

Seminários de Doutorado

Análise de mediação em epidemiologia: uma introdução

Rodrigo Citton P. dos Reis
citton.padilha@ufrgs.br

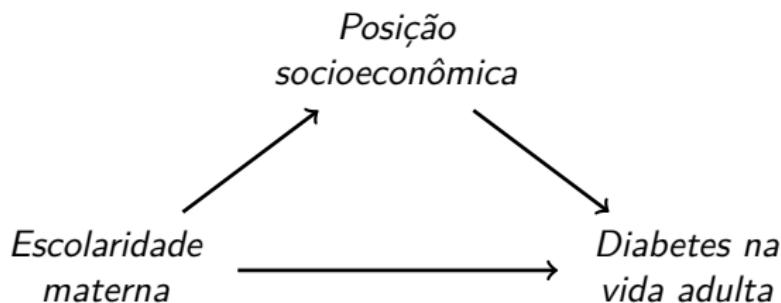
UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE MEDICINA
PROGRAMA DE PÓS-GRADUAÇÃO EM EPIDEMIOLOGIA

Porto Alegre, 2021



Introdução

Introdução



Como ocorre um efeito - o fenômeno da mediação

- ▶ Uma maneira de explicar uma relação de causa e efeito é explicar **como** a causa afeta o desfecho.
- ▶ Podemos descrever o **mecanismo** pelo qual isso ocorre.
- ▶ Esse mecanismo pode ser concebido como uma explicação de como a causa e certos estados iniciais levam a um estado final específico (o desfecho) por meio de um **processo** ou uma série de processos envolvendo diferentes **estágios intermediários**.

Como ocorre um efeito - o fenômeno da mediação

- ▶ Explicações (causais) desse tipo descrevem o que podemos e não podemos aprender sobre tais mecanismos e processos a partir da análise estatística de dados empíricos.
- ▶ Os métodos relacionados a este tipo de explicação enfocarão o cenário no qual um investigador acredita que um determinado **estado intermediário** pode ser responsável por parte ou a maior parte do efeito da causa no desfecho.
- ▶ Tais métodos tentam avaliar que parte do efeito da causa sobre o desfecho está de fato operando por meio desse intermediário em particular e que parte pode funcionar por meio de outros mecanismos ou **caminhos (trajetórias)**.

Como ocorre um efeito - o fenômeno da mediação

- ▶ O efeito da causa sobre o desfecho que opera por meio do intermediário de interesse é algumas vezes referido como um **efeito indireto** ou **efeito mediado**.
- ▶ O efeito da causa sobre o desfecho que não é por meio do intermediário de interesse é algumas vezes referido como o **efeito direto** ou **efeito não mediado**.

Efeito direto

- ▶ É importante ter em mente que tais efeitos são diretos apenas em relação ao intermediário de interesse;
- ▶ Provavelmente haverá outros intermediários ou mecanismos responsáveis por outros aspectos do efeito da causa no desfecho.

Como ocorre um efeito - o fenômeno da mediação

- ▶ O fenômeno pelo qual uma causa afeta um intermediário e a mudança no intermediário passa a afetar o desfecho é o que geralmente é referido como o fenômeno de “**mediação**”.
- ▶ O conjunto de técnicas pelas quais um pesquisador avalia a magnitude relativa desses efeitos diretos e indiretos é algumas vezes referido como “**análise de mediação**”.
- ▶ O próprio intermediário às vezes é chamado de “**mediador**” (**variável mediadora**).

Motivações para avaliar a mediação

- ▶ Em alguns casos, a motivação pode ser simplesmente **explicação e compreensão**.
- ▶ Um exemplo da **epidemiologia genética**: variantes genéticas foram encontradas associadas ao **comportamento de fumar** e ao **câncer de pulmão**.
 - ▶ **Questão:** as variantes afetaram o câncer de pulmão apenas porque afetaram o tabagismo e sabemos que fumar causa câncer de pulmão ou as variantes afetaram o câncer de pulmão por outros caminhos que não o fumo?

Motivações para avaliar a mediação

- ▶ O estudo da mediação também pode ajudar a **confirmar e refutar teoria**.
- ▶ Um exemplo da **epidemiologia do curso de vida**: foi constatado repetidamente que o **baixo status socioeconômico (SSE)** durante a infância está associado a desfechos adversos de saúde mais tarde na vida.
 - ▶ No entanto, permanece o debate se isso ocorre porque o baixo SSE durante a infância afeta o SSE adulto, o que por sua vez afeta a saúde do adulto (**um modelo de “trajetória social”**), ou se o SSE infantil afeta a saúde do adulto através de outros caminhos que não o SSE adulto (**um modelo de “efeitos latentes/periodo sensível”**), ou ambos.

