Marginal Structural Models and Causal Inference in Epidemiology

James M. Robins, Miguel A. Hernan & Babette Brumback

Thomas Fuchs

2006-07-04

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

ICOS - Institute of Computational Science
Pattern Analysis and Machine Learning Group

MSMs

Marginal structural models (MSMs) are causal models designed to adjust for time-dependent confounding in observational studies of time-varying treatments.

Estimation of Causal Effects

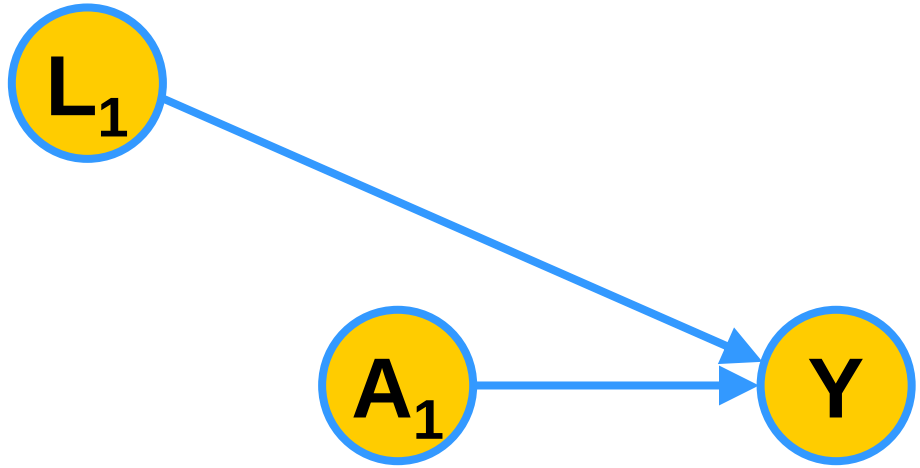
- **Traditionally done:**
  by modeling the probability of disease as a function of treatment and pretreatment covariates

- **Problem:**
  biased if time-varying covariates are simultaneously confounders and intermediates

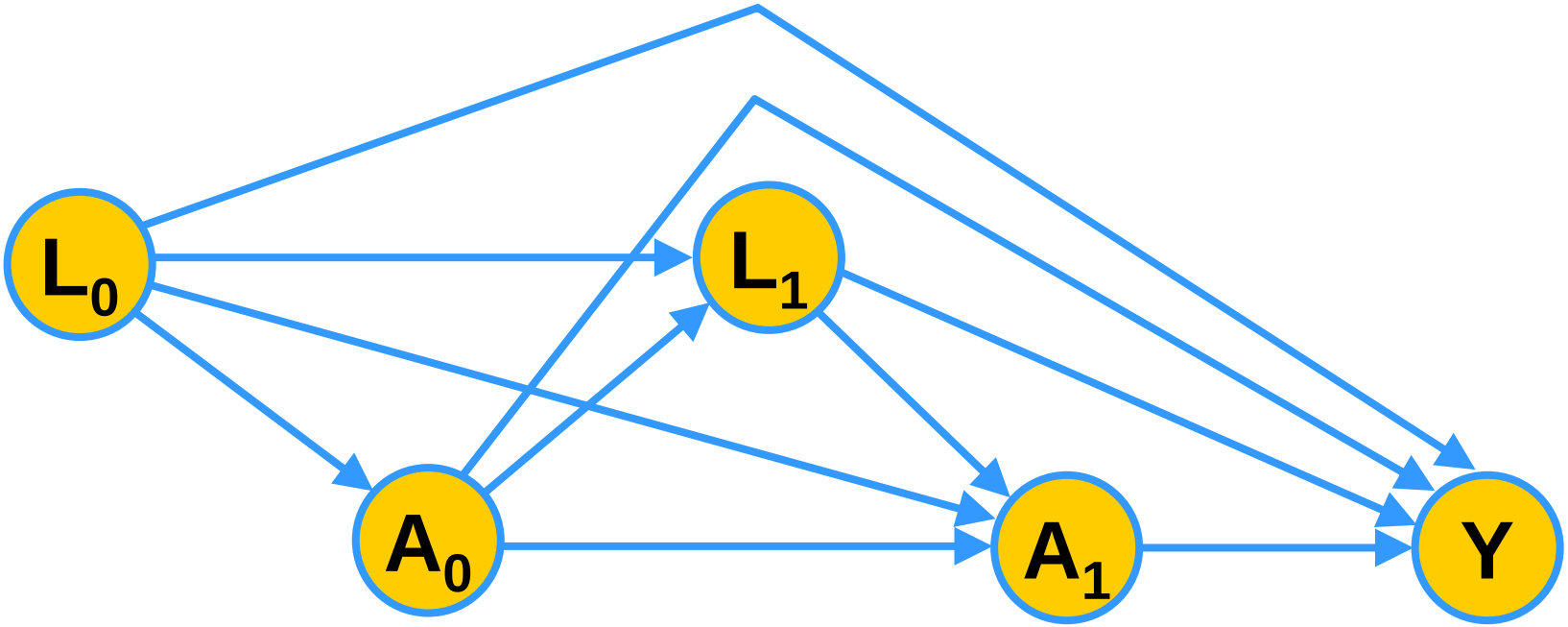
  i.e. covariates are predictors of outcome and also predict subsequent treatment, and past treatment history predicts resulting covariate level

