Structural Causal Models and Parameters

David Benkeser

September 20, 2016

Contents

I. Average treatment effect

Consider data generated under the following distribution that was used in class to illustrate optimal dynamic treatments.

$$\begin{aligned}
U_{W,1} &\sim \text{Bernoulli}(1/2) \\
U_{W,2} &\sim \text{Bernoulli}(1/2) \\
U_A &\sim \text{Normal}(0,1) \\
U_Y &\sim \text{Normal}(0,1) \ ,
\end{aligned}$$

and structural equations

$$\begin{aligned}
f_{W,1}(U_{W,1}) &= U_{W,1} \\
f_{W,2}(U_{W,2}) &= U_{W,2} \\
f_A(W_1, W_2, U_A) &= I(\text{expit}(W_1 - W_2 + U_A) > 0.5) \\
f_Y(W_1, W_2, A, U_Y) &= 2W_1 - 3W_1W_2A + W_2A - A + U_Y \ .
\end{aligned}$$

For the questions below that ask for an explanation of the observed phenomena, you are encouraged to write an explanation that would be understandable to an applied collaborator.

1.

(a) Numerically evaluate the true value of the average causal effect

$$\Psi(P_{U,X}) = E_{P_{U,X}}(Y_1) - E_{P_{U,X}}(Y_0)$$

by simulating from the intervened SCM, as was done in the lab. Then simulate a large data set from the observed SCM and compute the observed data parameter

$$\Psi(P_{U,X}) = E_{P_{U,X}}\{E_{P_{U,X}}(Y \mid A = 1, W_1, W_2) - E_{P_{U,X}}(Y \mid A = 0, W_1, W_2)\} \ .$$

Are the two quantities the same?

(b) Now compute the observed data parameter

$$E_{P_{U,X}}\{E_{P_{U,X}}(Y \mid A = 1, W_1) - E_{P_{U,X}}(Y \mid A = 0, W_1)\} \ .$$

Is this quantity the same as the one computed in part 1a? Why or why not?

2.

(a) Change the structural equation for Y to

$$f_Y(W_1, W_2, A, U_Y) = 2W_1 - A + U_Y \ ,$$

and numerically evaluate the true value of the average causal effect. Then simulate a large data set from the observed SCM and compute the observed data parameter

$$\Psi(P_{U,X}) = E_{P_{U,X}}\{E_{P_{U,X}}(Y \mid A = 1, W_1, W_2) - E_{P_{U,X}}(Y \mid A = 0, W_1, W_2)\} \ .$$

Are the two quantities the same?

(b) Now compute the observed data parameter

$$E_{P_{U,X}}\{E_{P_{U,X}}(Y \mid A = 1, W_1) - E_{P_{U,X}}(Y \mid A = 0, W_1)\} \ .$$

Is this quantity the same as the one computed in part 2a? Why or why not?

3. Change the structural equation for Y back to the original equation from question 1, but now change the distribution of $U_{W,1}$ and $U_{W,2}$ by setting

$$U_{W,1} \sim \text{Bernoulli}(1/4)$$
$$U_{W,2} \sim \text{Bernoulli}(3/4) \ .$$

Numerically evaluate the true value of the average causal effect. Is this effect the same as the one computed in part 1a? Why or why not?

4. Change the distribution of $U_{W,1}$ and $U_{W,2}$ back to the original distribution in question 1, but now change the structural equation for A by setting

$$f_A(W_1, W_2, U_A) = I(\text{expit}(W_1 - W_2 + U_A) > 0.25) \ .$$

Then simulate a large data set from the observed SCM and compute the observed data parameter

$$\Psi(P_{U,X}) = E_{P_{U,X}}[E_{P_{U,X}}\{Y \mid A = 1, W_1, W_2\} - E_{P_{U,X}}\{Y \mid A = 0, W_1, W_2\}] \ .$$

Is the value for this observed data parameter the same as the causal effect computed from the intervened SCM from part 1a. Why or why not?

II. Stochastic interventions

In this question, we will numerically confirm the identification results for stochastic interventions. Recall that a stochastic intervention on an SCM involves replacing the structural equation $f_A(W, U_A)$ with a random quantity A^* that has some distribution that we specify $G^*(\cdot|W)$.