Motivações para avaliar a mediação

- ▶ Em alguns cenários, podemos não ser capazes de **intervir** na exposição primária ou na causa do desfecho diretamente e, portanto, podemos estar interessados em saber se podemos eliminar um efeito prejudicial de uma exposição intervindo **em algum mecanismo particular** ou intermediário.
 - ▶ No exemplo de epidemiologia genética acima, não podemos intervir diretamente nas variantes genéticas, mas podemos estar interessados em quanto do efeito das variantes genéticas no câncer de pulmão poderíamos bloquear se pudéssemos intervir para eliminar o fumo.

Motivações para avaliar a mediação

- **Disparidades em saúde/Epidemiologia social:** podemos, por exemplo, encontrar diferenças em um desfecho de saúde entre grupos étnico-raciais. Obviamente, não podemos intervir na etnia e raça, mas podemos estar interessados na medida em que as disparidades de saúde entre grupos étnico-raciais podem ser reduzidas ou eliminadas se pudéssemos intervir para igualar os níveis de educação entre os grupos étnico-raciais.

Social Science & Medicine 226 (2019) 236–245



Contents lists available at [ScienceDirect](#)

Social Science & Medicine

journal homepage: www.elsevier.com/locate/socscimed



Methods for analytic intercategorical intersectionality in quantitative research: Discrimination as a mediator of health inequalities

Greta R. Bauer^{a,b,*}, Ayden I. Scheim^{a,c}

^a Epidemiology and Biostatistics, Schulich School of Medicine & Dentistry, Western University, London, Canada

^b Women's Studies and Feminist Research, Western University, London, Canada

^c Division of Infectious Diseases and Global Public Health, University of California San Diego School of Medicine, San Diego, USA



Motivações para avaliar a mediação

- ▶ Em um **ensaio clínico aleatorizado** que uma intervenção tem um efeito benéfico, **em média**, para a população do estudo.
 - ▶ Refinar a **intervenção** de modo a aumentar a magnitude do efeito.
 - ▶ Isso pode ser feito alterando/melhorando os componentes da intervenção que visam um determinado mecanismo para o desfecho.
 - ▶ É importante saber se, e em que medida, o mecanismo alvo é um caminho relevante da intervenção ao desfecho.
 - ▶ Se o mecanismo alvo explica uma grande parte do efeito, pode ser desejável refinar a intervenção para alcançar esse mecanismo.

Principais abordagens de análise

- ▶ Duas linhas principais na literatura para o estudo de mediação:
 - ▶ Abordagem clássica (Ciências sociais/psicometria) (BARON; KENNY, 1986; MACKINNON, 2008).
 - ▶ Abordagem geral (inferência causal) (PEARL, 2001; ROBINS; GREENLAND, 1992).

Abordagem clássica

Baron e Kenny (SEM)

- ▶ Suponha uma exposição X , um desfecho Y , um mediador M , e possíveis confundidoras pré-exposição C .
- ▶ BARON; KENNY (1986) propuseram uma abordagem para análise de mediação que vem sendo utilizado por muitos cientistas sociais.

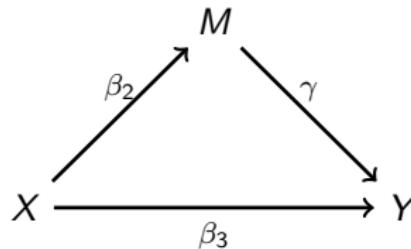
Baron e Kenny (SEM)

- ▶ Esta abordagem é baseada no seguinte sistema de equações estruturais lineares (*linear structural equation model – LSEM*):

$$Y = \alpha_1 + \beta_1 X + \epsilon_1 \quad (1)$$

$$M = \alpha_2 + \beta_2 X + \epsilon_2 \quad (2)$$

$$Y = \alpha_3 + \beta_3 X + \gamma M + \epsilon_3. \quad (3)$$



Baron e Kenny (SEM)

