[SER 2021 Workshop] Causal Mediation: Modern Methods for Path Analysis

Iván Díaz, Nima Hejazi, Kara Rudolph

updated: May 07, 2021

Contents

Welcome to SER!

This open source, reproducible vignette accompanies a half-day workshop on modern methods for *causal mediation analysis*, given at the SER 2021 Meeting[1] on Monday, 24 May 2021. While we encourage use of this bookdown site, for convenience, we have also made these workshop materials available in PDF[2].

0.1 About this workshop

Causal mediation analysis can provide a mechanistic understanding of how an exposure impacts an outcome, a central goal in epidemiology and health sciences. However, rapid methodologic developments coupled with few formal courses presents challenges to implementation. Beginning with an overview of classical direct and indirect effects, this workshop will present recent advances that overcome limitations of previous methods, allowing for: (i) continuous exposures, (ii) multiple, non-independent mediators, and (iii) effects identifiable in the presence of intermediate confounders affected by exposure. Emphasis will be placed on flexible, stochastic and interventional direct and indirect effects, highlighting how these may be applied to answer substantive epidemiological questions from real-world studies. Multiply robust, nonparametric estimators of these causal effects, and free and open source R packages (medshift[3] and medoutcon[4]) for their application, will be introduced.

To ensure translation to real-world data analysis, this workshop will incorporate hands-on R programming exercises to allow participants practice in implementing the statistical tools presented. It is recommended that participants have working knowledge of the basic

[1] https://epiresearch.org/annual-meeting/2021-meeting/workshop/
[2] https://code.nimahejazi.org/ser2021_mediation_workshop/ser2021mediation.pdf
[3] https://github.com/nhejazi/medshift
[4] https://github.com/nhejazi/medoutcon

notions of causal inference, including counterfactuals and identification (linking the causal effect to a parameter estimable from the observed data distribution). Familiarity with the R programming language is also recommended.

0.2 Workshop schedule

- 10:00A-10:30A: introductions/mediation set up
- 10:30A-11:00A: estimands and how to choose
- 11:00A-11:30A: discussion: how to choose in real-world examples
- 11:30A-12:00P: shift parameter introduction with application in lecture part
- 12:00P-12:15P break/discussion
- 12:15P-12:45P estimation for natural direct and indirect effects, interventional direct and indirect effects
- 12:45P-01:15P: practice R code for estimation
- 01:15P-01:30P: estimation for stochastic interventional direct and indirect effects
- 01:30P-01:50P: practice: code for estimation
- 01:50P-02:00P wrap up

NOTE: All times listed in Pacific Time.

0.3 About the instructors

Iván Díaz

I am an Assistant Professor at Weill Cornel Medicine. My research focuses on the development of non-parametric statistical methods for causal inference from observational and randomized studies with complex datasets, using machine learning. This includes but is not limited to mediation analysis, methods for continuous exposures, longitudinal data including survival analysis, and efficiency guarantees with covariate adjustment in randomized trials. I am also interested in general semi-parametric theory, machine learning, and high-dimensional data.

Nima Hejazi

I am a PhD candidate in biostatistics at UC Berkeley, working under the joint direction of Mark van der Laan and Alan Hubbard. My research interests fall at the intersection of causal inference and machine learning, drawing on ideas from non/semi-parametric estimation in large, flexible statistical models to develop efficient and robust statistical procedures for evaluating complex target estimands in observational and randomized studies. Particular areas of current emphasis include causal mediation/path analysis, outcome-dependent sampling designs, targeted loss-based estimation, and applications in vaccine efficacy trials. I am also passionate about statistical computing and open source software development for applied statistics.

Kara Rudolph

I am an Assistant Professor of Epidemiology at Columbia University. My research interests are in developing and applying causal inference methods to understand social and contextual influences on mental health, substance use, and violence in disadvantaged, urban areas of the United States. My current work focuses on developing methods for transportability and mediation, and subsequently applying those methods to understand how aspects of the school and peer environments mediate relationships between neighborhood factors and adolescent drug use across populations. More generally, my work on generalizing/ transporting findings from study samples to target populations and identifying subpopulations most likely to benefit from interventions contributes to efforts to optimally target available policy and program resources.

0.4 Reproduciblity

These workshop materials were written using bookdown[5], and the complete source is available on GitHub[6]. This version of the book was built with R version 4.0.5 (2021-03-31), pandoc[7] version r rmarkdown::pandoc_version(), and the following packages:

[5] http://bookdown.org/
[6] https://github.com/tlverse/tlverse-handbook
[7] https://pandoc.org/

package	version	source
bookdown	0.21.11	Github (rstudio/bookdown@33c4f70)
bslib	0.2.4.9003	Github (rstudio/bslib@e09af88)
dagitty	0.3-1	CRAN (R 4.0.5)
data.table	1.14.0	CRAN (R 4.0.5)
downlit	0.2.1	CRAN (R 4.0.5)
dplyr	1.0.5	CRAN (R 4.0.5)
ggdag	0.2.3	CRAN (R 4.0.5)
ggfortify	0.4.11	CRAN (R 4.0.5)
ggplot2	3.3.3	CRAN (R 4.0.5)
kableExtra	1.3.4	CRAN (R 4.0.5)
knitr	1.32	CRAN (R 4.0.5)
magick	2.7.1	CRAN (R 4.0.5)
medoutcon	0.1.0	Github (nhejazi/medoutcon@f8f14c4)
medshift	0.1.4	Github (nhejazi/medshift@f9e11a9)
mvtnorm	1.1-1	CRAN (R 4.0.5)
origami	1.0.3	CRAN (R 4.0.5)
pdftools	2.3.1	CRAN (R 4.0.5)
readr	1.4.0	CRAN (R 4.0.5)
rmarkdown	2.7.11	Github (rstudio/rmarkdown@e340d75)
skimr	2.1.3	CRAN (R 4.0.5)
sl3	1.4.3	Github (tlverse/sl3@5cddc6c)
stringr	1.4.0	CRAN (R 4.0.5)
tibble	3.1.1	CRAN (R 4.0.5)
tidyr	1.1.3	CRAN (R 4.0.5)

0.5 Setup instructions

0.5.1 R and RStudio

R and **RStudio** are separate downloads and installations. R is the underlying statistical computing environment. RStudio is a graphical integrated development environment (IDE) that makes using R much easier and more interactive. You need to install R before you install RStudio.

0.5.1.1 Windows

0.5.1.1.1 If you already have R and RStudio installed

- Open RStudio, and click on "Help" > "Check for updates". If a new version is available, quit RStudio, and download the latest version for RStudio.
- To check which version of R you are using, start RStudio and the first thing that appears in the console indicates the version of R you are running. Alternatively, you can type `sessionInfo()`, which will also display which version of R you are running. Go on the CRAN website[8] and check whether a more recent version is available. If so, please download and install it. You can check here[9] for more information on how to remove old versions from your system if you wish to do so.

0.5.1.1.2 If you don't have R and RStudio installed

- Download R from the CRAN website[10].
- Run the .exe file that was just downloaded
- Go to the RStudio download page[11]
- Under *Installers* select **RStudio x.yy.zzz - Windows XP/Vista/7/8** (where x, y, and z represent version numbers)
- Double click the file to install it
- Once it's installed, open RStudio to make sure it works and you don't get any error messages.

0.5.1.2 macOS / Mac OS X

0.5.1.2.1 If you already have R and RStudio installed

- Open RStudio, and click on "Help" > "Check for updates". If a new version is available, quit RStudio, and download the latest version for RStudio.

[8] https://cran.r-project.org/bin/windows/base/

[9] https://cran.r-project.org/bin/windows/base/rw-FAQ.html#How-do-I-UNinstall-R_003f

[10] http://cran.r-project.org/bin/windows/base/release.htm

[11] https://www.rstudio.com/products/rstudio/download/#download

- To check the version of R you are using, start RStudio and the first thing that appears on the terminal indicates the version of R you are running. Alternatively, you can type `sessionInfo()`, which will also display which version of R you are running. Go on the CRAN website[12] and check whether a more recent version is available. If so, please download and install it.

0.5.1.2.2 If you don't have R and RStudio installed

- Download R from the CRAN website[13].
- Select the .pkg file for the latest R version
- Double click on the downloaded file to install R
- It is also a good idea to install XQuartz[14] (needed by some packages)
- Go to the RStudio download page[15]
- Under *Installers* select **RStudio x.yy.zzz - Mac OS X 10.6+ (64-bit)** (where x, y, and z represent version numbers)
- Double click the file to install RStudio
- Once it's installed, open RStudio to make sure it works and you don't get any error messages.

