Antonella Basso

PHP 2610 Problem Set 5

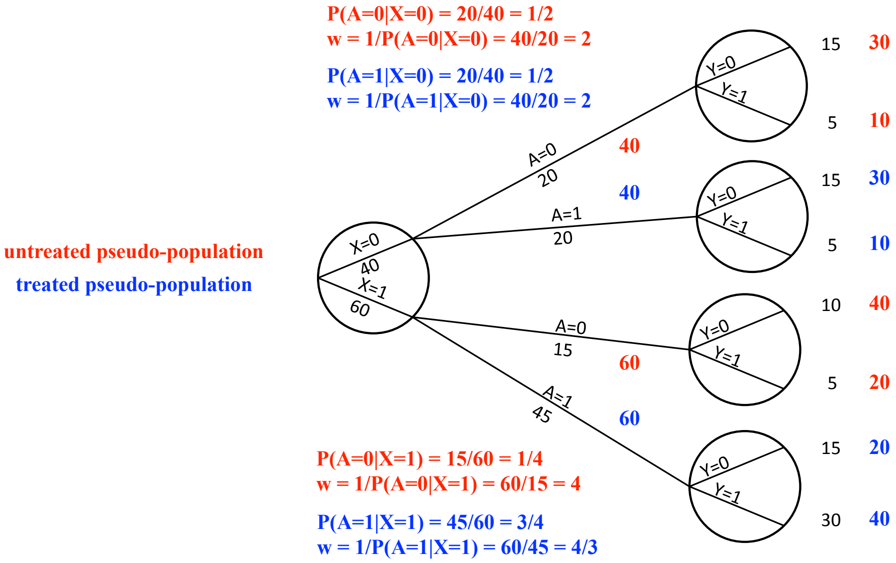
Due: December 2 by 11:59pm

Instructions: Please upload your answer to the Canvas course page as a pdf file. You can submit your answers in a separate pdf file (please be sure to properly mark the question number to your responses), or you can work on this pdf file, scan it, and upload it to the Canvas course page.

Late or missed assignments: Problem sets and the final report must be turned in online at or before the posted due date. Every one day (24 hours) of delay will result in a ten point (out of 100) downgrade.

Question 1 (30 points)

Consider the following treatment assignments under the ignorability condition: $Y^a \perp\!\!\!\perp A|X$. Derive $E(Y^{a=1}) - E(Y^{a=0})$ both using the (1) g-formula and the (2) IPW estimator. (3) Compare the two estimates.



P(A=0|X=0) = 20/40 = 1/2
w = 1/P(A=0|X=0) = 40/20 = 2

P(A=1|X=0) = 20/40 = 1/2
w = 1/P(A=1|X=0) = 40/20 = 2

untreated pseudo-population

treated pseudo-population

P(A=0|X=1) = 15/60 = 1/4
w = 1/P(A=0|X=1) = 60/15 = 4

P(A=1|X=1) = 45/60 = 3/4
w = 1/P(A=1|X=1) = 60/45 = 4/3

Answer:

(1) g-formula:

a) E[Y^(a=1)] = sumX{ E[Y|A=1, X=x] • P(X=x) }
= [E[Y|A=1, X=1] • P(X=1)] + [E[Y|A=1, X=0] • P(X=0)]
= [30/(30+15) • 60/(60+40)] + [5/(5+15) • (40/60+40)]
= [2/3 • 3/5] + [1/4 • 2/5]
= 2/5 + 1/10
= 0.5

b) E[Y^(a=0)] = sumX{ E[Y|A=1, X=x] • P(X=x) }
= [E[Y|A=0, X=1] • P(X=1)] + [E[Y|A=0, X=0] • P(X=0)]
= [5/(5+10) • 60/(60+40)] + [5/(5+15) • (40/60+40)]
= [1/3 • 3/5] + [1/4 • 2/5]
= 1/5 + 1/10
= 0.3

c) E[Y^(a=1)] - E[Y^(a=0)] = 0.5 - 0.3 = 0.2

(2) IPW estimator:

