Lab 7: Exam 1 Review

PHS2000B, Spring 2023

Sudipta Saha and Susan Rattigan

March 6, 2023

Topics Covered on Exam 1[1]

1. **Counterfactual thinking and DAG review**
2. **Propensity Scores**
3. **Sensitivity Analyses for Unmeasured Confounding**
4. **Measurement Error**
5. **Marginal Structural Models (MSMs) and Time-Varying Confounding**
6. **Interaction**
7. **Mediation**

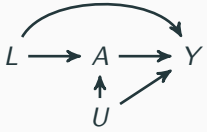
[1] Huge shoutout to Emma McGee who created a version of this review deck last year!

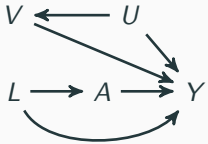
1. Counterfactual Thinking and DAGs

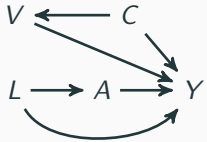
Review of Key Assumptions

- **Consistency**: Imagine that person Z is exposed to treatment $A = a$ in our observed data ("real life"). Consistency holds if, in a counterfactual world where we forced everyone's exposure level to $A = a$, person Z experiences the same outcome has they had under treatment $A = a$ in our observed data.
- **Exchangeability**: We say that treatment groups are exchangeable if the counterfactual outcome we would observe under any treatment level $A = a$ is independent of the actual treatment level the group received in our observed data.
  - We notate exchangeability as $A \perp\!\!\!\perp Y^a$.
- **Positivity**: There are individuals in all strata that we can define by level of exposure and stratification factors. That is, $Pr[A = a] > 0$ or $Pr[A = a|L = l]$ must hold for for all levels of A and, if relevant, all levels of L.

Reading Independencies / Exchangeabilities from DAGs

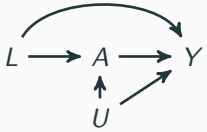




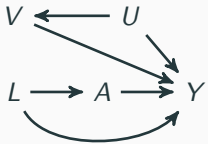


Reading Independencies / Exchangeabilities from DAGs

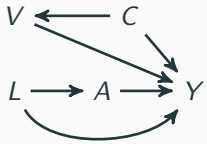
$Y^a \not\perp\!\!\!\perp A$; $Y^a \perp\!\!\!\perp A|L$



$Y^a \not\perp\!\!\!\perp A|L$



$Y^{a,v} \not\perp\!\!\!\perp A$ and $Y^{a,v} \not\perp\!\!\!\perp V$;
$Y^{a,v} \perp\!\!\!\perp A|L$ and $Y^{a,v} \not\perp\!\!\!\perp V|L$



$Y^{a,v} \not\perp\!\!\!\perp A$ and $Y^{a,v} \not\perp\!\!\!\perp V$;
$Y^{a,v} \perp\!\!\!\perp A|L$ and $Y^{a,v} \perp\!\!\!\perp V|C$;

Review of Effect Types

Marginal vs. Conditional Causal Effects: A **marginal causal effect** provides the effect in the *entire population* (e.g., $E[Y_{a=1} - Y_{a=0}]$). A **conditional causal effect** provides a causal effect that is conditional on some variable(s) (e.g., $E[Y_{a=1} - Y_{a=0}|V = v]$).

Types of Causal Effects: We discussed three common types of causal effects (i.e., casual estimands) that may be of interest. On the additive scale:

- **Average Treatment Effect (ATE)** $= \boldsymbol{E[Y_{a=1} - Y_{a=0}]}$. Quantifies the effect *had everyone been treated* vs. *had no one been treated.*
- **Average Treatment Effect in the Treated (ATT)** $= \boldsymbol{E[Y_{a=1} - Y_{a=0}|A = 1]}$. Quantifies the effect of intervening to *remove treatment from those who were treated.*
- **Average Treatment Effect in the Untreated (ATU)** $= \boldsymbol{E[Y_{a=1} - Y_{a=0}|A = 0]}$. Quantifies the effect of intervening to *give treatment to those who were untreated.*

1. Propensity Scores

Propensity Scores: Motivations for Use

We often assume that treatment groups are **conditionally exchangeable** within levels of a set of confounders C:

$$Y_a \perp\!\!\!\perp A|C$$

In theory, we could stratify by all these confounders and compute an estimate of the causal effect. However, it may be difficult (or undesirable) to stratify by all the confounders, especially when C is a vector of many variables $\implies$ **propensity score methods!**