0.5.1.3 Linux

- Follow the instructions for your distribution from CRAN[16], they provide information to get the most recent version of R for common distributions. For most distributions, you could use your package manager (e.g., for Debian/Ubuntu run `sudo apt-get install r-base`, and for Fedora `sudo yum install R`), but we don't recommend this approach as the versions provided by this are usually out of date. In any case, make sure you have at least R 3.3.1.
- Go to the RStudio download page[17]
- Under *Installers* select the version that matches your distribution, and install it with your preferred method (e.g., with Debian/Ubuntu `sudo dpkg -i rstudio-x.yy.zzz-amd64.deb` at the terminal).

[12] https://cran.r-project.org/bin/macosx/
[13] http://cran.r-project.org/bin/macosx
[14] https://www.xquartz.org/
[15] https://www.rstudio.com/products/rstudio/download/#download
[16] https://cloud.r-project.org/bin/linux
[17] https://www.rstudio.com/products/rstudio/download/#download

- Once it's installed, open RStudio to make sure it works and you don't get any error messages.

These setup instructions are adapted from those written for Data Carpentry: R for Data Analysis and Visualization of Ecological Data[18].

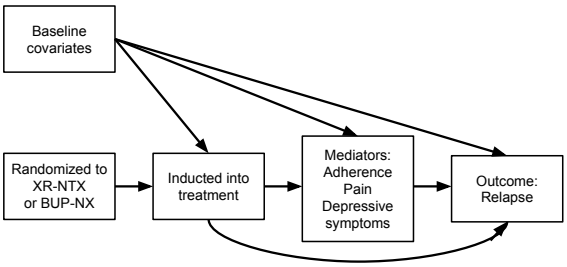
[18] http://www.datacarpentry.org/R-ecology-lesson/

Chapter 1

Preliminaries on causal mediation analysis

1.1 Motivating study

Do differences in the effects of treatment (comparing two medications for opioid use disorder, naltrexone vs buprenorphine) on risk of relapse operate through mediators of adherence, opioid use, pain, and depressive symptoms? (Rudolph et al., 2020)



1.2 What is causal mediation analysis?

- Like all causal analyses, *causal mediation analysis* is different from standard statistical mediation analyses
- Statistical mediation analyses merely assess associations between the variables
- Causal mediation analyses assess how the system behaves under interventions

- Causal mediation analysis is thus useful to understand mechanisms

1.2.1 An example of a non-causal mediation analysis (product of coefficients)

- Assume you are interested in the effect of a treatment A (naltrexone vs. buprenorphine) on an outcome Y (risk of relapse) through mediators M (adherence, opioid use, pain, depressive symptoms)
- We could fit the following models:

$$\mathbb{E}(Y \mid A = a) = \alpha_0 + \alpha_1 a \tag{1.1}$$

$$\mathbb{E}(M \mid A = a) = \gamma_0 + \gamma_1 a \tag{1.2}$$

$$\mathbb{E}(Y \mid M = m, A = a) = \beta_0 + \beta_1 m + \beta_2 a \tag{1.3}$$

- The product $(\gamma_1 \beta_1)$ has been proposed as a measure of the effect of A on Y through M
- Causal interpretation problems with this method:
  - What happens if there are confounders of the relation between treatment and outcome?
  - What happens if there are confounders of the relation between mediator and outcome?
  - What happens if there are confounders of the relation between treatment and mediator?
  - What happens if the confounders of mediator and outcome are affected by treatment? ### Example:
- Assume we have a pre-treamtment confounder of Y and M, denote it with W
- For simplicity, assume A is randomized
- We'll generate a really large sample from a data generating mechanism so that we are not concerned with sampling errors

```
n <- 1e7
w <- rnorm(n)
a <- rbinom(n, 1, 0.5)
m <- rnorm(n, w + a)
y <- rnorm(n, w - a)
```

- Note that the indirect effect (i.e., the effect through M) in this example is zero

- Let's see what the product of coefficients method would say:

```
lm_y <- lm(y ~ m + a)
lm_m <- lm(m ~ a)
## product of coefficients
coef(lm_y)[2] * coef(lm_m)[2]
#>         m
#> 0.50107
```

1.2.2 Statistical issues with this method:

- Assume the above confounding is not an issue
- Can I then interpret $(\gamma_1\beta_1)$ as the indirect effect?
- No: the regression models may be misspecified

1.3 Causal mediation models

In this workshopp we will use directed acyclic graphs to conceptualize the above confounding issues. We will focus on the two types of graph:

1.3.1 No intermediate confounders

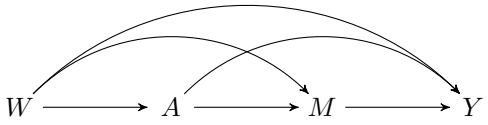


Figure 1.1: Directed acyclic graph under *no intermediate confounders* of the mediator-outcome relation affected by treatment

1.3.2 Intermediate confounders

The above graphs can be interpreted as a *non-parametric structural equation model* (NPSEM), also known as *structural causal model* (SCM):

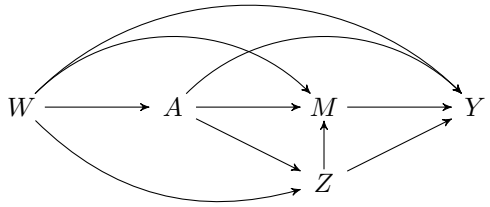


Figure 1.2: Directed acyclic graph under intermediate confounders of the mediator-outcome relation affected by treatment

$$W = f_W(U_W) \tag{1.4}$$
$$A = f_A(W, U_A) \tag{1.5}$$
$$Z = f_Z(W, A, U_Z) \tag{1.6}$$
$$M = f_M(W, A, Z, U_M) \tag{1.7}$$
$$Y = f_Y(W, A, Z, M, U_Y) \tag{1.8}$$

- Here $U = (U_W, U_A, U_Z, U_M, U_Y)$ is a vector of all unmeasured exogenous factors affecting the system
- The functions f are assumed fixed but unknown
- We posit this model as a system of equation that nature uses to geenrate the data at hand
- Therefore we leave the functions f unspecified (i.e., we do not know the true nature mechanisms)
- Sometimes we know something: e.g., if A is randomized we know $A = f_A(U_A)$ where U_A is the flip of a coin (i.e., independent of everything).

1.4 Counterfactuals

- We define all the effects of interest using *counterfactuals*
- Counterfactuals are hypothetical random variables that would have been observed in a world where we would be able to perform interventions on the random variables of interest
- Y_a is a counterfactual variable in a hypothetical world where $\mathbb{P}(A = a) = 1$ with probability one
- $Y_{a,m}$ is the counterfactual outcome in a world where $\mathbb{P}(A = a, M = m) = 1$

- M_a is the counterfactual variable representing the mediator in a world where $\mathbb{P}(A = a) = 1$.

1.4.1 How are counterfactuals defined?

- You can use counterfactual variables as *primitives*
- In the NPSEM framework, counterfactuals are quantities *derived* from the model.
- Take as example the DAG in Figure 1.2:

$$Y_a = f_Y(W, a, Z_a, M_a, U_Y) \tag{1.9}$$

$$Y_{a,m} = f_Y(W, a, Z_a, m, U_Y) \tag{1.10}$$

$$M_a = f_M(W, a, Z_a, U_M) \tag{1.11}$$

- You can also define *nested counterfactuals*
- For example, if A is binary, you can think of the following counterfactual

$$Y_{1,M_0} = f_Y(W, 1, Z_1, M_0, U_Y)$$

- Interpreted as *the outcome for an individual in a hypothetical world where treatment was given but the mediator was held at the value it would have taken under no treatment*
- Causal effects are defined in terms of the distribution of these counterfactuals
- I.e., causal effects give you information about what would happen *under intervention*

Chapter 2

Path-specific casual mediation effect types

- Controlled direct effects
- Natural direct and indirect effects
- Interventional direct and indirect effects

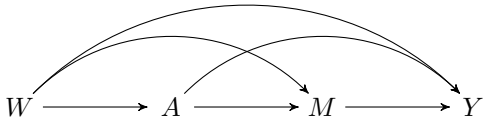


Figure 2.1: Directed acyclic graph under *no intermediate confounders* of the mediator-outcome relation affected by treatment

2.1 Controlled direct effects

- Set the mediator to a reference value $M = m$ uniformly for everyone in the population
- Compare $A = 1$ vs $A = 0$ with $M = m$ fixed

$$\psi_{\text{CDE}} = \mathbb{E}(Y_{1,m} - Y_{0,m})$$

2.1.1 Identification assumptions:

- Confounder assumptions:
  - $A \perp\!\!\!\perp Y_{a,m} \mid W$
  - $M \perp\!\!\!\perp Y_{a,m} \mid W, A$
- Positivity assumptions:
  - $\mathbb{P}(M = m \mid A = a, W) > 0 \; a.e.$
  - $\mathbb{P}(A = a \mid W) > 0 \; a.e.$