Treatment == Exposure

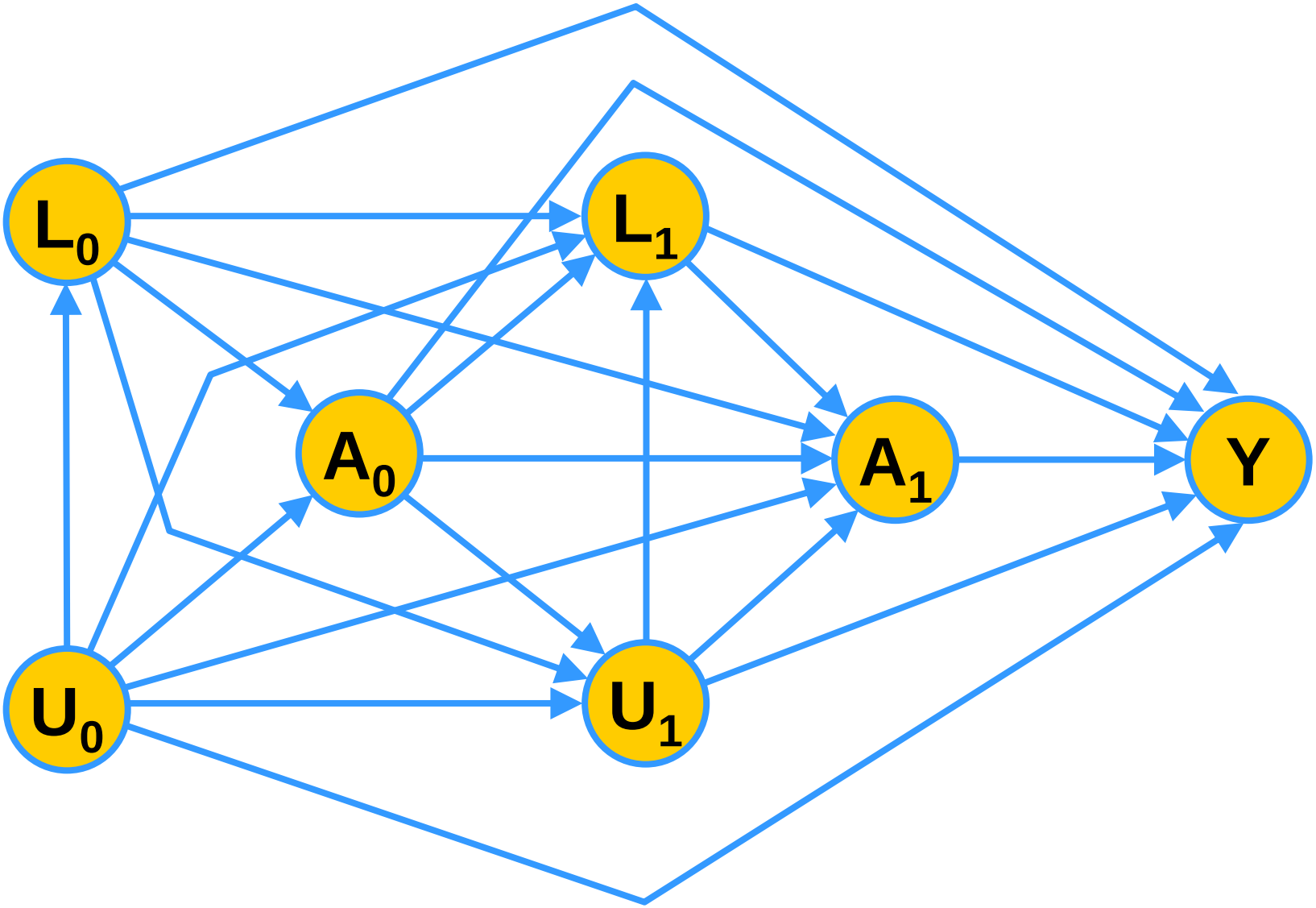
Causal Graphs



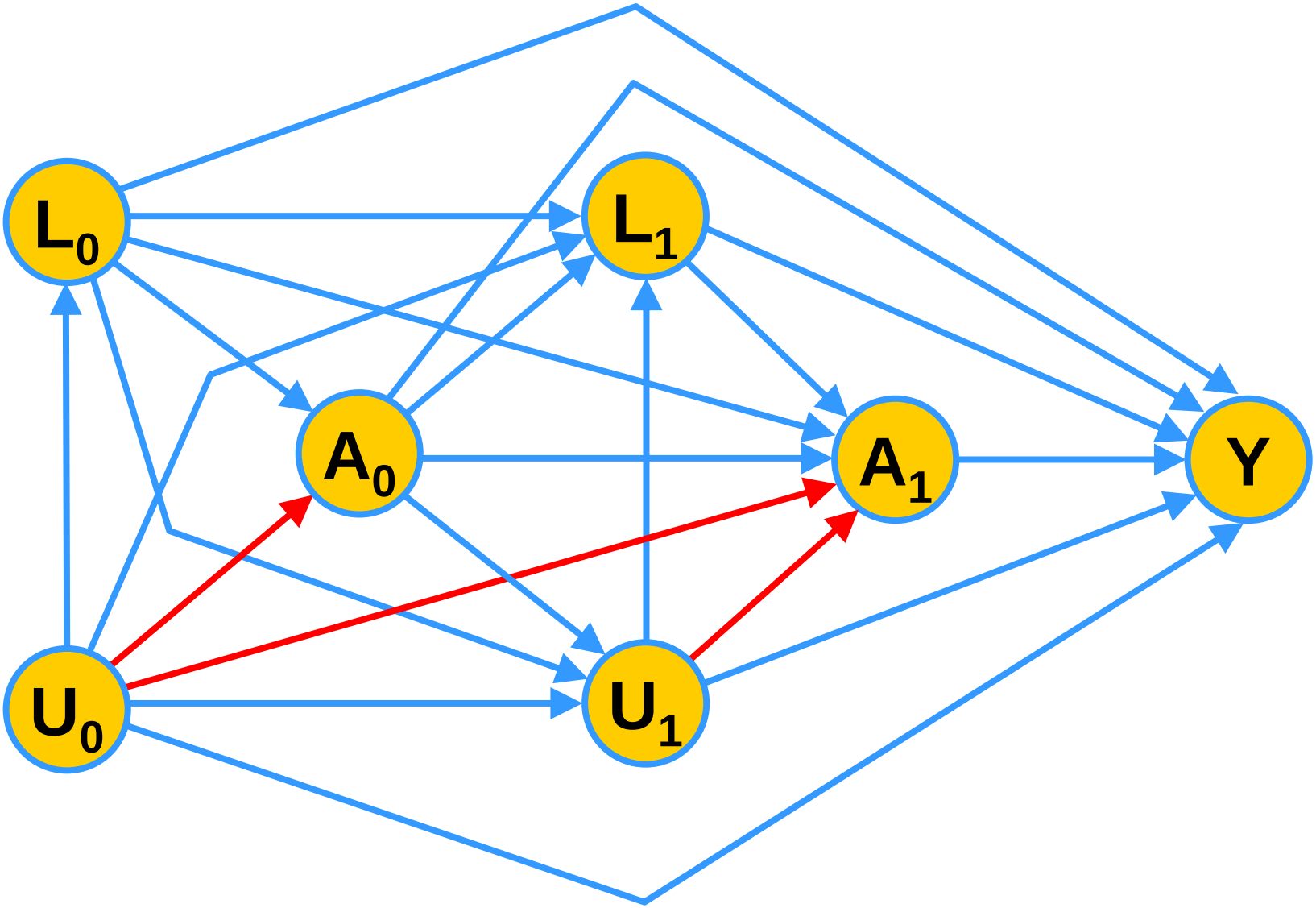
Causal Graphs



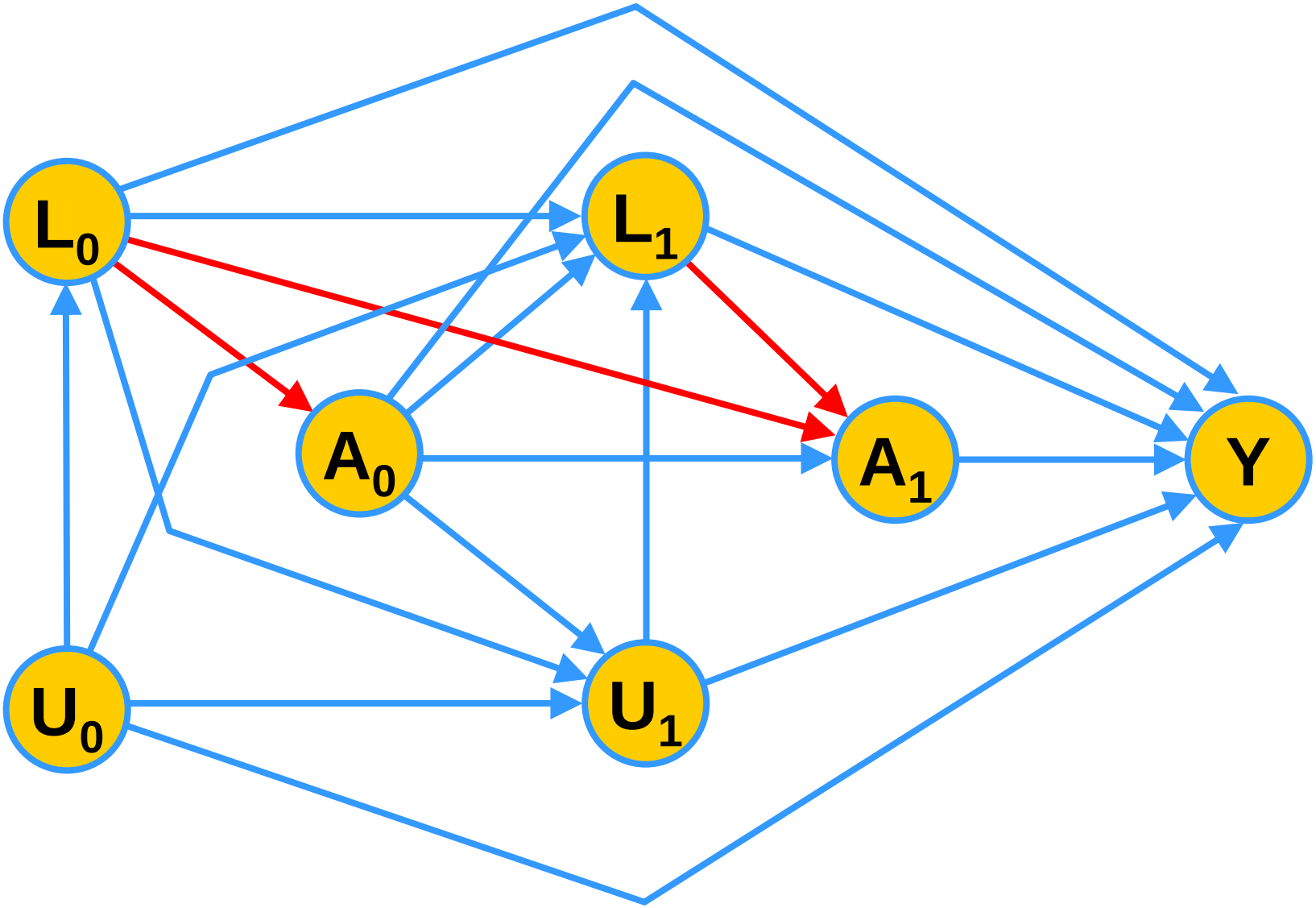
Causal Graphs



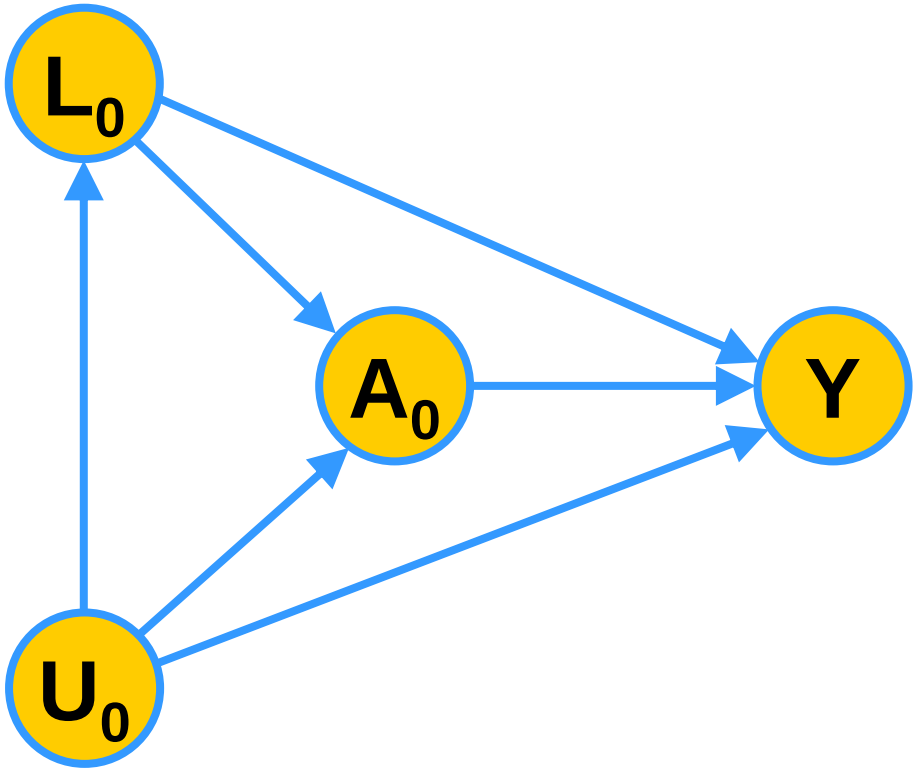
Causal Graphs



Causal Graphs

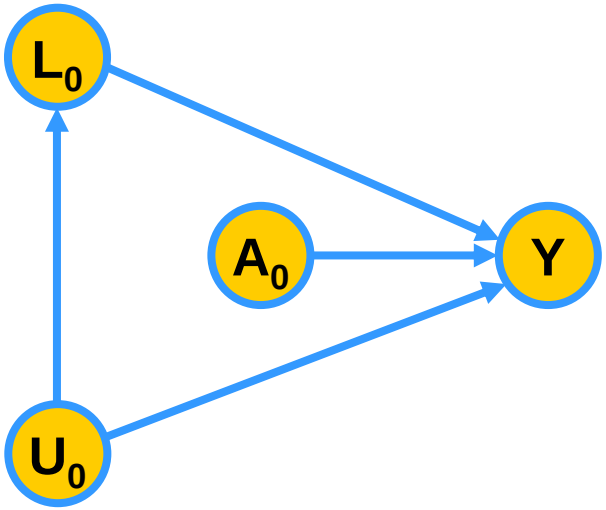


Point Treatment Study



A **sufficient condition for treatment to be unconfounded** is that, at each time k, among subjects with the same past treatment history A_0, …, A_{k-1}, the treatment A_k is unassociated with the past history of measured covariates L_0, …, L_k.

Point Treatment Study



Suppose: true causal graph

Neither measured nor unmeasured covariates confound the relation between treatment and outcome.

