Growth Mindset on Student Achievements

Minh Thu Bui and Vindyani Herath

MA592: Causal Inference

April 26, 2024

Overview

1.1. Background

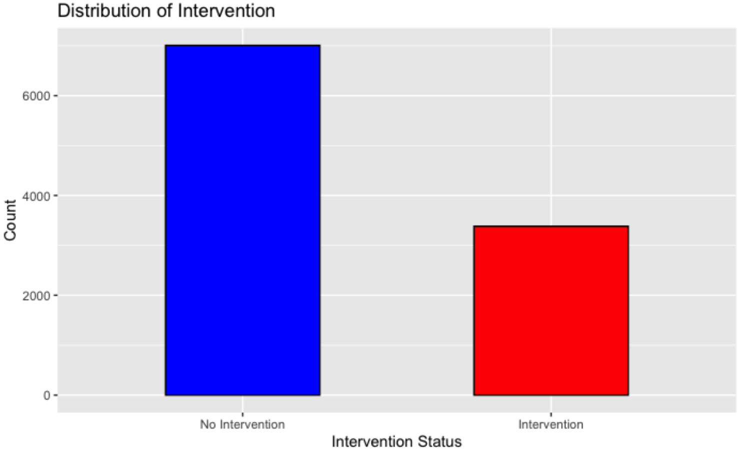
- A growth mindset is defined as the belief that intelligence can be developed and that people can become smarter through hard work, effective studying strategies, and help from the community (Athey and Wager (2019)).
- The objective of this study is to assess the effectiveness of mindset intervention (a nudge-like intervention) in enhancing student achievement.

1.2. Dataset Description

- The dataset contains observational data from 10,391 students across 76 high schools in the U.S.
- The outcome is a continuous variable for measure of achievement and the treatment is a binary variable indicating the receipt of the intervention.
- The predictors include students' self-reported expectations for success in the future, race/ethnicity, gender, first-generation status, urbanity of the school, school-level mean of student's mixed mindset, school achievement level, school racial/ethnic minority composition, school poverty concentration and, school size.

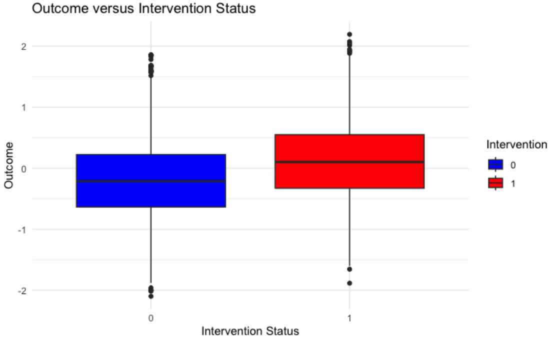
1.3. Distribution of Intervention

First, we want to see the distribution of people who received the intervention and who did not in the total dataset.



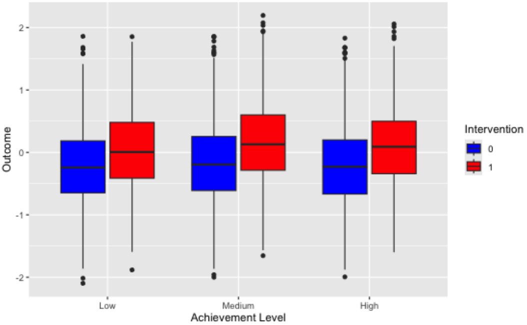
1.3.1: EDA - Outcome Versus Intervention Status

Next, we also want to investigate the relationship between the intervention status and the outcome.



1.3.4. EDA - Schools Achievement Level Versus Outcome

We check whether school achievement levels have a relationship with the intervention and outcome. The classification of schools' achievement levels is as follows: Low (25th percentile or lower), Middle (25th - 75th percentile), and High (75th percentile or higher).



2.1. Average Treatment Effect (ATE)

Average Treatment Effect can be used to evaluate the causal effect of treatment (Intervention) on outcome (Measure of achievement).

Method	**ATE**	**Standard error**	**95% bootstrap confidence interval**
OR Estimator	0.2788	0.0187	(0.2420, 0.3161)
Inverse Probability Weighting	-0.3348	0.0385	(-0.4309, -0.2832)
Hajek Estimator	-0.2119	0.0206	(-0.2397,-0.1907)
Doubly Robust Estimator	-0.3415	0.0740	(-0.5151, -0.2148)

Sensitivity Analysis

Table 1: Sensitivity analysis for the average causal effect.

