CSE 312

Foundations of Computing II

Lecture 28: Differential Privacy

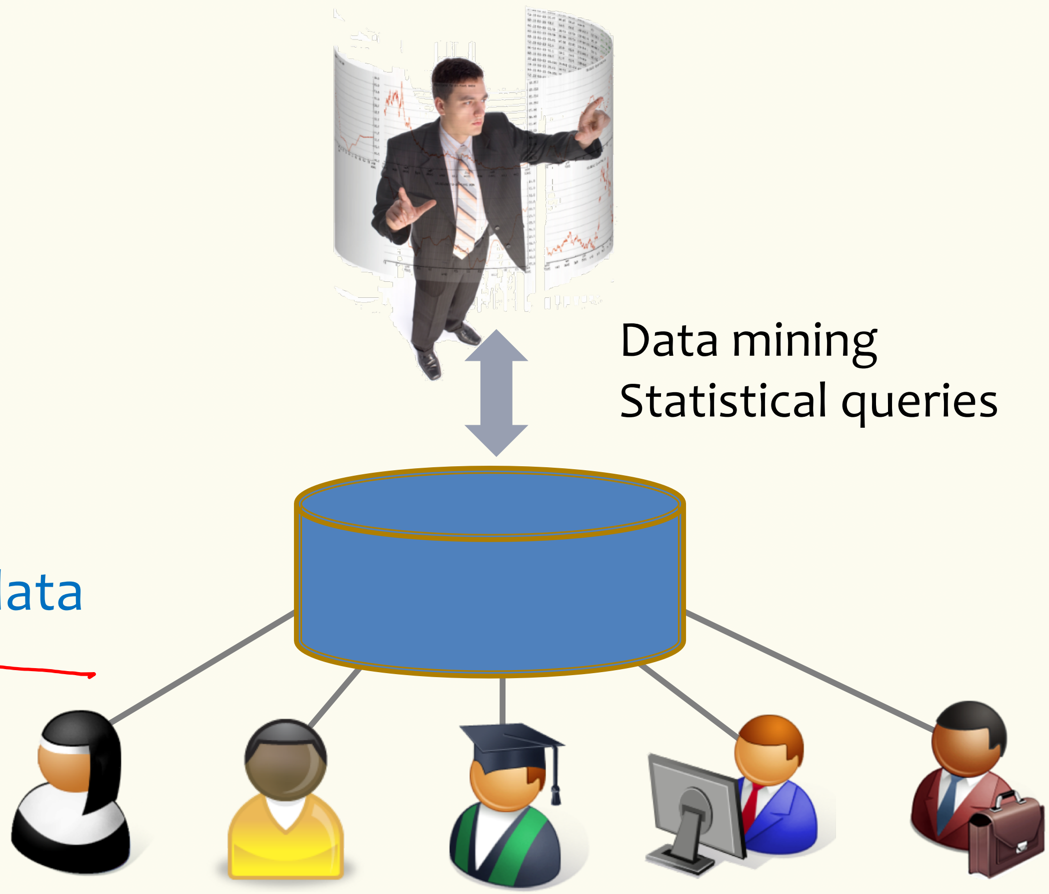
PAUL G. ALLEN SCHOOL
OF COMPUTER SCIENCE & ENGINEERING

Rachel Lin, Hunter Schafer

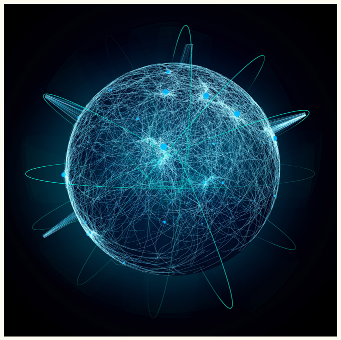
Slide Credit: Based on Stefano Tessaro's slides for 312 19au
incorporating ideas from Alex Tsun's and Anna Karlin's slides for 312 20su and 20au

