Applied Quantitative Analysis II
Lab 3

Eric G. Zhou

Spring 2020

NYU Wagner School of Public Service
NYU Grossman School of Medicine

AQA II: Lab 3

General reminder

This is the 3rd week of the semester. Guess what?

- Tiredness starts kicking in; just submitted your first assignment and feel like you used all you ever know.
- Stuck with a difficult course? Peer support is critical. Easy way or the hard way?
- This is the very time for growth and learning!

Keep it up!

AQA II: Lab 3

Contents

- Practical guide on designing and analyzing RCT
- Propensity Score Matching model

Practical guide on designing and analyzing RCT

Recap on this design

- Gold standard for many practical research, including drug trials, policy interventions, development research (2019 Nobel Economic Prize)
- The best internal validity you could get out of any design, except for theoretical studies.
- What about any disadvantages?

Practical guide on designing and analyzing RCT

Recap on this design

- Gold standard for many practical research, including drug trials, policy interventions, development research (2019 Nobel Economic Prize)
- The best internal validity you could get out of any design, except for theoretical studies.
- What about any disadvantages?
- Expensive, political issues, implementation issues, high cost to follow-up, limited external validity, etc.

Practical guide on designing and analyzing RCT

In theory, we can write down the estimation equation of RCT as

$$Y_{\text{outcome}} = \beta T_{\text{treatment}} + \epsilon \tag{1}$$

Since $Cov(T, \epsilon) = 0$, β is an unbiased estimator of the treatment effect.

What about its variance?

β_{ols} is the best linear unbiased estimator, also the most efficient estimator when the error term is normally distributed. Even if the error term is not normal, β_{ols} is still the best estimator, why?

Practical guide on designing and analyzing RCT

Let's derive the variance of β, which will reveals a lot about the properties of analyzing a RCT.

The variance of a random variable is $(E[X] - \bar{X})^2$, remember this part $(X'X)^{-1}X'\epsilon$?

Now let's replace X with T, so the variance of β is

$$Var(\beta|T) = E[(\beta - \hat{\beta})(\beta - \hat{\beta})|T] \tag{2}$$

$$Var(\beta|T) = E[(T'T)^{-1}T'ee'T'(T'T)^{-1}|T] \tag{3}$$

$$Var(\beta|T) = (T'T)^{-1}T'E[ee'|T]T'(T'T)^{-1} \tag{4}$$

$$Var(\beta|T) = \sigma^2(T'T)^{-1} \tag{5}$$

σ is the variance for the error term.

Practical guide on designing and analyzing RCT

The equation (5) $Var(\beta|T) = \sigma^2(T'T)^{-1}$ have two implications

Take a guess, what might they be?

Practical guide on designing and analyzing RCT

The equation (5) $Var(\beta|T) = \sigma^2(T'T)^{-1}$ have two implications

First implication, if you want to reduce the variance, i.e., standard error, you need to reduce σ that is the unexplained part of your regression estimate.

How? Easiest way is to include more relevant control variables.

Second implication, if your treatment variable is binary. How can you design the optimal ratio of treatment and control group?

The answer is 1:1. Think about the variance of Bernoulli Distribution $Var(X) = p(1-p)$, therefore easy to see $1 - p^2$ is minimized at $p = q = 1$.

Practical guide on designing and analyzing RCT

The equation (5) $Var(\beta|T) = \sigma^2(T'T)^{-1}$ have two implications

First implication, if you want to reduce the variance, i.e., standard error, you need to reduce σ that is the unexplained part of your regression estimate.

How?

Practical guide on designing and analyzing RCT

The equation (5) $Var(\beta|T) = \sigma^2(T'T)^{-1}$ have two implications

First implication, if you want to reduce the variance, i.e., standard error, you need to reduce σ that is the unexplained part of your regression estimate.

How? Easiest way is to include more relevant control variables.

Second implication, if your treatment variable is binary. How can you design the optimal ratio of treatment and control group?

Practical guide on designing and analyzing RCT

The equation (5) $Var(\beta|T) = \sigma^2(T'T)^{-1}$ have two implications

First implication, if you want to reduce the variance, i.e., standard error, you need to reduce σ that is the unexplained part of your regression estimate.

How? Easiest way is to include more relevant control variables.

Second implication, if your treatment variable is binary. How can you design the optimal ratio of treatment and control group?

The answer is 1:1. Think about the variance of Bernoulli Distribution $Var(X) = p(1-p)$, therefore easy to see $1 - p^2$ is minimized at $p = q = 1$.

Practical guide on designing and analyzing RCT

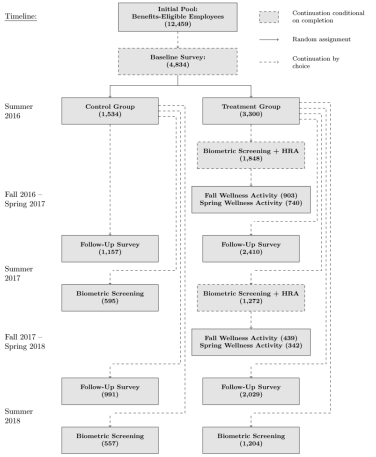

FIGURE I
Experimental Design of the Illinois Workplace Wellness Study

AQA II: Lab 3

Practical guide on designing and analyzing RCT

Question: How do you know if your randomization is successful?

Practical guide on designing and analyzing RCT

Question: How do you know if your randomization is successful?

One way to go is compare your baseline covariates Xs, to see if they are significantly different from each other. In theory they should be balanced.

Question: What if the covariates do not balance?

Practical guide on designing and analyzing RCT

Question: How do you know if your randomization is successful?

One way to go is compare your baseline covariates Xs, to see if they are significantly different from each other. In theory they should be balanced.

Question: What if the covariates do not balance?

You could either redo the randomization, or just control for those unbalanced covariates.

Propensity Score Matching

The idea is simple, if you cannot randomize the independent variable, but you want to study of relationship as if it's an experiment.

What do you do?

You run a (usually) logistic regression of your treatment variable on all of the observable covariates – you try to predict or measure the predilection of receiving treatment.

$$P(x) = \frac{e^X}{1 + e^X}$$

Then you run a regression of your Y on the treatment using the $P(x)$ as your weight. Using the law of conditional expectations, we could prove that the treatment effect is unbiased if the **Conditional Independent Assumption(CIA)** is met.

Propensity Score Matching

The CIA states that conditional all the observables(X), the outcome is independent from the treatment, i.e., $(Y \perp T)|X$

This is a hella strong assumption. You cannot convince any editors easily if you do not have a list of 100 relevant covariates, and you are expected to justify each.

But there are people using this on a daily basis – the success of RCT created a monster. The Harvard Pharm group likes this method in clinical settings very much.

Also, if you don't buy what I say here, check out this article:
`https://gking.harvard.edu/files/gking/files/psnot.pdf`

Do not use it unless you are told to.