

Introdução à Inferência Causal

DAGs e quantias de interesse

Fernando Meireles
IESP-UERJ

Agosto, 2019

Inferir causalidade

Como vimos na aula anterior, sustentar afirmações causais, do tipo $X \rightarrow Y$, depende principalmente de pressupostos (e um X).

- É possível usar nosso conhecimento para justificar quais unidades podem ser comparadas (algo injustificável na maioria dos casos);
- Ou estatística para comparar valores esperados.

Efeito causal

O problema fundamental da inferência causal

- Um efeito causal nada mais é do que a diferença entre dois resultados potenciais;
- Em uma mesma unidade i , apenas um é observado -- o outro é um contrafactual.

Efeito causal individual

Se considerarmos que $Y_i(1)$ é o resultado potencial da unidade i quando ela recebe tratamento; e $Y_i(0)$ o resultado potencial quando ela não recebe, então

O efeito causal individual é

$$Y_i(1) - Y_i(0)$$

D_i sendo um indicador de tratamento ($D_i = 1$ quando a unidade i é tratada), eu observo $Y_i(1)$ quando $D_i = 1$.

Usando factuais e pressupostos

Podemos
comparar
resultados
potenciais
observados de
unidades
diferentes, mas
precisamos de
pressupostos

- Unidades são homogêneas (comum nas ciências naturais e em algumas tradições nas humanas);
- O estatístico, que tem como ideal um experimento;

Ideal experimental

Em um experimento, o tratamento é por definição independente dos resultados potenciais. Com isso, uma diferença entre valores esperados me permite calcular o efeito causal médio:

$$E[Y_i(1)|T_i = 1] - E[Y_i(0)|T_i = 0]$$

Posso usar apenas valores observados:

$$E[Y_i|T_i = 1] - E[Y_i|T_i = 0]$$

Estimador simples

Usamos
diferenças de
médias para
obter um **efeito
causal médio**

- O mecanismo causal atuando pode ter diferentes formas funcionais (o *framework* é flexível);
- A diferença de valor esperado serve para identificar efeitos causais médios.

Identificação

O que significa identificar um efeito causal?

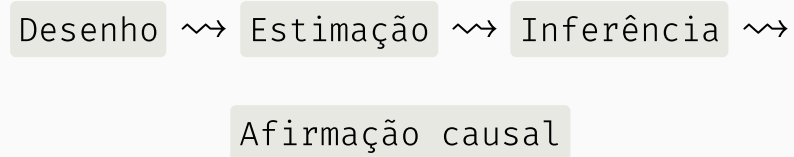
Identificar um efeito não depende necessariamente do tamanho da amostra, mas de:

- uma fonte de variação exógena no mecanismo de seleção (i.e., $\{Y(1), Y(0)\} \perp\!\!\!\perp D$) plausível;
- De técnica para explorá-la (um estimador).

\rightsquigarrow Um efeito causal é identificado quando temos um pressuposto que justifica $\{Y(1), Y(0)\} \perp\!\!\!\perp D$ e um estimador.

Inferência baseada em desenhos

A visão do conjunto



DAGs

Grafos Acíclicos Diretos

Antes de passarmos para outras quantias e aplicações práticas, vamos nos deter em algo que pode ser útil:

- Grafos causais (Knight and Winship, 2013; Pearl and others, 2009).

Relações diretas

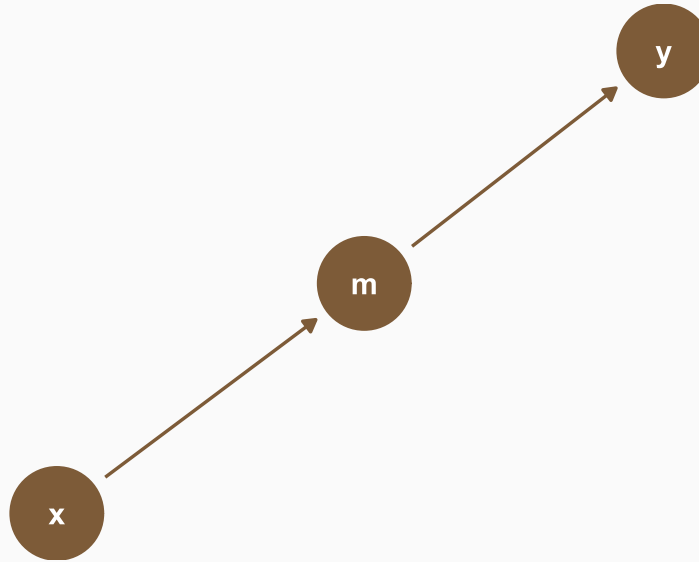
DAGs

estabelecem
relações **causais**
entre variáveis

- Elementos (ou nós, como X) representam variáveis;
- \leftarrow e \rightarrow indicam causalidade direta (lê-se $X \rightarrow Y$ como X causa Y)
- Setas apontam para uma única direção (daí o *directed*);
- Cadeias causais têm setas que apontam todas para a mesma direção;
- Cadeias diferentes não são ligadas diretamente (mas podem vir a ser).

