Binary binary model

Ricky

August 20, 2020

Overview

ATE

- The treatment effect for individual i can be simply dened as in equation
$$TE = y_{1i} - y_{0i}$$
Let y_{0i} denote the value of the outcome variable if individual i is not treated $(x_i = 0)$ and y_{1i} indicate the value of the outcome variable if i is treated $(x_i = 1)$
- Then, the ATE is the expected value of individual treatments as in equation:
$$ATE = E[y_{1i} - y_{0i}]$$
- However, it is impossible to observe both outcomes in the real world.

- The only observable outcome y_i is depend on the x_i. Thus, if the treatment x i for each individual is randomly assigned to individuals, then ATE becomes:

$$A\hat{T}E = E[y_i|x_i = 1] - E[y_i|X_i = 0] = E[y_1 i - y_0 i]$$

Introduction

In these slides, main contexts are followed from Chester and Rosen (2013)

- 2SLS provides a convenient way to compute point estimator and can be regarded as LATE. However, if we interest other treatment effect such as ATE, LATE is quite far away from ATE under binary binary model setting.
- The nonparametric IV model provdies set identification that estimate bounds on parameters.
- model: $Y_1 = h(Y_2, W, U)$,
  where Y_1 and Y_2 are endogenous variables with $\{0,1\}$ binary outcomes, W is exogenous variable, U is unobserved heterogeneity
- Incomplete model: no specific function about how Y_2 is generated (specific example: $Y_2 = \gamma + \delta Z + V$)

The Nonparametric Binary Response Model

Nonparametric threshold-crossing model for binary Y_1:

$$Y_1 = 1[p(Y_2, W) < U]$$

- where U follows uniform distribution and independent restriction $U \perp (W, Z)$
- Discrete outcomes are common in econometric, but they tend to place structure on determining Y_2 (complete model) in order to acquire point estimators. And these methods can be regarded as one special case of our model since our model is incomplete.

Identification Analysis

- To identify ATE, the level set $U_h(y_1, y_2, w) = \{u : y_1 = h(y_2, w, u)\}$ has no singleton sets due to discrete outcomes, e.g.
  $U_h(y_1, y_2, w) = [0, p(y_2, w)]$ if $y_1 = 0$,
  $U_h(y_1, y_2, w) = [p(y_2, w), 1]$ if $y_1 = 1$,
- The objective of interest are structural function h and condition distribution of U given (w, z), i.e. $G_{U|WZ}$
- CRS shows that $F_{Y|WZ}$ with $(h, G_{U|WZ})$ follow inequalities:

  $$\forall S \in C(w, z), G_{U|WZ}(S|w, z) \geq P(U_h(Y_1, Y_2, w) \subseteq S|w, z)$$

  where $C(w, z)$ is the collection of all unions of sets on support of $U_h(Y_1, Y_2, w)$ given Z
- e.g. $C(w, z)$ comprise $\{[0, p(y_2, w)], [p(y_2, w), 1]\}$.
- Besides, since U is uniform distribution and does not depend on (W, Z), so $G_{U|Z}$ is also uniform distribution $(0, 1)$

- Define a set $A_p(y_2, w) = \{y_2^* : p(y_2^*, w) \leq p(y_2, w)\}$, and $A_p(y_2, w)^c$ is complement of $A_p(y_2, w)$
- Hence, we know that $p(y_2, w) \geq P(y_1 = 0 \cup p(y_2^*, w) \leq p(y_2, w))$ and $1 - p(y_2, w) \geq P(y_1 = 1 \cup p(y_2^*, w) \geq p(y_2, w))$
- Finally, set identification on binary binary model comes out from:

$$\begin{aligned} sup_z(P(Y_1 = 0 \cup Y_2 \in A_p(y_2)|z)) &\leq p(y_2) \\ &\leq inf_z(1 - P(Y_1 = 1 \cup Y_2 \in A_p(y_2)^c) \end{aligned} \tag{1}$$

Binary Binary model

- Set $Y_1, Y_2 = \{0, 1\}$ and consider only one explanatory variable Y_2, so $p(y_2, w) = p(y_2)$
- $\rho_0 = 1 - p(0)$: the mean potential Y_1 outcome when Y_2 is zero.
- $\rho_1 = p(0) - p(1)$: the difference in mean potential Y_1 outcomes comparing $Y_2 = 1$ and $Y_2 = 0$, i.e. ATE.

- set identification:
  If $\rho_1 \leq 0$, then $p(y_2 = 0) \leq p(y_2 = 1)$ and:
  $0 \in A_p(y_2 = 0)$ and $\{0, 1\} \in A_p(y_2 = 0)^c$ when $y_2 = 0$
  $\{0, 1\} \in A_p(y_2 = 1)$ and $1 \in A_p(y_2 = 1)^c$ when $y_2 = 1$
  Thus,
  $f_{00} \leq p(y_2 = 0) \leq 1 - f_{10} - f_{11},$
  $f_{00} + f_{01} \leq p(y_2 = 1) \leq 1 - f_{11}$
  Finally,
  $f_{10} + f_{11} \leq \rho_0(1 - P(y_2 = 0)) \leq 1 - f_{00}$
  $f_{11} \leq \rho_0 + \rho_1(1 - P(y_2 = 1)) \leq 1 - f_{00} - f_{01} = f_{10} + f_{11}$
- And so on $\rho_1 \geq 0$ case...

Empirical results

- identified set is tend to be large and disconnected under weak IV (same-sex instrument),while it become one-dim and small under strong IV.
- For each point in identified set (rho_0, rho_1), there is a unique conditional distribution $U|Y_2$. Recall...
  $G_{U|Z} \sim U(0, p(y_2))$ if $y_1 = 0$ or
  $G_{U|Z} \sim U(p(y_2), 1)$ if $y_1 = 1$

Conclusion

- If we are interested in other treatment effect other than LATE under binary binary model, 2SLS may not be a good candidate to approximate the result. Hence, we use set identification as sensity analysis for ATE, and compare the result with 2SLS and OLS

Demo time...