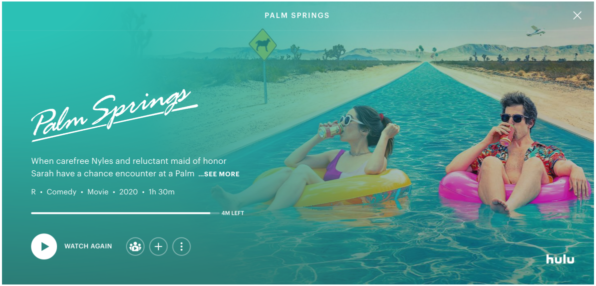
Observational Studies:
Handling Confounding
MATH 348, Vassar College, Spring 2024



Hulu, www.hulu.com

Prof. Lee Kennedy-Shaffer

April 1, 2024

Outline

Surveys vs. Observational Studies

Surveys

- Describe the world as it is
- Identify correlations, associations, discrepancies, disparities

Surveys vs. Observational Studies

Surveys

- Describe the world as it is
- Identify correlations, associations, discrepancies, disparities

Observational Studies

- Describe the world *as it could be*
- Identify causal relationships and mechanisms

Example Question

Is there gender discrimination in Vassar's faculty salaries?

(Potentially) Key Question

Disparity vs. *Causality*

Causal Inference

Definition

A framework of methods and assumptions that determine when an association/correlation parameter can be interpreted as a causal estimand.

Causal Inference

Definition

A framework of methods and assumptions that determine when an association/correlation parameter can be interpreted as a causal estimand.

Various frameworks and philosophies:

- Boyle and Hobbes on experimental proof
- Immanuel Kant and David Hume
- Karl Popper and Ludwig Wittgenstein
- Koch's Postulates
- Austin Bradford Hill

Statistical Notions of Causality

- Sufficient and necessary causes
- Rothman's Causal Pies
- Rubin's Potential Outcomes
- Pearl's Structural Models
- Robins, Rotnitzky, Hernán, Tchetgen-Tchetgen, and Richardson's Graphical Models and g-Formula

Necessary Elements

- **Positivity**: each unit has some positive probability of either exposure

Necessary Elements

- **Positivity**: each unit has some positive probability of either exposure
- **Consistency/SUTVA**: a well-defined exposure for each individual

Necessary Elements

- **Positivity**: each unit has some positive probability of either exposure
- **Consistency/SUTVA**: a well-defined exposure for each individual
- **Exchangeability**: in the absence of exposure, the expected outcomes in each arm would be equal

DAGs in Two Minutes

1. Write all possible contributing causes to either the exposure or outcome in temporal order

DAGs in Two Minutes

1. Write all possible contributing causes to either the exposure or outcome in temporal order
2. Draw arrows in the direction of causality anywhere you can't rule it out

DAGs in Two Minutes

1. Write all possible contributing causes to either the exposure or outcome in temporal order
2. Draw arrows in the direction of causality anywhere you can't rule it out
3. Put a box around any measured factors accounted for in the model

DAGs in Two Minutes

1. Write all possible contributing causes to either the exposure or outcome in temporal order
2. Draw arrows in the direction of causality anywhere you can't rule it out
3. Put a box around any measured factors accounted for in the model
4. See if there are any indirect (backdoor) paths from exposure to outcome

Example DAG

Example Study

Is there an effect of gender on Vassar faculty salaries?

Threats to Exchangeability

- Reverse causality: the outcome affects the (measured) exposure
- Confounding: a common cause of the exposure and outcome
- Selection bias: stratifying on/selecting on a common consequence of the exposure and outcome
- Mediation/effect modification: a factor that is on the causal pathway from the exposure to outcome and/or changes the magnitude of the effect
- More complex relationships, time-varying effects, etc.

Prospective Studies

Definition

Study units are identified and recruited to the study before the outcome would occur.

Prospective Studies

Definition

Study units are identified and recruited to the study before the outcome would occur.

Includes *longitudinal studies*, which follow individuals and measure the outcome repeatedly over time.

Retrospective Studies

Definition

Study units are identified and recruited to the study after the outcome has occurred.

Retrospective Studies

Definition

Study units are identified and recruited to the study after the outcome has occurred.

Includes *case-control* or *outcome-dependent* studies, where individuals are recruited based on their outcome value.

Ad-Hoc Analyses

Many ad-hoc analyses of existing data are essentially observational studies. Need to be especially careful about:

- *Selection bias*
- *Measurement error*
- *Rationalization over theory*

Outline

Observational Studies vs. RCTs

RCTs are often called the *gold standard* of causal inference. The key benefit of RCTs is establishing exchangeability through randomization.

So our first concern for observational studies is creating exchangeability by handling threats to exchangeability.

Confounding: Definition and DAG

Definition

A confounder causes both the exposure and the outcome. Also called: lurking variable, extraneous determinant. Related to endogeneity.