1. Confundidores pré-exposição podem ser incluídos no modelo.
2. BARON; KENNY (1986) sugeriram que a existência de efeito de mediação pode ser testado ao ajustar separadamente os três modelos lineares e testar as hipóteses (i) $\beta_1 = 0$, (ii) $\beta_2 = 0$, e (iii) $\gamma = 0$.
 - ▶ Se todas as hipóteses são rejeitadas, então $\beta_2 \times \gamma$ (*“método do produto dos coeficientes”*) pode ser interpretado como um efeito de mediação.
 - ▶ *Efeito direto de X em Y:* β_3 .
3. Ao substituir a equação (2) na equação (3) obtemos

$$Y = (\alpha_3 + \alpha_2\gamma) + (\beta_3 + \beta_2\gamma)X + (\gamma\epsilon_2 + \epsilon_3),$$

e portanto, a equação (1) é redundante e testar $\beta_1 = 0$ é desencessário, pois o efeito de mediação pode ser não zero, mesmo quando o efeito total é zero.

Baron e Kenny (SEM)

4. Para uma interpretação significante estas definições dependem de suposições que raramente são discutidas, como:
 - ▶ Não existência de confundidores não mensurados;
 - ▶ Correta especificação do modelo paramétrico (por exemplo, linearidade, ausência de interação $X - M$).
5. Estas definições são específicas deste modelo e não existem resultados equivalentes para desfechos e/ou mediadoras binárias/categóricas.
6. A natureza linear do modelo é chave e muito restritiva.

Abordagem geral (inferência causal)

Inferência causal

- ▶ A abordagem formal adotada recentemente com relação a questões de mediação é a dos “**desfechos potenciais**” ou estrutura “**contrafactual**” para inferência causal, que vem sendo amplamente empregada no trabalho metodológico em **estatística, epidemiologia, economia, sociologia, psicologia, educação, ciência da computação** e outras disciplinas.
- ▶ A estrutura fornece uma **notação formal e técnica** para conceituar a causalidade.
- ▶ Isso é feito principalmente concebendo o que poderia ter ocorrido **se** alguma ação ou estado fosse diferente do que era.

Inferência causal



Inferência causal

- ▶ Se algum desfecho tivesse sido diferente se alguma exposição ou ação fosse diferente do que foi, então diríamos que a exposição ou ação causa ou afeta o desfecho.
- ▶ Na literatura estatística, a notação formal para esta abordagem contrafactual foi descrita por NEYMAN (1923) no contexto de experimentos agrícolas aleatorizados.
 - ▶ A estrutura foi posteriormente desenvolvida por RUBIN (1974) e ampliada para estudos observacionais.
 - ▶ A estrutura foi posteriormente estendida para múltiplas exposições e exposições que variam ao longo do tempo por ROBINS (1986) e foi relacionado a representações gráficas de causalidade por SPIRITES; GLYMOUR; SCHEINES (1993) e PEARL (1995).

Inferência causal

- ▶ **Modelo causal de Rubin:** $Y_i(x)$ é o valor pontencial do desfecho caso X_i tivesse sido fixado ao valor x .
 - ▶ Considerando dois tratamentos/níveis de exposição (tratado vs. controle), $x = 0, 1$; o **efeito causal individual** é definido por $Y_i(1) - Y_i(0)$.
- ▶ **Problema fundamental da inferência causal:** o resultado que teria ocorrido no estado contrafactual em que realizamos uma ação diferente da que foi de fato executada está essencialmente ausente ou desconhecido.
 - ▶ Isso é o que torna a inferência causal desafiadora com dados empíricos.

Inferência causal

Roteiro

O roteiro da inferência causal consiste em (pelo menos) três etapas:

1. Uma **linguagem formal** para definir inequivocamente conceitos causais.
 - ▶ Desfechos potenciais, contrafactuals, operador *do()*
2. **Suposições causais** para a **identificação** dos efeitos causais.
 - ▶ **Diagramas causais (DAGs)** são uma ferramenta para exibir nossas suposições causais
3. **Métodos de análise** (isto é, métodos estatísticos) que podem nos ajudar a tirar conclusões causais mais confiáveis a partir dos dados disponíveis.

Inferência causal

- ▶ Embora seja difícil tirar conclusões sobre os efeitos causais para indivíduos específicos, às vezes é possível fazer inferências sobre esses efeitos em média para uma população.
- ▶ A **aleatorização** da ação/exposição/intervenção pode ajudar a garantir que os grupos que recebem as diferentes ações ou intervenções sejam comparáveis em média e, portanto, que qualquer diferença nos resultados entre os grupos que recebem diferentes ações ou intervenções seja atribuível à própria ação.

