

Regression Discontinuity Design

April 3, 2025

Motivation

- RD research design exploits precise knowledge of the rules determining treatment
- Examples
 - certain regulations applies to firms with at least 15 employees
 - school programs mandatory for students score below threshold
 - eligibility of medical services through medicare restricted by age cutoff
- In general two types
 - Sharp (SRD)
 - Fuzzy (FRD)

Sharp Regression Discontinuity Design

- Recall the Rubin Causal Model (RCM) where

$$Y_i = (1 - W_i)Y_i(0) + W_iY_i(1)$$

where $W_i \in \{0, 1\}$ is the treatment received.

- Both X_i, Z_i are observed characteristics and not affected by treatment.
- In the case of SRD, W_i is a deterministic function of X (forcing variable)

$$W_i = I\{X_i \geq c\}$$

- We look at discontinuity in the conditional expectation of the outcome

$$\lim_{x \downarrow c} E[Y_i | X_i = x] - \lim_{x \uparrow c} E[Y_i | X_i = x] \equiv E[Y_i(1) - Y_i(0) | X_i = c]$$

which is the average treatment effect (but very local at the cutoff).

Assumptions

- As we have emphasized in the earlier ATE (regression, propensity score, matching) methods, two assumptions are crucial.
- First is Unconfoundedness:

$$Y_i(0), Y_i(1) \perp W_i | X_i$$

Here this condition holds trivially, since the treatment is deterministic conditional on X .

- What about Overlap? $0 < P(W_i = 1 | X_i = x) < 1$.

Assumptions

- As we have emphasized in the earlier ATE (regression, propensity score, matching) methods, two assumptions are crucial.
- First is Unconfoundedness:

$$Y_i(0), Y_i(1) \perp W_i | X_i$$

Here this condition holds trivially, since the treatment is deterministic conditional on X .

- What about Overlap? $0 < P(W_i = 1 | X_i = x) < 1$. This condition is fundamentally violated. In SRD all the treatment probabilities are either 0 or 1.
- This says we will have extrapolate - but very locally is enough if we have sufficient data, i.e. as long as we observe units with covariate values arbitrarily close to c .

Fuzzy Regression Discontinuity Design

- We only briefly describe this method here. In this case, the probability of receiving treatment has a jump at the threshold but not necessarily from 0 to 1.
- Remember that FRD is IV ! More formally

$$ATT = \frac{\lim_{x \downarrow c} E[Y|X = x] - \lim_{x \uparrow c} E[Y|X = x]}{\lim_{x \downarrow c} E[W|X = x] - \lim_{x \uparrow c} E[W|X = x]}$$

- Imbens and Lemieux (2008) shows that this is equivalent to the treatment effect of those who are *induced* to treatment status by the cutoff (i.e. compliers) locally at c
- As also clear from this interpretation, the Unconfoundedness assumption is unlikely to hold for a FRD: some unobservables other than X must have affected individual unit's treatment status on cutoff c , and assume those independent of outcomes is usually problematic.

The Implementation of SRD

- Imbens and Lemieux (2008) suggests Graphical analyses.
- First, outcomes by forcing variable
 - For some binwidth h , and for some number of bins K_0 and K_1 to the left/right of cutoff values, construct bins $(b_k, b_{k+1}]$, $b_k = c - (K_0 - k + 1)h$
 - Calculate average outcome in the bin:
$$\bar{Y}_k = \frac{1}{N_k} \sum_{i=1}^N Y_i \cdot I\{b_k < X_i \leq b_{k+1}\}$$
 - Check whether there is a jump at c .
- Can also plot other covariates Z_i in similar fashion against forcing variable

The Implementation of SRD

- The regression version of these essentially relies on approximation of $\lim_{x \downarrow c} E[Y(1)|X = x]$ and $\lim_{x \uparrow c} E[Y(0)|X = x]$
- Either local linear regression (in a small neighborhood to c) or series regression (i.e. high order polynomials) are suggested.
- We will next illustrate using an example from Lee et al

Lee Moretti Butler

- Question: Do voters affect or elect policies?
 - Competing for votes induces politicians to move toward the center
 - Or voters merely elect policies
- Use data from the U.S. House (1946-1995)
- Look at how Representatives' roll-call voting behavior is affected by exogenous changes in their electoral strength.
- Investigate districts in which the previous election cycle (say, 1992) is determined by a “coin flip” (i.e. really close to 50% cutoff), generating a electoral strength for the incumbent party in current cycle (say, 1994).
- Examine the voting “scores” of the winners (say, in 1995-96) of these elections in current cycle (i.e. 1994).
- the identified effect has two components
 - First component - due to incumbent advantage, the winner is more likely to be a democrat (say). Then naturally voting scores are more liberal. (How voters elect policy). This can be estimated from data.
 - The rest - how candidates respond to competition.