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

Iván Díaz, Nima Hejazi, Kara Rudolph

updated: April 26, 2021

Contents

List of Tables

List of Figures

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]

[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/

Causal Mediation Analysis

0.6 Mediation models

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

0.6.1 No intermediate confounders

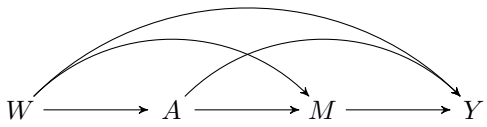


Figure 1: Directed acylcic graph under *no intermediate confounders* of the mediator-outcome relation affected by treatment

0.6.2 Intermediate confounders

0.7 Counterfactuals

In what follows, we will 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. For example, 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 $\mathbb{P}(A = a, M =$

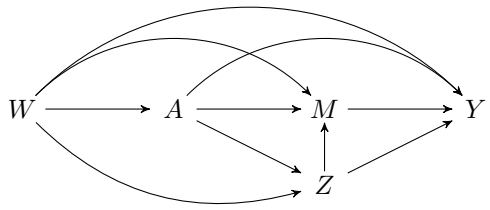


Figure 2: Directed acylcic graph under intermediate confounders of the mediator-outcome relation affected by treatment

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

In this workshop we use counterfactual variables as *primitives*, but we note that in other causal inference frameworks, such as structural equation models, counterfactuals are quantities *derived* from the model.

Path-specific casual mediation effect types

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

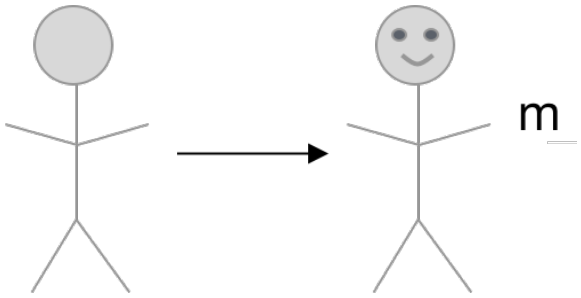
0.8 Controlled direct effects

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

- Set $M = m$ uniformly for everyone in the population
- Compare $A = 1$ vs $A = 0$ with $M = m$ fixed
- Confounder assumptions:
  - $A \perp\!\!\!\perp Y_{a,m} \mid W$
  - $M \perp\!\!\!\perp Y_{a,m} \mid W, A, Z$
- Positivity assumptions:
  - $\mathbb{P}(M = m \mid Z, A = a, W) > 0 \; a.e.$
  - $\mathbb{P}(A = a \mid W) > 0 \; a.e.$

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?

Natural direct and indirect effects

Natural direct effect (NDE): $$\psi_{\text{NDE}} = \mathbb{E}(Y_{1,M_0}) - \mathbb{E}(Y_{0,M_0})$$

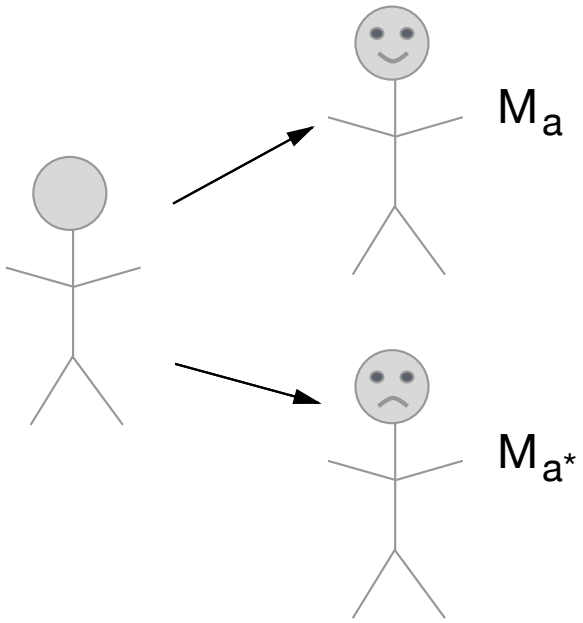
Natural indirect effect (NIE): $$\psi_{\text{NIE}} = \mathbb{E}(Y_{1,M_1}) - \mathbb{E}(Y_{1,M_0})$$

The NDE can also be written as: $\mathbb{E}_W \sum_m \{\mathbb{E}(Y_{1,m} \mid W) - \mathbb{E}(Y_{0,m} \mid W)\}\mathbb{P}(M_0 = m \mid W)$

- Weighted average of controlled direct effects at each level of m.
- If no interaction between A and M on Y, then CDE = NDE.