Setting

- Medical data
- Query logs
- Social network data
- …



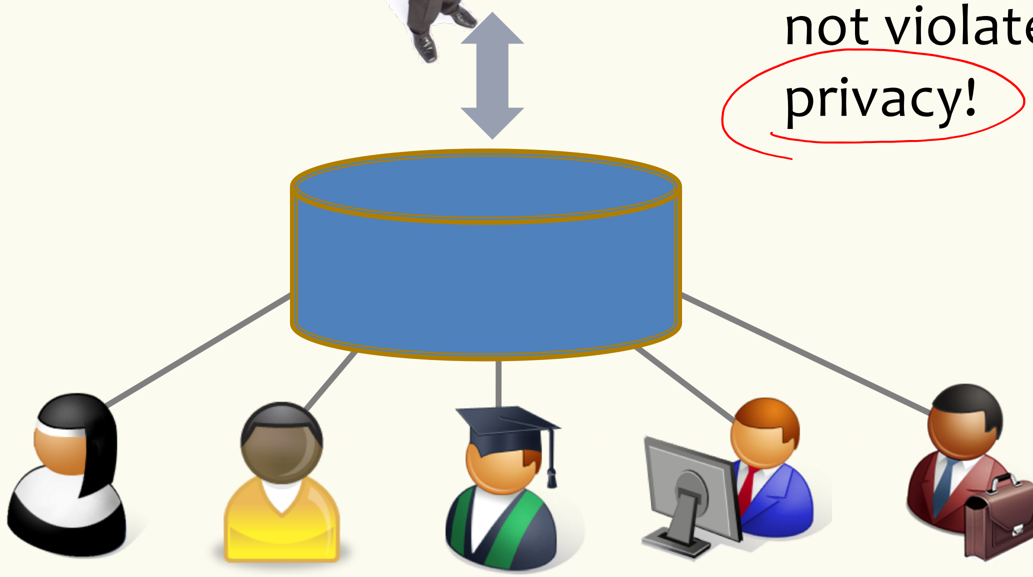
Setting – Data Release



Internet

Main concern: Do not violate user privacy!

Publish:
Aggregated data, e.g., outcome of medical study, research paper, …



Example – Linkage Attack

[Sweeney '00]

- The Commonwealth of Massachusetts Group Insurance Commission (GIC) releases 135,000 records of patient encounters, each with 100 attributes
  - Relevant attributes removed, but ZIP, birth date, gender available
  - Considered "safe" practice
- Public voter registration record
  - Contain, among others, name, address, ZIP, birth date, gender

"Linkage"

- Allowed identification of medical records of William Weld, governor of MA at that time
  - He was the only man in his zip code with his birth date …

+More attacks! (cf. Netflix grand prize challenge!)