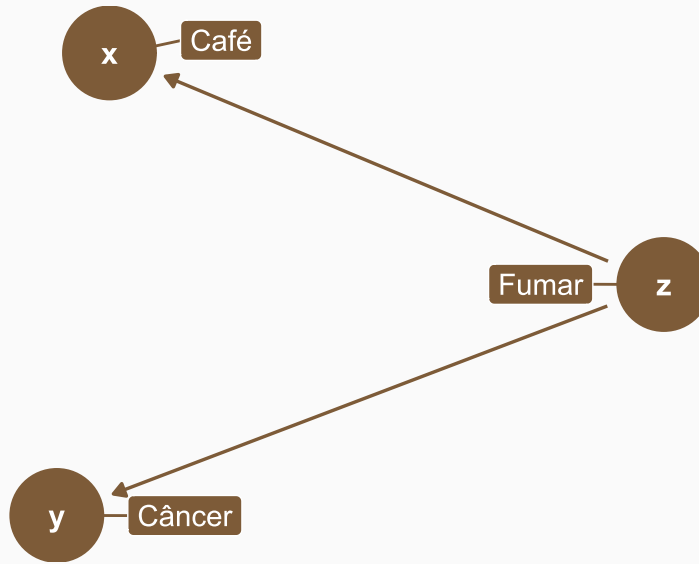
Exemplos de DAGs

Cadeia causal



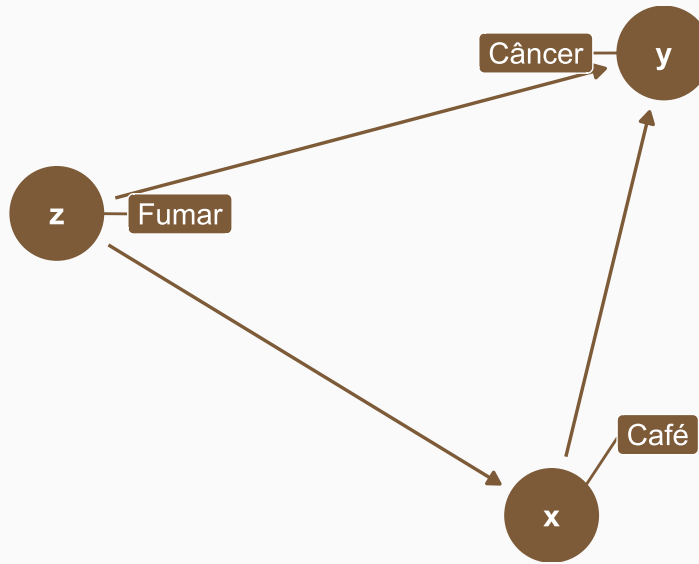
Caminhos não causais

Fatores conectados podem não ter relação causal



Confundidores

Cadeias podem abrir portas para viés



Porta dos fundos

Ou back door

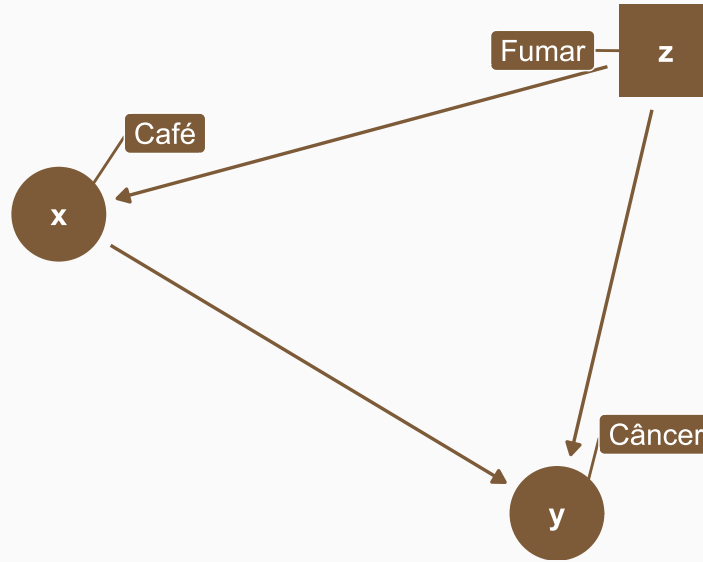
Fumar é um confundidor: ele é relacionado tanto com Y quanto com D e, portanto, abre uma porta que deixa escapar o efeito direto.

Parte da associação que vemos entre Y e D , em outras palavras, é espúria.

Se eu observo isso, posso controlar condicionar no confundidor (regressão, *matching*, ponderação).

