6.871: PSet 4 (Full)

Collaboration Policy

- Your solution document must be created completely on your own.
- You are free to collaborate via discussions with other students. Please acknowledge all collaborators on your final write-up.

Submission: Gradescope & Dates

As with prior homeworks, please submit your final write-up via Gradescope by 5:59 p.m. on April 14th, 2020.

Causality: Translating to Math

Overview

In this problem, we will present you with several free text scenarios. For each scenario, you must answer first whether or not causal inference is required in this scenario, and, if so, you must identify the relevant covariates (X), treatments (T), outcomes (Y), and any hidden confounders (H) that pose particular concern in this setting.

Learning Goals: We hope that this problem will help you develop the ability to recognize causal problems in the wild, to be able to translate them into the mathematical framework discussed in lectures, and to think critically about any missing confounders.

Problem

0. **Example:** You note that ice cream sales are strongly correlated to drowning rates in the United States, and decide to test if this is a causal phenomenon based on a large observational dataset at a monthly, county-specific level consisting of quantities of flavors of ice cream ordered, the

quantity of hot-dogs eaten, the average socioeconomic status of residents in that county, and number of drownings.

Solution: Yes, this does require causal inference, as you wish to understand the causal link between ice cream sales and drowning rates. T would be whether ice cream of various flavors were eaten. Y would be the number of drownings observed. X would be the socioeconomic data about the county and the # of hot-dogs eaten. A critical *hidden* confounder that might damage this analysis is the average temperature that month, which likely both inspires ice cream consumption and swimming.

1. You have collected physiological timeseries data (in particular, EEG traces, heart rate, SpO2, and blood pressure) from patients during a set of surgical procedures which are controlled by the application of anesthetic drug D at various dose levels continuously throughout the procedure. While most patients survive the procedure, a small minority do not. You want to design a model to automatically control the drug dosage level over time to maximize chance of patient survival.

2. You work for a hospital and have just been assigned to help a new payer (e.g., insurance agent) satisfy an odd request. They like to receive advanced estimates of billing codes they will eventually be charged for their patients, and have requested that you provide them with an estimate of each patient's final billing code after only their first 24 hours. Contractually, this payer is prohibited from using these estimates to affect whether they accept or dispute claims, but instead uses them to more proactively assess their revenue streams. To accomodate this request, you plan to build a model based on retrospective data to take the first 24 hours of a patient's stay and predict the final ICD10 codes they received at discharge time.

3. You work for a hospital which, like many hospitals, receives penalties from major payers if patients are readmitted within 30 days of discharge. To help prevent this, you decide to train a model based on past patient data, ingesting all EHR data up to discharge and predicting whether or not the patient will be readmitted within 30 days. Ultimately, you hope to use this model to audit and inform discharge decisions.

Computing ATE and CATE

Overview

Adapted from problem by Willie Boag and Irene Chen, in turn adapted from Uri Shalit and Rom Gutman

In this problem, you will work with a simulated experiment and dataset to formulate an explicit causal graph and work through calculating CATE and ATE by hand.

Learning Goals: In this problem, we hope to give you a hands-on familiarity with ATE and CATE with real data in a simulated context.

Problem

The data in Table 1 comes from an experiment run from 2004-2008 at Sunnydale Hospital. In the data, you have binary indicators for: prior education (Z), whether the surgeon had ≥ 100 successful surgeries in 2004 (X), whether the surgeon enrolled in the fellowship from 2005-2006 (T), whether the surgeon had ≥ 150 successful surgeries in 2007 (Y), and whether the surgeon cumulatively had ≥ 500 successful surgeries in their lifetime by 2008 (W).

We know the following:

(i) The number of successful surgeries in 2004 depends solely on the prior education.

(ii) Whether a doctor is selected to the surgery fellowship program depends on their prior education and number of successful surgeries in 2004.

(iii) The number of successful surgeries in 2007 depends on the fellowship, prior education, and number of successful surgeries in 2004.

(iv) The cumulative number of successful surgeries by 2008 is directly based on how many successful surgeries a doctor performs in 2004 and 2007.

Your task is as follows:

1. Draw the causal graph that describes the above experiment.

Z	X	T	Y	W
0	1	1	1	0
0	0	1	1	0
0	0	1	1	0
1	1	1	1	1
1	0	1	1	1
1	0	1	1	1
1	1	1	0	0
1	0	1	0	0
0	0	1	0	0
1	1	0	1	1
1	0	0	1	0
1	0	0	1	0
1	1	0	0	1
1	1	0	0	1
0	1	0	0	0
1	0	0	0	0
0	0	0	0	0
0	0	0	0	0

Table 1: Sample data for problem 2.