Measure causal effect
of A_0 on Y on different scales:

cRD: crude Risk Difference

cRR: crude Risk Ratio

cOR: crude Odds Ratio

$$\mathbf{cRD} = \mathbf{pr}[Y=1 \mid A_0=1] - \mathbf{pr}[Y=1 \mid A_0=0]$$

Potential Outcome

- Each person ***i*** has 2 responses:
  - One that would be observed if they were treated (Y_{1i})
  - One that would be observed if they were not treated (Y_{0i}).

- Since we can never observe the same person simultaneously as a case and as a control, we can never observe both potential outcomes

Counterfactual Variables

$Y_{a^0=1}$ denotes a subject's outcome if treated

$Y_{a^0=0}$ denotes a subject's outcome if not treated

Both can **NOT** be observed at the same time for any given subject!

$$\text{causal } \mathbf{RD} = \mathbf{pr}\left[Y_{a^0=1}=1\right] - \mathbf{pr}\left[Y_{a^0=0}=1\right]$$

Counterfactuals

- In the causal inference framework, the full range of treatment-specific outcomes for a given unit is referred to as the set of **counterfactuals** in the sense that only one treatment-outcome pair can be observer for any unit.

- All possible Y_a's for each subject are called **counterfactuals** because only one can be observed, with the rest "counter to the facts."

- *"The concept of **counterfactuals** itself, which is at the root of causality, appears to be an abstract concept which initially can seem esoteric."*

