Lecture: Defining DP (Contd.) and the Laplace Mechanism

Last time, we provided the definition of differential privacy and discussed a simple mechanism, Randomized Response, which allows for a DP release of an entire 0/1-valued dataset. Furthermore, we discovered a fundamental tension b/w achieving DP (say, via RR), and maximizing "utility," which was then measured via an ℓ_2-norm (e.g., how close is an estimator of $\sum_{i=1}^{N} x_i$, obtained as a function of $(Y_1, \dots, Y_N)$, to $\sum_{i=1}^{N} x_i$?).

In fact, via a key property of DP, the estimator of $\sum_{i=1}^{N} x_i$ that you had constructed, is by itself DP [we'll see this soon!]. But first, we shall carry out some sanity checks:

(i) If $\epsilon = 0$, in the defn. of DP, then $\forall S \subseteq \mathcal{Y}$ and $\underline{x} \sim \underline{x}'$, we require

$\underline{x}, \underline{x}'$ are neighbours

$$\Pr[A(\underline{x}) \in S] - \Pr[A(\underline{x}') \in S] \le 0,$$

or equivalently,

$$\overbrace{\sup_{S \subseteq \mathcal{Y}} \left(\Pr[A(\underline{x}) \in S] - \Pr[A(\underline{x}') \in S] \right)}^{(*)} \le 0.$$

The quantity in (*) is called the "total variational distance" b/w $A(\underline{x})$ and $A(\underline{x}')$, denoted as $d_{TV}(A(\underline{x}), A(\underline{x}'))$. Via standard arguments, we have that (assuming $\mathcal{Y}$ is discrete)

$$d_{TV}(A(\underline{x}), A(\underline{x}')) = \sum_{y:\, P_{A(\underline{x})}(y) \geq P_{A(\underline{x}')}(y)} \left(P_{A(\underline{x})}(y) - P_{A(\underline{x}')}(y) \right)$$

$$= \sum_{y:\, P_{A(\underline{x}')}(y) \geq P_{A(\underline{x})}(y)} \left(P_{A(\underline{x}')}(y) - P_{A(\underline{x})}(y) \right)$$

$$= \frac{1}{2} \sum_{y \in \mathcal{Y}} \left| P_{A(\underline{x})}(y) - P_{A(\underline{x}')}(y) \right| .$$

<u>HW*</u>: Attempt to prove the equalities above.

Hence, 0-DP requires that $d_{TV}(A(\underline{x}), A(\underline{x}')) = 0, \ \forall\, \underline{x} \sim \underline{x}'$

$$\Updownarrow$$

$$P_{A(\underline{x})} = P_{A(\underline{x}')} ,$$

which is a very strong notion of privacy.

(ii) If $\epsilon = \infty$, then we have no privacy, i.e., <u>any</u> (fairly regular) algorithm is ∞-DP.

Moreover, the following lemma holds.

Lemma 1: (i) If $A(\underline{x})$ is a continuous RV, for all $\underline{x}$, with associated pdf $f_{A(\underline{x})}$, then A is ϵ-DP iff $f_{A(\underline{x})}(t) \leq e^{\epsilon} \cdot f_{A(\underline{x}')}(t)$, for all $\underline{x} \sim \underline{x}'$, and all $t \in \mathcal{Y}$.

(ii) If $A(\underline{x})$ is a discrete RV, for all $\underline{x}$, then A is ϵ-DP iff
$$P_{A(\underline{x})}(y) \leq P_{A(\underline{x}')}(y), \text{ for all } \underline{x} \sim \underline{x}', \text{ and all } y \in \mathcal{Y}.$$

Proof: HW.

Achieving DP for aggregate statistics: The Laplace mechanism

Last time, we had encountered a DP mechanism (RR) for releasing an entire 0/1-valued dataset. In practice, however, a user/analyst is only interested in "aggregate statistics" (or simple functions) of a dataset $\mathcal{D} = (x_1, \ldots, x_N)$. We now present a "silver bullet" mechanism, the Laplace mechanism, for the DP release of such a statistic $f : \mathbb{D}_N \to \mathbb{R}^d$.

Def 1: The (global) sensitivity of $f : \mathbb{D}_N \to \mathbb{R}^d$ is defined as
$$\Delta_f \triangleq \max_{\underline{x} \sim \underline{x}' \in \mathbb{D}_N} \| f(\underline{x}) - f(\underline{x}') \|_1 .$$

Example: (i) $f = \frac{1}{N} \sum_{i=1}^{N} x_i$, with $x_i \in [a, b]$, $\forall\, i \in [N]$. We have
$$\Delta_f = \frac{b-a}{N} .$$

(ii) $\mathbb{D}_N = \mathcal{X}^N$, where $\mathcal{X} = \{0, 1, \dots, q-1\}$. We wish to release the "histogram" of $\underline{x} \in \mathbb{D}_N$, denoted as $f_{\text{Hist}}(\underline{x}) = \left(\sum_{i=1}^{N} 1\{x_i = a\} : a \in \mathcal{X}\right)$.

We then have $\Delta_f = 2$. [Why?]