- Propensity score (S) methods allow us to reduce the problem of **multivariate control** to **univariate control.** This is because of the following theoretical property:

$$Y_a \perp\!\!\!\perp A|C \implies Y_a \perp\!\!\!\perp A|S$$

- The propensity score also has **balancing properties:**

$$Pr(C = c|A = 1, S) = Pr(C = c|A = 0, S) \text{ i.e., } C \perp\!\!\!\perp A|S$$

Propensity Scores: Motivations for Use

- Propensity score (S) methods allow us to reduce the problem of **multivariate control** to **univariate control.** This is because of the following theoretical property:

$$Y_a \perp\!\!\!\perp A|C \implies Y_a \perp\!\!\!\perp A|S$$

Translation: Exchangeability conditional on a set of measured counfounders implies exchangeability conditional on the propensity score calculated using those confounders.

- The propensity score also has **balancing properties:**

$$Pr(C = c|A = 1, S) = Pr(C = c|A = 0, S) \text{ i.e., } C \perp\!\!\!\perp A|S$$

Translation: For a given value of the propensity score, treated and untreated individuals will have the same distribution of measured confounders C.

Propensity Scores: Definition and Estimation

The **propensity score** S is an individual's probability of *receiving treatment* A, conditional on some measured covariates C:

$$S = Pr(A = 1|C = c)$$

For a binary treatment, the propensity score can be estimated using a **logistic model.**

- **Step 1:** Fit a logistic model for treatment.

$$Logit(Pr(A = 1|C = c)) = \beta_0 + c^T\beta$$

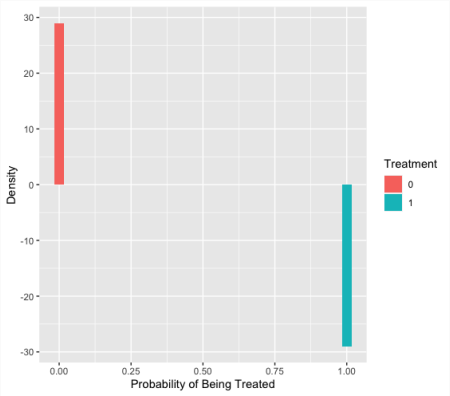
- **Step 2:** Estimate each individual's predicted probability of receiving treatment.

$$S = Pr(A = 1|C = c) = \frac{exp(\widehat{\beta}_0 + + c^T\widehat{\beta})}{1 + exp(\widehat{\beta}_0 + c^T\widehat{\beta})}$$

Propensity Scores: Methods

Method	Description	Effect Type	Pros and Cons
Stratification or Subclassification	Categorize individuals based on strata of S, then analyze data within those strata. Effects can also be combined across strata.	Conditional on S unless standardized. ATE if standardized by overall strata size. ATT if standardized by treated strata size. ATU if standardized by untreated strata size.	Pros: controls for all covariates while stratifying on only one variable. Cons: subject to residual bias (check covariate balance within strata of S!).
Matching	Match treated and untreated based on their values of S, then calculate effect measures in the matched data.	ATT if all treated are matched. ATU if all untreated are matched. Effect for region of common support if some treated/untreated unmatched.	Pros: analysis and presentation of results is simplified. Cons: loss of efficiency, interpretation conditional on matched subset, residual bias.
Weighting	Weight each individual by inverse of S (treated) or $1-S$ (untreated), then fit a regression model to the weighted data.	ATE if weighting treated by $\frac{1}{S}$ and untreated by $\frac{1}{1-S}$ (or stabilized equivalents). ATT or ATU if using different weights.	Pros: no residual bias, generalizable to time-varying treatments with feedback. Cons: can result in large SEs if some individuals have large weights.

Propensity Scores: Methods



Propensity Scores in Treated and Untreated (Density Histogram)

Based on this plot, which propensity score method(s) could we use?

a. Stratification
b. Matching
c. Weighting
d. None of the above

Propensity Scores: Strengths

Propensity score methods have some **strengths** as compared to traditional outcome regression, including:[2]

- Propensity score methods are helpful when it is **easier to model the exposure than the outcome** *(they do not have to impose covariate-outcome modeling assumptions).*
- With propensity score methods, modeling decisions for S can be made **prior to examining the outcome.** The region of common support is easily checked.
- If the **binary outcome is rare** and **exposure is common,** then propensity score approaches will generally perform better than a regular logistic regression.
- A subset of these methods (weighting) can be extended to **time-varying treatments** with **treatment-confounder feedback.**

[2] For a more complete discussion of the strengths and limitations of propensity score methods as compared to traditional outcome regression, see the "Regression vs. Propensity Scores" slides in the Propensity Scores lecture.