Under the above identification assumptions, the controlled direct can be identified:

$$\mathbb{E}(Y_{1,m} - Y_{0,m}) = \mathbb{E}\{\mathbb{E}(Y \mid A = 1, M = m, W) - \mathbb{E}(Y \mid A = 0, M = m, W)\}$$

- For intuition about this formula in R, let's continue with our toy example, where we now add an indirect effect

```
n <- 1e7
w <- rnorm(n)
a <- rbinom(n, 1, 0.5)
m <- rnorm(n, w + a)
y <- rnorm(n, w - a + m)
```

- First we fit a correct model for the outcome

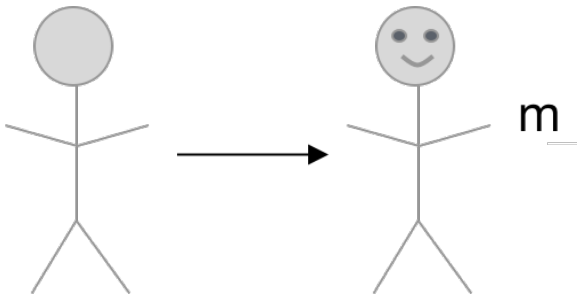
```
lm_y <- lm(y ~ m + a + w)
```

- Assume we would like the CDE at $m = 0$
- Then we generate predictions $\mathbb{E}(Y \mid A = 1, M = m, W)$ and $\mathbb{E}(Y \mid A = 0, M = m, W)$:

```
pred_y1 <- predict(lm_y, newdata = data.frame(a = 1, m = 0, w = w))
pred_y0 <- predict(lm_y, newdata = data.frame(a = 0, m = 0, w = w))
```

- Then we compute the difference between the predicted values $\mathbb{E}(Y \mid A = 1, M = m, W) - \mathbb{E}(Y \mid A = 0, M$ and average across values of W

```
## CDE at m = 0
mean(pred_y1 - pred_y0)
#> [1] -0.99974
```



2.1.2 Is this the estimand I want?

- Makes the most sense if can intervene directly on M
  - And can think of a policy that would set everyone to a single constant level $m \in \mathcal{M}$.
  - J. Pearl calls this *prescriptive*.
  - Can you think of an example?
  - Air pollution, rescue inhaler dosage, hospital visits
  - Does not provide a decomposition of the average treatment effect into direct and indirect effects

What if our research question doesn't involve intervening directly on the mediator?

What if we want to decompose the average treatment effect into its direct and indirect counterparts?

2.2 Natural direct and indirect effects

Still using the same DAG as above,

- Recall the definition of the nested counterfactual

$$Y_{1,M_0} = f_Y(W, 1, Z_1, M_0, U_Y)$$

- Interpreted as *the outcome for an individual in a hypothetical world where treatment was given but the mediator was held at the value it would have taken under no treatment*

- Recal that, because of the definition of counterfactuals
$$Y_{1,M_1} = Y_1$$

Then we can decompose the *average treatment effect* $E(Y_1 - Y_0)$ as follows

$$\mathbb{E}[Y_{1,M_1} - Y_{0,M_0}] = \underbrace{\mathbb{E}[Y_{1,M_1} - Y_{1,M_0}]}_{\text{natural indirect effect}} + \underbrace{\mathbb{E}[Y_{1,M_0} - Y_{0,M_0}]}_{\text{natural direct effect}}$$

- Natural direct effect (NDE): Varying treatment while keeping the mediator fixed at the value it would have taken under no treatment
- Natural indirect effect (NIE): Varying the mediator from the value it would have taken under treatment to the value it would have taken under control, while keeping treatment fixed

- If no effect modification of the effect of A on Y by M (i.e., no interaction between A and M), then CDE = NDE.

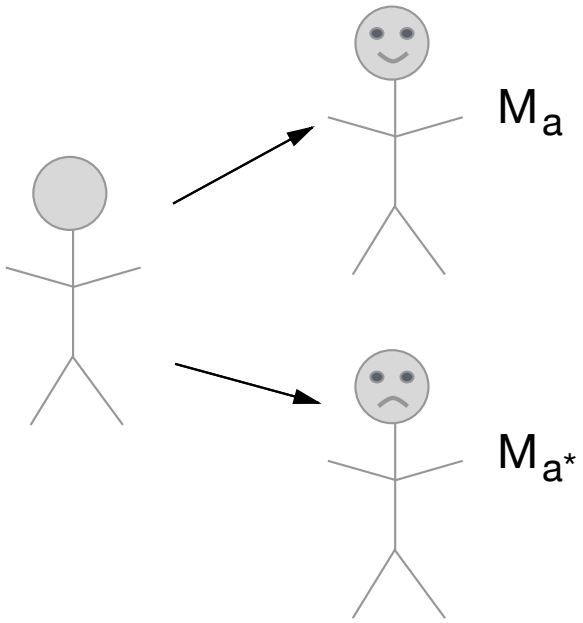
2.2.1 Identification assumptions:

- $A \perp\!\!\!\perp Y_{a,m} \mid W$
- $M \perp\!\!\!\perp Y_{a,m} \mid W, A$
- $A \perp\!\!\!\perp M_a \mid W$
- $M_0 \perp\!\!\!\perp Y_{1,m} \mid W$
- and positivity assumptions

Note: If the cross-world assumption holds (defined below), the NDE can also be written as

$$\mathbb{E} \sum_m \{\mathbb{E}(Y_{1,m} \mid W) - \mathbb{E}(Y_{0,m} \mid W)\} \mathbb{P}(M_0 = m \mid W)$$

- Weighted average of controlled direct effects at each level of $M = m$.



2.2.2 Identification formula:

- Under the above identification assumptions, the natural direct effect can be identified:

$$\mathbb{E}(Y_{1,M_0} - Y_{0,M_0}) = \mathbb{E}[\mathbb{E}\{\mathbb{E}(Y \mid A = 1, M, W) - \mathbb{E}(Y \mid A = 0, M, W) \mid A = 0, W\}]$$

- The natural indirect effect can be identified similarly.
- Let's dissect this formula in R:

```
n <- 1e7
w <- rnorm(n)
a <- rbinom(n, 1, 0.5)
m <- rnorm(n, w + a)
y <- rnorm(n, w - a + m)
```

- First we fit a correct model for the outcome

```
lm_y <- lm(y ~ m + a + w)
```

- Then we generate predictions
$$\mathbb{E}(Y \mid A = 1, M, W) \text{ and } \mathbb{E}(Y \mid A = 0, M, W)$$
with A fixed but letting M and W take their observed values

```
pred_y1 <- predict(lm_y, newdata = data.frame(a = 1, m = m, w = w))
pred_y0 <- predict(lm_y, newdata = data.frame(a = 0, m = m, w = w))
```

- Then we compute the difference between the predicted values
$$\mathbb{E}(Y \mid A = 1, M, W) - \mathbb{E}(Y \mid A = 0, M, W),$$

- and use this difference as a pseudo-outcome in a regression on A and W:
$$\mathbb{E}\{\mathbb{E}(Y \mid A = 1, M, W) - \mathbb{E}(Y \mid A = 0, M, W) | A = 0, W\}$$

```
pseudo <- pred_y1 - pred_y0
lm_pseudo <- lm(pseudo ~ a + m)
```

- Now we predict the value of this pseudo-outcome under $A = 0$, and average the result

```
pred_pseudo <- predict(lm_pseudo, newdata = data.frame(a = 0, w = w))
## NDE:
mean(pred_pseudo)
#> [1] -1.0011
```

2.2.3 Cross-world independence assumption

What does $M_0 \perp\!\!\!\perp Y_{1,m} \mid W$ mean?

- Conditional on W, knowledge of the mediator value in the absence of treatment, M_0, provides no information about the outcome under treatment, $Y_{1,m}$.
- Can you think of a data-generating mechanism that would violate this assumption?
- Example: in a randomized study, whenever we believe that treatment assignment works through adherence (i.e., almost always), we are violating this assumption (more on this later).

2.2.4 Is this the estimand I want?

- Makes sense to intervene on A but not directly on M.
- Want to understand a natural mechanism underlying an association/ total effect. J. Pearl calls this *descriptive*.
- NDE + NIE = total effect (ATE).
- Okay with the assumptions.

What if our data structure involves a post-treatment confounder of the mediator-outcome relationship (e.g., adherence)?

