DS 102 Homework 5

If you are handwriting your solution please make sure your answers are legible, as you may lose points otherwise.
Data science is a collaborative activity. While you may talk with others about the homework, please write up your solutions individually. If you do discuss the homework with others, please include their names on your submission.
Due on Gradescope by 9:29am, Tuesday April 21, 2020

1. (15 points) **Generalized Linear Model for Dilution Assay**

   Being able to reformulate problems as generalized linear models (GLMs) enables you to solve a wide variety of problems. If you haven't worked through Discussion 8 on GLMs, we'd recommend you review that material first. In particular, make sure you understand that formulating a GLM involves specifying 1) the output distribution and 2) the link function, and how to go about choosing those two components.

   In this problem, you'll retrace the footsteps of the statistician R. A. Fisher and develop one of the very first applications of GLMs. In a 1922 paper, Fisher formulated a GLM he used to estimate the unknown concentration ρ_0 of an infectious microbe in a solution.

   Without specialized technology to measure ρ_0 from the solution, Fisher envisioned the following procedure: we will progressively dilute the original solution, and after each dilution, we'll pour out some small volume v onto a sterile plate. If zero microbes land on the plate, it will remain sterile, but if any microbes land on a plate, they will grow visibly on it (we call this an "infected plate"). Fisher's idea was that by observing whether or not the plate is infected at each dilution, and by formulating the relationship between this data and ρ_0 as a GLM, we can estimate ρ_0 from this data.

   Specifically, let ρ_t denote the concentration at dilution t. Each time, we dilute the solution to be half its concentration:

   $$\rho_t = \frac{\rho_0}{2^t} \tag{1}$$

   for $t = 0, 1, \ldots$. When we pour out volume v of the solution onto the plate, and wait awhile to allow for microbe growth, we can observe whether a plate was infected (*i.e.*, has a non-zero number of microbes) or is sterile (*i.e.*, has zero microbes). Therefore, our data $Y_t \in \{0, 1\}$ is whether or not the plate is infected at each dilution.

   In the next few parts, we'll formulate a GLM that relates ρ_0 and t to the data Y_t. Estimating the parameters of this GLM allowed Fisher to then estimate ρ_0, as will become clear in the last part.

   (a) (2 points) At dilution t, the data $Y_t \in \{0, 1\}$ indicates whether or not the plate is infected. Let $\mu(t) := \mathbb{E}[Y_t]$ denote the chance that a plate is infected. Write down a plausible output distribution for Y_t, using $\mu(t)$.

(b) (5 points) At dilution t, we pour out volume v onto a plate, so the expected number of microbes on the plate is $\rho_t v$. The actual number of microbes is distributed as a Poisson random variable with this mean $\rho_t v$:

$$\#\text{ microbes on plate at dilution } t \sim \text{Poisson}(\rho_t v). \tag{2}$$

Using this fact, write out an expression for $\mu(t) := \mathbb{E}[Y_t]$. Start with

$$\mu(t) = \mathbb{P}(\text{plate is infected at dilution } t) \tag{3}$$
$$= 1 - \mathbb{P}(\text{there are 0 microbes on plate at dilution } t). \tag{4}$$

(c) (5 points) Find a link function g such that

$$g(\mu(t)) = \beta_0 + \beta_1 t \tag{5}$$

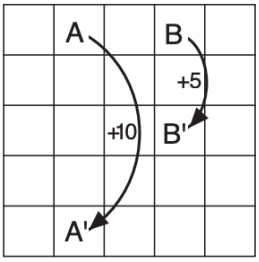
for some constants β_0 and β_1.

(d) (3 points) Choosing an appropriate output distribution and link function as we've done in Parts (a) and (c) completes the GLM specification. Now, suppose you've estimated β_0 and β_1 (*e.g.*, using maximum-likelihood estimation). Write down an estimate of ρ_0.

2. (25 points) **Policies and Value Functions** The figure below (left subplot) shows a grid representation of a simple finite Markov Decision Process (MDP). The cells of the grid correspond to states of the environment. At each cell, four actions are possible: north, south, east, and west. These actions deterministically move the agent one cell in the respective direction on the grid, and result in reward of 0, with the following exceptions:

- Actions that would take the agent off the grid leave its location unchanged, and result in a reward of 1.
- From state A, all actions yield a reward of +10 and take the agent to A'.
- From state B, all actions yield a reward of +5 and take the agent to B'.

Suppose the agent selects all four actions with equal probability in all states. The figure below (right subplot) shows the value function for this policy, for discounted rewards with $\gamma = 0.9$.



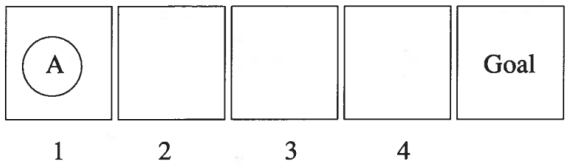
3.3	8.8	4.4	5.3	1.5
1.5	3.0	2.3	1.9	0.5
0.1	0.7	0.7	0.4	-0.4
-1.0	-0.4	-0.4	-0.6	-1.2
-1.9	-1.3	-1.2	-1.4	-2.0

(a) (10 points) Verify numerically that the center state has a value of +0.7 under this policy, assuming that the values of the adjacent squares are correct (+2.3, +0.4, -0.4, and +0.7 for the top-, right-, bottom-, and left-neighboring states, respectively). (These numbers are accurate only to one decimal place.) Hint: Use the Bellman equation.