2. Sensitivity Analyses for Unmeasured Confounding

Sensivity Analyses: Overview

Sensitivity analyses can help assess the extent to which an unmeasured variable (or variables) U would have to affect both exposure A and outcome Y in order for the observed A-Y association to be **attributed solely to confounding.** Sensitivity analysis can also aid in assessing a plausible **range of values** for the true causal effect.

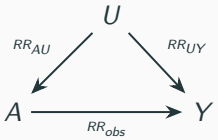
We covered **4 sensitivity analysis methods:**

1. **Cornfield Conditions**
2. **Bias Factors**
3. **E-values**
4. **Bias Factors with Prevalence Specification**

Sensivity Analyses: Cornfield Conditions

The **Cornfield Conditions** state that, in order for a single confounder to explain away an observed association, the association between i) the confounder and the exposure (RR_{AU}) and ii) the confounder and the outcome (RR_{UY}) must be at least as large as the observed association between the exposure and outcome (RR_{obs}):[3]

$$RR_{AU} \geq RR_{obs} \quad \text{and} \quad RR_{UY} \geq RR_{obs}$$



[3] Cornfield J et al. Smoking and Lung Cancer: Recent Evidence and a Discussion of Some Questions. JNCI. 1959;22:173- 203.

Sensitivity Analyses: Bias Factors

Bias Factor Equation 1 relates the observed and the true risk ratios. Let RR_{obs} represent the observed risk ratio, RR_{true} represent the true causal risk ratio, and BF represent the bias factor:}

$$RR_{obs} = BF \times RR_{true} \quad \Longleftrightarrow \quad BF = \frac{RR_{obs}}{RR_{true}}$$

Bias Factor Equation 2 provides a bound for the bias factor, which is represented by just two parameters, RR_{UY} and RR_{AU}:}

$$BF \leq \frac{RR_{UY} \times RR_{AU}}{RR_{UY} + RR_{AU} - 1}$$

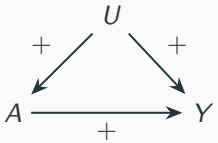
Note that it is in Bias Factor Equation 2 and Bias Factor Equation 2 only that we have the restrictive assumptions that RR_{AU}, RR_{UY}, and BF are all greater than 1.

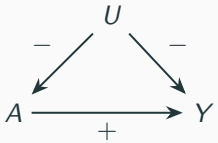
Sensitivity Analyses: Bias Factor with Protective Associations

Always check your rules of confounding! If you wish to explain away an association, you must have one of the following confounding structures.

$RR_{obs} > 1$

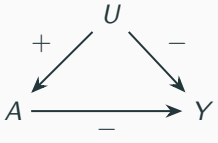
RR_{AU} and RR_{UY} that could explain away RR_{obs} must indicate the A-Y relationship is too high

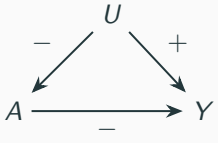




$RR_{obs} < 1$

RR_{AU} and RR_{UY} that could explain away RR_{obs} must indicate the A-Y relationship is too low





Sensitivity Analyses: E-values

The **E-value for Unmeasured Confounding** is a special case where $RR_{UY} = RR_{AU} = RR_{eq}$:[4]

$$\text{E-value} = RR_{eq} = RR_{obs} + \sqrt{RR_{obs}(RR_{obs} - 1)}$$

Interpretation: An unmeasured confounder U that is associated with both the [*outcome*] and the [*exposure*] by a risk ratio of [*insert E-value for point-estimate*], above and beyond the measured confounders, could explain away the estimated effect, but weaker confounding could not. To move the confidence interval to include the null, such a confounder would need to be associated with the outcome and exposure by risk ratios of at least [*insert E-value for CI*] each.

To estimate an **E-value for a protective association**, take the *inverse* of the RR first.

[4] VanderWeele TJ, Ding P. Sensitivity Analysis in Observational Research: Introducing the E-Value. Ann Intern Med. 2017;167(4):268-274.

