CS2080: Applied Privacy for Data Science
Introduction to Differential Privacy (cont.)

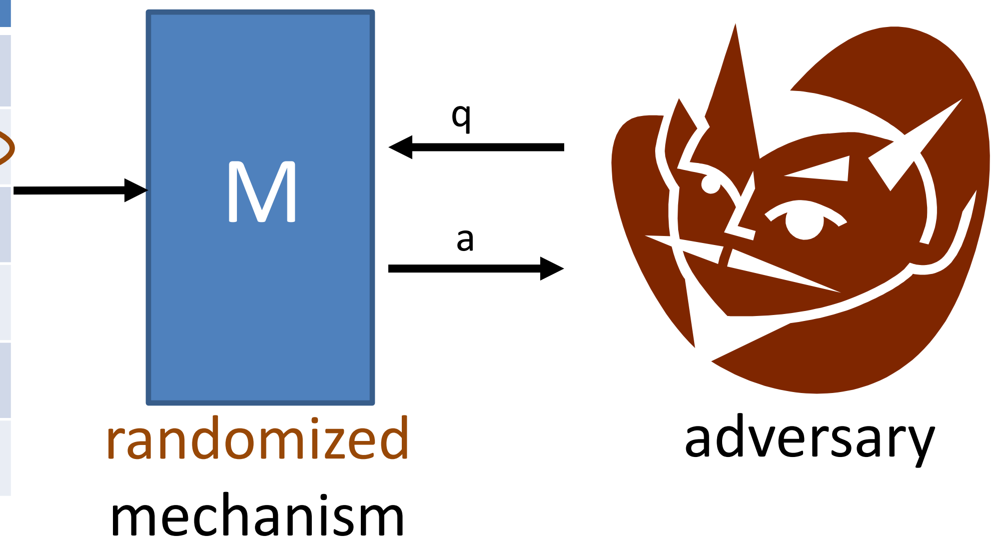
School of Engineering & Applied Sciences
Harvard University

February 12, 2025

DP for one query/release

[Dwork-McSherry-Nissim-Smith ’06]

Sex	Blood	⋯	HIV?
F	B	⋯	Y
F	A		Y
M	O	⋯	N
M	O	⋯	Y
F	A	⋯	N
M	B	⋯	Y



Def: M is ε-DP if for all x, x' differing on one row, and all q

$\forall$ sets T, $$\Pr[M(x,q) \in T] \le e^{\varepsilon} \cdot \Pr[M(x',q) \in T]$$

(Probabilities are (only) over the randomness of M.)

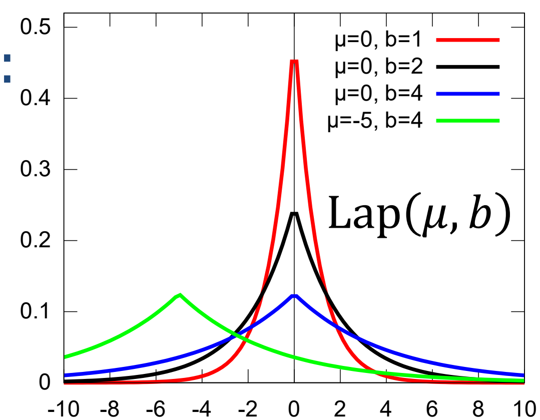
The Laplace Mechanism

[Dwork-McSherry-Nissim-Smith '06]