(b) (10 points) Prove that adding a constant c to all the rewards adds a constant v_c to the values of all states, and therefore does not affect the relative values of states under any policy (*i.e.*, the differences between states' values doesn't change). What is v_c in terms of c and γ?

(c) (5 points) In a few sentences, explain why state A has an expected return that is less than 10, its immediate reward, while state B is valued more than 5, its immediate reward.

3. (25 points) **Markov Decision Processes for Robot Soccer** A soccer robot A is on a fast break toward the goal, starting in position 1. From positions 1 through 3, it can either shoot (S) or dribble the ball forward (D). From 4 it can only shoot. If it shoots, it either scores a goal (state G) or misses (state M). If it dribbles, it either advances a square or loses the ball, ending up in state M.



In this MDP, the states are 1, 2, 3, 4, G, and M, where G and M are terminal states. The transition model depends on the parameter y, which is the probability of dribbling successfully (*i.e.*, advancing a square). Assume a discount of $\gamma = 1$. For $k \in \{1, 2, 3, 4\}$, we have

$$\begin{aligned} \mathbb{P}(G \mid k, S) &= \frac{k}{6} \\ \mathbb{P}(M \mid k, S) &= 1 - \frac{k}{6} \\ \mathbb{P}(k+1 \mid k, D) &= y \\ \mathbb{P}(M \mid k, D) &= 1 - y, \\ R(k, S, G) &= 1 \end{aligned}$$

and rewards are 0 for all other transitions.

(a) (5 points) Denote by V^π the value function for the specific policy π. What is $V^\pi(1)$ for the policy π that always shoots?

(b) (5 points) Denote by $Q^*(s, a)$ the value of a q-state (s, a), which is the expected utility when starting with action a at state s, and thereafter acting optimally. What is $Q^*(3, D)$ in terms of y?

(c) (10 points) Denote by $V_t^*(s)$ the value of a state s at iteration t, which is the expected utility when starting in state s and acting optimally. Using $y = \frac{3}{4}$, complete the first two iterations ($t = 1, 2$) of value iteration. Iteration 0 corresponds to having value 0 in every state: $V_0^*(1) = V_0^*(2) = V_0^*(3) = V_0^*(4) = 0$.
Hint: Recall that $V_{t+1}^*(s) = \max_{a \in A} \sum_{s'} \mathbb{P}(s'|s,a)(R(s,a,s') - V_t^*(s'))$.

(d) (5 points) For what range of values of y is $Q^*(3, S) \geq Q^*(3, D)$?

4. (20 points) **Using the Bootstrap to Evaluate Drug Bioequivalence**. When drug companies introduce new drugs, the FDA requires them to show that the new drug is *bioequivalent* to the current drug used to treat the same condition. Bioequivalence means that the effect of the new drug is not substantially different from the effect of the current drug. The way the effect is measured is application-dependent—here, we'll look at drugs that infuse a certain hormone into the blood. A drug's effect is therefore the amount of hormone in the blood after administering the drug.

To formally define bioequivalence, let the random variables O, N, P denote the effect of the old drug, the effect of the new drug, and the effect of a placebo, respectively. The FDA requirement for bioequivalence is that

$$|\theta| \leq 0.2 \tag{6}$$

where

$$\theta = \frac{\mathbb{E}[N - O]}{\mathbb{E}[O - P]}. \tag{7}$$

In this problem, you'll estimate θ from a dataset and use the bootstrap to determine, with a certain confidence, whether or not we have bioequivalence. Please submit all code and plots generated for this problem (for example, you can do the problem in a Jupyter notebook and save it as a PDF, if you'd like).

(a) (2 points) The CSV file `bioequivalence.csv` on the website contains the following data on the level of a hormone in 8 subjects' blood, after medications were administered.

subject	placebo	old	new
1	9243	17649	16449
2	9671	12013	14614
3	11792	19979	17274
4	13357	21816	23798
5	9055	13850	12560
6	6290	9806	10157
7	12412	17208	16570
8	18806	29044	26325

Download the data, and use it to compute the plug-in estimate $\hat{\theta}$ of θ.

(b) (15 points) Part (a) gave an estimate of θ, but by itself it doesn't capture the certainty we have in the estimate, so we can't use it to conclude that we have bioequivalence with a given confidence level. Instead, we'll compute a bootstrap confidence interval to do this.

(i) Implement a function `bootstrap_bioequivalence(N, O, P, B)` which takes in the following inputs:

- $N = (N_1, \ldots, N_n)$, an array of the effects of the new drug on n subjects
- $O = (O_1, \ldots, O_n)$, an array of the effects of the old drug on n subjects
- $P = (P_1, \ldots, P_n)$, an array of the effects of the placebo on n subjects
- B, an integer which is the number of bootstrap replicates

and outputs a length-B array of bootstrap replicates of $\hat{\theta}$.

(ii) Using `bootstrap_bioequivalence(N, O, P, B)`, compute $B = 10000$ bootstrap replicates of $\hat{\theta}$. Plot a histogram of these replicates, and label the x- and y-axes.

(iii) Using the replicates from (ii), compute a 95-percentile confidence interval for θ (make sure to include the code you use to compute this). Hint: Use the function `np.percentile`.

(c) (3 points) Based on Part (b), can we conclude the new drug and old drug are bioequivalent, at the 95% confidence level? That is, does the 95% confidence interval fall within the FDA requirement for bioequivalence?