Note for ML

Estimation Methods under Unconfoundedness

This section contains methods for estimating (heterogeneous) treatment effects, whose theoretical guarantees are valid only when all potential confounders/controls (factors that simultaneously had a direct effect on the treatment decision in the collected data and the observed outcome) are observed.

Two conditions of Instrumental variables

The setup of the model is as follows:

Y=g(T,X,W)+ϵ

where E[ε|X,W,Z]=h(X,W), so that the expected value of Y depends only on (T,X,W). This is known as the *exclusion restriction*. We assume that the conditional distribution F(T|X,W,Z) varies with Z. This is known as the *relevance condition*.

A screenshot of a diagram

AI-generated content may be incorrect.

1. **Backdoor criterion** (or more generally, adjustment sets): If all common causes of the action A and the outcome Y are observed, then the backdoor criterion implies that the causal effect can be identified by conditioning on all the common causes. This is a simplified definition (refer to Chapter 3 of the CausalML book for a formal definition).

│E[Y│do(A=a)]=EWE[Y|A=a,W=w]

where W refers to the set of common causes (confounders) of A and Y.

1. **Instrumental variable (IV) identification**: If there is an instrumental variable available, then we can estimate effect even when any (or none) of the common causes of action and outcome are unobserved. The IV identification utilizes the fact that the instrument only affects the action directly, so the effect of the instrument on the outcome can be broken up into two sequential parts: the effect of the instrument on the action and the effect of the action on the treatment. It then relies on estimating the effect of the instrument on the action and the outcome to estimate the effect of the action on the outcome.