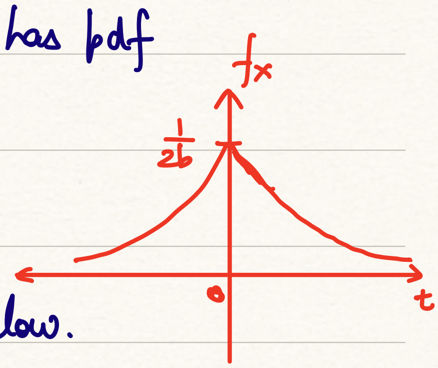
(iii) $\mathbb{D}_N = \mathcal{X}^N$, as above. What is the sensitivity of f = Median, i.e.,

$$f(\underline{x}) = \max\left\{a \in \mathcal{X} : \sum_{i=1}^{N} 1\{x_i \le a\} \le N/2\right\},$$

for $\underline{x} \in \mathcal{X}^N$? Answer: $\Delta_f = q-1$. [Why?]

Before we define the Laplace mechanism, we take a look at the (zero-mean) Laplace distribution $\text{Lap}(b)$. We have $X \sim \text{Lap}(b)$ if $f_X(t) = \frac{1}{2b} e^{-|t|/b}$, for $t \in \mathbb{R}$. The Laplace distribution with mean μ has pdf

$$f_{\text{Lap}(\mu, b)}(t) = \frac{1}{2b} \exp\left(-\frac{|t-\mu|}{b}\right).$$



We collect some properties of the Laplace distribution below.

<u>Proposition 1</u>: If $Z \sim \text{Lap}(b)$, we have

(a) $\mathbb{E}[|Z|] = b$,

(b) $\mathbb{E}[Z^2] = 2b^2$

(c) For every $t > 0$, $\Pr[|Z| > tb] \le \exp(-t)$.

Furthermore, if $\underline{Z} = (Z_1, \dots, Z_d) \overset{\text{iid}}{\sim} \text{Lap}(b)$, with $M = \max(|Z_1|, \dots, |Z_d|)$, then

(a) $\mathbb{E}[\|Z\|_1] = db$ [Why?]

(b) $\Pr[M \geq tb] \leq d \cdot \exp(-t)$ [Why?]

(c) $\mathbb{E}[M] \leq b(\ln d + 1)$. [How would you bound the LHS, using (b)?]

The Laplace mechanism simply adds (an independent) Laplace-distributed "noise" RV to the statistic $f(\mathcal{D})$. In particular, the Laplace mechanism M_ϵ^{Lap} returns (for $f: \mathbb{D}_N \to \mathbb{R}^d$),

$$M_\epsilon^{Lap}(\mathcal{D}) = f(\mathcal{D}) + \underline{Z},$$

where $\underline{Z} = (Z_1, \dots, Z_d) \overset{iid}{\sim} \mathrm{Lap}\left(\frac{\Delta_f}{\epsilon}\right)$.

<u>Theorem 1</u>: The Laplace mechanism M_ϵ^{Lap} is ϵ-DP.

<u>Proof</u>: Fix two neighbouring datasets $\underline{x}, \underline{x}'$. Let $f_{\underline{x}}$ and $f_{\underline{x}'}$ denote the pdfs of $M_\epsilon^{Lap}(\underline{x})$ and $M_\epsilon^{Lap}(\underline{x}')$, respectively. Via Lemma 1, it suffices to show that $f_{\underline{x}}(\underline{z})$ and $f_{\underline{x}'}(\underline{z})$ are at most a multiplicative factor of e^ϵ away. Let $\Delta = \Delta_f$.

Now, note that

$$\frac{f_{\underline{x}}(\underline{z})}{f_{\underline{x}'}(\underline{z})} = \frac{\prod_{i=1}^{d} \frac{\epsilon}{2\Delta} \exp\left(-\frac{\epsilon}{\Delta} |z_i - (f(\underline{x}))_i|\right)}{\prod_{i=1}^{d} \frac{\epsilon}{2\Delta} \exp\left(-\frac{\epsilon}{\Delta} |z_i - (f(\underline{x}'))_i|\right)}$$

$$= \exp\left(-\frac{\epsilon}{\Delta}\left(\|\underline{z}-f(\underline{x})\|_1 - \|\underline{z}-f(\underline{x}')\|_1\right)\right)$$

$$\leq \exp\left(\frac{+\epsilon}{\Delta}\,\|f(\underline{x})-f(\underline{x}')\|_1\right)$$ [Via the Δ-ineq for $\|\cdot\|_1$]

$$\leq \exp\left(\frac{\epsilon}{\Delta}\cdot\Delta\right) = e^{\epsilon}. \qquad \boxtimes$$

HW: Use Proposition 1 to establish the following "utility" bound on the Laplace mechanism, when $f(\underline{x}) = \frac{1}{N}\sum_{i=1}^{N} x_i$.

$$\left(\mathbb{E}\left[\left(M_{\epsilon}^{\text{Lap}}(\underline{x}) - f(\underline{x})\right)^2\right]\right)^{1/2} = \frac{\sqrt{2}}{n\epsilon}.$$

Clearly, from the above exercise, we see that setting $\epsilon \leq 1/n$ yields an RMSE of at least $\sqrt{2}$, when $|f(\underline{x})| \leq 1$, which is a very poor utility guarantee!