of a particular drug is highest, we build an **Optimal Policy Tree (OPT)** to prescribe the best treatment for patients using the same data set. In addition, we plan to modify the classical OPT as shown in the lecture to test it with multiple variations and added features.

We are planning to proceed as follows:

1. Training a baseline OPT as shown in class for prescribing estrogen treatments
2. Modify this tree by using the actual observed values for each patient for the treatment they got instead of using the newly predicted values
3. Using robust models to predict the targets and probabilities to receive a certain treatment for direct and IPW techniques respectively; finally using those in the doubly robust approach to get results with an OPT
4. Exploring a potential relationship between finding exceptional responders to a treatment and our results from the OPTs

4 Challenges

- Are there an implications to using a placebo treatment in OPTs?