Problem Set 4 (Part 1)

Q1 – Lack of Overlap?

Recall from lecture that certain causal assumptions, including ignorability (also called unconfoundedness) and common support (also called overlap or positivity), need to be met in order to identify the average treatment effect (ATE). Under these assumptions, we saw that we could use two different approaches to do this identification: covariate adjustment using the so-called "adjustment formula" or inverse probability weighting (IPW). In this problem, you will explore and gain intuition for why one of these causal assumptions, particularly overlap/positivity, is important for identifying and estimating the ATE.

To begin, we will define some notation. Let Y denote an outcome of interest and $T \in \{0,1\}$ denote a binary treatment. Furthermore, let $Y(t)$ denote the potential outcome of an individual under treatment, $T = t$. Finally, let X denote all other covariates. Recall the adjustment formula used to estimate the ATE,

$$\mathbb{E}_{x\sim p(x)}[Y(1) - Y(0)] = \mathbb{E}_{x\sim p(x)}[\mathbb{E}[Y|T=1, X=x] - \mathbb{E}[Y|T=0, X=x]] \tag{1}$$

Recall as well the formula for IPW,

$$\mathbb{E}_{x\sim p(x)}[Y(1) - Y(0)] = \mathbb{E}_{x\sim p(x|T=1)}\left[\frac{p(T=1)}{p(T=1|X)}Y\right] - \mathbb{E}_{x\sim p(x|T=0)}\left[\frac{p(T=0)}{p(T=0|X)}Y\right] \tag{2}$$

$$\widehat{ATE}_{\text{IPW}} = \frac{1}{n}\sum_{i=1}^{n}\frac{\mathbf{1}(t_i=1)}{\hat{p}(t_i=1|x_i)}y_i - \frac{1}{n}\sum_{i=1}^{n}\frac{\mathbf{1}(t_i=0)}{\hat{p}(t_i=0|x_i)}y_i, \tag{3}$$

where n is the total number of individuals in the cohort.

For this problem, we will work with a toy example. Suppose that $X \in \{-1, 0, 1\}$ and ignorability holds. In Table 1, we show the expected outcomes, $\mathbb{E}[Y|X,T]$, estimated using infinite data. However, there are some regions where $p(X,T) = 0$, e.g. $X = -1, T = 0$. Thus, even with infinite data, the outcomes are not estimable, and we put a "?" in those entries. Note that this implies that there is lack of overlap when $X = -1$ and $X = 0$.

	$T = 0$	$T = 1$
$X = -1$	?	0.1
$X = 0$	0.6	?
$X = 1$	0.5	0.4

Table 1: Entries represent $\mathbb{E}[Y|X,T]$ estimated from infinite number of samples

[4 points] 2.1 Assume that the true potential outcomes are linear as a function of X, as below,

$$\mathbb{E}[Y(1)|X] = \alpha_1 + \beta_1 \cdot X$$
$$\mathbb{E}[Y(0)|X] = \alpha_0 + \beta_0 \cdot X$$

Under this stuctural assumption, fill in Table 1 with the potential outcomes. What is the ATE?

[4 points] 2.2 Now assume that the true potential outcomes differ up to a constant effect, δ,

$$\mathbb{E}[Y(1)|X] = f(X) + \delta$$
$$\mathbb{E}[Y(0)|X] = f(X),$$

where f is a potentially non-linear function. As in the previous part, fill in Table 1 with the potential outcomes under this assumption. What is the ATE?

[3 points] 2.3 Graph the potential outcomes as a function of X under both structural assumptions from above. There should be four lines on your plot.

[3 points] 2.4 Given your answers to 2.1 and 2.2 as well as your graph, is the ATE identifiable in this setting? Why or why not? What does this imply in general about identifying the ATE when there is lack of overlap in some covariate regions?

[2 points] 2.5 Now, consider the IPW estimator given in Eq. (3). Suppose you evaluate the IPW estimator using an infinite number of samples from the data distribution specified in Table 1, with the additional specification of $p(T|X)$ as $p(T = 1|X = -1) = 1$, $p(T = 0|X = 0) = 1$, and $p(T = 1|X = 1) = 0.5$. Is the IPW estimator well-defined, i.e. will you get a number as an estimate for the ATE? If so, give one sentence on the correctness of the resulting estimate in lieu of your identifiability result from the previous part.