(Romain S. Neugebauer)

Linear Models

Causal RD, RR and OR expressed in terms of the parameters of a linear, log linear and linear logistic model:

$$\mathbf{pr}\left[Y_{a^0}=1\right] = \psi_0 + \psi_1 a_0$$

causal RD = $\psi 1$

$$\mathbf{log\ pr}\left[Y_{a^0}=1\right] = \theta_0 + \theta_1 a_0$$

causal RR = $\exp(\theta_1)$

$$\mathit{logit}\ \mathbf{pr}\left[Y_{a^0}=1\right] = \beta_0 + \beta_1 a_0$$

causal OR = $\exp(\beta_1)$

Saturated MSMs

Saturated MSMs

- **marginal**
  model the marginal distribution of the counterfactual random variables $Y_{a^0=1}$ and $Y_{a^0=0}$

- **structural**
  models for counterfactual variables are referred to as structural in econometric

- **saturated**
  has two variables an therefore no restriction on the possible values of the two unknown probabilities

Associational Models

Models for the observed data to calculate the crude RD, RR and OR.

$$\mathbf{pr}\left[\mathbf{Y=1} \mid \mathbf{A_0=a_0}\right] = \psi'_0 + \psi'_1 \mathbf{a_0}$$

$$\mathbf{log\ pr}\left[\mathbf{Y=1} \mid \mathbf{A_0=a_0}\right] = \theta'_0 + \theta'_1 \mathbf{a_0}$$

$$\textit{logit}\ \mathbf{pr}\left[\mathbf{Y=1} \mid \mathbf{A_0=a_0}\right] = \beta'_0 + \beta'_1 \mathbf{a_0}$$

The parameters of the associational models will differ from the parameters of the MSMs, except when treatment is unconfounden.

