Exercise Sheet: The Fundamental Problem of Causal Inference

This exercise illustrates why we cannot observe Individual Treatment Effects (ITE) and how different data-gathering strategies (Observational vs. Randomized) affect our (naive) estimate of the Average Causal Effect (ATE).

1 Full observed

In this hypothetical scenario, we suppose to observe the **Potential Outcomes** for 10 patients. $Y_i(1)$ is the recovery time (days) if the patient takes the medication, and $Y_i(0)$ is the recovery time if they do not.

Patient (i)	$Y_i(0)$	$Y_i(1)$	ITE ($Y_i(1) - Y_i(0)$)
1	14	10	
2	15	11	
3	13	9	
4	12	8	
5	11	7	
6	8	4	
7	7	3	
8	9	5	
9	6	2	
10	10	6	

- Calculate the ITE for each patient, then estimate the population **ATE**, ($\mathbb{E}[Y(1) - Y(0)]$).

2 Observational data (Confounding)

In reality, treatment assignment is often biased. Assume, in this example, that the doctor knows the severity of each patient, and takes his decision according to this condition (sicker patients - Severe - are more likely to receive the treatment). We have the following observations:

Patient (i)	Severity (X)	Treatment (T)	Observed Outcome (Y^{obs})
1	Severe	1	10
2	Severe	1	11
3	Severe	1	9
4	Severe	1	8
5	Severe	0	11
6	Mild	1	4
7	Mild	0	7
8	Mild	0	9
9	Mild	0	6
10	Mild	0	10

- Calculate the “Naive” ATE: $\hat{\tau}_{naive} = \mathbb{E}[Y|T = 1] - \mathbb{E}[Y|T = 0]$.
- Compare this to the truth. Why does the medicine look ineffective or harmful here?

3 Randomized data (RCT)

Imagine we randomly assign T regardless of Severity X. We would observe:

Group	Patients	Treatment (T)	Outcomes Y^{obs}
Treated	{1, 3, 5, 7, 9}	1	{10, 9, 7, 3, 2}
Control	{2, 4, 6, 8, 10}	0	{15, 12, 8, 9, 10}

- Calculate the Randomized ATE. Does randomization solve the FPCI?

4 Inverse Probability Weighting (IPW)

Using the biased data from **2**, assume we know the probability of receiving the treatment given the severity (the propensity scores):
$\mathbb{P}(T = 1|X = \text{Severe}) = e(\text{Severe}) = 0.8$, and $e(\text{Mild}) = 0.2$.

- Calculate the weight $w_i = \frac{T_i}{e_i} + \frac{1-T_i}{1-e_i}$ for Patient 1 (Severe, Treated) and Patient 6 (Mild, Treated).
- Why does Patient 6 receive a much higher weight than the treated Severe patients?

5 Lab Exercise: Causal Inference from scratch

Objective: Implement and compare ATE estimators (S-Learner, T-Learner, IPW, and AIPW) from scratch using Python's numpy and scikit-learn.

5.1 Data Generation

We consider N subjects. For each subject, we observe the following covariates:

- Age: $A \sim \mathcal{N}(50, 10)$
- Severity: $S \sim Bernoulli(0.4)$

We suppose that the true propensity score (determining treatment assignment, T) is given by:

$$e(A, S) = \frac{1}{1 + \exp(-(0.05(A - 50) + 1.5S - 0.5))}$$

Generate 2000 patients and their outcomes using the following code:

```python
import numpy as np
import pandas as pd

def generate_data(n):
    np.random.seed(42)
    age = np.random.normal(50, 10, n)
    severity = np.random.binomial(1, 0.4, n)
    # Propensity Score
    ps = 1 / (1 + np.exp(-(0.05 * (age - 50) + 1.5 * severity - 0.5)))
    treatment = np.random.binomial(1, ps)
    # Potential Outcomes
    y0 = 20 + 0.5 * age + 10 * severity + np.random.normal(0, 2, n)
    y1 = y0 - 5 + np.random.normal(0, 1, n)
    y_obs = np.where(treatment == 1, y1, y0)
    return pd.DataFrame({'A': age, 'S': severity, 'T': treatment, 'Y': y_obs})
```

What is the true average treatment effect?
The true underlying graph is as follows:

- $\{A, S\} \to Y$
- $\{A, S\} \to T$
- $T \to Y$

Is the treatment effect identifiable from the data? Explain why.

5.2 Implementation

By simply using linear regression and logistic regression (both available in `sklearn.linear_model`), implement and test over the synthetic dataset:

- The S-Learner
- The T-learner
- The IPW estimator
- The AIPW estimator

Which one is closer to the true ATE?
In your opinion, why do all methods work correctly in this example?