Time-Varying Exposures and Marginal Structural Models: Key Concepts Worksheet

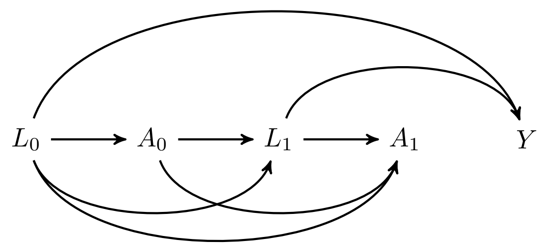
PHS 2000B: Lab 4

February 13, 2023

Traditional Regression Methods

Question 1:

In the DAG below, are these causal effects are identifiable using **traditional regression methods** alone? If yes, what model would you fit to estimate these effects?



1. $E[Y^{a_0=1} - Y^{a_0=0}|L_0]$

Yes, this effect is identifiable using traditional regression. If we're ok with estimating the full effect of A_0 through any and all pathways, then the only confounding we need to worry about is confounding by baseline confounders. No time-varying confounding is involved in this estimate. The following model could be used.

$$E[Y|A_0, L_0] = \beta_0 + \beta_1 A_0 + \beta_2 L_0$$

2. $E[Y^{a_0=1,a_1=1} - Y^{a_0=0,a_1=0}|L_0]$

No, this effect is not identifiable using traditional regression. A traditional regression model cannot appropriately control for time-varying confounding, which is present here. While L_1 may be on the causal pathway from A_0 to Y, and therefore should not be conditioned on, it also may confound the relationship between A_1 and Y. We therefore need to utilize a method which can control for L_1 without simply conditioning on it.

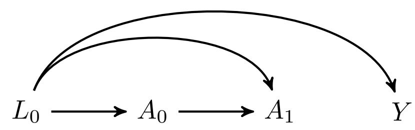
3. $E[Y^{a_1=1} - Y^{a_1=0}|L_1, L_0, A_0]$

Yes, this effect is identifiable using traditional regression. L_1, L_0, and A_0 constitute the set of confounders of the association between A_1 and Y. The following model can be used:

$$E[Y|L_1, A_0, L_0] = \beta_0 + \beta_1 A_1 + \beta_2 L_1 + \beta_3 A_0 + \beta_4 L_0$$

Question 2:

In the DAG below, are these causal effects are identifiable using **traditional regression methods?** If yes, what model would you fit to estimate these effects?



1. $E[Y^{a_0=1} - Y^{a_0=0}|L_0]$

Yes, this effect is identifiable using traditional regression. As in the DAG in question 1 above, we can estimate the effect of A_0 through all pathways. In this case, the only confounding we need to worry about is confounding by baseline confounders. The following model could be used.

$$E[Y|A_0, L_0] = \beta_0 + \beta_1 A_0 + \beta_2 L_0$$

2. $E[Y^{a_0=1,a_1=1} - Y^{a_0=0,a_1=0}|L_0]$

Yes, this effect is identifiable using traditional regression! In this DAG, there are no post-baseline confounders of our exposures A_0, A_1. As a result, the only confounder that needs to be controlled for is L_0; controlling for L_0 blocks no later causal effects. As a result, the following model could be used.

$$E[Y|A_1, A_0, L_0] = \beta_0 + \beta_1 A_1 + \beta_2 A_0 + \beta_3 L_0$$

3. $E[Y^{a_1=1} - Y^{a_1=0}|L_0, A_0]$

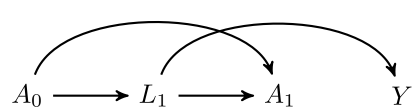
Yes, this effect is identifiable using traditional regression. L_0, and A_0 are the only confounders of the association between A_1 and Y. The following model can be used:

$$E[Y|L_1, A_0, L_0] = \beta_0 + \beta_1 A_1 + \beta_2 L_1 + \beta_3 A_0 + \beta_4 L_0$$

Sequential exchangeability

Question 1:

Does sequential exchangeability hold in the DAG below? If yes, write out the independences involved.

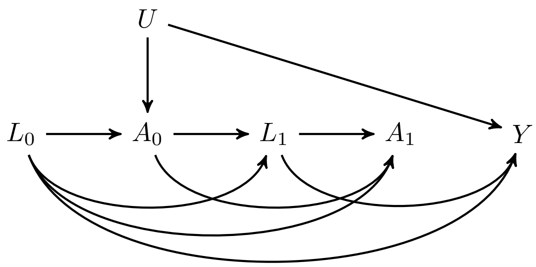


Yes, sequential exchangeability holds here. From this DAG, we can see that

$$Y^{a_0,a_1} \perp\!\!\!\perp A_0 \text{ and } Y^{a_0,a_1} \perp\!\!\!\perp A_1|L_1, A_0$$

Question 2:

Does sequential exchangeability hold in the DAG below, where U is an unmeasured variable? If yes, write out the independences involved.