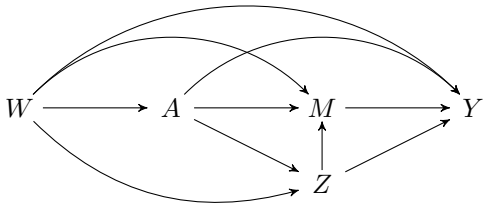
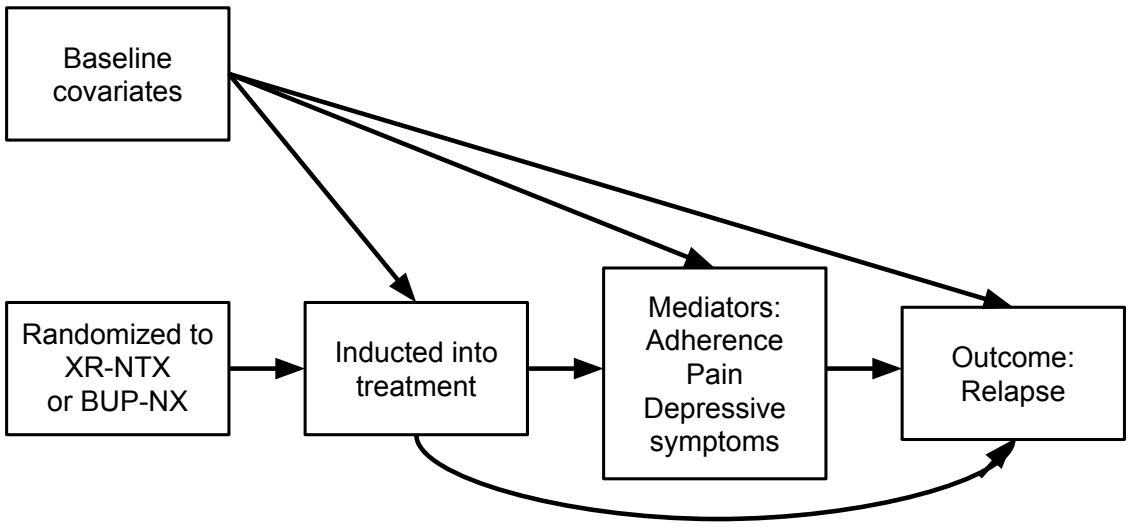


Figure 2.2: Directed acyclic graph under intermediate confounders of the mediator-outcome relation affected by treatment



2.2.5 Unidentifiability of the NDE and NIE in this setting

- In this example, natural direct and indirect effects are unidentifiable from observed data on (W, A, Z, M, Y).
- The technical reason for this is that the cross-world counterfactual assumption
$$Y_{1,m} \perp\!\!\!\perp M_0 \mid W$$
does not hold in the above directed acyiclic graph.
- Technically, the reason for this is that an intervention setting $A = 1$ (necessary for the definition of $Y_{1,m}$) induces a counterfactual variable Z_1.
- Likewise, an intervention setting $A = 0$ (necessary for the definition of M_0) induces a counterfactual Z_0.
- The variables Z_1 and Z_0 are correlated because they share unmeasured common causes.
- The variable Z_1 is correlated with $Y_{1,m}$, and the variable Z_0 is correlated with M_0, because they are counterfactual outcomes in the same hypothetical worlds.
- To see this in the definition of counterfactual from a causal structural model:
$$Y_{1,m} = f_Y(W, 1, Z_1, m, U_Y), \text{ and}$$
$$M_0 = f_M(W, 0, Z_0, U_M)$$
are correlated even after adjusting for W by virtue of Z_1 and Z_0 being correlated.

Intuitively:

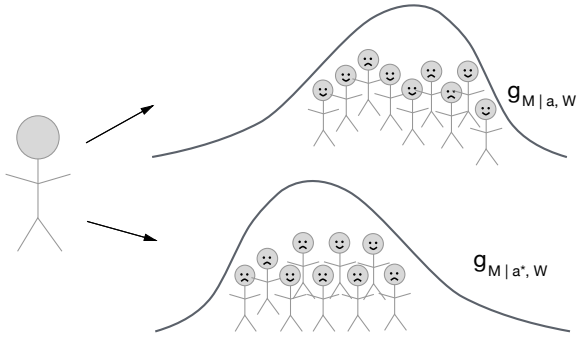
- Z is a confounder of the relation $M \to Y$, which requires adjustment
- But Z is on the pathway $A \to Y$, which precludes adjustment

Note: CDEs are still identified in this setting. They can be identified and estimated similarly to a longitudinal data sructure with a two-time-point intervention.

2.3 Interventional (in)direct effects

- Let G_a denote a random draw from the distribution of $M(a) \mid W$
- Define the counterfactual $Y(1, G_0)$ as the counterfactual variable in a hypothetical world where A is set $A = 1$ and M is set to $M = G_0$ with probability one.

- Define $Y(0, G_0)$ and $Y(1, G_1)$ similarly
- Then we can define:
$$\mathbb{E}[Y(1, G_1) - Y(0, G_0)] = \underbrace{\mathbb{E}[Y(1, G_1) - Y(1, G_0)]}_{\text{interventional indirect effect}} + \underbrace{\mathbb{E}[Y(1, G_0) - Y(0, G_0)]}_{\text{interventional direct effect}}$$
- Marginal PIDE: $\mathbb{E}(Y_{a,g_{M|a^\star,W}}) - \mathbb{E}(Y_{a^\star,g_{M|a^\star,W}})$
- Marginal PIIE: $\mathbb{E}(Y_{a,g_{M|a,W}}) - \mathbb{E}(Y_{a,g_{M|a^\star,W}})$
- Conditional PIDE: $\mathbb{E}(Y_{a,g_{M|Z,a^\star,W}}) - \mathbb{E}(Y_{a^\star,g_{M|Z,a^\star,W}})$
- Conditional PIIE: $\mathbb{E}(Y_{a,g_{M|Z,a,W}}) - \mathbb{E}(Y_{a,g_{M|Z,a^\star,W}})$
- Can you think of an example when you would want the conditional versions? Marginal versions?



Under the following identification assumptions, the population interventional direct and indirect effect is identified:
$$\mathbb{E}(Y_{a,G_{a'}}) = \mathbb{E}\left[\mathbb{E}\left\{\sum_z \mathbb{E}(Y \mid A = a, Z = z, M, W)P(Z = z \mid A = a, w) \mid A = a', W\right\}\right]$$
Identification assumptions:

- $A \perp\!\!\!\perp Y_{a,m} \mid W$
- $M \perp\!\!\!\perp Y_{a,m} \mid W, A, Z$
- $A \perp\!\!\!\perp M_a \mid W$
- and positivity assumptions.

Is this the estimand I want?

- Makes sense to intervene on A but not directly on M.
- Goal is to understand a natural mechanism underlying an association or total effect.
- Okay with the assumptions!

