Project Proposal: Observational studies with multi-level treatments

STAT 6390.001: Introduction to Causal Inference

Indrajith Wasala Mudiyanselage

1 Introduction

In an investigation of the causal effect of a treatment, one of the most common issues is that the groups of subjects in the study are not randomly assigned. Then there is no interference or manipulation of the research subjects and no control over treatment groups. These studies are called observational studies. Confounders are the variables associated with both the treatment and the outcome. When Confounders are presrnt, non-random group assignments could result in biased estimates of the treatment effect.

There are many popular literature to overcome this problem. Among those literature, Propensity score-based methods have been widely improved to adjust for confounders in observational studies to estimate causal treatment effect for binary treatments. Eventhough the initial work on propensity scores focused on the case of binary treatments, there are more recent work that generalizing the Propensity score-based methods to treatments with more than two levels (Multi-level treatment). In this project, my forcus is to introduce currenly availabel Propencity score based methods or improvements to the propencity scores to investigate the causal effect of Multi-level treatments.

2 Existing literature

2.1 Matching on the Estimated Propensity Score

Propensity score matching estimators are widely used in evaluation research to estimate average treatment effects. In this paper, the authors derive the large sample distribution of propensity score matching estimators by taking into account that the propensity score is

itself estimated before matching. But, the estimation of the propensity score first affects the large sample distribution of propensity score matching estimators. Therefore, the authors introduce adjustments to the large sample variances of propensity score matching estimators of the average treatment effect (ATE) and the average treatment effect on the treated (ATET). However, for the ATET estimator, the sign of the adjustment term depends on the data-generating process, and ignoring the estimation error in the propensity score may lead to misleading confidence intervals.

2.2 Causal Inference for Multi-level Treatments with Machine learned Propensity Scores

In this paper, the authors generalize Propensity score-based methods that have been developed to adjust for confounders in observational studies with binary treatments. In other words, improving existing binary treatment-based Propensity scores to the multi-level treatment case. They review the generalized causal inference framework and several propensity score estimation methods. Then conduct a comprehensive simulation study to evaluate the performance of multinomial logistic regression, generalized boosted models, random forest and data-adaptive matching scores for estimating propensity scores based on inverse probability of treatment weighting.

2.3 Propensity Score Matching and Subclassification in Observational Studies with Multi-level Treatments

This paper developed a new method for estimating average treatment effects in observational studies, in settings with multi-treatment levels by assuming unconfoundedness given pre-treatment variables. Some propensity-based methods do not naturally extend to the multi-level treatment. In these cases, the authors introduced the concept of weak unconfoundedness and the notion of the generalized propensity score.

2.4 The Prognostic Analogue of the Propensity Score

In this paper, the authors are comparing Propensity sores and Prognostic scores. Prognostic scores summarize covariates' association with potential responses while the Propensity score collapses the covariates of an observational study into a single measure summarizing their joint association with treatment conditions. Like propensity scores, stratification on prognostic scores brings to uncontrolled studies a concrete and desirable form of balance and reduce the dimensions the covariate. In designs for which unassisted propensity adjustment is difficult or infeasible, Prognostic scores will provide promising results.

3 Simulation setting

The simulation is a Monte Carlo study relative to the previously proposed estimators in the selected papers. To be specific, design the simulation with three treatment levels the covariates X_{1i}, X_{2i},and X_{3i} are multivariate normal with means zero, variances of (2, 1, 1) and covariances of (1,-1,-5). The select the sample size (N) to be 500. Then compare the proposed estimators in the papers over 1000 datasets.

4 A real example

In this project, I am planning to use the data from the REFLECTIONS (Real-world Examination of Fibromyalgia: Longitudinal Evaluation of Costs and Treatments) study. REFLECTIONS was a 12-month prospective observational study of patients being treated for fibromyalgia at 58 outpatient sites in the US and Puerto Rico. Patients had to be at least 18 years of age and initiating a new pharmacologic treatment for fibromyalgia. Data from patients were collected at baseline during a standard office visit and at 1, 3, 6, and 12 months post baseline via a computer assisted telephone interviews. In this analysis, three fibromyalgia medication cohorts will be selected; patients treated with an opioid (either monotherapy or with other medications), patients treated with tramadol but not an opioid, and patients

not treated with tramadol or an opioid (referred to as the Other cohort). The outcome variable utilized here is the total score of Fibromyalgia Impact Questionnaire (FIQ).

5 Timeline

Schedule your project explicitly. For example,

- 10/01 – 10/20 read the selected papers;
- 10/20 – 11/10 run the simulations;
- 11/10 – 11/24, analyze real data sets;
- 11/24 – 11/30, prepare the presentation;
- 11/30 – 12/08, write final report.

References

[1] A. Abadie and G. W. Imbens (2016). "Matching on the estimated propensity score". Econometrica, 84:781-807.

[2] Shu Yang, et al. (2016). "Propensity Score Matching and Subclassification in Observational Studies with Multi-level Treatments". Biometrics, 72(4), 1055-1065.

[3] Lin Lin, et al. (20119). "Causal inference for multi-level treatments with machine-learned propensity scores". Health Services and Outcomes Research Methodology 19:106–126.

[4] B. B. Hansen (2008). "The prognostic analogue of the propensity score". Biometrika, 95:481-488.