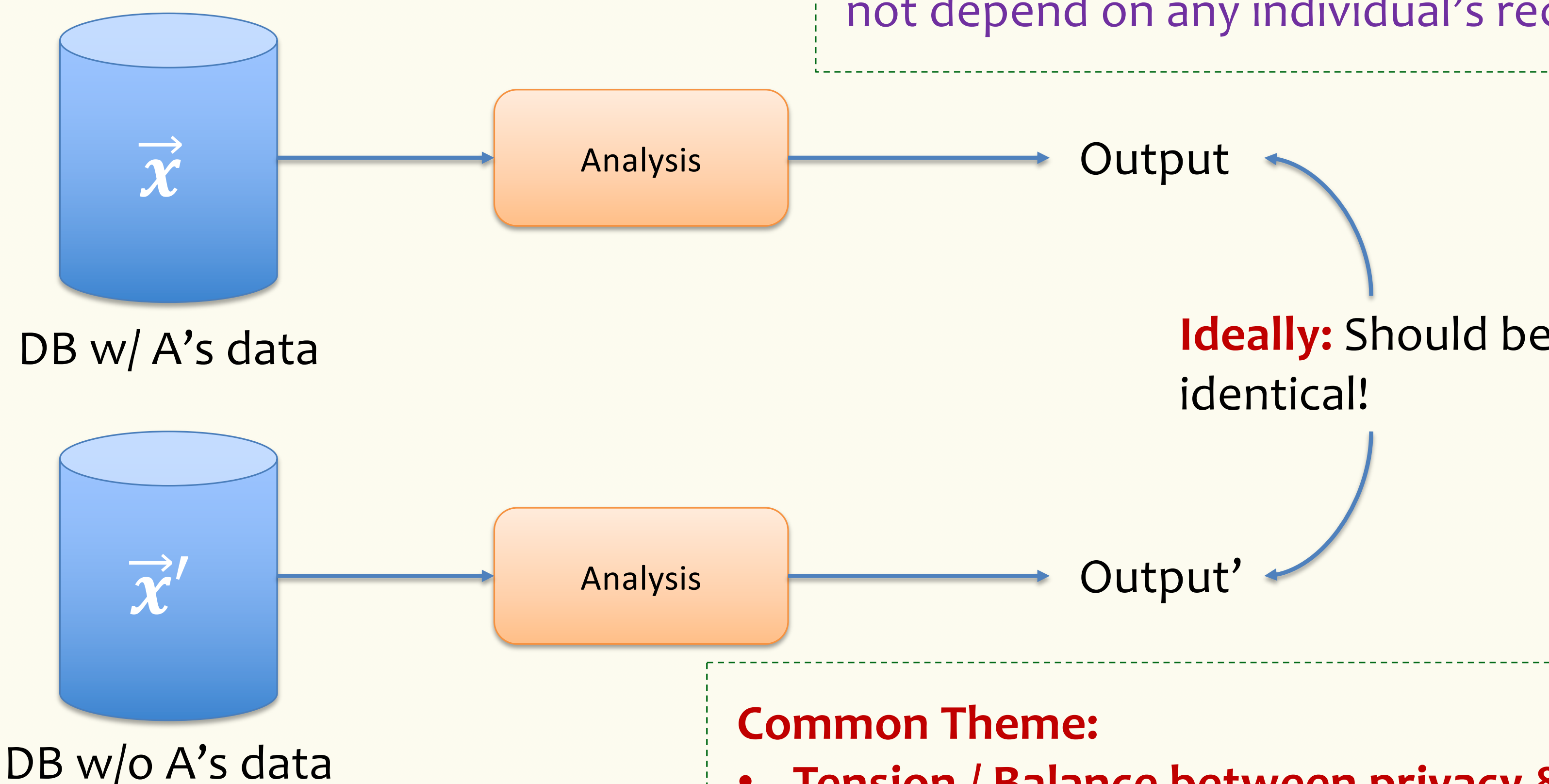
One way out? Differential Privacy

- A **formal definition** of privacy
  - Satisfied in systems deployed by Google, Uber, Apple, …
- Used by 2020 census
- Idea: *Any information-related risk to a person should not change significantly as a result of that person's information being included, or not, in the analysis.*
  - *Even with side information!*