a) E[Y^(a=1)] = E[Y|A=1, X]
= (40 + 10)/(40 + 20 + 10 + 30)
= 50/100
= 0.5

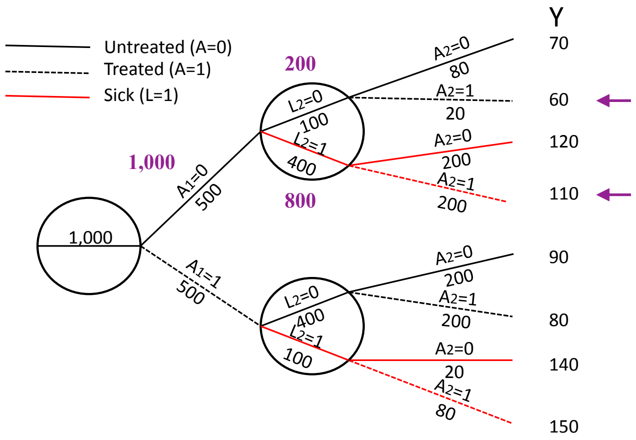
b) E[Y^(a=0)] = E[Y|A=0, X]
= (20 + 10)/(20 + 40 + 10 + 30)
= 30/100
= 0.3

c) E[Y^(a=1)] - E[Y^(a=0)] = 0.5 - 0.3 = 0.2

(3) Both approaches yield the same treatment causal effect estimate of 0.2.

Question 2 (40 points)

Q2.1. (10 points) Consider the following treatment assignments and assume the sequential ignorability. Calculate $E(Y^{a_1=0,a_2=1})$ using the inverse probability weighting.



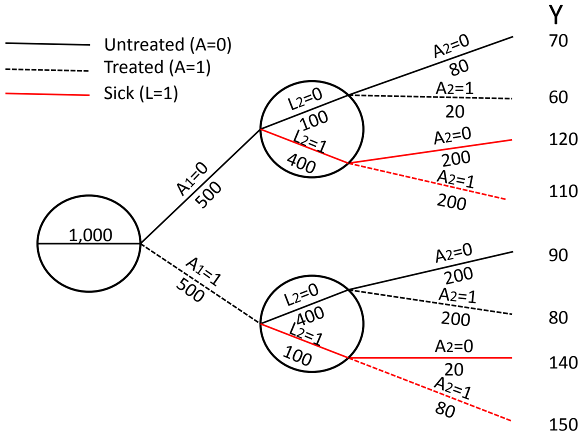
Answer:

E[Y^(a1=0, a2=1)] = E[Y|A1=0, A2=1, L2]

= [(60 • 200) + (110 • 800)]/1,000

= 100

Q.2.2. (30 points) Consider the following treatment assignment. Calculate $E(Y^{1,1-L_2})$ using (1) the g-formula and (2) the inverse probability weighted estimator under the dynamic treatment regimes version of the sequential ignorability assumption. (3) Compare these two estimates.



Answer:

(1) g-fromula:

E[Y^(1,1-L2)] = ∑ L2{ E[Y|A1=1, A2=1-L2, L2=l2] • P(L2=l2|A1=1) }
= (80 • 0.8) + (140 • 0.2)
= 92

(2) IPW estimator:

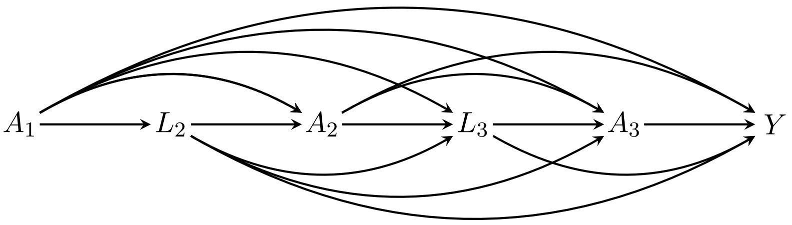
Let wi = 1/[P(Ai2=1-L2|Li2=l2, Ai1=a1) • P(Ai1=1)] for i=1,2,…,n.

E[Y^(1,1-L2)] = ∑ i { ∑ L2 { I(Ai1=1, Ai2=1-L2, Li2=l2) } • Yi • wi } /
∑ i { ∑ L2 { I(Ai1=1, Ai2=1-L2, Li2=l2) } • wi }
= [(200 • 80 • 4) + (20 • 140 • 10)]/[(200 • 4) + (20 • 10)]
= 92

(3) Both approaches yield the same treatment causal effect estimate of 0.2.