Identificação do efeito causal total médio

Supondo **consistência, não interferência e intercambiabilidade**, então o efeito causal total médio/populacional (ET) pode ser estimado a partir dos dados observados:

$$ET = E[Y_i(1) - Y_i(0)] = E[Y_i|X_i = 1] - E[Y_i|X_i = 0].$$

Inferência causal

- Com **dados observacionais** ainda podemos tentar controlar vários outros fatores que podem explicar as diferenças nos resultados entre os grupos de intervenção, exceto a própria intervenção.

Identificação do efeito causal total médio

Supondo **consistência, não interferência e intercambiabilidade condicional**, então o efeito causal total médio/populacional (ET) pode ser estimado a partir dos dados observados:

$$\begin{aligned} ET &= E[Y_i(1) - Y_i(0)] \\ &= \sum_{C_i} (E[Y_i|X_i = 1, C_i = c_i] - E[Y_i|X_i = 0, C_i = c_i]) \times \Pr(C_i = c_i), \end{aligned}$$

em que C_i representa um conjunto (vetor) de covariáveis (variáveis pré-tratamento/exposição).

Inferência causal

- ▶ Esse controle pode ajudar a atribuir diferenças nos resultados à causa ou ação em estudo, mas com dados observacionais na ausência de aleatorização, não se pode, em geral, ter certeza de que tal controle foi adequado.

Inferência causal

- ▶ A estrutura de desfechos potenciais não fornece uma caracterização formal de todos os aspectos da **causalidade**; e inúmeras questões sobre causalidade, como os critérios pelos quais identificamos a causa real de um evento, também são essencialmente deixadas sem solução pela estrutura de desfechos potenciais.
- ▶ O que a estrutura contrafactual permite principalmente é um conjunto de definições que fornecem critérios ou condições suficientes indicando que algum evento ou exposição foi a causa de outro - não necessariamente **a** causa, mas **uma** causa.

Inferência causal

- ▶ Esta abordagem contrafactual pode ser estendida para abordar questões de mediação, estendendo a notação contrafactual para incluir configurações contrárias aos fatos não apenas da exposição primária, mas também do mediador.

Análise de medição: abordagem causal

- ▶ No contexto da mediação, também haverá resultados potenciais para a variável mediadora.
- ▶ $M(x)$ denota o valor contrafactual do mediador M de um indivíduo se a exposição X fosse definida como o valor x .
 - ▶ Para uma exposição binária, teríamos novamente dois resultados potenciais para o mediador: $M(1)$ e $M(0)$, o que o mediador teria sido com ou sem a exposição, respectivamente.

Análise de medição: abordagem causal

- ▶ Também consideraremos os resultados potenciais para Y em intervenções hipotéticas tanto na exposição quanto no mediador.
- ▶ $Y(x, m)$ denota o valor contrafactual para Y de um indivíduo se X fosse definido como x e M fosse definido como m .
- ▶ Então, para cada indivíduo, temos um resultado potencial **para cada configuração da exposição e do mediador.**
 - ▶ No exemplo genético, $Y(x = 1, m = 10)$ seria o resultado que teríamos observado se o indivíduo tivesse a variante genética e fumado 10 cigarros por dia em média.
 - ▶ Da mesma forma, $Y(x = 0, m = 20)$ seria o resultado que teríamos observado se o indivíduo não tivesse a variante genética e fumou 20 cigarros por dia em média.

Análise de medição: abordagem causal

- ▶ Temos muitos resultados potenciais diferentes desta forma $Y(x, m)$, mas mais uma vez, só conseguimos observar um deles.
- ▶ Usando contrafactuals dessa forma, ROBINS; GREENLAND (1992) e PEARL (2001) deram então as seguintes definições para o que agora é chamado de “**efeitos diretos controlados**” e “**efeitos diretos e indiretos naturais**”.