Ideal Individual's Privacy

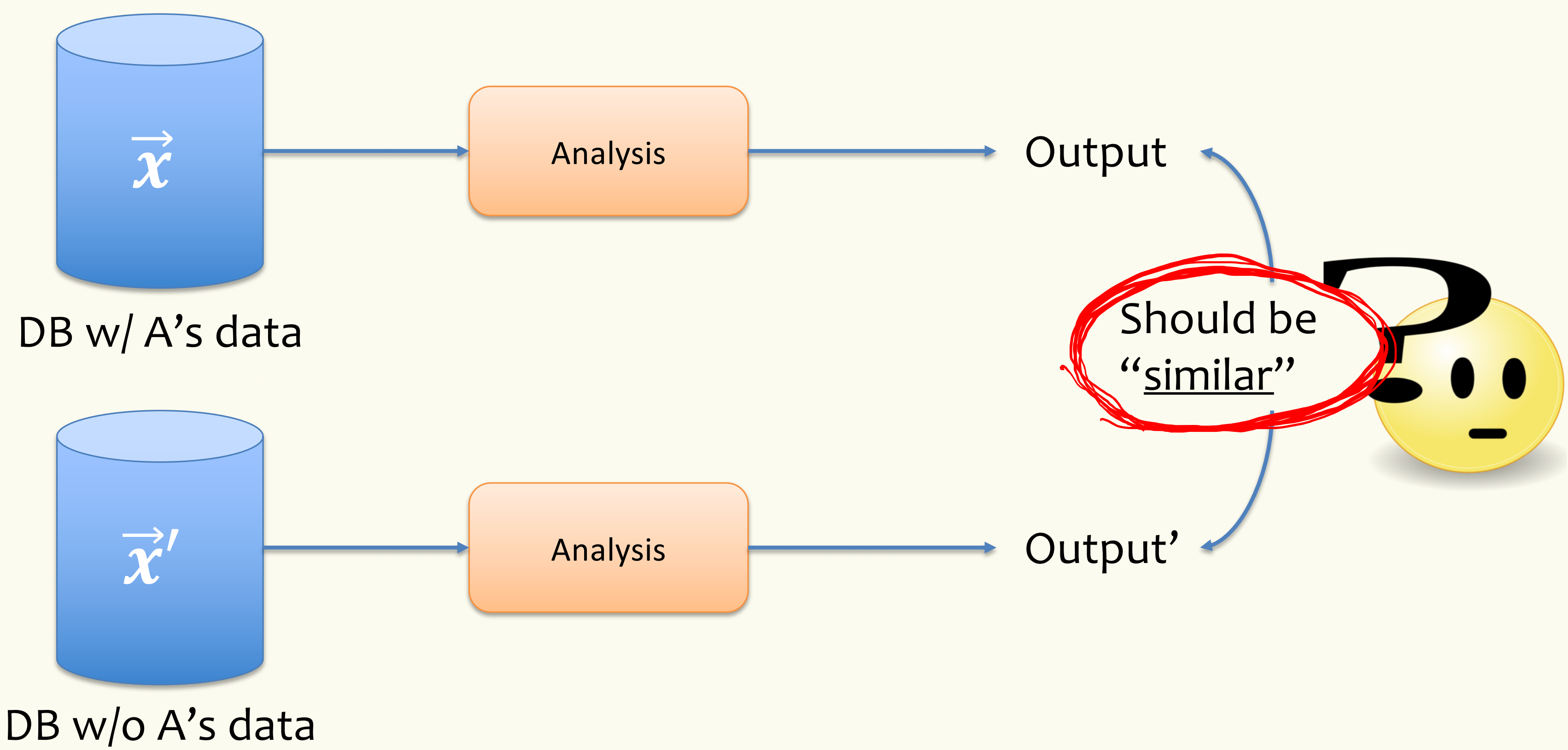
For every individual A whose record in DB

Very good for privacy.
But the output would be **useless** as it does not depend on any individual's record!



