Avaliação de Políticas Públicas II: Métodos não experimentais

Paula Pereda (USP)

October 23, 2020

Aula 2

Resumo da Aula 1

1. Abordagem clássica vs Abordagem Reduzida (MCR)
2. MCR: separa resultados potenciais do mecanismo de seleção
3. Modelo de Resultados Potenciais: W_i, Y_{i0} e Y_{i1}
4. Efeito Causal: $\tau_i = Y_{i1} - Y_{i0}$
5. Problema Fundamental: Y_i observado diferente Y_{i0}, Y_{i1}
6. Y_i observado ligado ao MS
7. Mecanismo de Seleção Regular e Irregular
8. Tipos de Efeitos: Momentos da distribuição do efeito causal

Causalidade com Dados Observacionais

Métodos para identificação do efeito causal com seleção nas observáveis:

$$W|X \perp Y(0), Y(1) \tag{1}$$

1. Análise de Regressão
2. Modelo de Seleção de Heckman
3. Métodos usando *propensity score*

Análise de Regressão

Análise de Regressão

- Do modelo de regressão:

$$Y_i = \beta_0 + \tau w_i + \varepsilon_i$$

- Temos que (Lousa)

$$E(Y_i|w_i = 1) - E(Y_i|w_i = 0) = \tau + [E(\varepsilon_i|w_i = 1) - E(\varepsilon_i|w_i = 0)]$$

- Viés de seleção para o tratamento:

$$[E(\varepsilon_i|w_i = 1) - E(\varepsilon_i|w_i = 0)]$$

- Hipótese de exogeneidade:

$$[E(\varepsilon_i|w_i = 1) = E(\varepsilon_i|w_i = 0)]$$

ou

$$Y_{0i}, Y_{1i} \perp W_i$$

Análise de Regressão

- A inclusão do vetor de covariadas (X_i) pode eliminar o viés de seleção

$$Y_i = \beta_0 + \tau w_i + X_i\theta + \varepsilon_i$$

$$E(Y_i|X_i, w_i = 1) - E(Y_i|X_i, w_i = 0) = \tau + [E(\varepsilon_i|X_i, w_i = 1) - E(\varepsilon_i|X_i, w_i$$

- Presença de X seria o motivo para que ε e w sejam não correlacionados (X prevê o mecanismo).
- Hipótese de exogeneidade:

$$[E(\varepsilon_i|w_i = 1, X_i) = E(\varepsilon_i|w_i = 0, X_i)]$$

ou

$$Y_{0i}, Y_{1i} \perp W_i | X_i$$

Análise de Regressão

Vantagens:

- Análise simples
- Possibilidade de incluir heterogeneidades no efeito (#obs do grupo): $Y_i = \beta_0 + \sum_s \tau_s w_{is} + X_i\theta + \varepsilon_i$
- Efeito médio

Críticas:

1. Dependência da forma funcional (especificação)
2. Falta de suporte comum: Aproximação local, problema de extrapolação dos resultados. Imbens e Rubin 2012: Para relevantes X, se $|\bar{X}_1 - \bar{X}_0| > \frac{1}{2}\hat{\sigma}_X$: regressão não remove viés associado a diferenças das covariadas.

Alternativas à Análise de Regressão

Alternativas:

- Para 1 e 2 -> Métodos baseados em *propensity score*: não exige forma funcional (em geral) e compara observações (não distribuições)
- Outro uso de análise de regressão -> Modelo de Seleção de Heckman: modela equação de seleção

Modelo de Seleção de Heckman

Modelo de Seleção de Heckman

Equação estrutural (mec. variável de resposta):

$$y_i = X_i\beta + w_i\tau + \varepsilon_i$$

Equação de seleção (mec. seleção):

$$w_i^* = Z_i\gamma + u_i$$

w_i^* é uma variável latente endógena $w_i = I(w_i^* > 0)$

$$\varepsilon_i^*, u_i \sim \mathcal{N}\begin{bmatrix} 0, & \sigma & \rho \\ 0 & \rho & 1 \end{bmatrix}$$

2 Regimes, tratamento e controle, portanto estimação em 2 estágios ou máxima verossimilhança. Lousa

Modelo de Seleção de Heckman

Função de Máxima Verossimilhança:

Para $w_i = 1$:

$$lnL_i = ln\Phi\left(\frac{(z_i\gamma + (y_i - X_i\beta - \tau))\rho/\sigma}{\sqrt{1-\rho^2}}\right) - 1/2\frac{(y_i - X_i\beta - \tau)^2}{\sigma^2} - ln\sqrt{2\pi\sigma^2}$$

Para $w_i = 0$:

$$lnL_i = ln\Phi\left(\frac{(-z_i\gamma - (y_i - X_i\beta - \tau))\rho/\sigma}{\sqrt{1-\rho^2}}\right) - 1/2\frac{(y_i - X_i\beta)^2}{\sigma^2} - ln\sqrt{2\pi\sigma^2}$$

É possível estimar τ com boas propriedades

Modelo de Seleção de Heckman

- ao propor equação de seleção, Heckman também propõe como estimar o propensity score (Stata: treatreg)
- lida com problema de seleção de forma indireta (VI)
- veremos usos mais flexíveis do propensity score
- vantagens: quando existe correlação entre e u
- sensível a erros de especificação

Análise usando *Propensity Score*

Análise usando *propensity score*

- Propensity-score (PS): probabilidade de seleção ao tratamento condicional a um vetor de covariadas, $P(w = 1|X)$ ou $e(X)$.
  - Após cond. X, não há nada sistemático na seleção (sobram aleatoriedades)
  - Ideia: Buscar covariadas que causam a maior diferença (**desbalanceamento**) entre os grupos
- Tipos de análises usando e(X):
  - Propensity score matching: Comparar Y de tratados e não tratados com mesmo e(X) (vários tipos de pareamento)
  - Propensity score subclassification/estratificado/grupo: Agrega unidades num S de e(X) (retem mais obs.)
  - Propensity score weighting: Uso do e(X) como peso (vem sendo o mais utilizado)
  - Estimadores duplo robustos: Regressão + e(X) (combinação de métodos)
  - Propensity score e métodos não paramétricos

Análise usando *propensity score*

- Rosenbaum e Rubin (1983) mostraram que se há seleção nas observáveis:

$$Y_i(1), Y_i(0) \perp w_i | X_i \tag{2}$$

Temos

$$Y_i(1), Y_i(0) \perp w_i | e(X_i) \tag{3}$$

Problema da dimensionalidade e o PS

- O vetor de K covariadas X pertence a $\Re^K$
- Veja que $e(X)$ que pertence a $\Re$.
- Redução de K dimensões a 1: vantagem do PS.
- $e(X)$ é uma medida de balanceamento que resume a informação do vetor X (indivíduos com mesmo e(X) tem a mesma distribuição de covariadas).
- Se vale seleção nas observáveis, elimina o viés de seleção condicional a e(X).

Propriedades do e(X)

1. e(X) balanceia diferenças entre tratados e não tratados (indivíduos com mesmo e(X) tem mesma distribuição de X)
2. seleção para o tratamento e as covariadas são condicionalmente independentes (condicional a e(x))
3. se há seleção nas observáveis, e(x) é o escore de balanceamento e $ATE|_{e(X)} = ATT|_{e(X)}$

Prop. 1 a 3 são a fundação das análises usando P.S.

Prop. 2: Balanceamento entre X's de tratados e controles

Teste de Balanceamento

- Teste Bivariado
- Teste de Distribuição (Kolmogorov-Smirnov)
- Teste Ranksum (Wilcoxon-Mann-Whitney)

Lousa