Análise de medição: abordagem causal

Efeito direto controlado

- ▶ O efeito direto controlado do tratamento X no desfecho Y comparando $X = 1$ com $X = 0$ e definindo M com m é definido por $Y(1, m) - Y(0, m)$ e **mede o efeito de X em Y não mediado por M** , ou seja, o efeito de X em Y depois de intervir para fixar o mediador em algum valor m .
 - ▶ No exemplo da genética, em uma escala de diferença, o efeito direto controlado, $Y(1, m) - Y(0, m)$, denotaria o efeito no câncer de pulmão, comparando a presença e a ausência da variante genética com cigarros por dia fixados no nível m .
 - ▶ Observe que este efeito direto pode variar com m .
- ▶ O efeito direto controlado médio para uma população é então denotado por $E[Y(1, m) - Y(0, m)]$ e o efeito direto controlado médio (*EDC*), condicional nas covariáveis $C = c$, é denotado por $EDC(m) = E[Y(1, m) - Y(0, m)|c]$.

Análise de medição: abordagem causal

Efeito direto natural

- ▶ O efeito direto natural da exposição X no desfecho Y comparando $X = 1$ com $X = 0$ intervindo para definir M para o que teria sido se a exposição fosse $X = 0$ é formalmente definido por $Y(1, M(0)) - Y(0, M(0))$.
- ▶ Essencialmente, o efeito direto natural assume que o intermediário M é definido como $M(0)$, o nível que teria sido para cada indivíduo se a exposição fosse 0, e então compara o efeito direto do tratamento (com o intermediário fixado para este nível $M(0)$).
 - ▶ No exemplo da genética, o efeito direto natural captura o efeito da exposição, comparando a variante presente com aquela ausente, no desfecho do câncer de pulmão, intervindo para definir o mediador, cigarros por dia, ao nível em que teria ocorrido na ausência do nível de exposição;
 - ▶ Isto é, se a variante genética estivesse ausente.

Análise de medição: abordagem causal

Efeito direto natural (continuação)

- ▶ Podemos definir o efeito direto natural médio (EDN) para a população como $E[Y(1, M(0)) - Y(0, M(0))]$ ou condicional nas covariáveis $C = c$, $EDN = E[Y(1, M(0)) - Y(0, M(0))|c]$.

Análise de medição: abordagem causal

Efeito indireto natural

- ▶ O efeito indireto natural comparando o mediador fixado em $M(1)$ versus $M(0)$ e intervir para definir a exposição em $X = 1$ é formalmente definido por $Y(1, M(1)) - Y(1, M(0))$.
- ▶ O efeito indireto natural assume que a exposição está definida para o nível $X = 1$ e, em seguida, compara o que teria acontecido se o mediador fosse definido para o que teria sido se a exposição fosse $X = 1$ versus o que teria acontecido se o mediador fosse definido para o que teria sido se a exposição fosse $X = 0$.
 - ▶ No exemplo da genética, o efeito indireto natural captura o efeito no câncer de pulmão comparando o que aconteceria se a variante genética estivesse presente, mas fixássemos cigarros por dia no nível que teria sido se a variante estivesse presente ou ausente.

Análise de medição: abordagem causal

Efeito indireto natural (continuação)

- ▶ Observe que para o efeito indireto natural ser diferente de zero, $M(1)$ e $M(0)$ devem ser diferentes um do outro; caso contrário, os dois contrafatuais no contraste, $Y(1, M(1)) - Y(1, M(0))$, seriam os mesmos e sua diferença seria zero.
 - ▶ Portanto, para que o efeito indireto natural seja diferente de zero, a exposição deve mudar o mediador e, então, essa mudança no mediador deve prosseguir para mudar o desfecho.
- ▶ Podemos definir o efeito indireto natural médio (EIN) para a população como $E[Y(1, M(1)) - Y(1, M(0))]$ ou condicional nas covariáveis $C = c$, $EIN = E[Y(1, M(1)) - Y(1, M(0))|c]$.

Análise de medição: abordagem causal

Efeitos total, direto e indireto naturais

- Com essas definições contrafatuais, também podemos ver que um efeito total pode ser decomposto em um efeito direto e indireto natural.

$$\begin{aligned} ET &= E[Y(1) - Y(0)] \\ &= E[Y(1, M(1)) - Y(0, M(0))] \\ &= E[(Y(1, M(1)) - Y(1, M(0))) + (Y(1, M(0)) - Y(0, M(0)))] \\ &= E[(Y(1, M(1)) - Y(1, M(0)))] + E[(Y(1, M(0)) - Y(0, M(0)))] \\ &= EIN + EDN. \end{aligned}$$

- Também podemos definir esses efeitos na **razão de riscos** ou na escala de **razão de chances**.