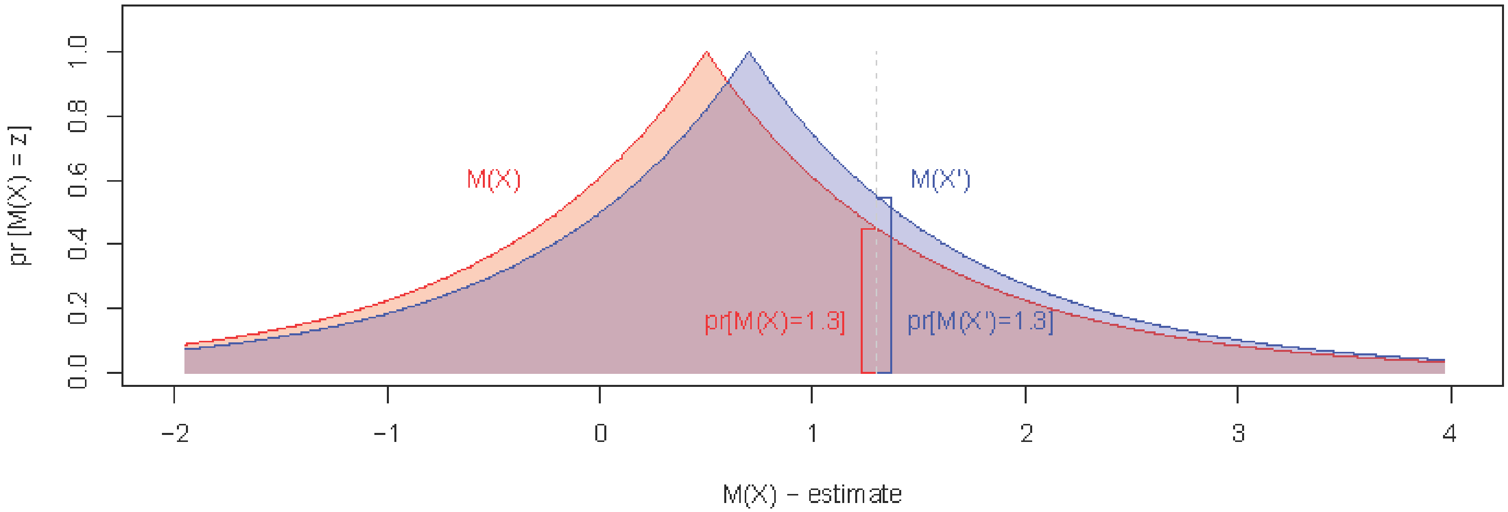
- Let $\mathcal{X}$ be a data universe, and $\mathcal{X}^n$ a space of datasets.
  - This is the Bounded DP setting: n known and public.
- For $x, x' \in \mathcal{X}^n$, write $x \sim x'$ if x and x' differ on ≤ 1 row.
- For a query $q : \mathcal{X}^n \to \mathbb{R}$, the global sensitivity is

$$\Delta q = \mathrm{GS}_q = \max_{x \sim x'} |q(x) - q(x')|.$$

- The Laplace distribution with scale b, $\mathrm{Lap}(b)$:
  - Has density function $f(y) = e^{-|y|/b}/2b$.
  - Mean 0, standard deviation $\sqrt{2} \cdot b$.



Theorem: $M(x, q) = q(x) + \mathrm{Lap}(\Delta q/\varepsilon)$ is ε-DP.



Two Laplace distributions, for two adjacent datasets x and x'. The definition of ϵ-differential privacy requires the ratio of $M(x)/M(x')$ is not greater than e^{ϵ} for all points along the x-axis. Thus for any realized output z (for example here, $z = 1.3$) we can not determine that x or x' were more likely to have produced z.

Calculating Global Sensitivity

1. $\mathcal{X} = \{0,1\}, q(x) = \sum_{i=1}^{n} x_i, \Delta q =$
2. $\mathcal{X} = \mathbb{R}, q(x) = \sum_{i=1}^{n} x_i, \Delta q =$
3. $\mathcal{X} = [0,1], q(x) = \text{mean}(x_1, x_2, \dots, x_n), \Delta q =$
4. $\mathcal{X} = [0,1], q(x) = \text{median}(x_1, x_2, \dots, x_n), \Delta q =$
5. $\mathcal{X} = [0,1], q(x) = \text{variance}(x_1, x_2, \dots, x_n), \Delta q =$

Q: for which of these queries is the Laplace Mechanism "useful"?

A:

Properties of the Definition

- Suffices to check pointwise: M is ε-DP if and only if
  $\forall x \sim x' \; \forall q \; \forall y \; \Pr[M(x,q) = y] \leq e^{\varepsilon} \cdot \Pr[M(x',q) = y]$.

- Preserved under post-processing: If M is ε-DP and f is any function, then $M'(x,q) = f(M(x,q))$ is ε-DP.