1. Provide an argument for why the interventions performed in question I are special cases of a stochastic interventions. Hint: think about what kind of function $G^*(\cdot|W)$ would need to be to result in the intervention performed in I.

2. Consider the SCM from Problem I, re-printed here:

$$U_{W,1} \sim \text{Bernoulli}(1/2)$$
$$U_{W,2} \sim \text{Bernoulli}(1/2)$$
$$U_A \sim \text{Normal}(0, 1)$$
$$U_Y \sim \text{Normal}(0, 1) \ ,$$

and structural equations

$$\begin{aligned} f_{W,1}(U_{W,1}) &= U_{W,1} \\ f_{W,2}(U_{W,2}) &= U_{W,2} \\ f_A(W_1, W_2, U_A) &= I(\text{expit}(W_1 - W_2 + U_A) > 0.5) \\ f_Y(W_1, W_2, A, U_Y) &= 2W_1 - 3W_1W_2A + W_2A - A + U_Y \ . \end{aligned}$$

Write a function to numerically evaluate $E_{P_{U,X}}(Y_{G_1^*})$, the average outcome under the intervention G_1^* defined as follows:

$$\begin{aligned} G_1^*(A^* = 1 \mid W_1 = 0, W_2 = 0) &= 0.25 \\ G_1^*(A^* = 1 \mid W_1 = 0, W_2 = 1) &= 0.50 \\ G_1^*(A^* = 1 \mid W_1 = 1, W_2 = 0) &= 0.50 \\ G_1^*(A^* = 1 \mid W_1 = 1, W_2 = 1) &= 0.75 \end{aligned}$$

3. Using the function written in Problem I.1a that simulates the observed data (i.e., no intervention), confirm the identification result that states

$$E_{P_{U,X}}(Y_{G_1^*}) = E_{P_{U,X}}[E_{G_1^*}\{E_{P_{U,X}}(Y \mid A, W)\}] \ .$$

Here, note that the inner expectation is taken with respect to the conditional distribution of Y given A and W implied by $P_{U,X}$, the middle expectation is taken with respect to the chosen conditional distribution for A given W, G_1^*, and the outer expectation is taken with respect to the marginal distribution of W.

4. Now consider the above SCM, but changing the structural equation f_A to the following:

$$f_A(W_1, W_2, U_A) = \{1 - I(W1 = 0)I(W_2 = 0)\}I(\text{expit}(W_1 - W_2 + U_A) > 0.5) \ .$$

Consider estimating the average causal effect as described in Problem I under this SCM. Can the observed data parameter

$$\Psi(P_{U,X}) = E_{P_{U,X}}[E_{P_{U,X}}\{Y \mid A = 1, W_1, W_2\} - E_{P_{U,X}}\{Y \mid A = 0, W_1, W_2\}] \ .$$

be used to describe the average causal effect? Why or why not?

5. Instead of estimating the causal effect, consider comparing the average counterfactual outcome under two interventions G_2^* and G_3^* defined as follows:

$$\begin{aligned} G_2^*(A^* = 1 \mid W_1 = 0, W_2 = 0) &= 0 \\ G_2^*(A^* = 1 \mid W_1 = 0, W_2 = 1) &= 1 \\ G_2^*(A^* = 1 \mid W_1 = 1, W_2 = 0) &= 1 \\ G_2^*(A^* = 1 \mid W_1 = 1, W_2 = 1) &= 1 \end{aligned}$$

and

$$\begin{aligned} G_3^*(A^* = 1 \mid W_1 = 0, W_2 = 0) &= 0 \\ G_3^*(A^* = 1 \mid W_1 = 0, W_2 = 1) &= 0 \\ G_3^*(A^* = 1 \mid W_1 = 1, W_2 = 0) &= 0 \\ G_3^*(A^* = 1 \mid W_1 = 1, W_2 = 1) &= 0 \end{aligned}$$

Numerically compute the value of this causal parameter

$$E_{P_{U,X}}(Y_{G_2^*}) - E_{P_{U,X}}(Y_{G_3^*}) \ .$$

6. What, if any, observed data quantity can be used to estimate this parameter? Confirm numerically that the observed data quantity you identify does indeed correspond with the counterfactual outcome.
7. Write a short summary of the implications of 4-7 on choosing causal parameters in settings with the potential for positivity violations.