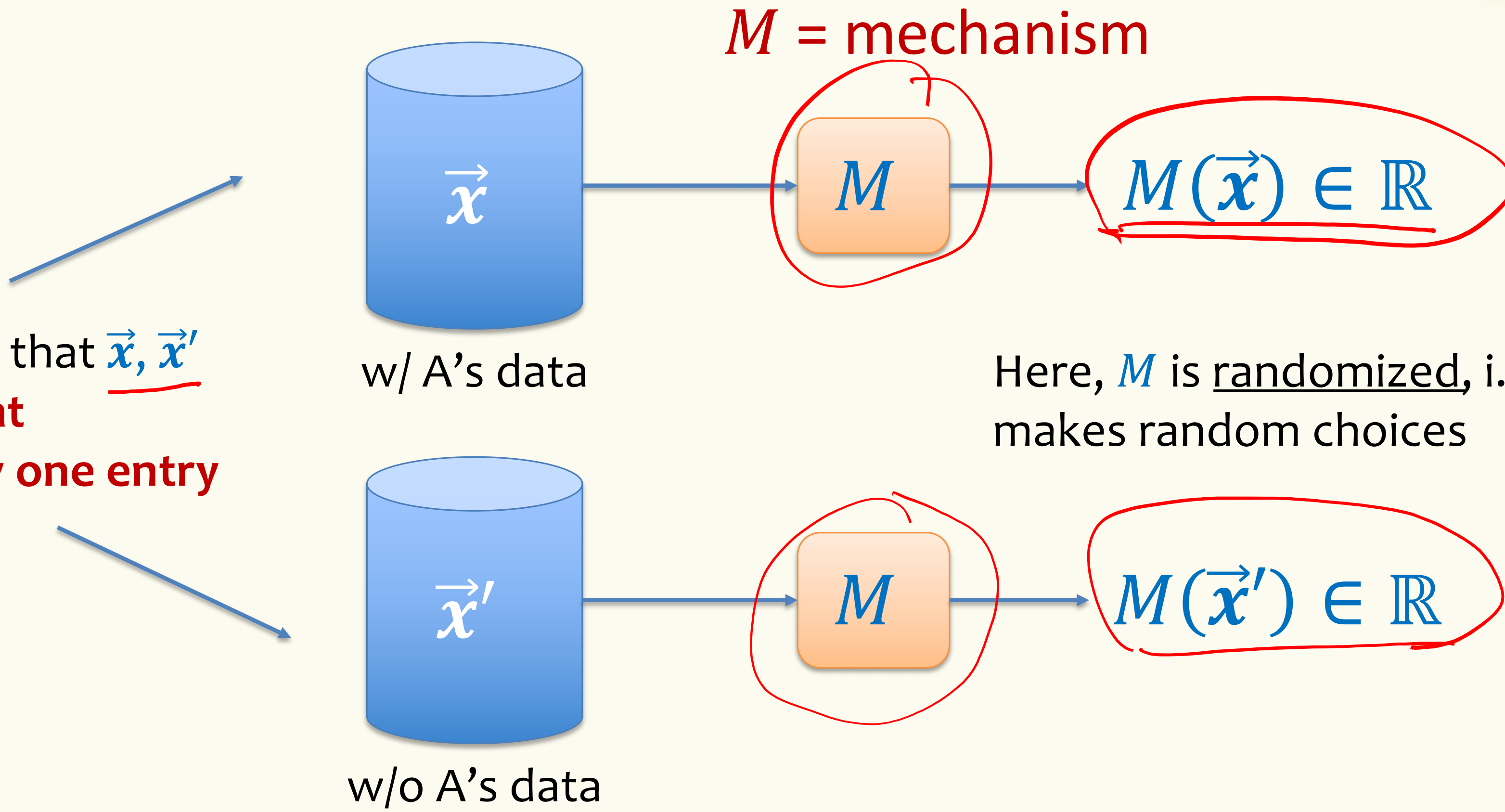
Common Theme:
- **Tension / Balance between privacy & utility**
- **Privacy is not a 0 / 1 property.**

More Realistic Privacy Goal



Setting – Formal

We say that $\vec{x}, \vec{x}'$ **differ at exactly one entry**

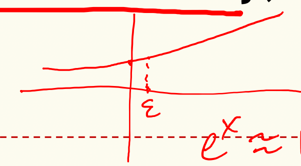


Here, M is randomized, i.e., it makes random choices

Setting – Mechanism

Definition. A mechanism M is ϵ-**differentially private** if for all subsets* $T \subseteq \mathbb{R}$, and for all databases $\vec{x}, \vec{x}'$ which differ at exactly one entry, e^x

$$\mathbb{P}(M(\vec{x}) \in T) \leq e^{\epsilon}\, \mathbb{P}(M(\vec{x}') \in T)$$