Análise de medição: abordagem causal

Efeitos controlado e naturais: resumo

- ▶ O $CDE(m)$ expressa o quanto o desfecho mudaria em média se o mediador fosse fixado no nível m **uniformemente na população**, mas o tratamento fosse alterado do nível $x = 0$ para o nível $x = 1$.
- ▶ O NDE captura qual seria o efeito da exposição sobre o desfecho **se desativássemos o caminho da exposição ao mediador**.
- ▶ O NIE captura o efeito da exposição sobre o desfecho que **opera mudando o mediador**.
 - ▶ Enquanto os efeitos diretos controlados são frequentemente de maior interesse na **avaliação de políticas** porque consideram qual seria o efeito da exposição se intervíssemos no mediador em toda a população, os efeitos direto e indireto natural podem ser de maior interesse na avaliação da ação de vários mecanismos e na importância dos diferentes caminhos e para a **decomposição do efeito**.

Análise de medição: abordagem causal

Suposições para identificação dos efeitos controlado e naturais

- S1.** Ausência de confundimento não mensurado na realação $X - Y$;
- S2.** Ausência de confundimento não mensurado na realação $M - Y$;
- S3.** Ausência de confundimento não mensurado na realação $X - M$;
- S4.** Ausência de confundidor na realação $M - Y$ afetado por X (ausência de confundidor intermediário).

Análise de medição: abordagem causal

Identificação dos efeitos controlado e naturais

- Quando as suposições S1-S4 são válidas, então o efeito direto controlado médio e os efeitos direto e indireto naturais médios, condicionais às covariáveis $C = c$, são identificados a partir dos dados observados por

$$\begin{aligned} EDC(m) &= E[Y(1, m) - Y(0, m)|c] \\ &= E[Y|X = 1, M = m, C = c] - E[Y|X = 0, M = m, C = c)] \\ EDN &= E[Y(1, M(0)) - Y(0, M(0))|c] \\ &= \sum_m \{E[Y|X = 1, M = m, C = c] - E[Y|X = 0, M = m, C = c]\} \Pr(M = m|X = 0, C = c) \} \\ EIN &= E[Y(1, M(1)) - Y(1, M(0))|c] \\ &= \sum_m E[Y|X = 1, M = m, C = c] \{\Pr(M = m|X = 1, C = c) - \Pr(M = m|X = 0, C = c)\} \end{aligned}$$

Análise de medição: abordagem causal

- ▶ Uma abordagem bastante comum para a estimação dos efeitos direto e indireto se dá através da especificação de um **modelo para o desfecho** $E[Y|x, m, c]$ e um **modelo para o mediador** $Pr(m|x, c)$ e então **calcular analiticamente** a expressão do lado direito das equações acima.
 - ▶ Se mudarmos os modelos, teremos que passar por um novo cálculo.
- ▶ No entanto, como as expressões acima são muito gerais e não pressupõem nenhum modelo estatístico específico, às vezes são chamadas de “**não paramétricas**”.

Análise de medição: abordagem causal

- ▶ Essa abordagem contrafactual é, portanto, **completamente geral** em termos dos modelos que pode acomodar.
 - ▶ Poderíamos especificar modelos diferentes envolvendo termos quadráticos, ou outras interações ou não linearidades adicionais, e ainda poderíamos usar as fórmulas acima para derivar os efeitos diretos e indiretos.
 - ▶ No entanto, cada vez que uma forma funcional diferente para o modelo é considerada, precisaríamos fazer uma nova derivação.

Análise de medição: abordagem causal

- ▶ Ainda, para alguns casos (de acordo com as especificações dos modelos) não existe forma fechada para a solução analítica.
 - ▶ Solução analítica aproximada;
 - ▶ Solução aproximada por simulação estocástica.
- ▶ Por fim, intervalos de confiança podem ser construídos pelo método delta ou *bootstrap*.