Confounding occurs in a study if there is at least one confounder not accounted for in the design/analysis.

Potential Confounders in School Study

Potential Outcomes

$$
\begin{aligned}
Y(1) &= \text{ potential outcome of study unit under exposure condition 1} \\
Y(0) &= \text{ potential outcome of study unit under exposure condition 0} \\
Y^{obs} &= \text{ actual outcome} \\
Z &= \text{ actual exposure value}
\end{aligned}
$$

Potential Outcomes

$$\begin{aligned} Y(1) &= \text{ potential outcome of study unit under exposure condition 1} \\ Y(0) &= \text{ potential outcome of study unit under exposure condition 0} \\ Y^{obs} &= \text{ actual outcome} \\ Z &= \text{ actual exposure value} \end{aligned}$$

- Under consistency, $Y^{obs} = Y(1)I(Z=1) + Y(0)I(Z=0)$.

Exchangeability in Potential Outcomes

For exchangeability, we need the two exposure groups to have equal expected *potential* outcomes:

$$E[Y(0)|Z=0] = E[Y(0)|Z=1]$$
$$E[Y(1)|Z=0] = E[Y(1)|Z=1]$$

Handling Confounding

1. Restrict population
2. Match or stratify on all confounders or a summary (e.g., propensity score)
3. Adjust for all confounders in statistical model (e.g., regression)

Restrict Population

Advantages:

- Clear applicability
- No statistical modeling needed

Restrict Population

Advantages:

- Clear applicability
- No statistical modeling needed

Disadvantages:

- Limited generalizability
- May limit sample size
- May still be *residual confounding*

Matching/Stratification

Advantages:

- Ensures balance in study population
- Lower variance, higher precision/power (if we actively recruit within strata)
- Allows for subgroup analyses

Matching/Stratification

Advantages:

- Ensures balance in study population
- Lower variance, higher precision/power (if we actively recruit within strata)
- Allows for subgroup analyses

Disadvantages:

- Logistically difficult
- Changes target population/generalizability
- Need many strata or a good model of propensity score

Adjust for Confounders in Model

Advantages:

- Logistically simple
- Allows for flexibility of modeling choices
- Can handle any residual confounding

Adjust for Confounders in Model

Advantages:

- Logistically simple
- Allows for flexibility of modeling choices
- Can handle any residual confounding

Disadvantages:

- Need to correctly specify model
- Can reduce power/increase variability (high weights)
- Need to collect many variables

Risks of Overadjusting

To get exchangeability, we need to restrict/stratify/adjust for *all confounders* but *no mediators, effect modifiers, or common consequences*.

Danger

Adjusting for too many variables can induce *selection bias*, *block the causal pathway*, or cause other biases.

DAG examples:

Outline

Consistency vs. Generalizability

Observational studies generally have less consistency in exposures than RCTs. This can bias the estimates/make the estimands less clear.

Consistency vs. Generalizability

Observational studies generally have less consistency in exposures than RCTs. This can bias the estimates/make the estimands less clear.

But their exposure levels may better reflect real-world conditions.

Bias vs. Variance

Observational studies have a higher risk of unmeasured or residual confounding than RCTs. Can cause bias through non-exchangeability.

Bias vs. Variance

Observational studies have a higher risk of unmeasured or residual confounding than RCTs. Can cause bias through non-exchangeability.

But they may be able to recruit many more individuals, reducing the variance.

Bias vs. Variance

Observational studies have a higher risk of unmeasured or residual confounding than RCTs. Can cause bias through non-exchangeability.

But they may be able to recruit many more individuals, reducing the variance.

Can quantify this in different ways, such as:

$$MSE(\hat{\theta};\theta) = E\left[\left(\hat{\theta}-\theta\right)^2\right] = Bias(\hat{\theta};\theta)^2 + Var(\hat{\theta})$$

Positivity vs. Generalizability

Observational studies may lack positivity for some strata/combinations of covariates, which limits ability to do inference.

Positivity vs. Generalizability

Observational studies may lack positivity for some strata/combinations of covariates, which limits ability to do inference.

But those may be structural zeros and even represent study units who would never take the intervention, so this could improve the generalizability.

Think Tradeoffs

- What does this study give us that we can't get from a RCT?
- How do the variance and generalizability compare to a similar RCT?
- What risks of bias exist?

Think Continuously, Not Binarily!

Remember

All of these tradeoffs, goals, and principles (except maybe ethics) exist on a spectrum, not a binary.

Don't let the perfect be the enemy of the good: does this study generate useful evidence even if it is imperfect?

But don't assume that something is always better than nothing.

Accurate and Clear Reporting

- Does the study accurately convey its strengths and weaknesses?
- Are limitations considered throughout, not just in a separate section?
- Can the audience get a clear takeaway that is both understandable and accurate?