2.4 Estimand Summary

Table 1. Mediation Estimand Definitions, Descriptions, and Assumptions

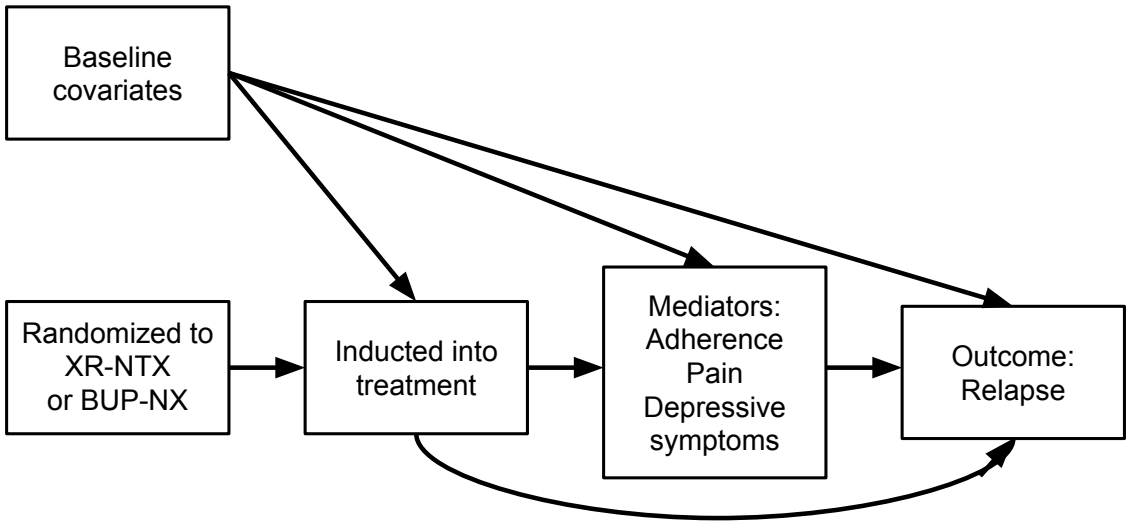
Estimand	Description	Identifying Assumptions in Addition to Positivi Consistency
Controlled direct effect $E(Y_{a,m}) - E(Y_{a^*,m})$	Difference in the expected value of Y setting A to a versus a^* and in both cases setting M to m	1. No unmeasured confounding between A a ($A \perp Y_{a,m} \mid W$). 2. No unmeasured confounding between M a ($M \perp Y_{a,m} \mid W, A$).
Natural direct effect $E(Y_{a,M_{a^*}}) - E(Y_{a^*,M_{a^*}})$ Natural indirect effect $E(Y_{a,M_a}) - E(Y_{a,M_{a^*}})$	Difference in the expected value of Y setting A to a versus a^* and in both cases letting M be the value that it would naturally be under a^* Difference in the expected value of Y in both cases setting A to a and contrasting M under a versus a^*	1. No unmeasured confounding between A a ($A \perp Y_{a,m} \mid W$). 2. No unmeasured confounding between M a ($M \perp Y_{a,m} \mid W, A$). 3. No unmeasured confounding of $A - M$ ($A \perp M_a \mid W$). 4. No measured or unmeasured posttreatme confounding of the $M - Y$ relationship ($M_{a^*} \perp Y_{a,m} \mid W$). 5. Y_a is equivalent to Y_{a,M_a}.
Interventional direct effect $E(Y_{a,g_{M\mid a^*,W}}) - E(Y_{a^*,g_{M\mid a^*,W}})$ Interventional indirect effect $E(Y_{a,g_{M\mid a,W}}) - E(Y_{a,g_{M\mid a^*,W}})$	Difference in the population average of Y setting A to a versus a^* and in both cases drawing the value of M from a distribution of M conditional on $A = a^*$ and the individual's set of covariate values, W Difference in the population average of Y in both cases setting A to a and contrasting drawing the value of M from a distribution of M conditional on $A = a$ versus $A = a^*$ and the individual's set of covariate values, W	1. No unmeasured confounding between A a ($A \perp Y_{a,m} \mid W$). 2. No unmeasured confounding between M a ($M \perp Y_{a,m} \mid W, A$). 3. No unmeasured confounding of $A - M$ ($A \perp M_a \mid W$).

Abbreviations: A, treatment; M, mediator; W, covariates; Y, outcome.

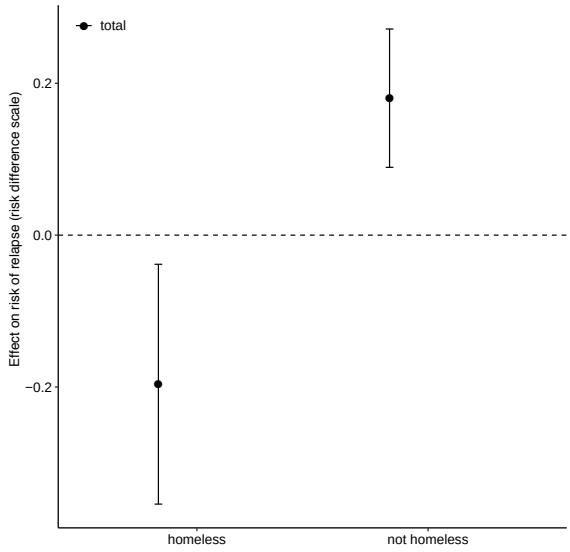
Chapter 3

How to choose an estimand: Real world example

3.1 Comparative effectivness of two medications for opioid use disorder (OUD)



Motivation: Opposite overall treatment effects for homeless versus nonhomeless participants.



3.1.1 Getting specific about the question

To what extent does the indirect effect through mediators of adherence, pain, and depressive symptoms explain the differences in treatment effects on OUD relapse for homeless and nonhomeless individuals?

What estimand do we want?

- Can we set $M = m$ (i.e., same value) for everyone?
- Are we interested in estimating indirect effects?

$\rightarrow$ So, *not* controlled direct effect.

- Do we have an intermediate confounder?
- Yes, and it's important.

$\rightarrow$ So, *not* natural (in)direct effects.

- So, we're left with the interventional direct and indirect effects.
- Do we want to estimate the path through treatment initiation (Z)?
- Yes, so, *not* the conditional versions of these effects.

- Estimands:
  - Direct effect: $\mathbb{E}(Y_{1,g_0} - Y_{0,g_0})$
  - Indirect effect: $\mathbb{E}(Y_{1,g_1} - Y_{1,g_0})$
- Need to incorporate multiple and continuous mediators

Chapter 4

Stochastic Direct and Indirect Effects

4.1 Definition of the effects

Consider the following directed acyclic graph.

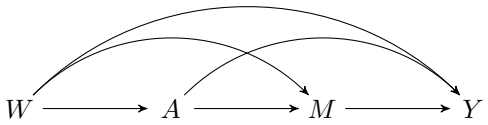


Figure 4.1: Directed acyclic graph under no intermediate confounders of the mediator-outcome relation affected by treatment

4.2 Motivation for stochastic interventions

- So far we have discussed controlled, natural, and interventional (in)direct effects
- These effects require that $0 < \mathbb{P}(A = 1 \mid W) < 1$
- They are defined only for binary exposures
- *What can we do when the positivity assumption does not hold or the exposure is continuous?*
- Solution: we can use stochastic effects

4.3 Definition of stochastic effects

There are two possible ways of defining stochastic effects:

- Consider the effect of an intervention where the exposure is drawn from a distribution
  - Example: [TO FILL IN]
- Consider the effect of an intervention where the post-intervention exposure is a function of the actually received exposure
  - Example: [TO FILL IN]
- In both cases $A \mid W$ is non-deterministic, thus the name *stochastic intervention*

Example: incremental propensity score interventions (IPSI) (Kennedy, 2018)

Definition of the intervention

- Assume A is binary, and $\mathbb{P}(A = 1 \mid W = w) = g(1 \mid w)$ is the propensity score
- Consider an intervention in which each individual receives the intervention with probability $g_\delta(1 \mid w)$, equal to
$$g_\delta(1 \mid w) = \frac{\delta g(1 \mid w)}{\delta g(1 \mid w) + 1 - g(1 \mid w)}$$
- e.g., draw the post-intervention exposure from a Bernoulli variable with probability $g_\delta(1 \mid w)$
- The value δ is user given
- Let A_δ denote the post-intervention distribution
- Some algebra shows that δ is an odds ratio comparing the pre- and post-intervantion distributions
$$\delta = \frac{\text{odds}(A_\delta = 1 \mid W = w)}{\text{odds}(A = 1 \mid W = w)}$$
- This gives the intervention a nice interpretation as *what would happen in a world where the odds of receiving treatment is increased by* δ
- Let Y_{A_δ} denote the outcome in this hypothetical world

4.3.1 Example: modified treatment policies

Definition of the intervention

4.3.2 Mediation analysis for stochastic interventions

- The total effect of an IPSI can be computed as a contrast of the outcome under intervention vs no intervention:
$$\psi = \mathbb{E}[Y_{A_\delta} - Y]$$
- Recall the NPSEM
$$W = f_W(U_W) \tag{4.1}$$
$$A = f_A(W, U_A) \tag{4.2}$$
$$M = f_M(W, A, U_M) \tag{4.3}$$
$$Y = f_Y(W, A, M, U_Y) \tag{4.4}$$
- From this we have $Y_{A_\delta} = f_Y(W, A_\delta, M_{A_\delta}, U_Y)$
- Thus, we have $Y_{A_\delta} = Y_{A_\delta, M_{A_\delta}}$ and $Y = Y(A, M(A))$
- Let us introduce the counterfactual $Y_{A_\delta, M}$, interpreted as the outcome observed in a world where the intervention on A is performed but the mediator is fixed at the value it would have taken under no intervention
- Then we can decompose the total effect into:
$$\mathbb{E}[Y_{A_\delta, M_{A_\delta}} - Y_{A, M_A}] = \underbrace{\mathbb{E}[Y_{A_\delta, M_{A_\delta}} - Y_{A_\delta, M_A}]}_{\text{stochastic natural indirect effect}} + \underbrace{\mathbb{E}[Y_{A, M_{A_\delta}} - Y_{A, M_A}]}_{\text{stochastic natural direct effect}}$$