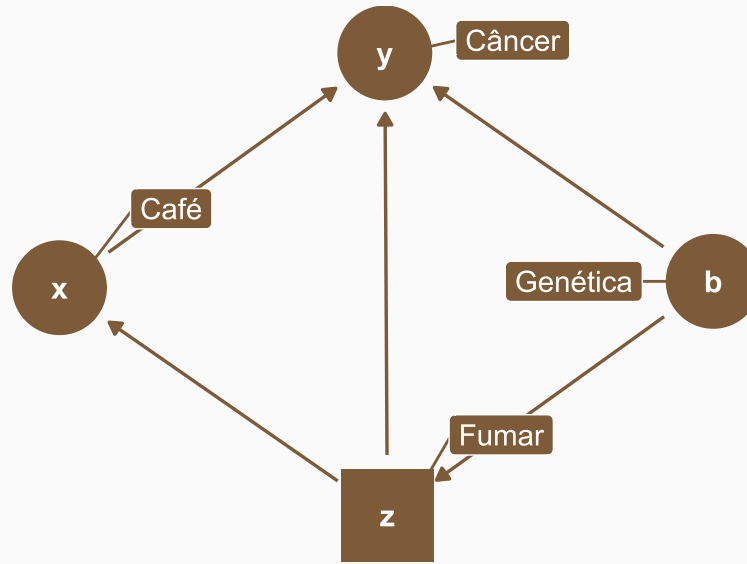
Confundidores

Fechando uma porta dos fundos



Confundidores

Fechando todas as portas dos fundos



Caminhos, não variáveis

Critério dos fundos

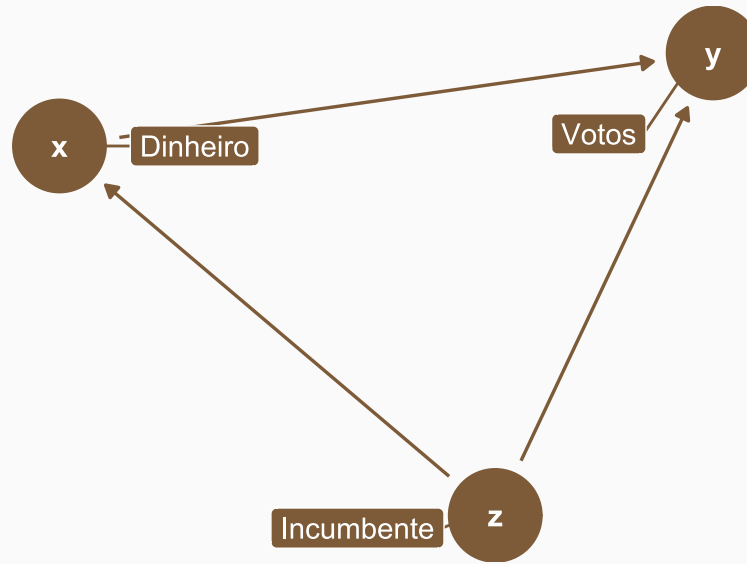
O que importa é o caminho, não necessariamente as variáveis.

Por isso, muitas vezes o importante não é controlar muitas variáveis em um modelo para evitar viés.

O capítulo 4 [Cunningham \(2018\)](#) cobre detidamente essas lições sobre DAGs, para quem quiser se aprofundar.

Viés clássico

Financiamento de campanhas e resultados eleitorais



Exercícios com R

Quantias de interesse

No modelo de **resultados potenciais**, estimamos parâmetros populacionais para **identificar efeitos causais**

- O primeiro que vimos foi o *efeito causal médio do tratamento*;
- Outro comum: o *efeito do tratamento sobre os tratados*;
- Mais um: o *efeito da intenção de tratar*.

Pesquisas que estimam aquelas quantias de interesse

- Gerber, Green, and Shachar (2003), que usou uma cartas ou visitas para incentivar eleitores americanos nas eleições de 1998 (intenção para tratar);
- Bertrand and Mullainathan (2004), que usou um *survey experiment* para identificar se nomes racializados recebem menos propostas de emprego.

Estimadores

Outros estimadores

Cobrimos a diferença entre médias, que um dos estimadores mais simples. O problema é que ele pode não servir em algumas aplicações

Outros comuns incluem:

- Regressão linear, útil para variáveis dependentes lineares e até binárias;
- Inferência randomizada (*aka* aproximações do teste exato de Fisher), não-paramétrico.

Materiais para praticar

Na pasta `4_scripts_basicos` vocês encontram o código em R para usar esses estimadores.

Para quem quiser se aprofundar neles, ver Dunning (2012).

Referências

- Bertrand, M. and S. Mullainathan (2004). "Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination". In: *American economic review* 94.4, pp. 991-1013.
- Dunning, T. (2012). *Natural experiments in the social sciences: a design-based approach*. Cambridge University Press.
- Gerber, A. S, D. P. Green, and R. Shachar (2003). "Voting may be habit-forming: evidence from a randomized field experiment". In: *American Journal of Political Science* 47.3, pp. 540-550.
- Knight, C. R. and C. Winship (2013). "The causal implications of mechanistic thinking: Identification using directed acyclic graphs (DAGs)". In: *Handbook of causal analysis for social research*. Springer, pp. 275-299.
- Pearl, J. and others (2009). "Causal inference in statistics: An overview". In: *Statistics surveys* 3, pp. 96-146.