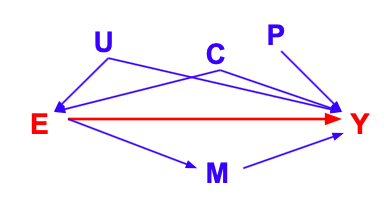
Propensity Score-Based Methods for Causal Inference

Module 4: Specifying Variables for the Propensity Score



**I. Module Objectives**

After we defining the research question in terms of a causal effect and describing the nature of available data (and its strengths and limitations), we need to identify variables to include in the propensity score. We also need to identify which variables might be missing from the data, and how those variables are part of the causal process. Causal graphs, which describe the possible relationships (and directions of those relationships) between variables of interest, are one useful tool for doing so.

By the end of this module, you will be able to:

1. Describe the motivation for, and process of using propensity score-based models.
2. Identify variables and draw a causal diagram to specify variables for propensity scores
3. Identify assumptions for making causal inferences

**II. Module Assignments**

**Required Assignments: (~26 minutes of videos + 11 pages to read)**

For an introduction to the challenge of confounding in observational studies, watch Module 3b (**~10 minutes**) of the videos from [Category 8 (on causal inference) from the PCORI Methodology Standards Academic Curriculum](https://www.pcori.org/research-results/about-our-research/research-methodology/methodology-standards-academic-curriculum-5). Then watch Module 5 (**~10 minutes**) from [Category 8](https://www.pcori.org/research-results/about-our-research/research-methodology/methodology-standards-academic-curriculum-5) for a very basic introduction to how those issues relate to common observational designs and the basic idea of propensity scores and instrumental variables.

To begin thinking about ways to quantify causality, including drawing a causal diagram, watch Module 4 (**~6.5 minutes**) from [Category 3 (on data integrity and rigorous analyses) from the PCORI Methodology Standards Academic Curriculum](https://www.pcori.org/research-results/about-our-research/research-methodology/methodology-standards-academic-curriculum-1).

Now that we have an understanding of the basic concepts necessary for studying methods for assessing causal relationships, it is also important to consider the assumptions being made for associated statistical methods. For a description of the key assumptions, read Chapter 3, Sections 3.1-3.5 (pages 25-36) from Hernan and Robins (in draft form, and freely available at <https://www.hsph.harvard.edu/miguel-hernan/causal-inference-book/> as of June, 2019): Hernán MA, Robins JM (2019). Causal Inference. Boca Raton: Chapman & Hall/CRC, forthcoming. ***For the reader without graduate level training in statistics***, you may want to skip the formulas and probability statements throughout the chapter.

**Optional Assignment: (~28 minutes)**

For more information on developing causal graphs and using the conditional probabilities and algorithms to discover causal relationships, watch the 30-Minute Tutorial on Causal Modeling and Discovery from the [Video Tutorials from the Center for Causal Discovery](https://www.ccd.pitt.edu/video-tutorials/).

**Additional Optional Assignment: These (longer) articles are also useful references:**

Two other articles by Rubin further emphasize the importance of design and illustrate the use of propensity scores in public health:

1. Rubin, D. B. (2008). For objective causal inference, design trumps analysis. *The Annals of Applied Statistics*, *2*(3), 808-840.
2. Rubin, D. B. (2001). Using propensity scores to help design observational studies: application to the tobacco litigation. *Health Services and Outcomes Research Methodology*, *2*(3-4), 169-188.

**III. Project Exercises**

Create a copy of this Google Doc or download the Module onto your computer and review the material offered above under Module Assignments before beginning these workbook exercises.

Thinking about what you learned in this module so far, begin developing the analysis plan for your project by answering the following questions:

1. Considering your research question and available data (from an existing data set, or from a data set that you plan to collect, as described in question #1 from Module 3), draw a causal graph to describe the relationship between the exposure, possible confounders, and the outcome of interest.

In your causal graph, make a distinction between measured and unmeasured confounders. Also draw your causal graph to include the following types of variables:

1. Predictors of the outcome that not also confounders of the relationship.
2. Mediators which are on the causal pathway between exposure and outcome.
3. Instrumental variables which predict the exposure mechanism, but are conditionally (given exposure) independent of outcome.

Your propensity score should include predictors, but not mediators or instruments.

1. Consider A) the causal graph from question #1 above and B) your responses to questions #2 and #3 of the Module 3, where you described the variables that you intend to use or collect, and the associated temporal associations. Based on those responses, reconsider your responses in the previous model and reassess whether those variables and temporal relationships are consistent with your causal graph. If not, revise those responses.
2. Reread Sections 3.1-3.5 of the Hernan and Robins text, and consider whether and describe how your data may induce any obvious violations of these assumptions. Also describe how you might change the analysis population to avoid violating assumptions.

As one example, consider the case of an experimental versus standard treatment. You may have subjects who have certain characteristics which will almost guarantee use of the experimental treatment, whereas others have characteristics of the standard treatment. This would violate the positivity assumption. To account for this, you may only consider subjects in your analysis/inclusion criteria where there is clinical indecision in the actual care.

[Link to go back to the Course Overview Document](https://docs.google.com/document/d/1UDTkp3rbhqdun7jvSvktaZmTtoUWOz_VUDQw3HIsElg/edit?usp=sharing)