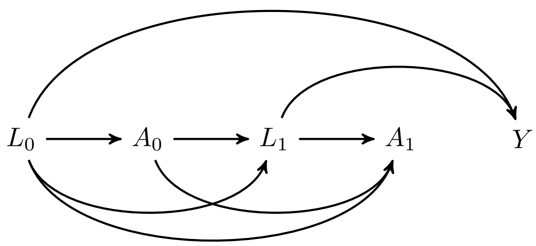


No, sequential exchangeability does **not** hold on this DAG. Although we have $Y^{a_0,a_1} \perp\!\!\!\perp A_1 | L_1, A_0$, confounding by the unmeasured covariate U means that $Y^{a_0,a_1} \not\perp\!\!\!\perp A_0$ and $Y^{a_0,a_1} \not\perp\!\!\!\perp A_0 | L_0$

Building Marginal Structural Models

Question 1:

Consider the DAG below. Write out the form that the marginal structural model should take under the following modeling assumptions.



1. What form should the MSM take if we assume no interaction between any variables on the DAG?

$$E[Y^{a_0,a_1}] = \beta_0 + \beta_1 a_0 + \beta_2 a_1$$

2. What form should the MSM take if we assume interaction between A_0 and A_1?

$$E[Y^{a_0,a_1}] = \beta_0 + \beta_1 a_0 + \beta_2 a_1 + \beta_3 (a_0 * a_1)$$

3. What form should the MSM take if we assume that baseline covariates interact with our exposures?

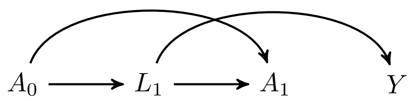
$$E[Y^{a_0,a_1} | L_0 = l_0] = \beta_0 + \beta_1 a_0 + \beta_2 a_1 + \beta_3 l_0 + \beta_4 (l_0 * a_0) + \beta_5 (l_0 * a_1)$$

4. What form should the MSM take if we assume that the effect of exposure is cumulative over time, but that the timing of exposure does not matter?

$$E[Y^{a_0,a_1}] = \beta_0 + \beta_1 [cum(a)] \text{ where } cum(a) = a_0 + a_1$$

Question 2:

For the following questions, consider estimating the MSM $E[Y^{a_0,a_1}] = \beta_0 + \beta_1 a_0 + \beta_1 a_1$ based on the following DAG.



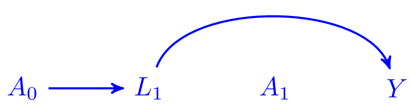
1. Define the **unstabilized** weights you should estimate to fit the MSM above. If you find it helpful, draw what the DAG will look like after weighting.

$$W_{t=0} = \frac{1}{Pr[A_0 = a_0]}$$

$$W_{t=1} = \frac{1}{Pr[A_1 = a_1 | L_1, A_0]}$$

$$W = \frac{1}{Pr[A_0 = a_0] \times Pr[A_1 = a_1 | L_1, A_0]}$$

Note that, after weighting each individual in the population by W, the DAG would look like:



2. What models would you need to fit to estimate these unstabilized weights?

For $W_{t=0}$: $logit[P(A_0 = 1)] = \beta_0$

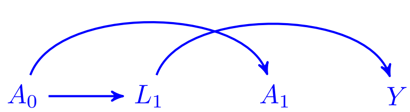
For $W_{t=1}$: $logit[P(A_1 = 1)|L_1, A_0] = \beta_0 + \beta_1 I[L_1 = 1] + \beta_2 I[A_0 = 1]$

3. Define the **stabilized** weights you should estimate to fit the MSM above. If you find it helpful, draw what the DAG will look like after weighting.

$$SW_{t=0} = \frac{Pr[A_0 = a_0]}{Pr[A_0 = a_0]} = 1$$

$$SW_{t=1} = \frac{Pr[A_1 = a_1 | A_0]}{Pr[A_1 = a_1 | L_1, A_0]}$$

$$SW = \frac{Pr[A_0 = a_0] \times Pr[A_1 = a_1 | A_0]}{Pr[A_0 = a_0] \times Pr[A_1 = a_1 | L_1, A_0]} = \frac{Pr[A_1 = a_1 | A_0]}{Pr[A_1 = a_1 | L_1, A_0]}$$



4. What models would you need to fit to estimate these stabilized weights?

No model is needed for time point $t = 0$.

For $W_{t=1}$ denominator: $logit[P(A_1 = 1)|L_1, A_0] = \beta_0 + \beta_1 I[L_1 = 1] + \beta_2 I[A_0 = 1]$

For $W_{t=1}$ numerator: $logit[P(A_1 = 1)|A_0] = \beta_0 + \beta_1 I[A_0 = 1]$