Privacy in Statistics and Machine Learning — Spring 2023
In-class Exercises for Lecture 5 (Differential Privacy Foundations II)
February 2, 2023

Adam Smith (based on materials developed with Jonathan Ullman)

Problems with marked with an asterisk (*) are more challenging or open-ended.

1. Consider the following two scenarios. For each one, decide whether the overall algorithm can be proven differentially private and justify your decision.

   (a) A biologist uses an ε-DP algorithm A_1 to release the approximate frequencies of d different diseases in the data set. She then selects the 10 diseases with *the highest reported frequencies in the output of* A_1, and uses a ε-DP algorithm to release an approximate version of all $\binom{10}{2}$ pairwise correlations between the selected diseases.

   (b) A biologist uses an ε-DP algorithm to release the approximate frequencies of d different diseases in the data set. She then selects the 10 diseases with *the highest true frequencies in the original data set*, and uses a ε-DP algorithm to release all $\binom{10}{2}$ pairwise correlations between the selected diseases.

2. (Group Privacy) You are reviewing a paper that claims a new, differentially-private version of Lloyd's algorithm. They claim to have experiments that show good performance on data sets of size 100 with epsilon = 0.005. Should you believe them? Why or why not?

3. (Exercise 1.4 from the notes) Sometimes it is much better to analyze an algorithm as a whole than to use the composition lemma. Consider the histogram example from Lecture 4, where $\mathcal{U}$ is written as a partition of disjoint sets $B_1, B_2, ..., B_d$, and we want to count how many records lie in each set. Viewed as one d-dimensional function, the histogram has global sensitivity 2. We could also view it as d separate functions $n_1, n_2, ..., n_d$, each with globabl sensitivity 1. How much noise would the Laplace mechanism add to these counts if we ran it spearately for each of the n_j with privacy budget divided equally among them? How does that compare to running the Laplace mechanism once on the joint function?

4. Analyze the name and shame algorithm (Exercise 3.3).

5. What happens if we try to run the Laplace mechanism with different noise distributions? Which of these distributions leads to an ε-DP mechanism? For simplicity, we'll focus on the 1-dimensional case were $f : \mathcal{U}^n \to \mathbb{R}$, and look at mechanisms of the form

$$A(\mathbf{x}) = f(\mathbf{x}) + \frac{GS_f}{\varepsilon} Z \qquad \text{where } Z \sim P \text{ and } P = ... \tag{1}$$

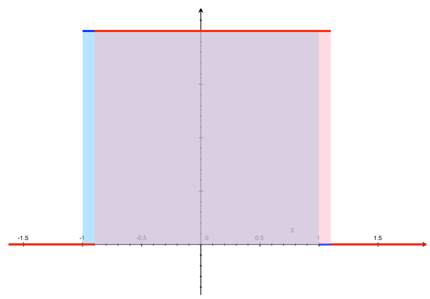
   (a) The uniform distribution on $[-1, 1]$ (density $h(y) = 1/2$ on $[-1, 1]$ and 0 elsewhere)

   (b) The Normal distribution $N(0, 1)$ (density $h(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2}$ for $y \in \mathbb{R}$)

(c) The Cauchy distribution (density $h(y) = \frac{1}{\pi(1+y^2)}$ for $y \in R$)

For which of the options above do we get an ε'-DP mechanism where ε' is finite (not that ε' need not be exactly equal to ε)?

Example: If we shift a copy of the uniform distribution by 0.1, we get the picture below. Are there events whose probability changes by a large multiplicative factor?



6. Prove Theorem 3.1 from the lecture notes (Exercise 3.2). (Comparing posterior distributions constructed assumign Alice's data are or are not used.)

7. **(*)** Can the Laplace mechanism be substantially improved? Answering that is complicated, but let's look at a sense in which the Laplace mechanism is basically optimal for 1-dimensional releases of functions with a given global sensitivity.

   Fix a function $f : \mathcal{U}^n \to \mathbb{R}$. Suppose that $1/\varepsilon$ is a positive integer, and there are two data sets $\mathbf{x}, \tilde{\mathbf{x}}$ that differ in $1/\varepsilon$ entries, and such that $|f(\mathbf{x}) - f(\tilde{\mathbf{x}})| = GS_f/\varepsilon$. Show that for every ε-DP algorithm A, for at least one the two data sets $\mathbf{x}$ and $\tilde{\mathbf{x}}$, the expected absolute value of the algorithm's error is $\Omega(GS/\varepsilon)$. That is, show that

$$\max\left\{\mathbb{E}\left(|A(\mathbf{x}) - f(\mathbf{x})|\right),\ \mathbb{E}\left(|A(\tilde{\mathbf{x}}) - f(\tilde{\mathbf{x}})|\right)\right\} \geq c \cdot \frac{GS_f}{\varepsilon}$$

   for some absolute constant c (e.g. $c = 1/100$ will do).

   Hint: You can simplify things a bit by using group privacy to show that $A(\mathbf{x}) \approx_{\varepsilon'} A(\tilde{\mathbf{x}})$ for $\varepsilon' = 1$.

   Hint 2: If A is a nonnegative random variable and $\Pr(A \geq \mu) \geq \frac{1}{100}$, then $\mathbb{E}(A) \geq \mu/100$ (by Markov's inequality).

   Hint 3: Look at the events that the algorithm's output is either *at least* $\frac{f(\mathbf{x})+f(\tilde{\mathbf{x}})}{2}$ or *at most* that quantity.