# CS208: Applied Privacy for Data Science
# Foundations of Differential Privacy (cont.)

James Honaker & Salil Vadhan

School of Engineering & Applied Sciences
Harvard University

February 25, 2019

CRCS Center for Research on Computation and Society

# Goals of Differential Privacy

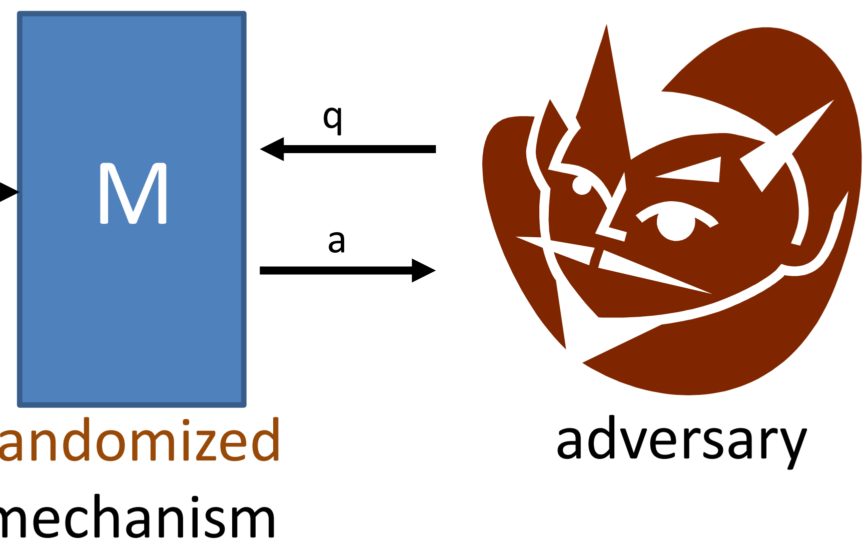
- Utility: enable "statistical analysis" of datasets
  - e.g. inference about population, ML training, useful descriptive statistics

- Privacy: protect individual-level data
  - against "all" attack strategies, auxiliary info.

# 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 $\varepsilon$-DP if for all D, D’ differing on one row, and all q

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

(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. (For now, we are treating $n$ as known and public.)
- For $x, x' \in \mathcal{X}^n$, write $x \sim x'$ if $x$ and $x'$ differ on at one row.
- For a query $q : \mathcal{X}^n \to \mathbb{R}$, the global sensitivity is
$$\mathrm{GS}_q = \max_{x \sim x'} |q(x) - q(x')|.$$
- The Laplace distribution with scale $s$, $\mathrm{Lap}(s)$:
  - Has density function $f(y) = e^{-|y|/s}/2s$.
  - Mean 0, standard deviation $\sqrt{2} \cdot s$.

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

# Calculating Global Sensitivity

1. $\mathcal{X} = \{0,1\}$, $q(x) = \sum_{i=1}^{n} x_i$, $\mathrm{GS}_q = 1$.
2. $\mathcal{X} = \mathbb{R}$, $q(x) = \sum_{i=1}^{n} x_i$, $\mathrm{GS}_q = \infty$. (useless)
3. $\mathcal{X} = [0,1]$, $q(x) = \mathrm{mean}(x_1, x_2, \ldots, x_n)$, $\mathrm{GS}_q = 1/n$.
4. $\mathcal{X} = [0,1]$, $q(x) = \mathrm{median}(x_1, x_2, \ldots, x_n)$, $\mathrm{GS}_q = 1$. (useless)
5. $\mathcal{X} = [0,1]$, $q(x) = \mathrm{variance}(x_1, x_2, \ldots, x_n)$, $\mathrm{GS}_q < 1/n$.

# Proof that the Laplace Mechanism is Differentially Private

# 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 value (with care).
  - Or use the Geometric Mechanism:
    - Ensure that $q(x)$ is always an integer multiple of $g$.
    - Define $M(x,q) = q(x) + g \cdot \mathrm{Geo}(\mathrm{GS}_q / g\varepsilon)$, where $\Pr[\mathrm{Geo}(s) = k] \propto e^{-|k|/s}$ for $k \in \mathbb{Z}$.