Sensitivity Analyses: Bias Factor with Prevalence Specification

Bias Factor with Prevalence Specification: We can calculate the **exact amount of bias** if we are willing to make a few additional assumptions.[5]

- Assume $Y_a \perp\!\!\!\perp A \mid (C, U)$ and U is a single binary confounder with the same risk ratio for Y among both the exposed and unexposed:

$$B_{mult}(c) = \frac{1+(\gamma-1)P(U=1|a_1,c)}{1+(\gamma-1)P(U=1|a_0,c)}$$

- $\gamma = \frac{P(Y=1|U=1,a,c)}{P(Y=1|U=0,a,c)}$ = risk ratio between U and Y, homogeneous with respect to exposure
- $P(U = 1 \mid a_1, c)$ = prevalence of confounder U among the exposed, conditional on C
- $P(U = 1 \mid a_0, c)$ = prevalence of confounder U among the unexposed, conditional on C

[5] Schlesselman JJ. Assessing effects of confounding variables. Am J Epidemiol. 1978;108:3-8.

3. Measurement Error

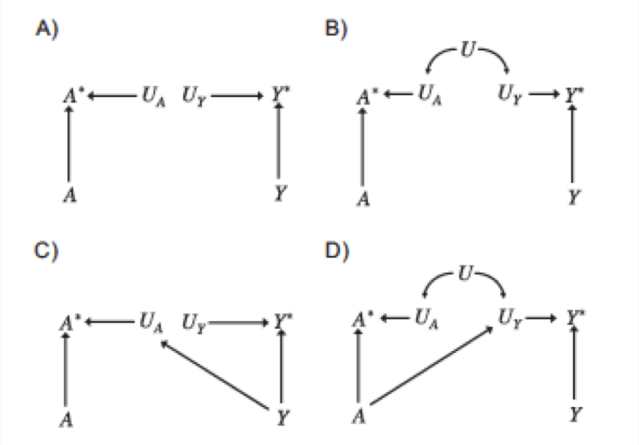
Measurement Error: Types

Let X and Y be correctly measured variables and X^* and Y^* be their mismeasured versions. We have the following **types of measurement error:**[6]

- **Non-differential**: given the true value X, the distribution of the mismeasured variable X^* is the same no matter the value of Y: $\Pr(X^*|X, Y) = \Pr(X^*|X)$.
- **Differential**: non-differential measurement error does not hold, i.e., Y provides additional information about the distribution of X^*, even if we know the value of X.
- **Independent**: the error in one variable doesn't tell you anything about the error in another variable: $\Pr(X^*|X, Y, Y^*) = \Pr(X^*|X, Y)$.
- **Dependent**: independent measurement error does not hold, i.e., knowing the value of mismeasured Y^* tells you something about X^*, conditional on both true values.

[6] Differential vs. non-differential applies whenever one or more variables are mismeasured. By contrast, independent vs. dependent only applies when there are at least two mismeasured variables.

Measurement Error: DAGs



Types of Measurement Error. A) Independent non-differential B) Dependent non-differential C) Independent differential D) Dependent differential[7]

[7] Hernán MA, Cole SR. 2009;170(8):959-962. VanderWeele TJ, Hernán MA. 2012;175(12):1303-10.

Measurement Error: Sensitivity, Specificity, PPV, NPV

Sensitivity (Se) and **specificity (Sp)**:

$$\textbf{Se: } 1 - \text{false negative probability} = \Pr(X^* = 1 | X = 1)$$
$$\textbf{Sp: } 1 - \text{false positive probability} = \Pr(X^* = 0 | X = 0)$$

Positive predictive value (PPV) and **negative predictive value (NPV)**:

$$\textbf{PPV: } \Pr(X = 1 | X^* = 1) = \frac{Se \cdot p_x}{Se \cdot p_x + (1 - Sp) \cdot (1 - p_x)}$$
$$\textbf{NPV: } \Pr(X = 0 | X^* = 0) = \frac{Sp \cdot (1 - p_x)}{(1 - Se) \cdot p_x + Sp \cdot (1 - p_x))}$$

Note that PPV and NPV depend on Se and Sp and on the prevalence of X, denoted p_x.