4.4 Identification of the effect of a stochastic intervention

Chapter 5

Preliminaries on semiparametric estimation

5.1 Why do we need new estimation tools?

- As we have seen all the mediation parameters that we consider can be seen as a function of the joint probability distribution of $O = (W, A, Z, M, Y)$
- For example, under identifiability assumptions the natural direct effect is equal to $\psi(\mathbb{P}) = \sum_{m,w} \Big[\mathbb{E}(Y \mid A = 1, M = m, W = w) - \mathbb{E}(Y \mid A = 0, M = m, W = w)\Big] \mathbb{P}(M = m \mid A = 0, W = w)$
- The notation $\psi(\mathbb{P})$ implies that the parameter is a function of $\mathbb{P}$
- This means that we can compute it for any distribution $\mathbb{P}$
- For example, if we know the true $\mathbb{P}(W, A, M, Y)$, we can comnpute the true value of the parameter by:
  - Computing the conditional expectation $\mathbb{E}(Y \mid A = 1, M = m, W = w)$ for all values (m, w)
  - Computing the probability $\mathbb{P}(M = m \mid A = 0, W = w)$ for all values (m, w)
  - Computing the probability $\mathbb{P}(W = w)$ for all values w
  - Computing the sum over all values (m, w)
- This is how you would compute the *true value* **if you knew** the true distribution $\mathbb{P}$

- But we can use the same logic for estimation:
  - Fit a regression to estimate $\mathbb{E}(Y \mid A = 1, M = m, W = w)$
  - Fit a regression to estimate $\mathbb{P}(M = m \mid A = 0, W = w)$
  - Estimate $\mathbb{P}(W = w)$ with the empirical distribution
- This is known as the g-computation estimator (more on this estimator later)

5.1.1 How can g-estimation be implemented in practice?

- There are two possible ways to do g-computation estimation:
  - Using parametric models for the above regressions
  - Using flexible data-adaptive regression (aka machine learning)

5.1.2 Pros and cons of parametric models

- Pros:
  - Easy to understand
  - Ease of implementation (standard regression software)
  - Can use the Delta method or the bootstrap for computation of standard errors
- Cons:
  - Unless W and M contain very few categorical variables, it is very easy to misspecify the models
  - This can introduce sizable bias in the estimators
  - This bias is highly problematic
    - We go through a thorough process to correctly specify our causal models to avoid bias
    - Overly simplisitc models introduce bias and squander those efforts
    - The bias can be small or large and you can never know from a single data analysis

5.1.3 Pros and cons of g-computation with data-adaptive regression

- Pros:
  - Easy to understand
  - Alleviate model-misspecification bias
- Cons:
  - Might be harder to implement depending on the regression procedure
  - No general approaches for computation of standard errors and confidence intervals

5.2 Semiparametric estimation - an alternative to solve these problems

5.2.1 Bias/variance tradeoff

- A lot of the recent literature in causal inference with data-adaptive regression uses the following ideas
- G-computation estimation with data-adaptive regression offers an incorrect bias/-variance trade-off
- Specifically, the bias of a g-computation estimator can often be expressed as
$$\psi(\mathbb{P}) - \psi(\mathbb{P}) \approx -\mathbb{E}[D(O;\mathbb{P})]$$
- The function $D(O;\mathbb{P})$ is called *the efficient influence function* (EIF)
- The EIF must be found on a case-by-case basis or each parameter $\psi(\mathbb{P})$
- For example, for estimating the standardized mean $\psi(P) = \mathbb{E}[\mathbb{E}(Y \mid A = 1, W)]$, we have
$$D(O,\hat{\mathbb{P}}) = \frac{A}{\hat{P}(A = 1 \mid W)}[Y - \hat{\mathbb{E}}(Y \mid A = 1, W)] + \hat{\mathbb{E}}(Y \mid A = 1, W) - \psi(\hat{\mathbb{P}})$$
- In this workshop we will not present the specific form of $D(O;\hat{\mathbb{P}})$ for all parameters that we use
- But the estimators we discuss and implement in the packages will be based on theser EIFs

5.2.2 Bias-correction of g-computation estimators

- There are at least two ways to use the EIF to perform a bias correction for a g-computation estimator
- The first one is the so-called *one step* estimator:

$$\psi(\hat{\mathbb{P}}) + \frac{1}{n}\sum_{i=1}^{n} D(O; \hat{\mathbb{P}}_i)$$

- The idea behind the one-step estimator is simple: subtract an estimate of the bias of the g-computation estimator
- The second approach is the *targeted maximum likelihood estimator* (TMLE)
- TMLE is based on the principle that it is possile to construct a data-adaptive estimator $\tilde{\mathbb{P}}$ such that

$$\frac{1}{n}\sum_{i=1}^{n} D(O; \tilde{\mathbb{P}}_i) = 0$$

- Thus, for this special data-adaptive estimate $\tilde{\mathbb{P}}$, the TMLE is actually just the g-computation estimator $\psi(\tilde{\mathbb{P}})$

Chapter 6

Estimating natural and interventional effects

6.1 Natural (in)direct effects

Recall:

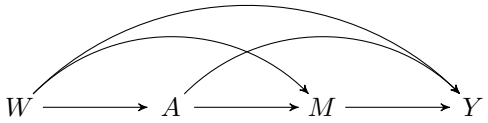


Figure 6.1: Directed acyclic graph under *no intermediate confounders* of the mediator-outcome relation affected by treatment

Assuming a binary A, we define the natural direct effect as:

$$NDE = E(Y_{1,M_0} - Y_{0,M_0})$$

,

and the natural indirect effect as:

$$NIE = E(Y_{1,M_1} - Y_{1,M_0})$$

.

6.1.1 Simple case for intuition

The observed data is $O = (W, A, M, Y)$

This SCM is represented in the above DAG and the following causal models:

$$\begin{aligned} W &= f_W(U_W) \\ A &= f_A(W, U_A) \\ M &= f_M(W, A, U_M) \\ Y &= f_Y(W, A, M, U_Y), \end{aligned}$$

where (U_W, U_A, U_M, U_Y) are exogenous random errors.

We assume - A is a single binary randomized treatment (and thus $A = f_A(U_A)$) - M is a single binary mediator - There are no restrictions on the distribution of W or Y

Recall that we need to assume the following to identify the above caual effects from our observed data:

- $A \perp\!\!\!\perp Y_{a,m} \mid W$
- $M \perp\!\!\!\perp Y_{a,m} \mid W, A$
- $A \perp\!\!\!\perp M_a \mid W$
- $M_0 \perp\!\!\!\perp Y_{1,m} \mid W$
- and positivity assumptions

6.1.2 How to estimate using G-computation

Let's take the NDE as an example:

1. Fit a regression of Y on M, A, W. Predict outcome values setting $A = 1$. We'll call the result $\bar{Q}_Y(M, 1, W)$. Predict outcome values setting $A = 0$. We'll call the result $\bar{Q}_Y(M, 0, W)$.
2. Take the difference $\bar{Q}_Y(M, 1, W) - \bar{Q}_Y(M, 0, W)$ and regress it on W among those for whom $A = 0$. This recovers the expected difference had all individuals been set to the control condition $A = 0$.
3. The sample mean of the predicted values gives the estimate.

6.1.3 How to estimate using the doubly robust methods that rely on the EIF

The EIC for the NDE (Ψ_{NDE}) is given by:

$$D^\star = \left\{ \frac{I(A=1)}{g(1|W)} \frac{Q(M|W,0)}{Q(M|W,1)} - \frac{I(A=0)}{g(0|W)} \right\} \times (Y - \bar{Q}_Y(M,A,W)) \tag{6.1}$$

$$+ \frac{I(A=0)}{g(0|W)} \{\bar{Q}_{diff} - E(\bar{Q}_{diff}|W,0)\} \tag{6.2}$$

$$+ E(\bar{Q}_{diff}|W,0) - \Psi_{NDE} \tag{6.3}$$

6.1.4 How to estimate using the one-step estimator (essentially A-IPTW)

1. Estimate
$$C_Y(Q_M, g)(O) = \left\{ \frac{\mathbb{I}(A=1)}{g(1 \mid W)} \frac{Q_M(M \mid 0, W)}{Q_M(M \mid 1, W)} - \frac{\mathbb{I}(A=0)}{g(0 \mid W)} \right\}.$$

   Breaking this down, $\frac{\mathbb{I}(A=1)}{g(1|W)}$ is the inverse probability weight for $A = 1$ and, likewise, $\frac{\mathbb{I}(A=0)}{g(0|W)}$ is the inverse probability weight for $A = 0$. The middle term is the ratio of the mediator density when $A = 0$ to the mediator density when $A = 1$.