Question 3 (30 points)

Consider the following DAG with binary A_j and L_j for $j = 1, 2, 3$:



Suppose that the following structural nested mean models are correct:

$$E(Y^{a_1,a_2=0,a_3=0} - Y^{a_1=0,a_2=0,a_3=0}) = \alpha_1 a_1$$
$$E(Y^{a_1,a_2,a_3=0} - Y^{a_1,a_2=0,a_3=0} \mid L_2^{a_1} = l_2, A_1 = a_1) = a_2(\beta_1 + \beta_2 l_2 + \beta_3 a_1 + \beta_4 a_1 l_2)$$
$$E(Y^{a_1,a_2,a_3} - Y^{a_1,a_2,a_3=0} \mid (L_2^{a_1}, L_3^{a_2}) = (l_2, l_3), (A_1, A_2) = (a_1, a_2)) = a_3(\gamma_1 + \gamma_2 l_2 + \gamma_3 l_3 + \gamma_4 a_1 a_2 + \gamma_5 a_2 l_3)$$

Q3.1. (5 points) According to the model above, what would be the effect of A_2 when a_3 is set to zero for those with $A_1 = 1$ and $L_2^{a_1=1} = 1$?

Answer: **ß1 + ß2 + ß3 + ß4**

Q3.2. (5 points) Suppose that we want to test if the effect of A_2 when a_3 is set to zero for those with $a_1 = 0$ would be the same between those with $L_2^{a_1=0} = 1$ and $L_2^{a_1=0} = 0$. Then which null hypothesis should be considered?

a. $H_0: \ \alpha_1 = 0$
b. $H_0: \ \beta_1 = 0$
[✓ c.] $H_0: \ \beta_2 = 0$
d. $H_0: \ \beta_3 = 0$
e. $H_0: \ \gamma_4 = 0$

Answer:

Q3.3. (8 points) What does $\gamma_3 = 0$ imply (choose all that apply)?

a. the effect of A_3 when $(a_1, a_2) = (1, 0)$ would be the same between those with $(l_2, l_3) = (1, 1)$ and those with $(l_2, l_3) = (0, 1)$

[✓ b.] the effect of A_3 when $(a_1, a_2) = (1, 0)$ would be the same between those with $(l_2, l_3) = (1, 1)$ and those with $(l_2, l_3) = (1, 0)$

c. the effect of A_3 when $(a_1, a_2) = (0, 0)$ would be the same between those with $(l_2, l_3) = (1, 1)$ and those with $(l_2, l_3) = (0, 1)$

[✓ d.] the effect of A_3 when $(a_1, a_2) = (0, 0)$ would be the same between those with $(l_2, l_3) = (1, 1)$ and those with $(l_2, l_3) = (1, 0)$

e. the effect of A_3 when $(a_1, a_2) = (1, 1)$ would be the same between those with $(l_2, l_3) = (1, 1)$ and those with $(l_2, l_3) = (1, 0)$

Answer:

Q 3.4. (6 points) Suppose that now we want to estimate $E(Y^{a_1=0,a_2=1,a_3=1})$ using the g-formula. Which models are not necessarily to be fitted (choose all that apply)?

a. $E(Y \mid L_2 = l_2, L_3 = l_3, A_1 = 0, A_2 = 1)$ for $l_2, l_3 \in \{0, 1\}$

b. $Pr(L_3 = l_3 \mid L_2 = l_2, A_1 = 0, A_2 = 1)$ for $l_2, l_3 \in \{0, 1\}$
c. $Pr(L_2 = l_2 \mid A_1 = 0)$ for $l_2 \in \{0, 1\}$
[✓ d.] $Pr(A_3 = 1 \mid L_2 = l_2, L_3 = l_3, A_1 = 0, A_2 = 2, A_1 = 0)$ for $l_2, l_3 \in \{0, 1\}$
[✓ e.] $Pr(A_2 = 1 \mid L_2 = l_2, A_1 = 0)$ for $l_2 \in \{0, 1\}$

Answer:

Q 3.5. (6 points) Suppose that now we want to estimate $E(Y^{a_1=0,a_2=l_2,a_3=1-l_3})$ using the marginal structural model approach with stabilized weights. Which models are not necessarily to be fitted (choose all that apply)?

a. $E(Y \mid L_2 = l_2, L_3 = l_3, A_1 = 0, A_2 = 1)$ for $l_2, l_3 \in \{0, 1\}$
b. $Pr(A_2 = l_2 \mid A_1 = 0, L_2 = l_2)$ for $l_2 \in \{0, 1\}$
[✓ c.] $Pr(L_2 = l_2 \mid A_1 = 0)$ for $l_2 \in \{0, 1\}$
d. $Pr(A_2 = l_2 \mid A_1 = 0)$ for $l_2 \in \{0, 1\}$
e. $Pr(A_3 = 1 - l_3 \mid A_1 = 0, A_2 = l_2, L_2 = l_2, L_3 = l_3)$ for $l_2, l_3 \in \{0, 1\}$

Answer: