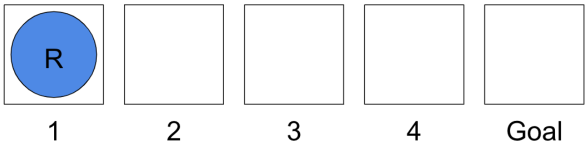
Overview

Submit your writeup including all code and plots as a PDF via Gradescope. We recommend reading through the entire homework beforehand and carefully using functions for testing procedures, plotting, and running experiments. Taking the time to test, maintain, and reuse code will help in the long run!

Data science is a collaborative activity. While you may talk with others about the homework, please write up your solutions individually. If you discuss the homework with your peers, please include their names on your submission. Please make sure any handwritten answers are legible, as we may deduct points otherwise.

1 Markov Decision Process for Robot Soccer

A soccer robot R 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 Markov Decision Process (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) (3 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) (4 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) (3 points) For what range of values of y is $Q^*(3, S) \geq Q^*(3, D)$? Interpret your answer in plain English.

2 Differential Privacy

When evaluating the frequency of stigmatized, private, illegal, or embarrassing activities, the **randomized response** approach is a simple yet effective way of achieving differential privacy. In this problem, you'll explore and show some useful properties of randomized responses.

Suppose we are interested in assessing drug use in Berkeley, and to this end we conduct a survey to estimate the fraction of the population that regularly uses drugs. However, this is a sensitive question: people may not feel comfortable at disclosing this type of information.

In the survey, each participant is asked: "Do you normally use drugs at least once a week?". To achieve randomized responses, the participant is then instructed to respond as follows:

1. Flip a fair coin (*i.e.*, 50% chance of heads and 50% chance of tails).

2. If tails, then respond truthfully.

3. If heads, then flip the coin again. Respond "Yes" if heads, and "No" if tails.

(a) (2 points) For a given participant, suppose $b \in \{0, 1\}$ is the value of the true answer to the question, where $b = 1$ is "Yes" and $b = 0$ is "No". As the surveyors, we don't observe b. Instead, due to the randomized response we observe $b' \in \{0, 1\}$, where $b' = 1$ means the participant responded "Yes" and $b' = 0$ means the participant responded "No" (note that both $b \in \{0, 1\}$ and $b' \in \{0, 1\}$).

i) Suppose that for a given participant, $b = 1$. Write down the distribution that their survey response b' is drawn from. Explain your reasoning.

ii) Alternatively, suppose $b = 0$. Write down the distribution that their survey response b' is drawn from. Explain your reasoning.

(b) (2 points) Randomized responses encourage participants to respond truthfully: in this example, if they respond "No" to the drug use question, there's always plausible deniability (*i.e.*, even if they responded "No", the true answer could be "Yes").

To formalize this idea, let $q \in [0, 1]$ denote the true fraction of the population that frequently uses drugs. Compute the probability $\mathbb{P}(b = 0 \mid b' = 0)$ that the true answer is "No", given that the survey response is "No". Your answer should be in terms of q and constants.

(c) (2 points) Despite the fact that each participant's response may or may not be truthful, the neat thing is that we can still use the responses in aggregate to estimate q. Let

$$b_i \sim \text{Bernoulli}(q)$$

for $i = 1, \ldots, n$ denote the true answers of n survey participants we randomly sampled from the population. Let $b_i', i = 1, \ldots, n$ denote the participants' survey responses, which are also random (based on b_i, following the conditional distribution you described in Part (a)). Suppose we compute the sample mean of the n survey responses, $\frac{1}{n}\sum_{i=1}^{n} b_i'$. Show that

$$\mathbb{E}_{\{b_i, b_i'\}_{i=1}^{n}}\left[\frac{1}{n}\sum_{i=1}^{n} b_i'\right] = 1/4 + (1/2)q, \tag{1}$$

where the expectation is over both the b_i and the b_i'.

Hint: Use the Tower property (iterated expectation).

(d) (2 points) Suppose you take the sample mean of the survey responses and get a value $A = \frac{1}{n}\sum_{i=1}^{n} b_i'$. Using the fact in Equation (1), propose an estimator for q, the true fraction of the population that frequently uses drugs. Your solution should be in terms of A and constants.

(e) (0 points) (Optional) Performing the survey with randomized responses, then computing the sample mean, can be thought of as a function $\mathcal{A} : \{0,1\}^n \to \mathbb{R}$ that we call the "randomized response algorithm". The function $\mathcal{A}$ takes as input a dataset of true answers $\mathcal{D} = \{b_1, \ldots, b_n\}$ and outputs a random value (the sample mean $\frac{1}{n}\sum_{i=1}^{n} b_i'$, where the b_i' have the conditional distributions you found in Part (a)).

The output $\mathcal{A}(\mathcal{D})$ can take on any value in the set $\mathcal{R} = \{0, \frac{1}{n}, \frac{2}{n}, \ldots, 1\}$. Reviewing the definition of ϵ-differential privacy, the randomized response algorithm $\mathcal{A}$ is called **ϵ-differentially private** if

$$\mathbb{P}(\mathcal{A}(\mathcal{D}_1) \in S) \leq \exp(\epsilon)\mathbb{P}(\mathcal{A}(\mathcal{D}_1) \in S) \tag{2}$$

for any set $S \subseteq \mathcal{R}$, and any two datasets $\mathcal{D}_1, \mathcal{D}_2$ of size n that differ by only one datapoint (*i.e.*, they contain all the same true answers b_i except for one, which has the value 0 in one dataset but 1 in the other). Note that the probability is over the randomness in the algorithm (*i.e.*, the randomness in the survey responses given fixed $\mathcal{D}$), not randomness in $\mathcal{D}$ which we consider fixed.

For a dataset of size $n = 1$ (a single survey participant), show that the randomized response algorithm is $(\log 3)$-differentially private.

Hint: For $n = 1$, $\mathcal{R} = \{0, 1\}$, which has three possible subsets.