Foundations of Causal Inference:
Understanding and Misunderstanding Randomized Controlled Trials

Rao Sihang 15220162202309

Brief Summary

Recently, random controlled trials are not only common in medicine, but also are increasingly popular in the social sciences, such as in development economics. Many researchers hold the view that RCT,with unbiased estimation and no prior knowledge, is superior to other methods in terms of casusal inference. However, in practice, the value of RCT is overvalued with lots of limitation.The best method depends on the specific questions and purposes we wants to investigate. And there are lots of other things need to be done to ensure the estimation of RCT from a particular sample useful and serviceable.

Estimation of ATE

$$\overline{Y}_1 - \overline{Y}_2 = \overline{\beta_1} + \sum_{j=1}^{J} \gamma_j(\overline{x}_{1ij} - \overline{x}_{0ij})$$

$$E[\overline{Y}_1] - E[\overline{Y}_2] = E[\overline{Y}_1 - \overline{Y}_2] = \overline{\beta_1} + (\overline{S_1} - \overline{S_2})$$

Average treatment effect is difference average outcome between treatment group and control group.Usually, the estimation of ATE is simply the differnece between means in this two groups, which is an unbiased estimatior theoretically. However, in practice, there are other causes correlated to outcomes across these two groups leading to non-zero balance. With perfect balance (these two groups are identical), the difference bettwen the two means is exactly equal to ATE. How to get a perfect balance as much as possible?

- With prior knowledge of the other causes to control them.
- Matching. Finding a matched subject with similar causes for each subject.
- Randomization, which is the method used in RCT.

Lots of people think estimation of ATE is unbiased and reliable without prior knowledge on causes in RCT after randomly generating treatment group and control group. However, in practice, there are several problems.

- Even though the estimation is unbiased, it is only the truth for trial sample, not the truth for general population.

- The mean of other causes in treatment group and control group become close only in large sample. And even with very large sample sizes, if there is a large number of causes, balance on each cause may be infeasible. Stratification is a method to improve balance before randomization, but this procedure needs prior knowledge.
- Randomization without any prior knowledge avoids potential disagreements, but negatively prevents cumulative scientific progress.
- When the two samples have different variances or the distribution of treatment effects is not symmetric, test on the differencce won't work well.

- Post-treatment bias, the difference in two groups after randomiztion will affect outcomes. The reasons might be lack of blinding, different places,time, and other practical factors.

The usage of results

RCT can conrtibute to (a) simple extrapolation and simple generalization, (b) drawing lessons about the population enrolled in the trial, (c) extrapolation with adjustment, (d) estimating what happens if we scale up, (e) predicting the results of treatment on the individual, and (f) building and testing theory.

1. Different support factors in different settings will affect ATE, and what we need to understand is the mechanism of treatment.
2. The failure of replication in other settings can make us analyze and understand deeply the reasons and mechanisms.
3. Post-experimental stratification can be used to estimate the ATE in a new context, or to correct estimates to the parent population when the trial sample is not a random sample of the parent.
4. Combine theory and RCT in empirical study. Use the data collected in RCT to fit a structure model that can be used in more general setting.
5. The intervention might work differently at scale.

Comments

This paper provides lots of quotes and examples in various fields to illustrate the potential problem in estimation of ATE and the practical value of random controlled trials.

Generally, Deatons thinks RCT is overvalued by some researchers, (1) the practical useness is only limited in small regions and sample investigated; (2) In reality, lots of potential problem in process of operating causes the failure in random. (3) There is no mature test mechanism in RCT.

Opinions vary from person to person. Opponents focus on questioning the drawbacks of RCT, but proponents, like me, tend to attach importance in values of RCT, which provides a relatively free setting to do research. Nothing is perfect, and everything keeps developing. Of course, it is necessary for us to admit those limitation of RCT and develop more suitable designs to improve them.