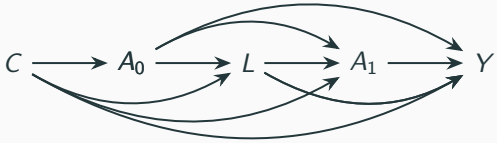
5. Marginal Structural Models (MSMs) and Time-Varying Confounding

MSMs: Why Traditional Regression Approaches Fail

Causal Inference Principle I: If C is a *common cause* of A and Y then we must control for it.

Causal Inference Principle II: If a variable is an *intermediate* between A and Y, we should not control for it.

- These two principles may conflict when exposure A is **time-varying.** For example, say we are interested in the joint effect of both A_0 and A_1 on Y in the following DAG:



Traditional regression approaches fail in this setting as a result of **treatment-confounder feedback (confounders for future exposure are affected by prior exposure**). One solution is to use a **Marginal Structural Model (MSM).**

MSMs: Fitting an MSM

Step 1: Create **inverse probability of treatment weights (IPTW)** for each individual.

- Fit a regression model for the treatment at time 0, conditional on its past (prior confounders and/or treatments). E.g.:

$$logit(Pr(A_0 = 1|C = c)) = \beta_0 + \beta_1 C$$

- Use this model to generate predicted probabilities of receiving the treatment level each individual *actually received* at time 0.
- Take the inverse of these probabilities to obtain the weight for treatment at time 0. E.g.:

$$w_0^i = \frac{1}{\Pr(A_0 = a_0^i | C = c^i)}$$

- Repeat for each time point. Assign each individual the product of their weights across all time points. E.g.:

$$W = \frac{1}{\Pr(A_0 = a_0^i | C = c^i)} \times \frac{1}{\Pr(A_1 = a_1^i | C = c^i, A_0 = a_0^i, L = l^i)}$$

MSMs: Fitting an MSM

Step 2: Fit a **regression model** to the **weighted study data.**

- Fit a regression model to the dataset where each individual is weighted by their IPTW. Use a sandwich/robust variance estimator or bootstrapping to obtain valid SEs.E.g.:

$$E[Y \mid A_0 = a_0, A_1 = a_1] = \beta_0 + \beta_1 A_0 + \beta_2 A_1$$

- Under the assumptions of **sequential exchangeability, consistency, and positivity**,[8] the parameters from this weighted regression are consistent estimates for the parameters of the corresponding MSM. E.g.:

$$E[Y_{a_0 a_1}] = \beta_0 + \beta_1 a_0 + \beta_2 a_1$$

Under consistency, positivity, and the following sequential exchangeability assumption:

$$Y_{a_0 a_1} \perp\!\!\!\perp A_0 | C \text{ and } Y_{a_0 a_1} \perp\!\!\!\perp A_1 | A_0, C, L$$

[8] If we fit any parametric models, then we must also assume no model misspecification.

6. Interaction

Mind your language: Effect Heterogeneity vs. Causal Interaction

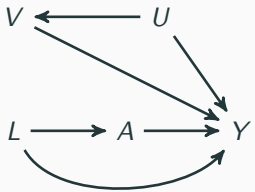
If we are interested in interaction, it is important to clearly delineate the goal of our inferences. Do we aim to identify a **heterogeneous effect** or a **causal interaction?**

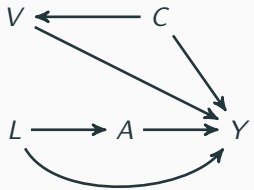
Type	Causal Question	Counterfactual	Confounding	Relevance
Effect Heterogeneity	Does the effect of one exposure A differ across strata of another exposure V?	Interested in Y^a, and requires $Y^a \perp\!\!\!\perp A$	Must control for confounders of the A-Y relationship only.	Relevant for choosing subgroups who should be targeted for an intervention.
Causal Interaction	Does the effect of one exposure A differ if I intervene to fix a second exposure, V, to a different value?	Interested in $Y^{a,v}$, and requires $Y^{a,v} \perp\!\!\!\perp A$ **and** $Y^{a,v} \perp\!\!\!\perp V$	Must control for confounders of the A-Y relationship AND of the V-Y relationship.	Relevant for deciding about joint interventions on two exposures.

Note: Our causal assumptions have implications for relevance, *not* the other way round. Feasibility of joint intervention does not determine whether something is causal interaction or effect heterogeneity.

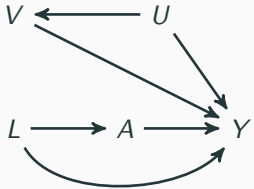
Interaction: Effect Heterogeneity vs. Causal Interaction

In which DAG(s) can we identify causal interaction?

