CAS751: IT Metods in Trustworthy ML Homework 3 - **Due: 11/24/2024**

Apprixmate Differential Privacy

Instructor: Shahab Asoodeh

1. (**25 points**) Let $D = \{x_1, \ldots, x_n\} \in [0,1]^n$ be a dataset and q is the average query, meaning $q_D = (\sum_{i=1}^n x_i)/n$. Consider the mechanism

$$Z_D = \begin{cases} 1, & \text{with probability } q(D), \\ 0, & \text{with probability } 1 - q(D). \end{cases}$$

   Determine the approximate privacy parameters for this mechanism.

2. (**25 points**) Let D be dataset of size n and q be a counting query. Consider the *uniform* mechanism which adds uniform noise to q_D, that is

$$Z_D = q_D + N,$$

   where $N \sim \texttt{Uniform}[-\lambda, \lambda]$ is uniformly distributed on the interval $[-\lambda, \lambda]$ for some $\lambda > 0$. How large must λ be to satisfy (ε, δ)DP? When $\delta < \frac{1}{n}$, will this mechanism produce useful information?

3. (**Bonus 25 points**) Suppose dataset $D = (X_1, \ldots, X_n)$ is a dataset consisting of n i.i.d. random variables drawn from Bernoulli(p) with a given value of p. Moreover, suppose $\mathsf{M} : \{0,1\}^n \to \mathcal{Y}$ is an (ε, δ)-DP mechanism and $A : \mathcal{Y} \to \{0,1\}^n$ is an adversary that seeks to reconstruct the dataset D from the output of M. Prove that the expected fraction of bits (i.e., coordinates) that the adversary successfully reconstructs is not much larger than the trivial bound of $\max\{p, 1-p\}$ (which can be achieved by guessing the all-zeroes or all-ones dataset). Specifically:

$$\mathbb{E}\left[\frac{\#\{i \in \{1,2,\ldots,n\} : A(\mathsf{M}(D))_i = X_i\}}{n}\right] \leq e^{\varepsilon} \cdot \max\{p, 1-p\} + \delta.$$

   Here by $A(\mathsf{M}(D))_i$, we mean the ith coordinate of $A(\mathsf{M}(D))$.