Privacy in Statistics and Machine Learning — Spring 2021
In-class Exercises for Lecture 9 (Approximate Differential Privacy)
February 25 & 26, 2021

Adam Smith and Jonathan Ullman

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

1. Consider the following mechanism NS_δ. On input $\mathbf{x} = (x_1, ..., x_n)$, for each i from 1 to n, it generates

$$Y_i = \begin{cases} (i, x_i) & \text{w.p. } \delta, \\ \perp & \text{w.p. } 1-\delta. \end{cases} \tag{1}$$

   and outputs $(Y_1, \ldots, Y_n)$. Here $\perp$ is just a special symbol meaning "no information." Show that NS_δ satisfies $(0, \delta)$-DP. ***Discussion Topic:*** Given that NS_δ satisfies $(0, \delta)$-DP, do you think that (ε, δ)-DP is a suitable definition of privacy when $\delta \geq 1/n$?

2. Suppose we add *uniform* noise to a count query, that is, we release $A_\lambda(\mathbf{x}) = f(\mathbf{x}) + U_{[-\lambda,\lambda]}$ where f counts how many records staistfy some condition, and $U_{[-\lambda,\lambda]}$ is uniformly distributed in the interval $[-\lambda, \lambda]$. How large must λ be to satisfy (ε, δ)-DP? Do both ε and δ matter in setting λ? When $\delta < 1/n$, will this mechanism produce useful information?

3. Show that for every $\varepsilon, \delta > 0$, there is a mechanism A such that

   (a) for every pair of neighboring data sets $\mathbf{x}, \mathbf{x}'$, for every $y \in \mathcal{Y}$,

$$\mathbb{P}\left(A(\mathbf{x}) = y\right) \leq e^\varepsilon \cdot \mathbb{P}\left(A(\mathbf{x}') = y\right) + \delta$$

   (that is, the DP condition holds for singleton events $E = \{y\}$), but

   (b) A does *not* satisfy (ε', δ')-DP for *any* finite ε' and $\delta' < 1$.

   To avoid technical issues with continuous ranges and density functions, your mechanism should output a discrete value, so that the probabilities in the above description are all well defined. ***Hint:*** Try starting with some mechanism A that is blatantly not private, then modify it to make it satisfy the first condition.

4. For $\lambda, \tau > 0$, the *truncated Laplace distribution* $\mathsf{Lap}(\lambda, \tau)$ is defined by the density function

$$p_{\lambda,\tau}(y) = \begin{cases} \frac{1}{Z} e^{-|y|/\lambda} & |y| \leq \tau \\ 0 & |y| > \tau \end{cases} \tag{2}$$

   where $Z = \int_{-\tau}^{\tau} e^{-|y|/\lambda} dy$ is a normalizing constant. Prove that for any real-valued statistic f, the mechanism

$$A(\mathbf{x}) = f(\mathbf{x}) + \mathsf{Lap}\left(\frac{\Delta}{\varepsilon}, \frac{\Delta}{\varepsilon} \cdot \log(1/\delta)\right) \tag{3}$$

   satisfies $(\varepsilon, O(\delta))$-differential privacy, where Δ is the global sensitivity of f.

5. (*) Prove the basic (non-adaptive) composition theorem for (ε, δ)-differential privacy: if A_1 is a mechanism that is $(\varepsilon_1, \delta_1)$-DP and A_2 is a mechanism that is $(\varepsilon_2, \delta_2)$-DP, then the composed mechanism $A(\mathbf{x}) = (A_1(\mathbf{x}), A_2(\mathbf{x}))$ is $(\varepsilon_1 + \varepsilon_2, \delta_1 + \delta_2)$-DP.

   Note: You may find that your first attempt yields a weaker bound, namely that the composed protocol is $(\varepsilon_1 + \varepsilon_2, \delta_1 + e^{\varepsilon_1}\delta_2)$-DP (or something similar). That's ok, but see if you can prove the tighter bound above.

6. **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.]