Homework 2

Introduction to Causal Inference 097400

Spring 2025

Submission Date: June 10, 2025

Goal:

This homework is designed to help you practice the tools we have learned in Tutorials 3-5, propensity and potential outcome estimation. You will build S-learners, T-Learners and perform matching, IPW on data to estimate ATE or ATT. You will also measure uncertainty.

Data: hw2_data.csv

The data has been simulated and has the following columns:

Features – Columns that begin with X, (X0…X29). These may or may not be confounders. For this homework, you are not tasked with modelling the causal problem (identification). You may assume that all the features are confounders.

All the features are continuous and there are no NaN values.

Treatment/intervention – Binary (0/1) in a column named *T*.

Outcome – Continuous outcome in a column named *Y*.

Instructions:

You have received a python notebook, hw2.ipynb, that has been *partially* filled out. There are three components to the notebook: following instructions, coding and answering questions.

1. **Following instructions** – the notebook has comments that instruct you to show various things or place holders for running functions
   For example, ATE estimation

   ```
   Provide the point-estimation of the ate on the validation set

   1  calculate_s_learner_ate()
   ```

   Or a placeholder to show the overlap graph that appears in Tutorials 3-5 and allows us to understand if there was overlap.

   ```
   check_overlap()
   ```

   You need to follow the instructions/run the placeholders.
2. **Coding** – The notebook has functions that have not been implemented. These functions currently have “pass” as their code.

For example, the overlap graph

```python
def check_overlap(propensity_model, X_train, T_train):
    """
    Show histograms of the propensity on the T=0 and the T=1 group like in the tutorial
    Use this to ascertain if there is a violation of overlap.

    You must show the graph.
    :param propensity_model:
    :param X_train:
    :param T_train:
    :return:
    """
    pass
```

Or building confidence intervals for S-learners

```python
def s_learner_confidence_intervals(s_learner_model, X_train, y_train, X_val, y_val):
    """
    Receives an s-learner model and returns 95% confidence intervals.
    You do not need to use every parameter passed to this method and you can add additional parameters.

    You must either print, visualize the confidence intervals or return them as a list.
    :param s_learner_model:
    :param X_train:
    :param y_train:
    :param X_val:
    :param y_val:
    :return:
    """
    pass
```

You need to implement these functions. You may change the function's inputs and should return what is necessary per the instructions.

3. **Answering Questions** – Throughout the notebook, there are questions that check your understanding of the material and challenge you to think critically about your findings.
   For example, understanding what the evaluation methods of the propensity score mean

> Q1. Explain what the Brier score represents? What is the best possible Brier score? What is the Brier score if the model always guesses 0.5?
> Q2. Explain what the calibration plot shows? What is the best possible calibration plot?

Or asking you to explain if you think there is overlap or not

> Q5. Explain why there is or is not overlap in this sample?
>
> For this exercise do not trim the data, even if you deemed it necessary.

You need to answer these questions either in the notebook itself or in a separate pdf file. (See later).

Important Comments:

1. *You may add to the code, but you cannot remove code.* For example, if you want to show confident intervals in a graph like we did in the tutorial, you may (but you are not required). However, you cannot erase the placeholder functions or not implement a function with "pass" in it.
2. *You may change the input of functions but you must return what is stated in the comment for the function. You may find that some functions receive input that you do not use. That is fine.*
3. *Some of the questions are subjective and some are objective*. The objective questions have correct answers. The subjective questions may depend on your modelling or may be a matter of opinion. They will be graded on your reasoning depending on your results.

Submission requirements:

You must submit a completed version of the hw2.ipynb notebook. This must include "following the instructions" and "coding", as described above. *You should submit your notebook and either a pdf or html file.*

You may answer the questions within the notebook or *attach a separate pdf file* that includes the questions followed by the answers. (Q1, A1, Q2, A2, etc.)

If your answer depends on a visual or other information in the notebook, please include the visual or the information in the pdf. (For example, include a screenshot of the overlap graph in your answer to Q5.)

Do not submit a zip file

Objective vs. Subjective Questions:

Q1. Explain what the Brier score represents? What is the best possible Brier score? What is the Brier score if the model always guesses 0.5?
Q2. Explain what the calibration plot shows? What is the best possible calibration plot?
Q3. The sklearn implementation of calibration_plot (see Tutorial) has a parameter "n_bins". Explain this parameter. What happens if you use too many bins (for example 100 on this dataset) or too little bins (for example 1)?
Q4. Choose a propensity model and explain why you chose that model?
Q5. Explain why there is or is not overlap in this sample?
Q6. Using a Lasso model, check the ATE over different regularization parameters. What happens to the ATE with the S-Learner? Why?
Q7. Using a Lasso model, check the ATE over different regularization parameters. Is this similar or different to what we saw earlier with the S-Learner as we increased regularization? If it was different, explain what causes this.
Q8. Show the Balancing Table on features X16, X17, X18, X19, X20 on the training data. Is the data balanced over these features? Explain. Use the convention of SMD below 0.1 to judge.
Q9. Explain your matching algorithm. Are there advantages and/or disadvantages to your matching algorithm?
Q10. Did the balance improve after matching? Did the overlap improve after matching? Explain.
Q11. Let's imagine that each row in the data represents an employee. T represents if the employee has undertaken a training program. And Y is the amount of money (in thousands of dollars) that each person has cost or earned the company after going through the program. Would you recommend the program?
Explain your decision by looking at the confidence intervals that you received from S-Learner, T-Learner, IPW and matching.
Note that the ATT from matching is not the same as ATE for the others.