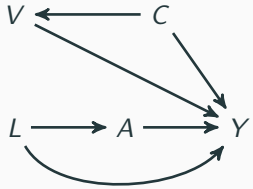




Interaction: Effect Heterogeneity vs. Causal Interaction



If U is unmeasured, we cannot identify the causal effect of V on Y, and so in the presence of statistical interaction, we can only say there is effect heterogeneity (assuming that the causal effect of A on Y is identifiable).



If we control for confounding by C, we can identify the causal effect of V on Y. So, in the presence of statistical interaction, we can say that there is causal interaction (assuming that the causal effect of A on Y is also identifiable).

Interaction: Additive vs. Multiplivative Scales

In most cases, **interaction is scale dependent.** If both exposures A and V have an effect on (or are associated with) the outcome, then there must be an interaction on at least one scale!

Scale	Description	Measure	Positive vs. Negative
Additive	"How much bigger as a difference is the effect of two exposures together than the sum of the effects of each exposure separately?"	$\boldsymbol{p_{11} - p_{10} - p_{01} + p_{00}}$ Where p_{av} denotes the risk of the outcome for those with $A = a$ and $V = v$	Positive: $p_{11} - p_{10} - p_{01} + p_{00} > 0$ Negative: $p_{11} - p_{10} - p_{01} + p_{00} < 0$
Multiplicative	"How much bigger as a ratio is the effect of two exposures together than the product of the effects of each exposure separately?"	$\boldsymbol{RR_{11}/(RR_{10} \times RR_{01})}$ Where $RR_{11} = p_{11}/p_{00}$, $RR_{10} = p_{10}/p_{00}$, and $RR_{01} = p_{01}/p_{00}$	Positive: $RR_{11}/(RR_{10} \times RR_{01}) > 1$ Negative: $RR_{11}/(RR_{10} \times RR_{01}) < 1$

Interaction: Other Important Measures of Interaction

Relative Excess Risk due to Interaction (RERI) can be used to estimate **additive interaction** from a multiplicative model:

$$\begin{aligned} \text{RERI} &= p_{11}/p_{00} - p_{10}/p_{00} - p_{01}/p_{00} + p_{00}/p_{00} \\ &= RR_{11} - RR_{10} - RR_{01} + 1 \end{aligned}$$

If the outcome is rare, then *ORs* can be used in place of *RRs*.

Rothman's Attributable Proportion: How much of the **risk of disease** when both exposures are present is due to interaction?

$$AP = RERI/RR_{11}$$

Both of these are **additive measures of interaction.** Values > 0 indicate positive additive interaction, while values < 0 indicate negative additive interaction.

Interaction: Other Important Measures of Interaction

The **Attributable Proportion for Effects** (joint effect decomposition) captures the proportion of the **effect** of both exposures that is due to the:

$$\text{First exposure only} : (p_{10} - p_{00})/(p_{11} - p_{00}) = (RR_{10} - 1)/(RR_{11} - 1)$$
$$\text{Second exposure only} : (p_{01} - p_{00})/(p_{11} - p_{00}) = (RR_{01} - 1)/(RR_{11} - 1)$$
$$\text{Interaction of the two exposures} : (p_{11} - p_{10} - p_{01} + p_{00})/(p_{11} - p_{00}) = RERI/(RR_{11} - 1)$$

These three quantities will **sum to 100%.** In contrast to Rothman's Attributable proportion, this interaction measure essentially *subtracts out the risk of disease when neither exposure is present from the denominator.*

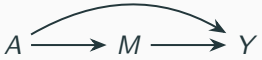
See the lecture notes and lab handout for additional details on **qualitative interaction** and **mechanistic/sufficient cause interaction.**

7. Mediation

Mediation: Overview

In some research settings, we may be interested in the following questions:

- **To what extent is the effect of some exposure A on some outcome Y mediated by an intermediate variable M?**
- **To what extent is the effect of the exposure on the outcome direct?**



Mediation analyses help answer these and other related questions.