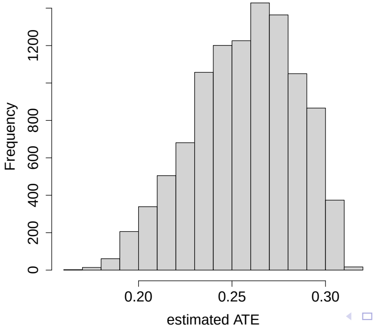
	0.50	**0.59**	**0.67**	**0.77**	**1.00**	**1.30**	**1.50**	**1.70**	**2.00**
0.50	0.13	0.14	0.14	0.15	0.16	0.17	0.17	0.17	0.18
0.59	0.15	0.16	0.16	0.17	0.18	0.19	0.19	0.19	0.20
0.67	0.17	0.18	0.18	0.19	0.20	0.21	0.21	0.21	0.22
0.77	0.19	0.20	0.21	0.21	0.22	0.23	0.23	0.24	0.24
1.00	0.25	0.26	0.26	0.27	0.28	0.29	0.29	0.29	0.29
1.30	0.32	0.33	0.33	0.34	0.35	0.36	0.36	0.36	0.37
1.50	0.37	0.38	0.38	0.39	0.40	0.41	0.41	0.41	0.41
1.70	0.41	0.42	0.43	0.44	0.45	0.45	0.46	0.46	0.46
2.00	0.49	0.49	0.50	0.51	0.52	0.52	0.53	0.53	0.53

3.1. Causal Forests (Wager and Athey (2018))

- Causal forests are a causal inference learning method that is an extension of Random Forests.
- A causal forest is simply the average of a large number of causal trees, where the trees differ due to sub-sampling (Athey and Imbens (2019)).
- In causal forests, data is split in order to maximize the difference across splits in the relationship between an outcome variable and a "treatment" variable.
- A variant of Augmented Inverse Probability Weighting is used to estimate the average treatment effect.

3.2. Causal Forest with `grf()`

1. We first draw a subsample of clusters (not of all observations).
2. Then we draw k samples at random from each cluster.
3. We perform causal forest with the `grf()` package. To account for potential correlations within each cluster, we only consider an observation i to be out-of-bag if its cluster was not drawn in step (1).



3.3. Causal Forests : Estimating ATE

If we train a causal forest on students without clustering by school, we obtain markedly different results from before. The confidence interval for the average treatment effect is now roughly half as long as before, and there appears to be unambiguously detectable heterogeneity according to the `test_calibration` function.

Method	ATE	Standard error	95% bootstrap confidence interval
Causal Forests with Clusters	0.2473	0.0202	(0.208, 0.286)
Causal Forests w/o Clusters	0.2530	0.0114	(0.231, 0.275)

3.4. Test Calibration for Heterogeneity

To test for treatment heterogeneity, we use the package `test_calibration` of the forest. If the coefficient of the 'differential.forest.prediction' is significantly greater than 0, then we can reject the null of no heterogeneity.
With clusters:

```
                               Estimate Std. Error t value Pr(>t)
mean.forest.prediction         1.003681   0.082837  12.116 <2e-16 ***
differential.forest.prediction 0.197951   0.675577   0.293 0.3848
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Without clusters:

```
                               Estimate Std. Error t value    Pr(>t)
mean.forest.prediction         1.006475   0.044922 22.4049 < 2.2e-16 ***
differential.forest.prediction 0.533790   0.133047  4.0121 3.031e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

3.5. Variable Importance

The following were found to be the most important variables used in creating causal forests.

Variable	Weighted Sum
Mean of students' fixed mindsets	0.1462
Student's self reported expectation	0.1251
School achievement level	0.1171
School size	0.1151
School racial/ethnic minority composition	0.1062

*Weighted sum is how many times a feature was split on at each depth in the forest.

4.1. Summary of ATE estimates

Average Treatment Effect can be used to evaluate the causal effect of treatment (Intervention) on outcome (Measure of achievement).

Method	**ATE**	**Standard error**	**95% bootstrap confidence interval**
OR Estimator	0.2788	0.0187	(0.2420, 0.3161)
Inverse Probability Weighting	-0.3348	0.0385	(-0.4328, -0.2815)
Hajek Estimator	-0.2119	0.0206	(-0.2397,-0.1907)
Doubly Robust Estimator	-0.3415	0.0740	(-0.5151, -0.2148)
Causal Forests with Clusters	0.2473	0.0202	(0.208, 0.286)
Causal Forests w/o Clusters	0.2530	0.0114	(0.231, 0.275)

References

Athey, S. and Imbens, G. W. (2019). Machine learning methods that economists should know about. *Annual Review of Economics*, 11:685–725.

Athey, S. and Wager, S. (2019). Estimating treatment effects with causal forests: An application. *Observational studies*, 5(2):37–51.

Wager, S. and Athey, S. (2018). Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, 113(523):1228–1242.

Thank You.