Estimating Q_M is a really hard problem when M is high-dimensional. But, since we have the ratio of these conditional densitities, we can reparamterize using Bayes rule to get something that is easier to compute:

$$\frac{\mathbb{P}(A = 0 \mid M, W) g(0 \mid W)}{\mathbb{P}(A = 1 \mid M, W) g(1 \mid W)}.$$

1. We estimate $g_{A|W}(W) = P(A = a \mid W)$ from a logistic regression of A on W, generating predicted probabilities that $A = 1$ for $g(1 \mid W)$ and $A = 0$ for $g(0 \mid W)$.

2. We estimate $\mathbb{P}(A = a \mid M, W)$ from a logistic regression of A on M, W, generating predicted probabilities that $A = 1$ for and $A = 0$.

```
amodel <- "a ~ w "
mmodel <- "m ~ a + w"
amodel <- "a ~ m + w"
ymodel <- "y ~ m + a*w"

# make gm
afit <- glm(formula = amodel, family = "binomial", data = obsdat)
```

```
mfit <- glm(formula = mmodel, family = "binomial", data = obsdat)

a1 <- predict(afit, newdata = data.frame(w = obsdat$w), type = "respon
a0 <- 1 - a1

am1 <- predict(amfit,
  newdata = data.frame(w = obsdat$w, m = obsdat$m),
  type = "response"
)
am0 <- 1 - am1
cy <- (am0 * a0) / (am1 * a1)
```

$\bar{Q}_Y(M, a', W) - \bar{Q}_Y(M, a^\star, W)$

3. To obtain an estimate of $\bar{Q}_{diff} = \bar{Q}_Y(M, 1, W) - \bar{Q}_Y(M, 0, W)$, predict values of Y from a regression of Y on M, A, W, setting $A = 1$ and $A = 0$, giving $\hat{Y}(m, 1, w)$ and $\hat{Y}(m, 1, w)$.

```
qyinit <- cbind(
  predict(glm(
    formula = ymodel, family = "binomial",
    data = data.frame(cbind(datw, a = a, m = m, y = y))
  ),
  newdata = data.frame(cbind(datw, a, m)), type = "response"
  ),
  predict(glm(
    formula = ymodel, family = "binomial",
    data = data.frame(cbind(datw, a = a, m = m, y = y))
  ),
  newdata = data.frame(cbind(datw, a = 0, m)), type = "response"
  ),
  predict(glm(
    formula = ymodel, family = "binomial",
    data = data.frame(cbind(datw, z = z, m = m, y = y))
  ),
  newdata = data.frame(cbind(datw, a = 1, m)), type = "response"
  )
)
```

```
qbardiff <- qyinit[, 3] - qyinit[, 2]
```

5. We then regress $\hat{\bar{Q}}_{diff}(M, A, W)$ on W among those with $A = 0$ and take the mean.

```
margqdiff_fit <- glm(qbardiff ~ w,
  data = data.frame(
    qbardiff = qbardiff[a == 0],
    w = w[a == 0]
  )
)
margqdiff <- predict(margqdiff_fit,
  newdata = data.frame(qbardiff = qbardiffup, w = w)
)
meanmargqdiff<- mean(margqdiff)
```

6. Plug in the estimates of these nuisance parameters to the EIC, set it equal to 0, and solve. This gives an estimate of the NDE.

```
os_nde <- cy*(y-qyinit[,1]) + (a0/am0)*(qbardiff - meanmargqdiff) + meanma
```

6.2 Interventional direct and indirect effects

Recall that in the presence of a intermediate confounder natural (in)direct effects are not identified

We define the interventional direct effect as:

$$\psi_{\text{PIDE}} = \mathbb{E}(Y_{a',g_{M|a^\star,W}} - Y_{a\star,g_{M|a^\star,W}}),$$

and the interventional indirect effect as:

$$\psi_{\text{PIIE}} = \mathbb{E}(Y_{a',g_{M|a',W}} - Y_{a',g_{M|a^\star,W}}).$$

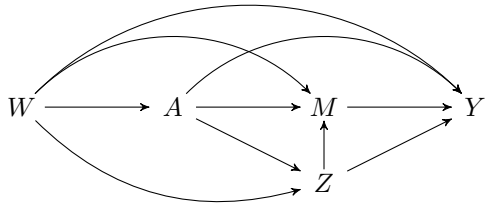


Figure 6.2: Directed acyclic graph under intermediate confounders of the mediator-outcome relation affected by treatment

6.3 Simple case for intuition

Consider a simple data structure $O = (W, A, Z, M, Y)$. This SCM is represented in the above DAG and the following causal models:

$$\begin{aligned} W &= f_W(U_W) \\ A &= f_A(W, U_A) \\ Z &= f_Z(W, A, U_Z) \\ M &= f_M(W, A, Z, U_M) \\ Y &= f_Y(W, A, Z, M, U_Y), \end{aligned}$$

where $(U_W, U_A, U_Z, U_M, U_Y)$ are exogenous random errors. We assume A is a single binary treatment, Z is a single binary intermediate confounder, M is a single binary mediator. There are no restrictions on the distribution of W or Y.

$g_{M|a',W}$ represents the counterfactual, conditional distribution of M, as described by VanderWeele (2016):

$$g_{M|A,W}(m, a^\star, W) \equiv g_{M|a^\star,W}(W) = \sum_{z=0}^{1} \mathbb{P}(M = 1 \mid Z = z, W)\mathbb{P}(Z = z \mid A = a^\star, W).$$

In what follows, we are going to assume that $g_{M|A,W}(m, a^\star, W)$ is known, estimated from observed data, which we call $\hat{g}_{M|a^\star,W}$. This is going to slightly alter the usual identification assumptions such that we no longer need to assume exchangeability of A and the counterfactual M values. This means the remaining assumptions are the same as those for controlled direct effects.

6.3.1 Estimation using G-Computation

The estimand $E(Y_{a',\hat{g}_{M|a^\star,W}})$ can be identified via sequential regression, which provides the framework for the G-computation-based estimator. The procedure is as follows

1. Fit a regression of Y on M, Z, W. Predict outcome values under under $M = m$. We'll call the result $\bar{Q}_Y(M, Z, W)$.
2. Integrate out M under our stochastic intervention $\hat{g}_{M|a^\star,W}$. We can do this by evaluating $\mathbb{E}(Y \mid M = m, Z = z, W)$ at each m and multiplying it by the probability that $M = m$ under $\hat{g}_{M|a^\star,W}$, summing over all m. We'll call the results $\bar{Q}^g_M(Z, W)$.
3. Integrate out Z and set $A = a'$. Again, we can do this by evaluating the predicted values from Step 2, setting $A = a'$, and at each z, multiplying the prediction by the probability that $Z = z$ under $A = a'$. We'll call the result $\bar{Q}^{a'}_Z(W)$.
4. Taking the sample mean (marginalizing over W) gives the parameter estimate.

6.3.2 Estimate with doubly robust methods based on the EIF

We are showing the one-step estimator (generalization of A-IPW)

The EIF for the parameter $\Psi(P)(a', \hat{g}_{M|a^\star,W})$, where, again, $\hat{g}_{M|a^\star,W}$ is assumed known, is given by:

$$D^\star(a', \hat{g}_{M|a^\star,W}) = \sum_{k=0}^{2} D^\star_k(a', \hat{g}_{M|a^\star,W}), \text{ where}$$

$$D^\star_0(a', \hat{g}_{M|a^\star,W}) = \bar{Q}^{a'}_{Z(W)} - \Psi(P)(a', \hat{g}_{M|a^\star,W})$$

$$D^\star_1(a', \hat{g}_{M|a^\star,W}) = \frac{I(A = a')}{\mathbb{P}(A = a' \mid W)}(\bar{Q}^{\hat{g}}_M(Z, W) - \bar{Q}^{a'}_{Z(W)})$$

$$D^\star_2(a', \hat{g}_{M|a^\star,W}) = \frac{I(A = a')\{I(M = 1)\hat{g}_{M|a^\star,W} + I(M = 0)(1 - \hat{g}_{M|a^\star,W})\}}{\mathbb{P}(A = a'} \times (Y - \bar{Q}_{Y(M,Z,W)}).$$

