Privacy in Statistics and Machine Learning — Spring 2021
Homework 2: Due Friday, March 5, 2021

Adam Smith and Jonathan Ullman

Collaboration and Honesty Policy Reminder: Collaboration in the form of discussion is allowed. However, all forms of cheating (copying parts of a classmate's assignment, plagiarism from papers or old posted solutions) are NOT allowed. A rough rule of thumb: you should be able to walk away from a discussion of a homework problem with no notes at all and write your solution on your own. Finding answers to problems on the Web or from other outside sources (these include anyone not enrolled in the class) is forbidden.

- *You must write up each problem solution by yourself without assistance, even if you collaborate with others to solve the problem.*
- You must identify your collaborators. If you did not work with anyone, you should write "Collaborators: none."
- Asking and answering questions in every forum the class provides (on Piazza, in class, and in office hours) is encouraged!
- Even though look up answers is forbidden, using the web to find alternative explanations of concepts you need for the homework *is* allowed, and encouraged. For example, you can look up background on probability and linear algebra, documentation for particular programming languages, etc.

Problems to be handed in

1. **Medians.** Suppose we want to find the median of a list of real numbers $\mathbf{x} = (x_1, ..., x_n)$ that lie in the set $\{1, ..., R\}$.

   Consider an instantiation of the exponential mechanism based on the following score function: For every $y \in \{1, ..., R\}$, let

   $$q(y; \mathbf{x}) = -\left| \sum_{i=1}^{n} sign(y - x_i) \right|$$

   where

   $$sign(z) = \begin{cases} 1 & \text{if } z > 0\,, \\ 0 & \text{if } z = 0\,, \\ -1 & \text{if } z < 0\,. \end{cases}$$

   Note that this score is 0 exactly when y is a valid median for $\mathbf{x}$.

   (a) Show that q has sensitivity at most 1 when neighboring data sets are allowed to differ by the insertion or deletion of one entry.

   (b) Let A_ε be the algorithm one gets by instantiating the exponential mechanism with score q, parameter ε and output set $\mathcal{Y} = \{1, ..., R\}$. Show that there is a constant $c > 0$ such that: for every data set $\mathbf{x}$, for every R and $\varepsilon < 1$, and for every $\beta \in (0, 1)$, the probability that $A_\varepsilon(\mathbf{x})$ samples a value y with $|rank_{\mathbf{x}}(y) - n/2| > c \cdot \frac{\ln(R)+\ln(1/\beta)}{\varepsilon}$ is at most β. Here $rank_{\mathbf{x}}(y) \in \{0, 1, ..., n\}$ is the position y would have in the sorted order of $\mathbf{x}$.

[*Hint:* How does $rank_{\mathbf{x}}(\cdot)$ relate to $q(\cdot; \mathbf{x})$? Look at the ratio between the probability mass of a true median and the probability mass of an element with very low or high rank.]

2. **Implementing the Median Algorithm.** Implement the report-noisy-max version of this algorithm. Your code should take a set x of real values as input along with parameters R and ε. The first step should be to round all entries to the nearest integer in $\{1, ..., R\}$. The output should be a single real number.

   Test your code on data drawn from the following distributions, using $\varepsilon = 0.1$ and $n \in \{50, 100, 500, 2000, 10000\}$:

   - Gaussian: $\mathcal{N}(R/4, R^2/10)$ for $R \in \{100, 1000, 10000\}$.
   - Poisson: $Poi(50)$ for $R \in \{100, 1000, 10000\}$.
   - Bimodal: $R = 1,000$; each data point uniform over two values $\{R - k, R + k\}$ for $k = 10, 100, 200$.

   For each setting of parameters (data distribution, R and n), sample 50 data sets from the distribution, and run the algorithm 10 times on each. Collect the following statistics:

   - Average error in rank: how far from $n/2$ is the rank of the output?
   - Standard deviation of error in rank
   - Standard deviation of error in rank *among runs on the same data set.*

3. **Report Noisy Max.** Prove that report noisy max with Laplace noise is differentially private (Exercise 2.1 from the notes for Lecture 6).

4. **Histograms.** Consider the following algorithm for releasing histograms.

Algorithm 1: Stable Histogam$(\mathbf{x}; \varepsilon, \delta)$

Input: $\mathbf{x}$ is a multi-set of values in $\mathcal{U}$.

1 **for** *every* $z \in \mathcal{U}$ *that appears in* $\mathbf{x}$ **do** $\tilde{c}_z = \#\{i : x_i = z\} + \mathsf{Lap}(1/\varepsilon)$s

2 Release the set of pairs $\{(z, \tilde{c}_z) : \tilde{c}_z > \tau\}$ where $\tau = 1 + \frac{\ln(1/\delta)}{\varepsilon}$.

   (a) Show that for any domain $\mathcal{U}$, Algorithm 1 is (ε, δ)-differentially private when neighboring data sets are allowed to differ by the insertion or deletion of one value.

   Hint: The delicate part of this result is that we add noise only to counts of non-empty bins. (For example, if we were counting how many people live on each square mile of land in Alaska, most of the bins would be empty, but others would have lots of people.) There are two kinds of adjacent data sets: those where the set of nonempty bins changes, and those where it does not. You may need the following simple concentration bound for Laplace random variables: If $Y \sim \mathsf{Lap}(\lambda)$, then for every $t > 0$, we have $\Pr(Y > \lambda t) \leq \frac{1}{2}\exp(-t)$.

   (b) Prove that the Stable Histograms algorithm is not $(\varepsilon', 0)$ differentially private for any finite positive value ε'. [*Hint*: Give two neighboring data sets and a histogram y such that y is a possible output for only one of the two data sets.]