Differential privacy

**Definition**
Given $\epsilon, \delta \geq 0$, a probabilistic query $Q: X^n \rightarrow R$ is $(\epsilon, \delta)$-differentially private iff for all adjacent database $b_1, b_2$ and for every $S \subseteq R$:

$$\Pr[Q(b_1) \in S] \leq \exp(\epsilon) \Pr[Q(b_2) \in S] + \delta$$
Noise on Input vs Noise on Output

Input

Local Computation

Global Computation

\[
q(d_1) \\
\vdots \\
q(d_n)
\]

\[
\frac{1}{n} \sum_{i=0}^{n} q(d_i)
\]
Algorithm 2 Pseudo-code for the Laplace Mechanism

1: function \textsc{LapMech}(D, q, \epsilon)
2: \hspace{1em} Y \overset{\$}{\leftarrow} \text{Lap}(\frac{\Delta q}{\epsilon})(0)
3: \hspace{1em} return \; q(D) + Y
4: end function
Algorithm 14 Pseudo-code for the Gaussian Mechanism

1: function \textsc{GaussMech}(D, q, \epsilon)
2: \quad Y \leftarrow \text{Gauss}(0, \frac{2\ln(\frac{1.25}{\delta})(\Delta_2 q)^2}{\epsilon^2})
3: \quad \text{return } q(D) + Y
4: end function
Exponential Mechanism:

\[ M_E(x, u, \mathcal{R}) \]

return \( r \in \mathcal{R} \) with prob.

\[
\frac{\exp\left( \frac{\varepsilon u(x, r)}{2\Delta u} \right)}{\sum_{r' \in \mathcal{R}} \exp\left( \frac{\varepsilon u(x, r')}{2\Delta u} \right)}
\]
Question: What do all of these algorithms have in common?
Global Sensitivity

**Definition 1.8** (Global sensitivity). The *global sensitivity* of a function $q : \mathcal{X}^n \rightarrow \mathbb{R}$ is:

$$\Delta q = \max \left\{ |q(D) - q(D')| \mid D \sim_1 D' \in \mathcal{X}^n \right\}$$
Exponential Mechanism:

\[ M_E(x, u, \mathcal{R}) \]

return \( r \in \mathcal{R} \) with prob.

\[
\frac{\exp\left(\frac{\varepsilon u(x, r)}{2\Delta u}\right)}{\sum_{r' \in \mathcal{R}} \exp\left(\frac{\varepsilon u(x, r')}{2\Delta u}\right)}
\]

where

\[
\Delta u = \max_{r \in \mathcal{R}} \max_{x \sim_1 y} \left| u(x, r) - u(y, r) \right|
\]
Global sensitivity

**Question:** What is an example of a query with excessively high global sensitivity?
Median

Let’s consider the median \( \text{Med}(D) \) for \( D \in \{0, \ldots, 100\}^n \)

**Question:** What is the sensitivity of Med?

Let’s consider the datasets:

\[
(0,0,0,0,0,100,100,100,100,100,100,100) \\
and \\
(0,0,0,0,0,0,100,100,100,100,100,100)
\]

This is the worst case, but adding noise proportional to this destroys utility.
**Definition 1.8** (Global sensitivity). The *global sensitivity* of a function $q : \mathcal{X}^n \rightarrow \mathbb{R}$ is:

$$\Delta q = \max \left\{ |q(D) - q(D')| \ \bigg| \ D \sim_1 D' \in \mathcal{X}^n \right\}$$

**Definition 1.14** (Local sensitivity). The *local sensitivity* of a function $q : \mathcal{X}^n \rightarrow \mathbb{R}$ at $D \in \mathcal{X}^n$ is:

$$\ell \Delta q(D) = \max \left\{ |q(D) - q(D')| \ \bigg| \ D \sim_1 D', \ D' \in \mathcal{X}^n \right\}$$
Calibrating noise to the local sensitivity

We may add noise proportional to the local sensitivity (LS).

Unfortunately, this does not guarantee privacy.

Suppose that for a given D we have LS(D)=0 but that we also have D~D’ with LS(D’)=10^9.

We will see that we can do anyway better than GS.
Some methods

- Smooth Sensitivity
- Propose-Test-Release
- Releasing Stable Values
Smooth Sensitivity

Definition 2.2 (Smooth sensitivity). For $\beta > 0$, the $\beta$-smooth sensitivity of $f$ is

$$S^*_{f,\beta}(x) = \max_{y \in D^n} \left( L S_f(y) \cdot e^{-\beta d(x,y)} \right).$$

Definition 2.1 (A Smooth Bound on $LS$). For $\beta > 0$, a function $S : D^n \to \mathbb{R}^+$ is a $\beta$-smooth upper bound on the local sensitivity of $f$ if it satisfies the following requirements:

$$\forall x \in D^n : \quad S(x) \geq L S_f(x); \quad (1)$$

$$\forall x, y \in D^n, d(x, y) = 1 : \quad S(x) \leq e^\beta \cdot S(y). \quad (2)$$

What kind of noise can we add?