Análise de medição: abordagem causal

Exemplo: normalidade + linearidade

$$M = \alpha_2 + \beta_2 X + \epsilon_2$$

$$Y = \alpha_3 + \beta_3 X + \gamma M + \epsilon_3$$

$$\begin{aligned} EDN &= \int_m \{(\alpha_3 + \beta_3 + \gamma m) - (\alpha_3 + \gamma m)\} \times f_M(m|X=0) dm \\ &= \beta_3 \int_m f_M(m|X=0) = \beta_3. \end{aligned}$$

$$\begin{aligned} EIN &= \int_m (\alpha_3 + \beta_3 + \gamma m) \{f_M(m|X=1) - f_M(m|X=0)\} dm \\ &= \alpha_3 + \beta_3 + \gamma E[M|X=1] - (\alpha_3 + \beta_3 + \gamma E[M|X=1]) \\ &= \gamma(\alpha_2 + \beta_2) - \gamma\alpha_2 = \gamma\beta_2. \end{aligned}$$

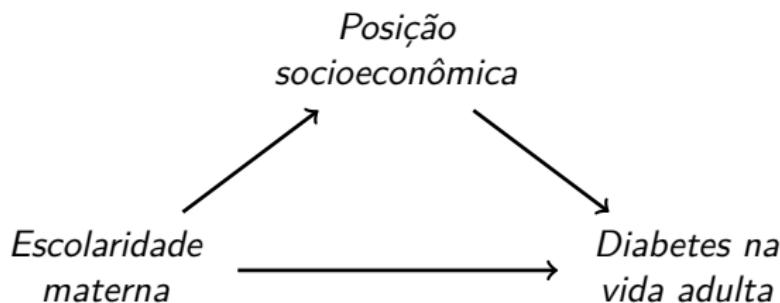
Pacotes estatísticos

Pacotes para análise de mediação causal

- ▶ Pacotes em R
 - ▶ mediation (TINGLEY *et al.*, 2014)
 - ▶ medflex (STEEN *et al.*, 2017)
- ▶ SAS
 - ▶ PROC CAUSALMED
- ▶ Stata
 - ▶ paramed (EMSLEY; LIU, 2013)

Exemplo

Escolaridade materna e diabetes



Escolaridade materna e diabetes

- ▶ Dados de 13.629 participantes (1.497 (11%) com Diabetes) de Onda 2 do Estudo Longitudinal de Saúde do Adulto (ELSA-Brasil).
- ▶ Escolaridade materna: 1º grau completo ou mais = 11.551 (87%); Nunca frequentou a escola = 1.765 (13%).
- ▶ Posição socioeconômica (vida adulta): escore criado a partir de variáveis como renda, escolaridade atual e natureza da ocupação.

$$M = \alpha_2 + \beta_2 X + \epsilon_2$$

$$\Pr(Y = 1) = \Phi(\alpha_3 + \beta_3 X + \gamma M).$$

- ▶ *EDN* e *EIN* estimados por simulação (pacote mediation).

Escolaridade materna e diabetes

Table 1: Mediação simples: Primeiro grau incompleto ou mais (0) vs. Nunca frequentou a escola (1).

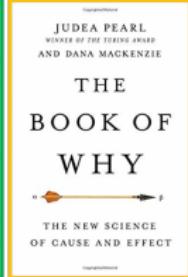
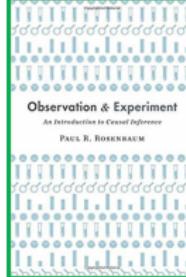
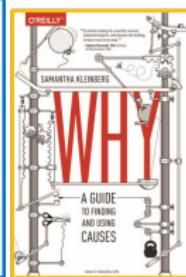
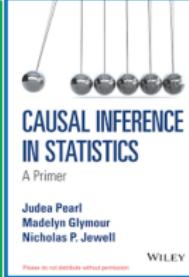
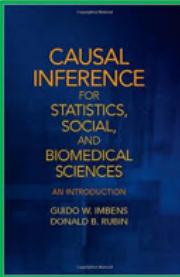
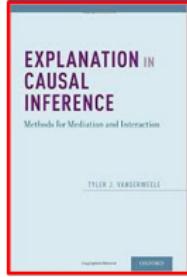
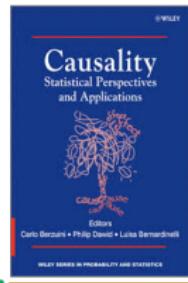
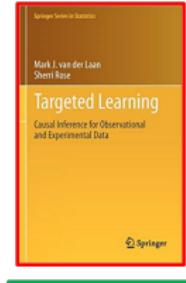
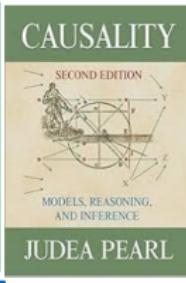
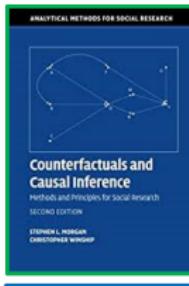
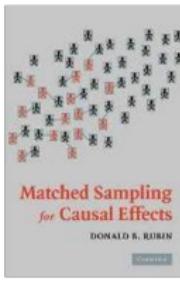
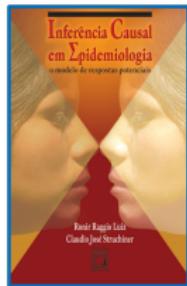
Efeitos causais	Estimativa	IC 95% Inferior	IC 95% Superior
Efeito Indireto Natural	0.0352	0.0287	0.0428
Efeito Direto Natural	0.0172	0.0024	0.0351
Efeito Total	0.0524	0.0354	0.0691
Prop. Mediada	0.6857	0.5012	0.9326