2. Calculate the Average Treatment Effect (ATE) of the fellowship (T) on 2007 success (Y). Use covariate adjustment and empirically estimate the probabilities/expectations from the observed data. Recall that

$$\text{ATE} = \mathbb{E}\left[Y_1 - Y_0\right]$$

where, for this causal graph,

$$\mathbb{E}\left[Y_t\right] = \sum_{z\in\{0,1\}}\sum_{x\in\{0,1\}} P(Z=z)P(X=x|Z=z)\mathbb{E}\left[Y|X=x,Z=z,T=t\right]$$

3. Calculate the Conditional Average Treatment Effect (CATE) of the fellowship (T) on 2007 success (Y) for patients without prior education ($Z=0$). Recall that

$$\text{CATE} = \mathbb{E}\left[Y_1 - Y_0|Z=0\right] = \mathbb{E}\left[Y_1|Z=0\right] - \mathbb{E}\left[Y_0|Z=0\right],$$

where, for this causal graph,

$$\mathbb{E}\left[Y_t|Z=0\right] = \sum_{x\in\{0,1\}} P(X=x|Z=0)\mathbb{E}\left[Y|X=x,Z=0,T=t\right]$$

Reinforcement Learning & Causal Inference

Overview

In this problem, we'll explore reinforcement learning from two lenses. First, on a very simple problem, you'll work through the Q-Learning algorithm to design an optimal treatment policy over a simplified patient state space. Then, we'll ask you think about how you would solve this from a causal inference perspective, recognizing that, in fact, the broader framework of Q-Learning has a natural interpretation as iterative causal inference over increasingly long time-horizons.

Learning Goals: To gain a working familiarity with Q-Learning & Reinforcement Learning in general, and to understand the connection between these techniques and causal inference.

Problem

Suppose we have a treatment decision over possible treatments $T \in \{A, B\}$. The patient's state X, encoded by their "severity level," is encoded as $X \in \{1, 2, 3, M\}$, where M indicates mortality. For example a patient's treatment trajectory may look like this:

$$(X_1 = 2, T_1 = A) \rightarrow (X_2 = 3, T_2 = A) \rightarrow (X_3 = M)$$

Assume that this is a Markov decision process, and that we have transition probabilities as given in Table 2:

T_t \ X_t	1	2	3	M
A	$\begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{8} & \frac{1}{8} \end{bmatrix}$	$\begin{bmatrix} 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$	-
B	$\begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} & 0 \end{bmatrix}$	$\begin{bmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{8} & \frac{1}{8} \end{bmatrix}$	$\begin{bmatrix} \frac{1}{8} & \frac{1}{2} & \frac{1}{4} & \frac{1}{8} \end{bmatrix}$	-

Table 2: Probability distribution over the next state $X_{t+1} \in \{1, 2, 3, M\}$ given we observe treatment T_t while in state X_t.

1. Without doing any formal RL or causal inference, argue that the best treatment policy π^* is the following deterministic function from states to actions:

$$\pi^* : \begin{cases} 1 & \mapsto A \\ 2 & \mapsto A \\ 3 & \mapsto B \end{cases}$$

2. Now, we'll formalize this derivation using Q-Learning. Assume that mortality has a reward of -1, and initialize all Q-values as $Q_0(s,a) = 0$ for all s, a. We additionally assume no discount factor, i.e. $\gamma = 1$, and a learning rate $\alpha = 1$.[1] Note that M is a terminal state – once a patient reaches state M, no further actions are taken or rewards collected.

   Until the resultant policy reaches the optimal policy identified in the previous question (which will also be a stable point in Q-Learning), compute the next iteration of $Q_i(s,a)$ and the corresponding policy at that iteration, $\pi^{(i)} : s \mapsto \operatorname{argmax}_a Q_i(s,a)$. Represent your Q function by a table with a row per treatment and a column per state. Don't resolve any ties in the argmax within $\pi^{(i)}$ – just indicate that there is a tie in your solution.

   Hint: Recall that

$$Q_{k+1}(s,a) = Q_k(s,a) + \alpha \left(\sum_{s' \in X} P^a_{ss'} \left(r^a_{ss'} + \gamma \max_{a' \in \{A,B\}} Q_k(s',a') \right) - Q(s,a) \right),$$

   as seen in Lecture 16's slides[2], slide 44.

[1] Note that by setting $\alpha = 1$ and using iteration not over *observed samples* from this trajectory but from the *exact* MDP, we're really doing a form of Q-Value Iteration here, which will converge.

[2] https://mlhcmit.github.io/slides/lecture16.pdf

3. How do you see the above process as related to causal inference? In particular, if we were to perform causal inference considering only one treatment decision to minimize the probability of mortality in the immediate next time-step, what policy would we design? Does this look at all like any sub-part of the Q-learning update equations?