# Properties of the Definition

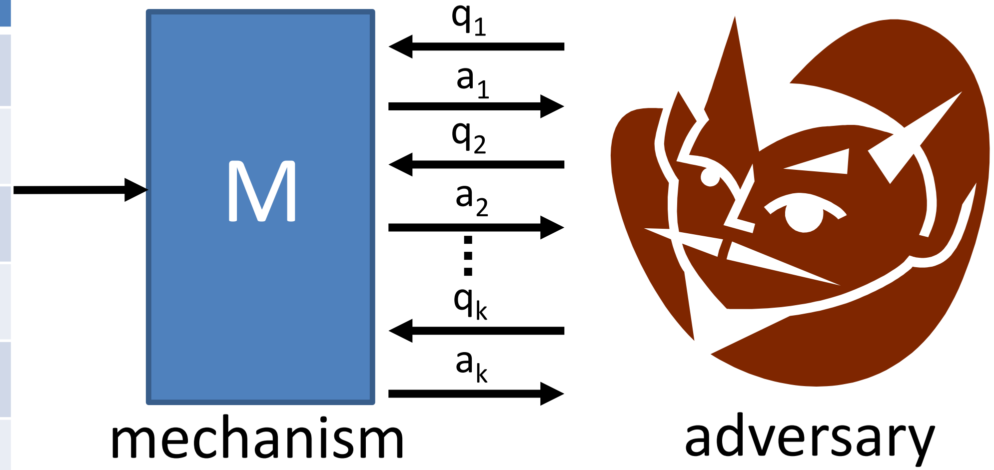
- Suffices to check pointwise: $M$ is $\epsilon$-DP if and only if
$$\forall x \sim x', \forall q, \forall t \;\; \Pr[M(x,q) = t] \le e^{\epsilon} \cdot \Pr[M(x',q) = t]$$
Replace with densities for continuous distributions

- Closed under post-processing: if $M$ is $\epsilon$-DP and $f$ is any function, then $M'(x,q) = f(M(x,q))$ is also $\epsilon$-DP.

- (Basic) composition: If $M_i$ is $\epsilon_i$-DP for $i = 1, \dots, k$, then
$$M\big(x, (q_1, \dots, q_k)\big) = \big(M_1(x, q_1), \dots, M_k(x, q_k)\big)$$
is $(\epsilon_1 + \cdots + \epsilon_k)$-DP.
  - Use independent randomness for $k$ queries.
  - Holds even if $q_i'$s are adaptively chosen by an adversary.

# Composition & Privacy Budgeting

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




**Thm:** If M is $\varepsilon$-DP if for one query, then it is $k\varepsilon$-DP for $k$ queries.

- To maintain global privacy loss at most $\varepsilon_g$, can set $\varepsilon = \varepsilon_g/k$ and stop answering after $k$ queries.
- More queries $\Rightarrow$ Smaller $\varepsilon$ $\Rightarrow$ Less accuracy.
  Some query-accuracy tradeoff is necessary! (why?)

# Composition for Algorithm Design

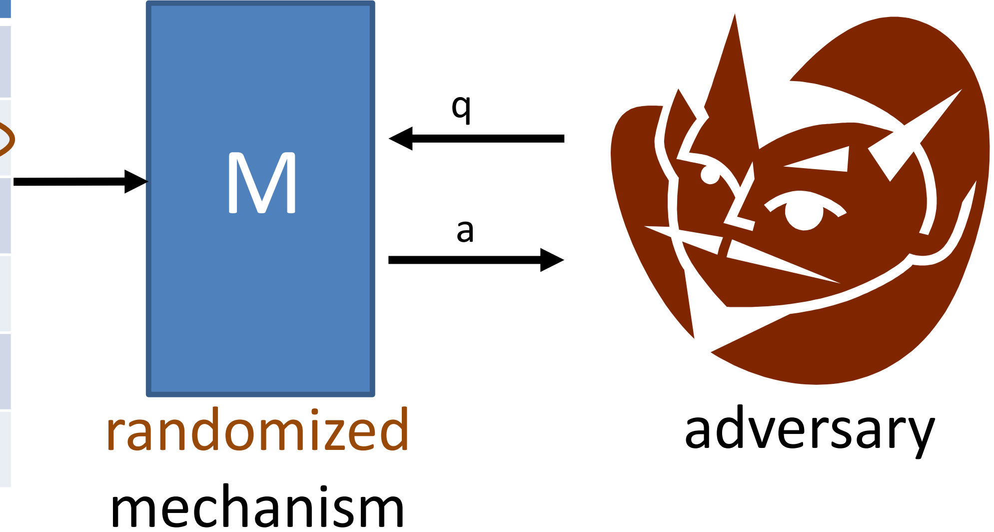
Composition and post-processing allow designing more complex differentially private algorithms from simpler ones.

Example:

- Many machine learning 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. (The "Statistical Query Model.")
- Can answer each query by adding Laplace noise.
- By composition and post-processing, trained model is DP and safe to output.