0.8.1 Identification assumptions:

- $A \perp\!\!\!\perp Y_{a,m} \mid W$



- $M \perp\!\!\!\perp Y_{a,m} \mid W, A$
- $A \perp\!\!\!\perp M_a \mid W$
- $M_{a^\star} \perp\!\!\!\perp Y_{a,m} \mid W$
- and positivity assumptions

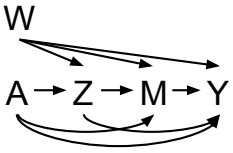
What does $M_0 \perp\!\!\!\perp Y_{1,m} \mid W$ mean?

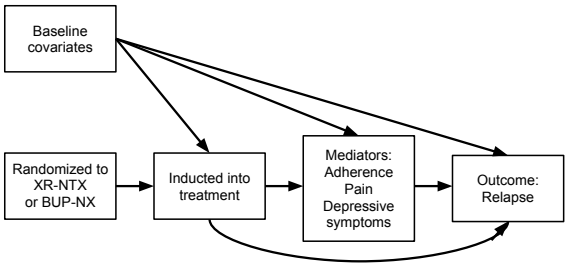
- Conditional on W, knowledge of M in the absence of treatment A provides no information of the effect of A on Y.
- Can you think of a data-generating mechanism that would violate this assumption?
- Whenever we believe that treatment assignment works through adherence (i.e., almost always), we are violating this assumption.

0.8.2 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)?



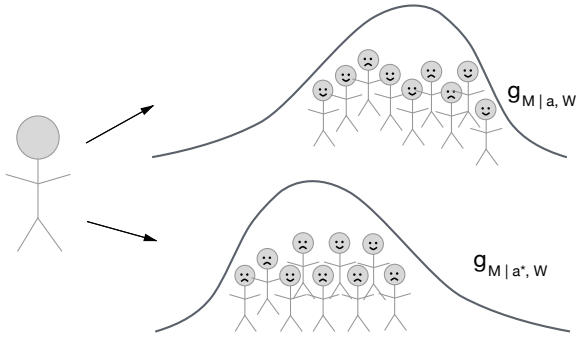


Interventional direct/ indirect effects

- Fully conditional on past
- Conditional SDE: $\mathbb{E}(Y_{a,g_{M|Z,a^\star,W}}) - \mathbb{E}(Y_{a^\star,g_{M|Z,a^\star,W}})$
- Conditional SIE: $\mathbb{E}(Y_{a,g_{M|Z,a,W}}) - \mathbb{E}(Y_{a,g_{M|Z,a^\star,W}})$
- Marginal SDE: $\mathbb{E}(Y_{a,g_{M|a^\star,W}}) - \mathbb{E}(Y_{a^\star,g_{M|a^\star,W}})$
- Marginal SIE: $\mathbb{E}(Y_{a,g_{M|a,W}}) - \mathbb{E}(Y_{a,g_{M|a^\star,W}})$
- Note that $g_{M|Z,a^\star,W}$, $g_{M|a^\star,W}$ represents stochastic intervention on the mediator, where value m is drawn with probability $\mathbb{P}(M = m \mid Z, A = a^\star, W = w)$, $\mathbb{P}(M = m \mid A = a^\star, W = w)$, respectively
- Can you think of an example when you would want the conditional versions? Marginal versions?

0.8.3 Identification assumptions:

- $A \perp\!\!\!\perp Y_{a,m} \mid W$
- $M \perp\!\!\!\perp Y_{a,m} \mid W, A$



- $A \perp\!\!\!\perp M_a \mid W$
- 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!