- (Basic) composition: If M_i is ε_i-DP for $i = 1, \ldots, k$, then
  $$M'\big(x, (q_1, \ldots, q_k)\big) = (M_1(x, q_1), \ldots, M_k(x, q_k))$$
  is $(\varepsilon_1 + \cdots + \varepsilon_k)$-DP
  - Use independent randomness for the k queries
  - Holds even if q_i's are chosen adaptively

Interpreting the Definition

- Whatever an adversary learns about me, it could have learned from everyone else's data.
- Mechanism cannot leak "individual-specific" information.
- Above interpretations hold regardless of adversary's auxiliary information or computational power.
- Protection against MIAs: let $X = (X_1, \ldots, X_n)$ be a r.v. distributed on $\mathcal{X}^n$ and $X_{-i} = (X_1, \ldots, X_{i-1}, \perp, X_{i+1}, \ldots, X_n)$ be X with Alice's data removed. Then for every MIA A,

$$\Pr\left[A\left(M(X)\right) = \text{"In"}\right] \leq e^{\varepsilon} \cdot \Pr\left[A\left(M(X_{-i})\right) = \text{"In"}\right]$$

Varying the Data Domain and Privacy Unit

- Unbounded DP (n not publicly known):
  - Datasets: multisets x from a data universe $\mathcal{X}$
    - Can represent as histogram $h_x: \mathcal{X} \to \mathbb{N}$, $h_x(i) =$ # copies of i
  - Adjacency: $x \sim x'$ if $|x \Delta x'| \leq 1$ (add/remove 1 record)
    - Equivalently $\sum_{i \in \mathcal{X}} |h_x(i) - h_{x'}(i)| \leq 1$

- Social Networks:
  - Datasets: graphs G
  - Adjacency: $G \sim G'$ if
    - differ by ≤ 1 edge (edge privacy), OR
    - differ by ≤ 1 node and incident edges (node privacy)

    Q: which is better for privacy?

Real Numbers Aren't

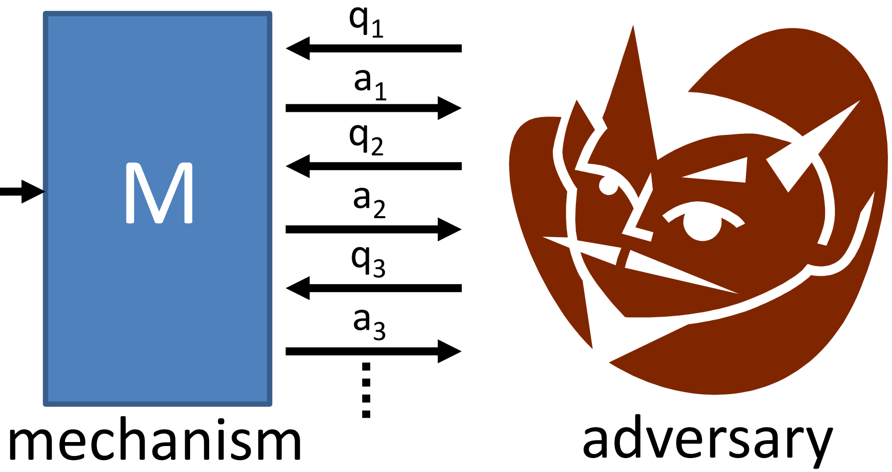
[Mironov `12]

- Digital computers don't manipulate actual real-numbers
  - Floating-point implementations of the Laplace Mechanism can have $M(x, q)$ and $M(x', q)$ disjoint → privacy violation!

- Solutions:
  - Round outputs of M to a discrete number (with care).
  - Or use the Geometric Mechanism:
    - Ensure that $q(x)$ is always an integer multiple of γ.
    - Define $M(x, q)$=$q(x) + \gamma \cdot \mathrm{Geo}(\Delta q / \gamma \varepsilon)$,
      where $\mathrm{Pr}[\mathrm{Geo}(b) = k] \propto \exp(-|k|/b)$ for $k \in \mathbb{Z}$.

DP for Interactive Mechanisms

Sex	Blood	⋯	HIV?
F	B	⋯	Y
F	**A**		**Y**
M	O	⋯	N
M	O	⋯	Y
F	A	⋯	N
M	B	⋯	Y



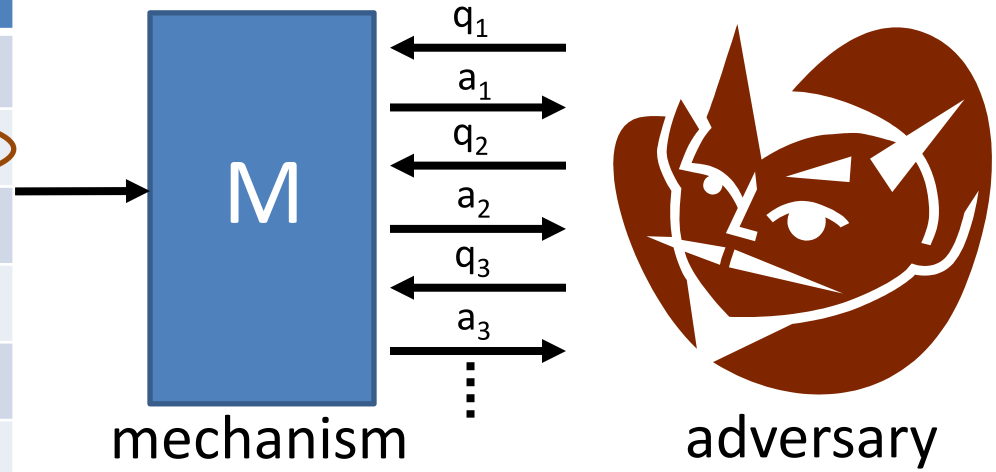
1st Attempt: for all $x \sim x'$, all $q_1, \ldots, q_t$, all T