Confounded Treatment

Suppose:

Treatment is confounded by L but assuming we have no unmeasured confounders. Then

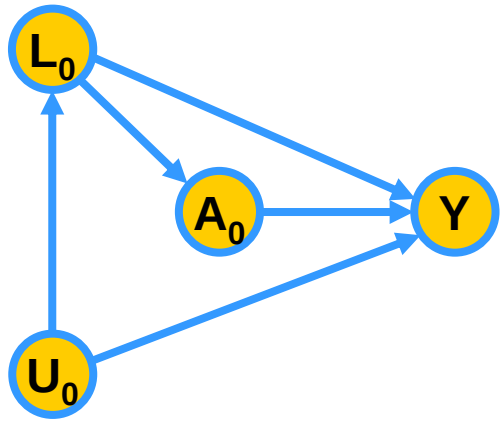
$\beta_0 \neq \beta_0'$, $\beta_1 \neq \beta_1'$, …

Goal:

Unbiased estimates of the parameters ψ_1, θ_1, β_1

Solution:

Weighted Analysis



IPTW

Each subject ***i*** is assigned a weight w_i equal to the inverse of the conditional probability of receiving his or her own treatment.

$$w_i = \frac{1}{pr[A_0 = a_{0i} \mid L_0 = l_{0i}]}$$

l_{0i} is the observed value of the variable **L** for subject ***i***

Estimation of Weights

The true weights are unknown, but can be estimated from the data with a preliminary logistic regression:

$$logit\ \ pr[A_0 = 1 | L_0 = l_0] = \alpha_0 + \alpha_1 l_0$$

Now we can estimate $\mathbf{w_i}$

$$w_i = 1 + \exp(\hat{\alpha}_0 + \hat{\alpha}_1 l_{0i})$$

IPTW

That means:

If there are no unmeasured confounders given data on L_0, one can control confounding (due to L_0) by modifying the crude analysis by weighting each subject ***i*** by w_i.
The denominator of w_i is the probability that a subject had his or her own observed treatment.

Inverse **P**robability of **T**reatment **W**eighting

IPTW Pseudopopulation

Creation of a **pseudopopulation**.

Consists of w_i copies of each subject **i**.

With the following properties:

- A_0 is unconfounded by the measured covariates L_0.
- $Pr[Y_{a^0=1}=1]$ and $Pr[Y_{a^0=0}=1]$ are the same as in the true study population so that the causal **RD, RR, OR** are the same in both populations. Therefore they can be unbiasedly estimated by a standard crude analysis in the pseudopopulation.

Stabilized Weights

Problem:
Components of $\mathbf{L_0}$ are strongly associated with $\mathbf{A_0}$.

$$w_i = \frac{1}{pr[A_0 = a_{0i} \mid L_0 = l_{0i}]}$$

may vary greatly between subjects and can result in extremely large values for a few subjects.

Stabilized Weights

The stabilized weights are less variable and the estimates are still unbiased.

$$sw_i = \frac{pr[A_0 = a_{0i}]}{pr[A_0 = a_{0i} \mid L_0 = l_{0i}]}$$

The probabilities can be estimated similar to before with a polytomous model in the multilevel case.

Time-Dependent Treatment

Standard regression methods adjust for covariates by including them in the model as regressors.

They fail when treatment is time varying because:

- L_k may be a confounder for later treatment
- L_k may also be affected by earlier treatment

Solution:
Adjust for the time-dependent covariates L_k by using them to calculate the weights sw_i rather than by adding the L_k to the regression model as regressors.

Extensions

- **Multilevel Treatment**
  A_0 is a categorical variable
- **Censoring**
  Loss of Follow-Up
- **Confidence Intervals**
  robust Wald intervals

Causality Assumptions

- The given information (data) is accurate
- The measured covariates L are sufficient to adjust for confounding (and selection bias due to loss of follow up)
- The MSMs are correctly specified

Advantage:

MSMs do not require the absence of time-dependent confounding by variables affected by previous treatment.

Discussion

Discussion