Estimand Summary

Table 1. Mediation Estimand Definitions, Descriptions, and Assumptions

Estimand	Description	Identifying Assumptions in Addition to Positivity and 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 and Y ($A \perp Y_{a,m} \mid W$). 2. No unmeasured confounding between M and Y ($M \perp Y_{a,m} \mid W, A$).
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 and Y ($A \perp Y_{a,m} \mid W$). 2. No unmeasured confounding between M and Y ($M \perp Y_{a,m} \mid W, A$). 3. No unmeasured confounding of $A - M$ ($A \perp M_a \mid W$). 4. No measured or unmeasured posttreatment 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 and Y ($A \perp Y_{a,m} \mid W$). 2. No unmeasured confounding between M and Y ($M \perp Y_{a,m} \mid W, A$). 3. No unmeasured confounding of $A - M$ ($A \perp M_a \mid W$).

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

The Interventional Direct and Indirect Effects

0.9 Definition of the effects

Consider the following directed acyclic graph.

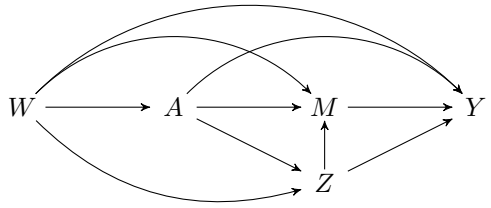


Figure 3: Directed acylcic graph under intermediate confounders of the mediator-outcome relation affected by treatment

Here, A is the treatment of interest, M is the mediator of interest, W is a pre-treatment variable containing confounders of A on M and Y,and Z is a post-treatment variable contaning confounders of the mediator and the exposure which are affected by treatment.

0.9.1 Example

[TO FILL IN]

0.9.2 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 aboce directed acyiclic graph. Intuitively, 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. Thus, to achieve $Y(1, m)$ independent of $M(0)$, it would be necessary to adjust for either $Z(1)$ or $Z(0)$. This is impossible to do since these variables are unmeasured.

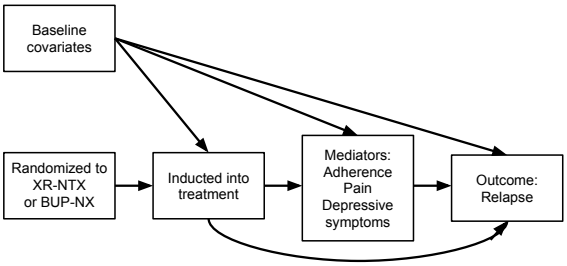
0.9.3 Recovering direct and indirect effects

Even though estimation of the NDE and NIE is not possible in the presence of confounders of the mediation-outcome relation affected by treatment, it is possible to redefine the effects in a way such that they are identifiable. Specifically:

- 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 porbability one.
- Define $Y(0, G(0))$ and $Y(1, G(1))$ similarly
- Then we can define:
$$\mathbb{E}[Y(1, G(1)) - Y(0, G(0))] = \underbrace{\mathbb{E}[Y(1, G(1)) - Y(1, G(0))]}_{\text{interventional indirect effect}} + \underbrace{\mathbb{E}[Y(1, G(0)) - Y(0, G(0))]}_{\text{interventional direct effect}}$$

How to choose an estimand: Real world example

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



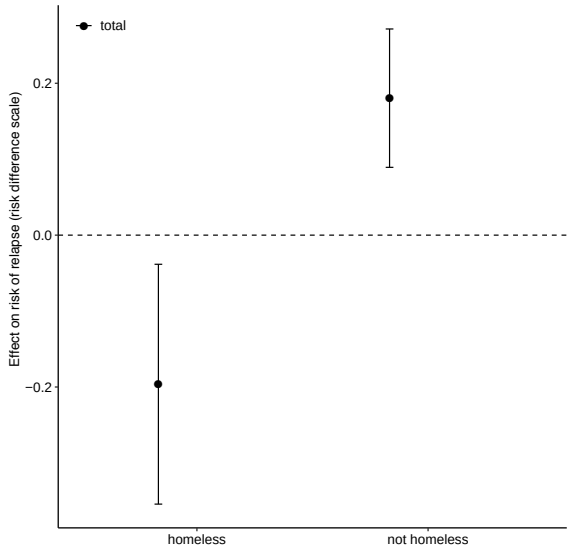
Motivation: Opposite overall treatment effects for homeless versus nonhomeless participants.

0.10.1 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?



→ So, *not* controlled direct effect.

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

→ So, *not* natural (in)direct effects.

