# Inverse Probability Weighting: Intuition in Two Periods

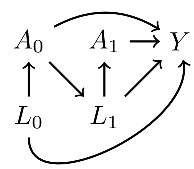
Info 6751: Causal Inference in Observational Settings

SOLUTIONS to Class Exercise. 18 October 2022.

## The causal problem and data

You teach a class of $n = 16$ elementary students. At each time point $t$, the student can either be below grade level ($L_t = 1$) or not below grade level ($L_t = 0$). You can assign extra support $A_t = 1$ (e.g., reading practice with a parent volunteer) to some but not all of the students; the untreated students have $A_t = 0$.

We will study this process over two time periods: $t = 0$ and $t = 1$. The outcome would be something defined after these periods, but for our purposes there will be no outcome. We will focus on how you would weight, which would apply to any outcome you might care about.




Here is how the causal process unfolds:

- At time 0, you see that 50% of the class is below grade level: $P(L_0 = 1) = 0.5$
- At every time period $t$, treatment is assigned so that
  - those below grade level $L_t = 1$ receive extra support $A_{t+1} = 1$ with probability 1
  - those not below grade level $L_t = 0$ receive extra support $A_{t+1} = 1$ with probability 0.5
- At every time period $t$, treatment affects the next confounder
  - if you receive support $A_t = 1$, then you will not fall behind $P(L_{t+1} = 1) = 0$
  - if you do not receive support $A_t = 0$, then you have a 50% chance of falling behind: $P(L_{t+1} = 1) = 0.5$

The table below shows data from this causal process.

% latex table generated in R 4.2.0 by xtable 1.8-4 package % Mon Oct 17 23:15:30 2022

| ID | L0 | A0 | L1 | A1 | w0 | w1 | w |
|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 0 | 1 | | | |
| 2 | 1 | 1 | 0 | 0 | | | |
| 3 | 1 | 1 | 0 | 1 | | | |
| 4 | 1 | 1 | 0 | 0 | | | |
| 5 | 1 | 1 | 0 | 1 | | | |
| 6 | 1 | 1 | 0 | 0 | | | |
| 7 | 1 | 1 | 0 | 1 | | | |
| 8 | 1 | 1 | 0 | 0 | | | |
| 9 | 0 | 1 | 0 | 1 | | | |
| 10 | 0 | 1 | 0 | 0 | | | |
| 11 | 0 | 1 | 0 | 1 | | | |
| 12 | 0 | 1 | 0 | 0 | | | |
| 13 | 0 | 0 | 1 | 1 | | | |
| 14 | 0 | 0 | 0 | 1 | | | |
| 15 | 0 | 0 | 1 | 1 | | | |
| 16 | 0 | 0 | 0 | 0 | | | |

Table 1: Weighting Two Time Periods: A Table to Fill In

## 1. Estimating $E(Y \mid do(A_0 = 1))$

Suppose we want to estimate the expected outcome under an intervention to set $A_0 = 1$. We plan to reweight the 12 units observed with $A_0 = 1$. What weight should we put on each of these 6 units?

Note that $L_0$ is a sufficient conditioning set.

**Solution:** See `w0` in "Table of Weights in Two Time Periods.' '

## 2. Estimating $E(Y \mid do(A_1 = 1))$

Suppose we want to estimate the expected outcome under an intervention to set $A_1 = 1$. We plan to reweight the 9 units observed with $A_1 = 1$. What weight should we put on each of these 6 units?

Note that $L_1$ is a sufficient conditioning set.

**Solution:** See `w1` in "Table of Weights in Two Time Periods.' '

## 3. Estimating $E(Y \mid do(A_0 = 1, A_1 = 1))$

Suppose we want to estimate the expected outcome under an intervention to set $A_0 = 1$ and $A_1 = 1$. We plan to reweight the 6 units observed in this condition. What weight should we put on each of these 6 units?

**Solution:** See `w` in "Table of Weights in Two Time Periods.' '

% latex table generated in R 4.2.0 by xtable 1.8-4 package % Mon Oct 17 23:15:31 2022

## 4. Check: Is it right?

Using your weight from 1.3,

- take a weighted average of $L_0$. Does your result match the factual distribution of $L_0$?
- take a weighted average of $L_1$. Does your result match the factual distribution of $L_1$?

Think about why you would get the answers above.

**Solution.** In the population, the mean of $L_0$ is 0.5. In the pseudo-population, the mean of $L_0$ is 0.5. These match; an intervention to set $A_0$ to the value 1 would leave $L_0$ unchanged.

In the population, the mean of $L_1$ is 0.125. In the pseudo-population, the mean of $L_1$ is 0. These do not match; an intervention to set $A_0 = 1$ causes $L_1 = 0$ for everyone.

| ID | L0 | A0 | L1 | A1 | w0 | w1 | w |
|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 0 | 1 | 1 | 2 | 2 |
| 2 | 1 | 1 | 0 | 0 | 1 | | |
| 3 | 1 | 1 | 0 | 1 | 1 | 2 | 2 |
| 4 | 1 | 1 | 0 | 0 | 1 | | |
| 5 | 1 | 1 | 0 | 1 | 1 | 2 | 2 |
| 6 | 1 | 1 | 0 | 0 | 1 | | |
| 7 | 1 | 1 | 0 | 1 | 1 | 2 | 2 |
| 8 | 1 | 1 | 0 | 0 | 1 | | |
| 9 | 0 | 1 | 0 | 1 | 2 | 2 | 4 |
| 10 | 0 | 1 | 0 | 0 | 2 | | |
| 11 | 0 | 1 | 0 | 1 | 2 | 2 | 4 |
| 12 | 0 | 1 | 0 | 0 | 2 | | |
| 13 | 0 | 0 | 1 | 1 | | 1 | |
| 14 | 0 | 0 | 0 | 1 | | 2 | |
| 15 | 0 | 0 | 1 | 1 | | 1 | |
| 16 | 0 | 0 | 0 | 0 | | | |

Table 2: Table of Weights in Two Time Periods: Solution of Weights for Estimating Outcomes Under Treatment. w0 is for estimmating E(Y | do(A0 = 1)), w1 is for estimating E(Y | do(A1 = 1)), and w is for estimating E(Y | do(A0 = 1, A1 = 1))