Instructions

- **Due:** Tuesday, February 3 in class
- Turn in a handwritten or printed PDF in class. Ensure it contains your name and email.
- You may use any resources in completing this homework. These include other students, AI tools, textbooks, and course notes.
- While completing the assignment, follow these instructions for each subproblem:
  - Make sure you understand the question.
  - Try to solve it alone.
  - If you use outside help, make sure you understand and can explain the solution.
- **Collaborator Acknowledgement:** After each problem, list any people/tools/resources you relied on. Write a few sentences reflecting on what insights you required help with.

Problems

Problem 1. Better Composition for Randomized Response In class we saw the ε-DP algorithm Randomized Response $\mathrm{RR}_\varepsilon : \{0,1\} \to \{0,1\}$, which returns its input with probability $\frac{e^\varepsilon}{1+e^\varepsilon}$ and otherwise flips its input. Consider its privacy loss: for $y \in \{0,1\}$,

$$L^{0\to 1}_{\mathrm{RR}_\varepsilon}(y) = \ln\left(\frac{\Pr[\mathrm{RR}_\varepsilon(0) = y]}{\Pr[\mathrm{RR}_\varepsilon(1) = y]}\right).$$

When $Y \sim \mathrm{RR}_\varepsilon(0)$ is a random variable, we call $L^{0\to 1}_{\mathrm{RR}_\varepsilon}(Y)$ the *privacy loss random variable* (PLRV). This is an important quantity in the study of differential privacy.

Let $A(x) = (\mathrm{RR}_\varepsilon(x), \ldots, \mathrm{RR}_\varepsilon(x))$ be T independent repetitions of RR_ε. Basic composition says that A satisfies ε'-DP for $\varepsilon' = T\varepsilon$. In this problem, we will show how one can achieve a smaller ε' by switching to approximate DP.

(a) Write the PLRV for $A(0)$ as a sum of independent and identically distributed random variables. For a single random variable, compute (i) its expectation and (ii) the minimum and maximum values it can achieve.

(b) Derive a $1-\delta$ high-probability upper bound on the PLRV. You can use the following concentration inequality.

Lemma (Hoeffding Bound)**.** *Let $Z_1, \ldots, Z_T$ be independent random variables such that for all $1 \le t \le T$, $\mathbb{E}[Z_t] = \mu$ and $\Pr[a \le Z_t \le b] = 1$. Then for any $\alpha > 0$*

$$\Pr\left[\sum_{t=1}^{T} Z_t \ge T\mu + \alpha\right] \le \exp\left(-\frac{2\alpha^2}{T(b-a)^2}\right).$$

(c) State this as an approximate DP guarantee: for any $\delta > 0$, $A(x)$ satisfies (ε', δ)-DP for $\varepsilon' = \cdots$.

(d) Set $\varepsilon = 0.1$ and $\delta = 10^{-6}$. For what T does our bound yield a smaller ε' than the result of basic composition?

Problem 2. Private Histograms In this problem we will analyze two DP histogram algorithms. Histograms are a fundamental tool and also serve as a subroutine in many DP algorithms.

We receive data points $x_1, \ldots, x_n \in \mathcal{X}$ and a collection of "bins" $\{S_j\}_{j \in \mathcal{J}}$, disjoint subsets of $\mathcal{X}$. Here $\mathcal{J} \subseteq \mathbb{N}$ is a (possibly infinite) index set. We wish to privately approximate the count $c_j = \#\{i \in [n] : x_i \in S_j\}$ for each bin.

You can use the following tail bound for zero-mean Laplace random variables: for any $t, b > 0$, if $Z \sim \mathrm{Lap}(b)$ then $\Pr[|Z| \geq t] = \exp(-|t|/b)$.

(a) We saw in class that adding noise $\mathrm{Lap}(1/\varepsilon)$ to a sum of 0/1 inputs preserves ε-DP. This can be seen as a one-bin histogram algorithm: we release noisy count $\tilde{c} = c + Z$ for $Z \sim \mathrm{Lap}(1/\varepsilon)$.

Analyze its accuracy: for any $\beta > 0$, with probability at least $1 - \beta$ we have $|\tilde{c} - c| \leq \cdots$.

(b) Prove that, for k bins, adding independent $\mathrm{Lap}(2/\varepsilon)$ noise to each bin preserves ε-DP.

(c) Analyze the accuracy of this k-bin algorithm: with probability at least $1 - \beta$, for all $j \in [k]$ we have $|\tilde{c}_j - c_j| \leq \cdots$. (Hint: union bound.)

(d) The above guarantees become vacuous when run on histograms with an infinite number of bins. Algorithm 1 avoids this difficulty. Show that it satisfies (ε, δ)-DP.

(Hint: when introducing a new data point, it can either be added to an empty bin or a non-empty bin. Treat these two cases differently.)

(e) (Informal, No Single Correct Answer) Suppose we are given i.i.d. samples from a normal distribution $\mathcal{N}(\mu, 1)$ but have no prior knowledge about μ's location. Suppose how we could use Algorithm 1 to privately produce a rough approximation of μ.

How might one use similar ideas to produce a location estimate for a d-dimensional Gaussian $\mathcal{N}(\mu, \mathbb{I}_d)$? Discuss possible drawbacks of your approach.

Algorithm 1 Approx DP Histogram

Input: Data $x_1, \ldots, x_n \in \mathcal{X}$, disjoint bins $\{S_j\}_{j \in \mathcal{J}}$, privacy parameters $\varepsilon, \delta \geq 0$.
Returns: Noisy bin counts

for $j \in \mathcal{J}$ **do**
  $c_j \leftarrow \#\{i \in [n] : x_i \in S_j\}$
end for
for $j \in \mathcal{J}$ with $c_j > 0$ **do**
  $\tilde{c}_j \leftarrow c_j + Z_j$, where $Z_j \sim \mathrm{Lap}(2/\varepsilon)$
end for
$\tau \leftarrow \frac{2\log(1/\delta)}{\varepsilon} + 1$
return $\{(j, \tilde{c}_j) : \tilde{c}_j \geq \tau\}$