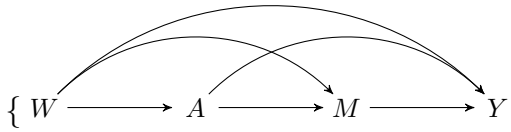
- So, we're left with the interventional direct and indirect effects.
- Do we want to estimate the path through treatment initiation (Z)?
- Yes, so, *not* the conditional versions of these effects.
- Estimands:
  - Direct effect: $\mathbb{E}(Y_{1,g_0} - Y_{0,g_0})$
  - Indirect effect: $\mathbb{E}(Y_{1,g_1} - Y_{1,g_0})$
- Need to incorporate multiple and continuous mediators

Stochastic Direct and Indirect Effects

0.11 Definition of the effects

Consider the following directed acyclic graph.

\begin{figure}



}

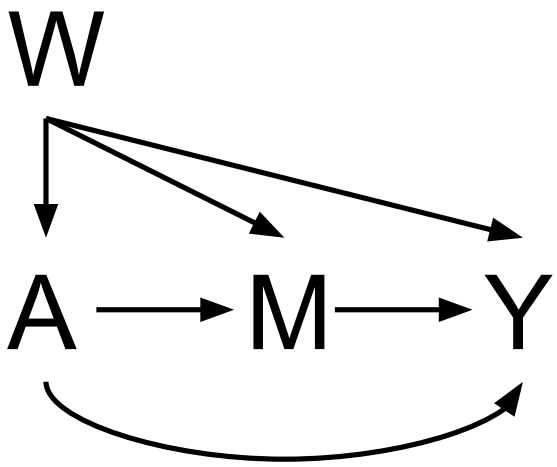
\caption{Directed acylcic graph under *no intermediate confounders* of the mediator-outcome relation affected by treatment} \end{figure} ## Motivation for stochastic interventions

- So far we have discussed effects for binary exposures
- Controlled direct effects and natural effects require that $0 < \mathbb{P}(A = 1 \mid W) < 1$

Estimation of natural (in)direct effects, interventional (in)direct effects

0.12 Natural direct and indirect effects

Recall:



Assuming a binary A, we define the natural direct effect as:
$NDE = E(Y_{1,M_0}) - E(Y_{0,M_0})$

,

and the natural indirect effect as:
$NIE = E(Y_{1,M_1}) - E(Y_{1,M_0})$

.

0.12.1 Simple case for intuition

$$O = (W, A, M, Y)$$

This SCM is represented in the above DAG and the following causal models: $W = f(U_W)$, $A = f(W, U_A)$, $M = f(A, W, U_M)$, and $Y = f(M, A, W, U_Y)$, where (U_W, U_A, U_M, U_Y) are exogenous random errors.

We assume A is a single binary randomized treatment, 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: - $A \perp\!\!\!\perp Y_{a,m} \mid W$ - $M \perp\!\!\!\perp Y_{a,m} \mid W, A$ - $A \perp\!\!\!\perp M_a \mid W$ - $M_{a^\star} \perp\!\!\!\perp Y_{a,m} \mid W$ - and positivity assumptions

0.12.2 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.

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

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

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

$$+ \frac{I(A=0)}{g(0|W)} \{\bar{Q}_{diff} - E(\bar{Q}_{diff}|W, 0)\} \tag{2}$$

$$+ E(\bar{Q}_{diff}|W, 0) - \Psi_{NDE} \tag{3}$$

0.12.4 How to estimate using TMLE

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

Underneath the hood, the counterfactual outcome difference $\bar{Q}_{\text{diff}}$ and $P(A \mid Z, W)$, the conditional probability of A given Z and W, are used in constructing the auxiliary covariate for TML estimation. These nuisance parameters play an important role in the bias-correcting *TMLE-update step*.

