

EAE5837-Microdesenvolvimento e Economia Política

Raphael Corbi

Universidade de São Paulo

Agosto 2017

Quatro Perguntas Básicas do Projeto de Pesquisa

1. Qual a relação causal de interesse?

- ▶ útil para prever as consequências de políticas públicas
- ▶ conceito de *contrafactual*
- ▶ *relação de interesse teórico (parâmetro do modelo)*
- ▶ *ex: impacto de educação em salário (mudança no custo da universidade)*
- ▶ *unidade de análise: indivíduo, firmas, cidades, países*
- ▶ *ex: instituições democráticas e crescimento econômico (Acemoglu, Johnson, Robinson, 2001)*

Quatro Perguntas Básicas do Projeto de Pesquisa

2. Qual o experimento ideal para capturar a relação causal?

- ▶ experimentos ideais são geralmente hipotéticos
- ▶ zero restrição orçamentária e ética
- ▶ escolha da tópicos promissores e na formulação precisa da pergunta
- ▶ ex: alocação aleatória de instituições no dia da independência
- ▶ algumas perguntas não possuem experimento ideal

Quatro Perguntas Básicas do Projeto de Pesquisa

2. Qual o experimento ideal para capturar a relação causal?

- ▶ exemplo: sexo/raça e discriminação
- ▶ Bertrand and Mullainathan (2004) - resposta de empregadores à CVs com nomes que soam negros/brancos (e.g. Lakisha and Emily)
- ▶ Fryer and Levitt (2004)
- ▶ algumas perguntas não possuem experimento ideal
- ▶ idade ao início da escola vs maturidade vs tempo-na-escola

Quatro Perguntas Básicas do Projeto de Pesquisa

3. Qual a estratégia de identificação?

- ▶ experimentos ideais quase nunca existem
- ▶ dados observados (não-aleatórios) aproximam experimento natural
- ▶ parte da informação observada é aleatória
- ▶ ex: mês de nascimento e lei de escolaridade compulsória (Angrist Krueger, 2001)

- ▶ **first-best**: experimento ideal (alocação aleatória)
- ▶ **second-best**: experimento natural
 - ▶ IV, diff-in-diff, RDD, propensity score matching
 - ▶ validação interna: *instrumento afeta somente tratamento?*

Quatro Perguntas Básicas do Projeto de Pesquisa

4. Qual o modo da inferência estatística?

- ▶ depende da população estudada, amostragem
- ▶ hipóteses na construção do erros-padrão
- ▶ ex: microdados do censo vs dados agrupados
 - ▶ clustering (classroom), serial-correlation (DID)
 - ▶ nickel bias (dynamic pane with finite T)

*"T-stat looks too good.
Use robust standard errors
significance gone."*

Problema de Seleção

- ▶ Pergunta: **Pronto-Socorro torna pacientes mais saudáveis?**
- ▶ hospitais oferecem tratamento, mas também contato com doentes
- ▶ dados da National Health Interview Survey
- ▶ 2 perguntas: foi ao pronto-socorro? condição de saúde 1 a 5?

	Sample size	Mean Health Status	Std Error
No Hospital	90,049	3.93	0,003
Hospital	7,774	3.21	0.014
Difference	—	0.72	0.000

Problema de Seleção - formalizando...

Abordagem dos Resultados Potenciais

- ▶ Status de tratamento observado $D_i = 0, 1$
- ▶ Resultado observado Y_i
- ▶ Resultado potencial $Y_{0,i}, Y_{1,i}$
- ▶ Efeito Tratamento indivíduo: $Y_{1,i} - Y_{0,i}$

$$Y_i = \begin{cases} Y_{1,i} & \text{if } D_i = 1 \\ Y_{0,i} & \text{if } D_i = 0 \end{cases}$$

$$Y_i = Y_{0,i} + (Y_{1,i} - Y_{0,i})D_i$$

Problema de Seleção - formalizando...

$$\underbrace{E[Y_i|D_i = 1] - E[Y_i|D_i = 0]}_{\text{diferença observada em saúde média}} = \underbrace{E[Y_{1,i}|D_i = 1] - E[Y_{0,i}|D_i = 1]}_{\text{efeito tratamento médio nos tratados}} + \underbrace{E[Y_{0,i}|D_i = 1] - E[Y_{0,i}|D_i = 0]}_{\text{viés de seleção}}$$

- ▶ viés de seleção mascara o efeito causal do tratamento
- ▶ grande objetivo da pesquisa empírica é superar este viés