Mediation: Standard (non-causal) Approaches

Difference method:

$$E[Y \mid A{=}a, C{=}c] = \phi_0 + \phi_1 a + \phi_2' c$$

$$E[Y \mid A{=}a, M{=}m, C{=}c] = \theta_0 + \theta_1 a + \theta_2 m + \theta_2' c$$

$$\text{Direct Effect} = \theta_1$$

$$\text{Indirect Effect} = \phi_1 - \theta_1$$

Product method:

$$E[M \mid A{=}a, C{=}c] = \beta_0 + \beta_1 a + \beta_2' c$$

$$E[Y \mid A{=}a, M{=}m, C{=}c] = \theta_0 + \theta_1 a + \theta_2 m + \theta_2' c$$

$$\text{Direct Effect} = \theta_1$$

$$\text{Indirect Effect} = \beta_1 \theta_2$$

What are potential problems with the standard approaches to mediation analysis?

Mediation: Standard (non-causal) Approaches

Difference method:

$$E[Y \mid A{=}a, C{=}c] = \phi_0 + \phi_1 a + \phi_2' c$$

$$E[Y \mid A{=}a, M{=}m, C{=}c] = \theta_0 + \theta_1 a + \theta_2 m + \theta_2' c$$

$$\text{Direct Effect} = \theta_1$$

$$\text{Indirect Effect} = \phi_1 - \theta_1$$

Product method:

$$E[M \mid A{=}a, C{=}c] = \beta_0 + \beta_1 a + \beta_2' c$$

$$E[Y \mid A{=}a, M{=}m, C{=}c] = \theta_0 + \theta_1 a + \theta_2 m + \theta_2' c$$

$$\text{Direct Effect} = \theta_1$$

$$\text{Indirect Effect} = \beta_1 \theta_2$$

What are potential problems with the standard approaches to mediation analysis?

Problem #1: Mediator-outcome confounders have traditionally been ignored.

Problem #2: Exposure-mediator interactions, if present and ignored, can lead to bias.

Mediation: Causal Mediation Effects

We discussed **three** causal mediation effects: the **controlled direct effect (CDE)**, the **natural direct effect (NDE)**, and the **natural indirect effect (NIE)**. On the additive scale:

Effect	Notation	Interpretation
$CDE(m)$	$E[Y_{am} - Y_{a^*m}]$	Effect of switching exposures from a to a^* if everyone is **forced to take mediator level** m
NDE	$E[Y_{aM_{a^*}} - Y_{a^*M_{a^*}}]$	Effect of switching exposures from a to a^* if we force everyone's mediator to the level that **it would naturally take under exposure** a^*
NIE	$E[Y_{aM_a} - Y_{aM_{a^*}}]$	Effect of **switching the mediator** from the level it would naturally take under exposure a^* to the level it would naturally take under exposure a if we force everyone to take treatment a.

Mediation: Assumptions

Identification of causal mediation effects generally requires several **no unmeasured confounding assumptions (NUCA):**

1. $Y_{a,m} \perp\!\!\!\perp A \mid C$: No unmeasured exposure-outcome confounders
2. $Y_{a,m} \perp\!\!\!\perp M \mid (C, A)$: No unmeasured mediator-outcome confounders
3. $M_a \perp\!\!\!\perp A \mid C$: No unmeasured exposure-mediator confounders
4. $Y_{a,m} \perp\!\!\!\perp M_{a*} \mid C$: No mediator-outcome confounder affected by exposure

- Only assumptions **1 and 2** are necessary for identifying the **CDE(m)**.
- In general, **all 4** assumptions are needed to identify the **NIE and NDE**.
- If there is no exposure-mediator interaction, then the **CDE(m) = NDE!**

Mediation: Causal Mediation Calculations

One approach to estimate causal mediation effects is to fit **regression models** for the outcome and the mediator. For example, on the additive scale:[9]

$$E[M \mid A = a, C = c] = \beta_0 + \beta_1 a + \beta_2' c$$
$$E[Y \mid A = a, M = m, C = c] = \theta_0 + \theta_1 a + \theta_2 m + \theta_3 am + \theta_4' c$$

If all 4 NUCA assumptions hold and these models are correctly specified, then the *CDE*, *NDE* and *NIE* are given by:

$$CDE(m) = (\theta_1 + \theta_3 m)(a - a^*)$$
$$NDE = (\theta_1 + \theta_3(\beta_0 + \beta_1 a^* + \beta_2' c))(a - a^*)$$
$$NIE = (\theta_2\beta_1 + \theta_3\beta_1 a)(a - a^*)$$

[9] Note that here we fit linear regression models for both the mediator and the outcome. However, this approach also extends to other types of mediators and outcomes (e.g., binary). See lecture slides for further details.