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 = "response")
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]
```

4. Estimate $\hat{\epsilon}$ by setting ϵ as the intercept of a weighted logistic regression model of Y with $logit(\hat{\bar{Q}}_Y(M, A, W))$ as an offset and weights $\hat{C}_Y$.

5. The estimates of $\bar{Q}_Y(M, 1, W)$ and $\bar{Q}_Y(M, 0, W)$ are updated by $\hat{\bar{Q}}^{\star}_Y(M, A, W) = \hat{\bar{Q}}_Y(\epsilon_n)(M, A, W)$. This gives an updated difference: $\hat{\bar{Q}}^{\star}_{diff}(M, A, W)$.

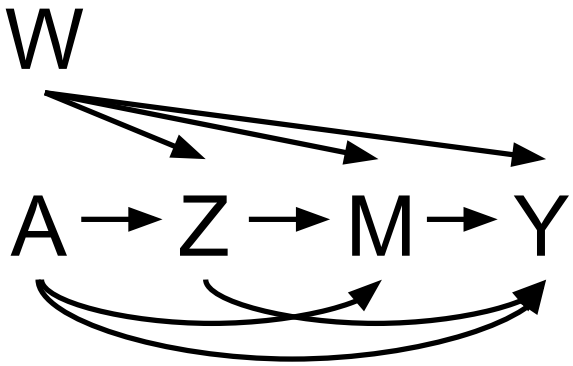
```
epsilon <- coef(glm(y ~ 1,
  weights = cy, offset = (qlogis(qyinit[, 1])),
  family = "quasibinomial"
))
qyupa0 <- plogis(qlogis(qyinit[, 2]) + epsilon)
qyupa1 <- plogis(qlogis(qyinit[, 3]) + epsilon)
qdiffup <- qyupa1 - qyupa0
```

6. We then regress $\hat{\bar{Q}}^{\star}_{diff}(M, A, W)$ on W among those with $A = 0$. Taking the empirical mean of the predicted values gives us the TML estimate of the NDE.

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

0.13 Interventional direct and indirect effects

Recall:



We define the interventional direct effect as:
$\psi_{\text{PIDE}} = \mathbb{E}(Y_{a',g_{M|a^\star,W}}) - \mathbb{E}(Y_{a\star,g_{M|a^\star,W}})$,

and the interventional indirect effect as:
$\psi_{\text{PIIE}} = \mathbb{E}(Y_{a',g_{M|a',W}}) - \mathbb{E}(Y_{a',g_{M|a^\star,W}})$.

0.14 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:

$$\begin{aligned} W &= f(U_W) \\ A &= f(U_A) \\ Z &= f(A, W, U_Z) \\ M &= f(Z, A, W, U_M) \\ Y &= f(M, Z, A, W, U_Y), \end{aligned}$$

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 a stochastic draw from the counterfactual, conditional distribution of M, as described by **?**:

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

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.

0.14.1 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.

0.14.2 Estimate with doubly robust methods based on the EIF

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:

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

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

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

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

0.14.3 Estimate using TMLE

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 to be used for the initial targeting step:

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

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

6. Estimate $\hat{\epsilon}$ by setting ϵ as the intercept of a weighted logistic regression model of Y with logit$(\hat{\bar{Q}}_Y(M, Z, W))$ as an offset and weights $\hat{h}_1(a)$. (Note that this is just one possible TMLE.)

7. The estimate of $\bar{Q}_Y(M, Z, W)$ is updated by $\hat{\bar{Q}}^{\star}_Y(M, Z, W) = \hat{\bar{Q}}_Y(\epsilon_n)(M, Z, W)$.

```
# for E(Y_{1,gmastar})
epsilonma1g0 <- coef(glm(y ~ 1,
  weights = tmpdat$ha1gma0,
  offset = (qlogis(qyinit[, 1])),
  family = "quasibinomial", data = tmpdat
))
tmpdat$qyupm0a1g0 <- plogis(qlogis(tmpdat$qyinit[, 2]) + epsilonma1g0)
tmpdat$qyupm1a1g0 <- plogis(qlogis(tmpdat$qyinit[, 3]) + epsilonma1g0)

# for E(Y_{1,gma})
epsilonma1g1 <- coef(glm(y ~ 1,
  weights = tmpdat$ha1gma1,
  offset = (qlogis(qyinit[, 1])),
  family = "quasibinomial", data = tmpdat
))
tmpdat$qyupm0a1g1 <- plogis(qlogis(tmpdat$qyinit[, 2]) + epsilonma1g1)
tmpdat$qyupm1a1g1 <- plogis(qlogis(tmpdat$qyinit[, 3]) + epsilonma1g1)
```