[Nissim, Raskhodnikova, Smith '06]
Definition 2.5 (Admissible Noise Distribution). A probability distribution on $\mathbb{R}^d$, given by a density function $h$, is $(\alpha, \beta)$-admissible (with respect to $\ell_1$) if, for $\alpha = \alpha(\epsilon, \delta)$, $\beta = \beta(\epsilon, \delta)$, the following two conditions hold for all $\Delta \in \mathbb{R}^d$ and $\lambda \in \mathbb{R}$ satisfying $||\Delta||_1 \leq \alpha$ and $|\lambda| \leq \beta$, and for all measurable subsets $S \subseteq \mathbb{R}^d$:

**Sliding Property:**

$$
\Pr_{Z \sim h} \left[ Z \in S \right] \leq e^{\frac{\epsilon}{2}} \cdot \Pr_{Z \sim h} \left[ Z \in S + \Delta \right] + \frac{\delta}{2}.
$$

**Dilation Property:**

$$
\Pr_{Z \sim h} \left[ Z \in S \right] \leq e^{\frac{\epsilon}{2}} \cdot \Pr_{Z \sim h} \left[ Z \in e^\lambda \cdot S \right] + \frac{\delta}{2}.
$$

[Nissim, Raskhodnikova, Smith '06]
Lemma 2.6. Let $h$ be an $(\alpha, \beta)$-admissible noise probability density function, and let $Z$ be a fresh random variable sampled according to $h$. For a function $f : D^n \to \mathbb{R}^d$, let $S : D^n \to \mathbb{R}$ be a $\beta$-smooth upper bound on the local sensitivity of $f$. Then algorithm $A(x) = f(x) + \frac{S(x)}{\alpha} \cdot Z$ is $(\epsilon, \delta)$-differentially private.

For two neighbor databases $x$ and $y$, the output distribution $A(y)$ is a shifted and scaled version of $A(x)$. The sliding and dilation properties ensure that $\Pr[A(x) \in S]$ and $\Pr[A(y) \in S]$ are close for all sets $S$ of outputs.

[Nissim, Raskhodnikova, Smith '06]
Admissible Noise

Adding noise $O(SS_q^\varepsilon(x)/\varepsilon)$ (according to a Cauchy distribution) is sufficient for $\varepsilon$-differential privacy.

Laplace and Gauss give $(\varepsilon,\delta)$-DP

Computing the Smooth Sensitivity can be intractable.

[Nissim, Raskhodnikova, Smith ’06]
**Propose-test-release** Given $q : \mathcal{X}^n \to \mathbb{R}$, $\epsilon, \delta, \beta \geq 0$

1. Propose a target bound $\beta$ on local sensitivity.
2. Let $\hat{d} = d(x, \{x' : \text{LS}_q(x') > \beta\}) + \text{Lap}(1/\epsilon)$, where $d$ denotes Hamming distance.
3. If $\hat{d} \leq \ln(1/\delta)/\epsilon$, output $\perp$.
4. If $\hat{d} > \ln(1/\delta)/\epsilon$, output $q(x) + \text{Lap}(\beta/\epsilon)$. 
**Laplace Mechanism**

**Accuracy Theorem:** \( \text{let } r = \text{LapMech}(D, q, \epsilon) \)

\[
\Pr \left[ |q(D) - r| \geq \left( \frac{\Delta q}{\epsilon} \right) \ln \left( \frac{1}{\beta} \right) \right] = \beta
\]

\[
\text{Lap}(b, \mu)(X) = \frac{1}{2b} \exp \left( - \frac{|\mu - X|}{b} \right)
\]

\[
\Pr \left[ |X| \geq b t \right] = \exp(-t)
\]
Proposition 3.2 (propose-test-release [33]). For every query \( q : X^n \to \mathbb{R} \) and \( \varepsilon, \delta, \beta \geq 0 \), the above algorithm is \((2\varepsilon, \delta)\)-differentially private.

Proof. Consider any two neighboring datasets \( x \sim x' \). Because of the Laplacian noise in the definition of \( \hat{d} \) and the fact that Hamming distance has global sensitivity at most 1, it follows that

\[
\Pr[\mathcal{M}(x) = \bot] \in [e^{-\varepsilon} \cdot \Pr[\mathcal{M}(x') = \bot], e^{\varepsilon} \cdot \Pr[\mathcal{M}(x') = \bot]].
\]

(3)

Case 1: \( \text{LS}_q(x) > \beta \). In this case, \( d(x, \{ x'' : \text{LS}_q(x'') > \beta \}) = 0 \), so the probability that \( \hat{d} \) will exceed \( \ln(1/\delta)/\varepsilon \) is at most \( \delta \). Thus, for every set \( T \subseteq \mathbb{R} \cup \{ \bot \} \), we have:

\[
\Pr[\mathcal{M}(x) \in T] \leq \Pr[\mathcal{M}(x) \in T \cap \{ \bot \}] + \Pr[\mathcal{M}(x) \neq \bot] \\
\leq e^{\varepsilon} \cdot \Pr[\mathcal{M}(x') \in T \cap \{ \bot \}] + \delta \\
\leq e^{\varepsilon} \cdot \Pr[\mathcal{M}(x') \in T] + \delta,
\]

where the second inequality follows from (3), noting that \( T \cap \{ \bot \} \) equals either \( \{ \bot \} \) or \( \emptyset \).

Case 2: \( \text{LS}_q(x) \leq \beta \). In this case, \( |q(x) - q(x')| \leq \beta \), which in turn implies the \((\varepsilon, 0)\)-indistinguishability of \( q(x) + \text{Lap}(\beta/\varepsilon) \) and \( q(x') + \text{Lap}(\beta/\varepsilon) \). Thus, by (3) and Basic Composition, we have \((2\varepsilon, 0)\)-indistinguishability overall. \( \square \)