1. We estimate $g_{Z|a^\star,W}(W) = \mathbb{P}(Z = 1 \mid A = a^\star, W)$ from a logistic regression of Z on A, W setting $A = a^\star$.
2. We then estimate $g_{M|z,W}(W) = \mathbb{P}(M = 1 \mid Z = z, W)$ from a logistic regression of M on Z, W, setting $z = \{0, 1\}$.
3. We use these quantities to calculate $\hat{g}_{M|a^\star,W} = \hat{g}_{M|z=1,W}\hat{g}_{Z|a^\star,W} + \hat{g}_{M|z=0,W}(1 - \hat{g}_{Z|a^\star,W})$.

```
zmodel <- "z ~ a + w1 "
mmodel <- "m ~ z + w1"
ymodel <- "y ~ m + z*w1"

# make gm and get counterfactual predictions
zfit <- glm(formula = zmodel, family = "binomial", data = obsdat)
mfit <- glm(formula = mmodel, family = "binomial", data = obsdat)

za0 <- predict(zfit,
  newdata = data.frame(w1 = obsdat$w1, a = 0),
  type = "response"
)
za1 <- predict(zfit,
  newdata = data.frame(w1 = obsdat$w1, a = 1),
  type = "response"
)

mz1 <- predict(mfit,
  newdata = data.frame(w1 = obsdat$w1, z = 1),
  type = "response"
)
mz0 < -predict(mfit,
  newdata = data.frame(w1 = obsdat$w1, z = 0),
  type = "response"
)

gm0 <- (mz1 * za0) + (mz0 * (1 - za0))
gma1 <- (mz1 * za1) + (mz0 * (1 - za1))
```

4. To obtain an estimate of $\bar{Q}_Y(M, Z, W)$, predict values of Y from a regression of Y on M, Z, W, setting $m = 1$ and $m = 0$, giving $\hat{Y}(m = 1, z, w)$ and $\hat{Y}(m = 0, z, w)$.

```
tmpdat$qyinit <- cbind(
  predict(glm(
    formula = ymodel, family = "binomial",
    data = data.frame(cbind(datw, z = z, m = m, y = y))
  ),
```

```
  newdata = data.frame(cbind(datw, z = z, m = m)), type = "response"
  ),
  predict(glm(
    formula = ymodel, family = "binomial",
    data = data.frame(cbind(datw, z = z, m = m, y = y))
  ),
  newdata = data.frame(cbind(datw, z = z, m = 0)), type = "response"
  ),
  predict(glm(
    formula = ymodel, family = "binomial",
    data = data.frame(cbind(datw, z = z, m = m, y = y))
  ),
  newdata = data.frame(cbind(datw, z = z, m = 1)), type = "response"
  )
)
```

5. Estimate the weights :

$$h_1(a) = \frac{I(A=a)\{I(M=1)\hat{g}_{M|a^\star,W} + I(M=0)(1-\hat{g}_{M|a^\star,W})\}}{\mathbb{P}(A=a)\{I(M=1)g_{M|Z,W} + I(M=0)(1-g_{M|Z,W})\}}$$

```
psa1 <- I(a == 1) / mean(a)
psa0 <- I(a == 0) / mean(1 - a)
mz <- predict(glm(
  formula = mmodel, family = "binomial",
  data = data.frame(cbind(datw, z = z, m = m))
),
newdata = data.frame(cbind(datw, z = z)), type = "response"
)
psm <- (mz * m) + ((1 - mz) * (1 - m))

tmpdat$ha1gma1 <- ((m * gma1 + (1 - m) * (1 - gma1)) / psm) * psa1
tmpdat$ha1gma0 <- ((m * gm + (1 - m) * (1 - gm)) / psm) * psa1
tmpdat$ha0gma0 <- ((m * gm + (1 - m) * (1 - gm)) / psm) * psa0
```

6. We next integrate out M from $\bar{Q}_Y(M,Z,W)$. First, we estimate $\bar{Q}_{Y,n}(M,Z,W)$ setting $m=1$ and $m=0$, giving $\bar{Q}_Y(m=1,z,w)$ and $\bar{Q}_Y(m=0,z,w)$. Then, multiply these predicted values by their probabilities under $\hat{g}_{M|a^\star,W}(W)$ (for $a \in$

$\{a, a^\star\}$), and add them together (i.e., $\bar{Q}^{\hat{g}}_{M,n}(Z,W) = \hat{Q}^\star_Y(m=1,z,w)\hat{g}_{M|a^\star,W} + \hat{Q}^\star_Y(m=0,z,w)(1-\hat{g}_{M|a^\star,W})$).

```
tmpdat$Qma1g0 <- tmpdat$qyinit[, 2] * (1 - gm) + tmpdat$qyinit[, 3] *
tmpdat$Qma1g1 <- tmpdat$qyinit[, 2] * (1 - gma1) + tmpdat$qyinit[, 3]
tmpdat$Qma0g0 <- tmpdat$qyinit[, 2]* (1 - gm) + tmpdat$qyinit[, 3] * g
```

7. We now fit a regression of $\bar{Q}^{\hat{g}}_{M,n}(Z,W)$ on W among those with $A = a'$. We call the predicted values from this regression $\hat{\bar{Q}}^{a'}_Z(W)$.

```
Qzfita1g0 <- glm(
  formula = paste("Qma1g0", qmodel, sep = "~"),
  data = tmpdat[tmpdat$a == 1, ], family = "quasibinomial"
)
Qzfita1g1 <- glm(
  formula = paste("Qma1g1", qmodel, sep = "~"),
  data = tmpdat[tmpdat$a == 1, ], family = "quasibinomial"
)
Qzfita0g0 <- glm(
  formula = paste("Qma0g0", qmodel, sep = "~"),
  data = tmpdat[tmpdat$a == 0, ], family = "quasibinomial"
)
```

```
Qza1g0 <- predict(Qzfita1g0, type = "response", newdata = tmpdat)
Qza1g1 <- predict(Qzfita1g1, type = "response", newdata = tmpdat)
Qza0g0 <- predict(Qzfita0g0, type = "response", newdata = tmpdat)
```

(Note that if A were not randomly assigned, we would need to complete a second targeting step.)

8. Take the empirical mean of these predicted values $\Psi(P)(a', \hat{g}_{M|a^\star,W})$.

```
meanQza1m0 <- mean(Qza1g0)
meanQza1m1 <- mean(Qza1g1)
meanQza0m0 <- mean(Qza0g0)
```

9. Repeat the above steps for each of the interventions. For example, for binary A, we would execute these steps a total of three times to estimate:

    1. $\Psi(P)(1, \hat{g}_{M|1,W})$,
    2. $\Psi(P)(1, \hat{g}_{M|0,W})$, and
    3. $\Psi(P)(0, \hat{g}_{M|0,W})$.

10. Plug in all of the nuisance parameter estimates to the EIC, set it equal to zero and solve. This is the one-step estimate of the PIDE and PIIE.

11. The variance can be estimated as the sample variance of the EIF (defined above) divided by n.

```
# first get EIF
eic1a1g0 <- tmpdat$ha1gma0 * (tmpdat$y - tmpdat$qyinit[,1])
eic2a1g0 <- psa1 * (tmpdat$Qma1g0 - Qza1g0)
eic3a1g0 <- Qza1g0 - meanQza1m0
eica1g0 <- eic1a1g0 + eic2a1g0 + eic3a1g0

eic1a1g1 <- tmpdat$ha1gma1 * (tmpdat$y - tmpdat$qyinit[,1])
eic2a1g1 <- psa1 * (tmpdat$Qma1g1 - Qza1g1)
eic3a1g1 <- Qza1g1 - meanQza1m1
eica1g1 <- eic1a1g1 + eic2a1g1 + eic3a1g1

eic1a0g0 <- tmpdat$ha0gma0 * (tmpdat$y - tmpdat$qyinit[,1)
eic2a0g0 <- psa0 * (tmpdat$Qma0g0 - Qza0g0)
eic3a0g0 <- Qzupa0g0 - meanQza0m0
eica0g0 <- eic1a0g0 + eic2a0g0 + eic3a0g0

pide <- mean(eica1g0 - eica0g0)
piie <- mean(eica1g1 - eica1g0)

# estimands
pideeic <- eica1g0 - eica0g0
vareic <- var(pideeic) / nrow(tmpdat)

piieeic <- eica1g1 - eica1g0
varpiieeic <- var(piieeic) / nrow(tmpdat)
```

6.4 The general case

Actually, we would want to have the fixed parameter with the true, unknown $g_{M|a,W}$ and would like M to be continuous/multi-dimensional.

This is a pain to do by hand, but Nima made an easy-to-use package for all of us called medoutcon[1]! He will go through this next.

[1] https://github.com/nhejazi/medoutcon

Chapter 7

Estimating stochastic effects

[TO FILL IN]

Bibliography

Kennedy, E. H. (2018). Nonparametric causal effects based on incremental propensity score interventions. *Journal of the American Statistical Association*, (just-accepted).

Rudolph, K., Diaz, I., Hejazi, N., van der Laan, M., Luo, S., Shulman, M., Campbell, A., Rotrosen, J., and Nunes, E. (2020). Explaining differential effects of medication for opioid use disorder using a novel approach incorporating mediating variables. *Addiction*.

VanderWeele, T. J. (2016). Mediation analysis: a practitioner's guide. *Annual review of public health*, 37:17–32.