```
# for E(Y_{0,gmastar})
epsilonma0g0 <- coef(glm(y ~ 1,
  weights = tmpdat$ha0gma0,
  offset = (qlogis(qyinit[, 1])),
  family = "quasibinomial", data = tmpdat
))
tmpdat$qyupm0a0g0 <- plogis(qlogis(tmpdat$qyinit[, 2]) + epsilonma0g0)
tmpdat$qyupm1a0g0 <- plogis(qlogis(tmpdat$qyinit[, 3]) + epsilonma0g0)
```

8. We next integrate out M from $\bar{Q}^{\star}_{Y}(M, Z, W)$. First, we estimate $\bar{Q}^{\star}_{Y,n}(M, Z, W)$ setting $m = 1$ and $m = 0$, giving $\bar{Q}^{\star}_{Y}(m = 1, z, w)$ and $\bar{Q}^{\star}_{Y}(m = 0, z, w)$. 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$Qma1g0 <- tmpdat$qyupm0a1g0 * (1 - gm) + tmpdat$qyupm1a1g0 * gm
tmpdat$Qma1g1 <- tmpdat$qyupm0a1g1 * (1 - gma1) + tmpdat$qyupm1a1g1 * gma1
tmpdat$Qma0g0 <- tmpdat$qyupm0a0g0 * (1 - gm) + tmpdat$qyupm1a0g0 * gm
```

9. We now fit a regression of $\bar{Q}^{\hat{g},\star}_{M,n}(Z, W)$ 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.)

10. The empirical mean of these predicted values is the TML estimate of $\Psi(P)(a', \hat{g}_{M|a^\star,W})$.

```
tmlea1m0 <- sum(Qzupa1g0 * svywt) / sum(svywt)
tmlea1m1 <- sum(Qzupa1g1 * svywt) / sum(svywt)
tmlea0m0 <- sum(Qzupa0g0 * svywt) / sum(svywt)
```

11. 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. $\Psi(P)(0, \hat{g}_{M|0,W})$.
12. The PIDE can then be obtained by substituting estimates of parameters $\Psi(P)(a, \hat{g}_{M|a^\star,W}) - \Psi(P)(a^\star, \hat{g}_{M|a^\star,W})$ and the PIIE can be obtained by substituting estimates of parameters $\Psi(P)(a, \hat{g}_{M|a,W}) - \Psi(P)(a, \hat{g}_{M|a^\star,W})$.

```
nde <- tmlea1m0 - tmlea0m0
nie <- tmlea1m1 - tmlea1m0
```

13. The variance can be estimated as the sample variance of the EIF (defined above, substituting in the targeted fits) divided by n.

```
# first get EIF
tmpdat$qyupa1g0 <- plogis(qlogis(tmpdat$qyinit[, 1]) + epsilonma1g0)
tmpdat$qyupa1g1 <- plogis(qlogis(tmpdat$qyinit[, 1]) + epsilonma1g1)
tmpdat$qyupa0g0 <- plogis(qlogis(tmpdat$qyinit[, 1]) + epsilonma0g0)

eic1a1g0 <- tmpdat$ha1gma0 * (tmpdat$y - tmpdat$qyupa1g0)
eic2a1g0 <- psa1 * svywt * (tmpdat$Qma1g0 - Qzupa1g0)
eic3a1g0 <- Qzupa1g0 - tmlea1m0
eica1g0 <- eic1a1g0 + eic2a1g0 + eic3a1g0

eic1a1g1 <- tmpdat$ha1gma1 * (tmpdat$y - tmpdat$qyupa1g1)
eic2a1g1 <- psa1 * svywt * (tmpdat$Qma1g1 - Qzupa1g1)
eic3a1g1 <- Qzupa1g1 - tmlea1m1
eica1g1 <- eic1a1g1 + eic2a1g1 + eic3a1g1

eic1a0g0 <- tmpdat$ha0gma0 * (tmpdat$y - tmpdat$qyupa0g0)
eic2a0g0 <- psa0 * svywt * (tmpdat$Qma0g0 - Qzupa0g0)
eic3a0g0 <- Qzupa0g0 - tmlea0m0
eica0g0 <- eic1a0g0 + eic2a0g0 + eic3a0g0

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

nieeic <- eica1g1 - eica1g0
varnieeic <- var(nieeic) / nrow(tmpdat)
```

0.15 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[19]! He will go through this next.

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

Estimation of stochastic (in)direct effects

[TO FILL IN]

Bibliography