Overview

Submit your writeup, including any code, as a PDF via gradescope.[1] We recommend reading through the entire homework beforehand and carefully using functions for testing procedures, plotting, and running experiments. Taking the time to 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.

Simulation Study of Bandit Algorithms

In this problem, we evaluate the performance of two algorithms for the multi-armed bandit problem. The general protocol for the multi-armed bandit problem with K arms and n rounds is as follows: in each round $t = 1, \ldots, n$ the algorithm chooses an arm $A_t \in \{1, \ldots, K\}$ and then observes reward r_t for the chosen arm. The bandit algorithm specifies how to choose the arm A_t based on what rewards have been observed so far. In this problem, we consider a multi-armed bandit for $K = 2$ arms, $n = 50$ rounds, and where the reward at time t is $r_t \sim \mathcal{N}(A_t - 1, 1)$, i.e. $\mathcal{N}(0, 1)$ for arm 1 and $\mathcal{N}(1, 1)$ for arm 2.

(a) (4 points) Consider the multi-armed bandit where the arm $A_t \in \{1, 2\}$ is chosen according to the explore-then-commit algorithm (below) with $c = 4$. Let $G_n = \sum_{t=1}^{n} r_t$ denote the total reward after $n = 50$ iterations. Simulate the random variable G_n a total of $B = 2000$ times and save the values $G_n^{(b)}$, $b = 1, \ldots, B$ in a list. Report the (empirical) average pseudoregret $\frac{1}{B} \sum_{b=1}^{B} \left(50\mu^* - G_n^{(b)}\right)$ (where μ^* is the mean of the best arm) and plot a normalized histogram of the rewards.

Algorithm 1 Explore-then-Commit Algorithm

input: Number of initial pulls c per arm
for $t = 1, \ldots, cK$: **do**
  Choose arm $A_t = (t \bmod K) + 1$
end
Let $\hat{A} \in \{1, \ldots, K\}$ denote the arm with the highest average reward so far.
for $t = cK + 1, cK + 2, \ldots, n$: **do**
  Choose arm $A_t = \hat{A}$
end

(b) (4 points) Consider the multi-armed bandit where the arm $A_t \in \{1, 2\}$ is chosen according to the UCB algorithm (below) with $c = 4$, $n = 50$ rounds. Repeat the simulation in Part (a) using the UCB algorithm, again reporting the (empirical) average pseudoregret and the histogram of $G_n^{(b)}$ for $b = 1 \ldots B$ for $B = 2000$. How does the pseudoregret compare to your results from part (a)?

[1] In Jupyter, you can download as PDF or print to save as PDF

Note: If $T_A(t)$ denote the number of times arm A has been chosen (before time t) and $\hat{\mu}_{A,t}$ is the average reward from choosing arm A (up to time t), then use the upper confidence bound $\hat{\mu}_{A,T_A(t-1)} + \sqrt{\frac{2\log(20)}{T_A(t-1)}}$. Note also that this algorithm is slightly different than the one used in lab and lecture as we are using an initial exploration phase.

Algorithm 2 UCB Algorithm

input: Number of initial pulls c per arm
for $t = 1, \ldots, cK$: **do**
| Choose arm $A_t = (t \bmod K) + 1$
end
for $t = cK + 1, cK + 2 \ldots$: **do**
| Choose arm A_t with the highest upper confidence bound so far.
end

(c) (1 point) Compare the distributions of the rewards by also plotting them on the same plot and briefly justify the salient differences.

Private Mean Estimation

One of the most important techniques in data analysis and machine learning is mean estimation. It is used a subroutine in essentially every task. In this question, we will explore how to incorporate differential privacy into mean estimation. En route, we will explore the Laplace mechanism, which is one of the fundamental tools in building differentially private algorithms.

Let $S = \{X_1, \ldots X_n\}$ be i.i.d. samples from a Bernoulli distribution with unknown mean p. Recall, from HW5, that the sample mean

$$p_n(S) = \frac{1}{n}\sum_{x\in S} x \tag{1}$$

satisfies $|p_n - p| \leq cn^{-1/2}$ with probability 0.99 for some constant c.

In order to incorporate privacy, the main idea is to add noise to the estimator Equation (1). For the noise distribution, we will use the Laplace distribution, which has density given by

$$f_{\mu,b}(x) = \frac{1}{2b}\exp\left(-\frac{|x-\mu|}{b}\right).$$

We will denote this distribution as $\mathrm{Lap}(\mu, b)$. The mean of the distribution is μ and the variance is $2b^2$. The differentially private estimator is given by

$$\hat{p}_{\epsilon,n}(S) = p_n(S) + Y$$

where Y is sampled from $\mathrm{Lap}\left(0, \frac{1}{\epsilon n}\right)$. Here ϵ is a parameter that will control the privacy.

(a) (1 point) Let S_1 and S_2 be two data sets with n binary samples ($\{0,1\}$-valued) each. Additionally, also assume that S_1 and S_2 differ only in one item. More precisely, we can construct S_2 by removing one element from S_1 and adding another binary value (0 or 1).

Show that the sample means for the two sets are close. Specifically, show:

$$\left|p_n(S_1) - p_n(S_2)\right| \leq \frac{1}{n}. \tag{2}$$

This is referred to as p_n having sensitivity n^{-1}.

(b) (1 point) For any fixed S, explain why $\hat{p}_{\epsilon,n}(S)$ is distributed according to a Laplace distribution. What are the corresponding parameters?

(c) (2 points) First, we will show that the above estimator is still fairly accurate. Show that with probability 0.99 (over the sampling of the noise), for every S, we have

$$\left|p_n(S) - \hat{p}_{\epsilon,n}(S)\right| \leq \frac{20}{\epsilon n}.$$

You may find it especially useful to apply a concentration inequality we learned about in class.

(d) In this part, we will see that the mechanism is ϵ-differentially private. Let us recall the definition of differential privacy in this context. An estimator g is ϵ-differentially private if for all sets $A \subset \mathbb{R}$, we have

$$\Pr[g(S_1) \in A] \leq \exp(\epsilon) \cdot \Pr\left[g(S_2) \in A\right]$$

where S_1, S_2 are two data sets that differ only in one item.

(i) (3 points) Let $Y_1 \sim \text{Lap}(\mu_1, b)$ and $Y_2 \sim \text{Lap}(\mu_2, b)$. Show that

$$\Pr[Y_1 \in A] \leq \exp\left(\frac{|\mu_1 - \mu_2|}{b}\right) \cdot \Pr[Y_2 \in A].$$

This hints at why the Laplace distribution is particularly well suited for differential privacy.

Hint: Find a bound on the likelihood ratio, and relate that to the inequality above

(ii) (2 points) Using Equation (2) and earlier parts of the question, show that the estimator $\hat{p}_{\epsilon,n}$ is ϵ-differentially private.

(iii) (1 points) Put these steps together show that $\hat{p}_{\epsilon,n}$ is a ϵ-DP estimator for p with error

$$\left|p - \hat{p}_{\epsilon,n}\right| \leq O\left(\frac{1}{\sqrt{n}} + \frac{1}{n\epsilon}\right)$$

with probability 0.98 over the randomness of the sample and the mechanism.

(e) (1 point) Now, suppose that instead of Bernoulli, the individual samples X_i were real-valued random variables taking values in $[0, 5]$. Which part(s) of the analysis above (if any) would change? You don't need to redo the analysis.