Alocação Aleatória e o Problema de Seleção

- ▶ alocação aleatória do tratamento resolve o problema da seleção
- ▶ torna D_i independente de $Y_{0,i}, Y_{1,i}$

$$\begin{aligned} E[Y_i|D_i = 1] - E[Y_i|D_i = 0] &= E[Y_{1,i}|D_i = 1] - E[Y_{0,i}|D_i = 0] \\ &= E[Y_{1,i}|D_i = 1] - E[Y_{0,i}|D_i = 1] \\ &= E[Y_{1,i}] - E[Y_{0,i}] \end{aligned}$$

- ▶ efeito causal é igual a simples diferença de médias condicionais
- ▶ aleatorização resolve o grande problema da seleção
- ▶ porém há outros problemas...

Benchmark: Experimento Aleatório

- ▶ desenho de pesquisa mais confiável e influente
- ▶ idéia: **tratamento e controle são aleatórios**
- ▶ resolve o **problema de seleção!**

- ▶ e.g. *Perry treatment group, 1962-1993*
- ▶ *123 jovens negros escolhidos aleatoriamente*
- ▶ *moldes para o programa pré-escolar Head Start*
 - ▶ *reanálise dos dados: efeito somente sobre meninas (Anderson, 2008)*

Tennessee STAR Experiment

- ▶ qual o efeito do tamanho da classe sobre aprendizado?
- ▶ literatura não-experimental não acha efeito
- ▶ **seleção**: alocação dos alunos por sala não é aleatória

- ▶ **STAR experiment**: 1985-1986 com 11,600 crianças na pré-escola
- ▶ podem participar todas escolas com 3 turma por ano
- ▶ acompanha os alunos por 4 anos (até 3^a série)
- ▶ controle 1: 22-25 alunos sem assistente
- ▶ tratamento 2: 22-25 alunos com assistente
- ▶ tratamento 3: 13-17 alunos

Tennessee STAR Experiment

Aleatorização foi alcançada (pelo menos nas observáveis)?

Table 2.2.1: Comparison of treatment and control characteristics in the Tennessee STAR experiment

		Students who entered STAR in kindergarten			
	Variable	Small	Regular	Regular/Aide	Joint <i>P</i> -value
1.	Free lunch	.47	.48	.50	.09
2.	White/Asian	.68	.67	.66	.26
3.	Age in 1985	5.44	5.43	5.42	.32
4.	Attrition rate	.49	.52	.53	.02
5.	Class size in kindergarten	15.10	22.40	22.80	.00
6.	Percentile score in kindergarten	54.70	48.90	50.00	.00

Análise de Regressão de Experimentos

- ▶ dados observados reescritos como regressão:

$$Y_i = \underbrace{\alpha}_{E[Y_{0i}]} + \underbrace{\rho}_{Y_{1i}-Y_{0i}} D_i + \underbrace{\eta_i}_{Y_{0i}-E[Y_{0i}]}$$

- ▶ tirando a média condicional:

$$E[Y_i | D_i = 1] = \alpha + \rho + E[\eta_i | D_i = 1]$$

$$E[Y_i | D_i = 0] = \alpha + E[\eta_i | D_i = 0]$$

$$E[Y_i | D_i = 1] - E[Y_i | D_i = 0] = \underbrace{\rho}_{\text{treatment}} + \underbrace{E[\eta_i | D_i = 1] - E[\eta_i | D_i = 0]}_{\text{selection bias}}$$

- ▶ viés de seleção corresponde a correlação entre erro η_i e tratamento D_i

Tennessee STAR Experiment

Table 2.2.2: Experimental estimates of the effect of class-size assignment on test scores

Explanatory variable	(1)	(2)	(3)	(4)
Small class	4.82 (2.19)	5.37 (1.26)	5.36 (1.21)	5.37 (1.19)
Regular/aide class	.12 (2.23)	.29 (1.13)	.53 (1.09)	.31 (1.07)
White/Asian (1 = yes)	-	-	8.35 (1.35)	8.44 (1.36)
Girl (1 = yes)	-	-	4.48 (.63)	4.39 (.63)
Free lunch (1 = yes)	-	-	-13.15 (.77)	-13.07 (.77)
White teacher	-	-	-	-.57 (2.10)
Teacher experience	-	-	-	.26 (.10)
Master's degree	-	-	-	-0.51 (1.06)
School fixed effects	No	Yes	Yes	Yes
R ²	.01	.25	.31	.31