Considerações finais

Considerações finais

- ▶ Para afirmar que os efeitos são causais, não é suficiente o uso de efeitos causalmente definidos.
 - ▶ Exige um **modelo causal/teórico**.
- ▶ Suposições para identificação dos efeitos causais são não testáveis; necessidade de **análise de sensibilidade**.
- ▶ Modelos “mais realistas”:
 - ▶ Múltiplos mediadores;
 - ▶ Confundidoras intermediárias;
 - ▶ Mediação e interação (decomposições tripla e quádrupla).

Considerações finais



Referências bibliográficas

- BARON, R. M.; KENNY, D. A. The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. **Journal of Personality and Social Psychology**, [S. l.], v. 51, n. 6, p. 1173–1182, 1986. Disponível em: <https://doi.org/10.1037/0022-3514.51.6.1173>
- EMSLEY, R.; LIU, H. **PARAMED: Stata module to perform causal mediation analysis using parametric regression models**. Boston College Department of Economics, 2013. Disponível em: <https://ideas.repec.org/c/boc/bocode/s457581.html>. Acesso em: 18 nov. 2021.
- MACKINNON, D. P. **Introduction to Statistical Mediation Analysis**. [S. l.]: Routledge, 2008.
- NEYMAN, J. **On the Application of Probability Theory to Agricultural Experiments. Essay on Principles. Section 9**. 1923. phdthesis[s. l.], 1923.

Referências bibliográficas

PEARL, J. Causal Diagrams for Empirical Research. **Biometrika**, [S. I.], v. 82, n. 4, p. 669–688, 1995. Disponível em:
<https://doi.org/10.2307/2337329>. Acesso em: 18 nov. 2021.

PEARL, J. Direct and indirect effects. In: 2001, Seattle, Washington. **Proceedings of the Seventeenth conference on Uncertainty in artificial intelligence**. Seattle, Washington: [s. n.], 2001. p. 411–420.

ROBINS, J. M. A new approach to causal inference in mortality studies with a sustained exposure period – application to control of healthy worker survivor effect. **Mathematical Modeling**, [S. I.], v. 7, p. 1393–1512, 1986.

ROBINS, J. M.; GREENLAND, S. Identifiability and Exchangeability for Direct and Indirect Effects. **Epidemiology**, [S. I.], v. 3, n. 2, p. 143–155, 1992. Disponível em: <https://www.jstor.org/stable/3702894>. Acesso em: 18 nov. 2021.

Referências bibliográficas

RUBIN, D. B. Estimating causal effects of treatments in randomized and nonrandomized studies. **Journal of Educational Psychology**, [S. l.], v. 66, p. 688–701, 1974.

SPIRITES, P.; GLYMOUR, C.; SCHEINES, R. **Causation, Prediction, and Search**. New York, NY: Springer New York, 1993. (Lecture Notes em Statistics).v. 81 Disponível em:
<https://doi.org/10.1007/978-1-4612-2748-9>. Acesso em: 18 nov. 2021.

STEEN, J. et al. medflex: An R Package for Flexible Mediation Analysis using Natural Effect Models. **Journal of Statistical Software**, [S. l.], v. 76, n. 11, p. 1–46, 2017. Disponível em:
<https://doi.org/10.18637/jss.v076.i11>

TINGLEY, D. et al. mediation: R Package for Causal Mediation Analysis. **Journal of Statistical Software**, [S. l.], v. 59, n. 5, p. 1–38, 2014.
Disponível em: <http://www.jstatsoft.org/v59/i05/>