$$\Pr[M(\underbrace{x, q_1, \ldots, q_t}) \in T] \leq e^{\varepsilon} \cdot \Pr[M(\underbrace{x', q_1, \ldots, q_t}) \in T]$$

vectors of answers $a_1, \ldots, a_t$

DP for Interactive Mechanisms

Sex	Blood	⋯	HIV?
F	B	⋯	Y
F	A		Y
M	O	⋯	N
M	O	⋯	Y
F	A	⋯	N
M	B	⋯	Y



Better: for all $x \sim x'$, all adversarial strategies A

$$\mathrm{View}_A\big(A \leftrightarrow M(x)\big) \approx_\varepsilon \mathrm{View}_A(A \leftrightarrow M(x'))$$

Everything A sees (its internal randomness & query answers)

Equivalently: $\forall\, A \;\; \Pr[A \text{ outputs "In" after interacting w}/M(x)]$
$\leq e^{\varepsilon} \cdot \Pr[A \text{ outputs "In" after interacting w}/M(x')]$

Composition as an Interactive Mechanism

privacy-loss budget ε

dataset x

If $\varepsilon_1 \leq \varepsilon$ and M_1 is ε_1-DP,
then $a_1 = M_1(x)$, $\varepsilon = \varepsilon - \varepsilon_1$
else $a_1 = \perp$

If $\varepsilon_2 \leq \varepsilon$ and M_2 is ε_2-DP,
then $a_2 = M_2(x)$, $\varepsilon = \varepsilon - \varepsilon_2$
else $a_2 = \perp$

⋮

$\mathrm{PureDPFilter}_{\varepsilon}(x)$

M_1, ε_1

a_1

M_2, ε_2

a_2

⋮

data analysts

Theorem: $\mathrm{PureDPFilter}_{\varepsilon}$ is an ε-DP interactive mechanism.

Privacy Budgeting

privacy budget ε

dataset x

If $\varepsilon_1 \le \varepsilon$ and M_1 is ε_1-DP,
then $a_1 = M_1(x)$, $\varepsilon = \varepsilon - \varepsilon_1$
else $a_1 = \perp$

If $\varepsilon_2 \le \varepsilon$ and M_2 is ε_2-DP,
then $a_2 = M_2(x)$, $\varepsilon = \varepsilon - \varepsilon_2$
else $a_2 = \perp$

M_1, ε_1

a_1

M_2, ε_2

a_2

data analysts

- To answer k queries, can set each $\varepsilon_i = \varepsilon/k$.
- More queries $\Rightarrow$ smaller ε_i $\Rightarrow$ less accuracy per query.
- Some tradeoff #queries vs. accuracy necessary. (Q: why?)

Composition for Algorithm Design

Composition and post-processing allow designing more complex differentially private algorithms from simpler ones.

Example: The “Statistical Query Model” for ML

- Many ML algorithms (e.g. stochastic gradient descent) can be described as sequence of low-sensitivity queries (e.g. averages) over the dataset, and can tolerate noisy answers to the queries.
- Can answer each query by adding Laplace noise.
- By composition and post-processing, trained model is DP and safe to output.

Group Privacy & Setting ε

- Proposition: If M is ε-DP for individuals, then it is $k\varepsilon$-DP for groups of k individuals. That is, if x and x' differ on at most k individuals, then
$$\forall T \;\; \Pr[M(x) \in T] \leq e^{k\varepsilon} \cdot \Pr[M(x') \in T]$$

- Q: what are examples of "groups" for which this is useful?

- Consequence: need $n \gg 1/\varepsilon$ for any reasonable utility.

- Typical recommendation for a "good" privacy guarantee: $.01 \leq \varepsilon \leq 1$.