# Interpreting the Definition

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




**Def:** M is $\varepsilon$-DP if for all D, D' differing on one row, and all q

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

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

# 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.

But:

- No guarantee that adversary won’t infer sensitive attributes.
- No guarantee that subjects won’t be “harmed” by results of analysis.
- No protection for information that is not localized to a few rows.

# Group Privacy & Setting ε

- **Thm:** If M is $\varepsilon$-DP if for one query, then it is $k\varepsilon$-DP for $k$ groups of size $k$: for all $x, x'$ that differ on at most $k$ rows,
  $$\forall q\ \forall T\ \ \Pr[M(x,q) \in T] \leq e^{k\varepsilon} \cdot \ \Pr[M(x',q) \in T]$$
  - Meaningful privacy for groups of size $O(1/\varepsilon)$.

- **Cor:** Need $n \geq 1/\varepsilon$ for any reasonable utility.

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

# A Bayesian Interpretation

- Let $X = (X_1, \dots, X_n) \in \mathcal{X}^n$ be a random variable distributed according to an adversary's "prior beliefs" about a dataset, and let $X_{-i} = (X_1, \dots, X_{i-1}, \perp, X_i, \dots, X_n)$ have person $i$'s data removed or replaced with a dummy value in $\mathcal{X}$.

- Suppose $M : \mathcal{X}^n \to \mathcal{Y}$ is $\varepsilon$-DP, and let $y \in \mathcal{Y}$ be any possible output. Then for every $x_i \in \mathcal{X}$,

$$\underbrace{\Pr[X_i = x_i | M(X) = y]}_{\text{Posterior belief about person } i \text{ after seeing output } y} \in e^{\pm\varepsilon} \cdot \underbrace{\Pr[X_i = x_i | M(X_{-i}) = y]}_{\text{Posterior belief about person } i \text{ after seeing output } y \text{ if person } i\text{'s data wasn't used}}$$

- Explains choice of multiplicative distance in def of DP.

# Variants of the Definition

- When $n$ is not publicly known:
  - Datasets: multisets $D$ of elements of $\mathcal{X}$, can represent as a histogram $D \in \mathbb{N}^{\mathcal{X}}$, where $D_x =$ number of copies of $x$.
  - Neighbors: $D \sim D'$ iff $|D \Delta D'| = 1$ (add or remove an elt) In histogram notation: $|D \Delta D'| = \sum_x |D_x - D'_x| \stackrel{\text{def}}{=} \|D - D'\|_1$
- Social network data:
  - Datasets: graphs $G$, possibly with labels on nodes and edges
  - Neighbors v1: $G \sim G'$ if differ by modifying one edge
  - Neighbors v2: $G \sim G'$ if differ by modifying one node & incident edges.
  - Q: which choice provides better privacy protection?

# Approximate Differential Privacy

**Def:** M is $(\varepsilon, \delta)$-DP if for all D∼D', and all q

$\forall$ sets T, Pr[M(D,q)∈T] $\leq e^{\varepsilon}\cdot$ Pr[M(D',q)∈T]$+\delta$

- Intuitively: $\varepsilon$-DP with probability at least $1-\delta$.
- Picking a random person from dataset and publishing their data is $(0, 1/n)$-DP, so want $\delta \ll 1/n$ .
- Ideally set $\delta$ to be cryptographically small (e.g. $2^{-50}$).
- Satisfies postprocessing, basic composition (adding $\delta_i$'s).
- Group privacy for groups of size up to $O(1/\varepsilon)$.
- Does not suffice to check pointwise (need to consider sets $T$).

# Benefits of Approximate DP

- More mechanisms, e.g. Gaussian Mechanism:
$$M(x,q) = q(x) + \mathcal{N}\left(0,\sigma^2\right),$$
$$\text{for } \sigma = \frac{\mathrm{GS}_q}{\varepsilon} \cdot \sqrt{2\ln(2/\delta)}$$

- Advanced Composition Thm: If $M_i$ is $(\varepsilon,\delta)$-DP for $i = 1,\dots,k$ and $k < 1/\varepsilon^2$, then $\forall\delta > 0$
$$M\left(x,(q_1,\dots,q_k)\right) = \left(M_1(x,q_1),\dots,M_k(x,q_k)\right)$$
is $(\varepsilon', k\cdot\delta + \delta')$ −DP, for
$$\varepsilon' = O\left(\varepsilon \cdot \sqrt{k \cdot \log(1/\delta')}\right).$$