Homework 2

Deadline: 30th October, 2022

Saturday 15^{th} October, 2022

1. (Name and Shame Mechanism). Consider the following mechanism A, for a given input dataset $D = \{x_1, \cdots, x_n\}$, it generates
$$Y_i = \begin{cases} (i, x_i) \text{ w.p. } \delta \\ \text{nothing w.p. } 1-\delta \end{cases} \tag{1}$$
and outputs $Y = (Y_1, \cdots, Y_n)$. Show that A is $(0, \delta)$-DP.

2. (Noisy-max with Laplace Noise). In the class (Lecture 5), we have showed that adding the exponential noise $\exp(\frac{2\Delta}{\epsilon})$ in the Noisy-Max mechanism could preserve ϵ-DP. Now, instead of using the exponential distribution, we use $\text{Lap}(\frac{\Delta}{\epsilon})$ in the Noisy-Max mechanism. Try to show this is also ϵ-DP.

3. (Adding Uniform Noise) Suppose we add uniform noise to a count query $f : \{0,1\}^n \mapsto \mathbb{R}$ with $f(D) = \sum_{i=1}^n x_i$, that is, we release $A(D) = f(D) + Z$ where $Z \sim U_{[-\lambda,\lambda]}$, $U_{[-\lambda,\lambda]}$ is the uniform distribution on the interval $[-\lambda, \lambda]$. How large must λ be to satisfy (ϵ, δ)-DP? Do both ϵ and δ matter in setting? When $\delta < \frac{1}{n}$, will this mechanism produce useful information?

4. (Implementation of Noisy-max Mechanism and Exponential Mechanism) You can find a selection problem (you have to say the output space, the score function and its sensitivity) and try to implement the noisy-max mechanism and the exponential mechanism. Write a report on your findings.

5. **Differentially Private Top k Selection:** Suppose we have d candidates items and a score function $q : [d] \times \mathcal{X} \mapsto \mathbb{R}$. In the selection in Lecture 5 we aimed to find a single high-score item. Suppose we now want to find $k < \frac{d}{2}$ items. Given an algorithm that outputs a set of k items $S = A(D)$, we measure the error as follows: let $q_{(k)}(D)$ be the score function of the k-th best item, The error of the algorithm is
$$q_{(k)}(D) - \min_{j \in S} q(j; D). \tag{2}$$
What expected error guarantee can you prove for the algorithm that proceeds by repeating the exponential mechanism k times without replacement. Try to consider both of using the basic composition property of ϵ-DP and the advanced composition property of (ϵ, δ)-DP cases.

6. (Comparison with Gaussian and Laplace Mechanism) Generate a dataset $D = \{x_1, \cdots, x_n\}$ where each $x_i \in \{0,1\}^d$. Consider answering the average query $f(D) = \frac{1}{n}\sum_{i=1}^n x_i$ via Laplace and Gaussian mechanism. Implement these two mechanisms with variate n, d, ϵ (You cases must at least include $d = 1$ and $d \gg 1$). Write a report on your findings.