$e^x \approx 1+x$ x small

Dwork, McSherry, Nissim, Smith, '06

Think: $\epsilon = \frac{1}{100}$ or $\epsilon = \frac{1}{10}$

* Can be generalized beyond output in $\mathbb{R}$

Example – Counting Queries

- DB is a vector $\vec{\boldsymbol{x}} = (x_1, \dots, x_n)$ where $x_1, \dots, x_n \in \{0,1\}$
  - $x_i = 1$ if individual i has diseases
  - $x_i = 0$ means patient does not have disease or patient data wasn't recorded.
- Query: $q(\vec{\boldsymbol{x}}) = \boxed{\sum_{i=1}^{n} x_i}$

$q(x) = S$ $\quad$ $q(x') = S-1$

$\Pr(q(x) = S-1) = 0$ $\quad$ $P(q(x') = S-1) = 1$

Here: $\vec{\boldsymbol{x}}$ and $\vec{\boldsymbol{x}}'$ differ at one entry means they differ at one single coordinate, e.g., $x_i = 1$ and $x'_i = 0$

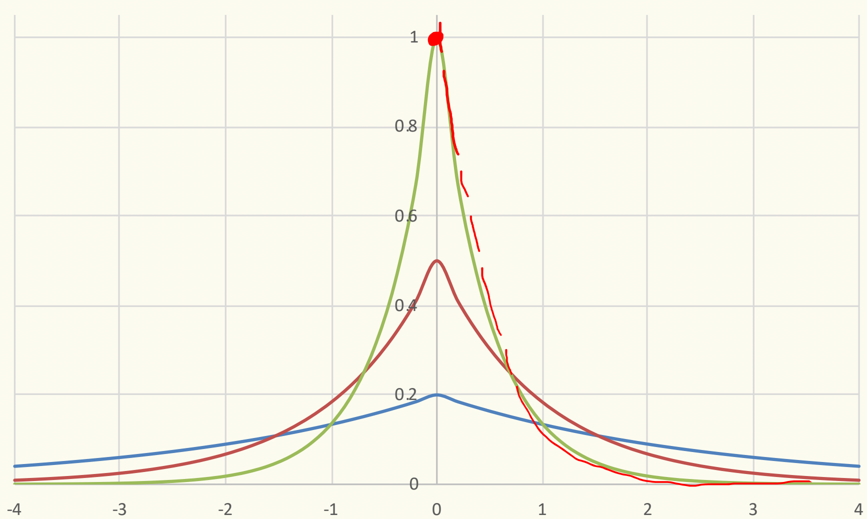
A solution – Laplacian Noise

Mechanism M taking input $\vec{\boldsymbol{x}} = (x_1, \ldots, x_n)$:

- Return $M(\vec{\boldsymbol{x}}) = \sum_{i=1}^{n} x_i + Y$

"Laplacian mechanism with parameter ϵ"

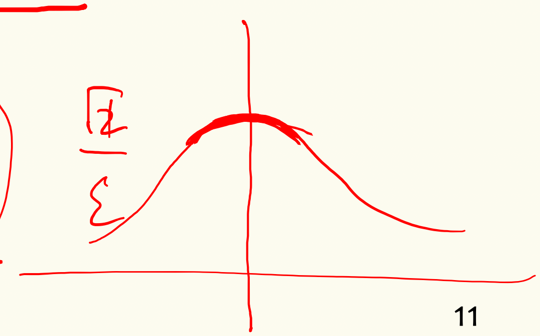
Here, Y follows a **Laplace distribution** with parameter ϵ



$$f_Y(y) = \frac{\epsilon}{2} e^{-\epsilon|y|}$$

$$\mathbb{E}(Y) = 0$$

$$\mathrm{Var}(Y) = \frac{2}{\epsilon^2}$$



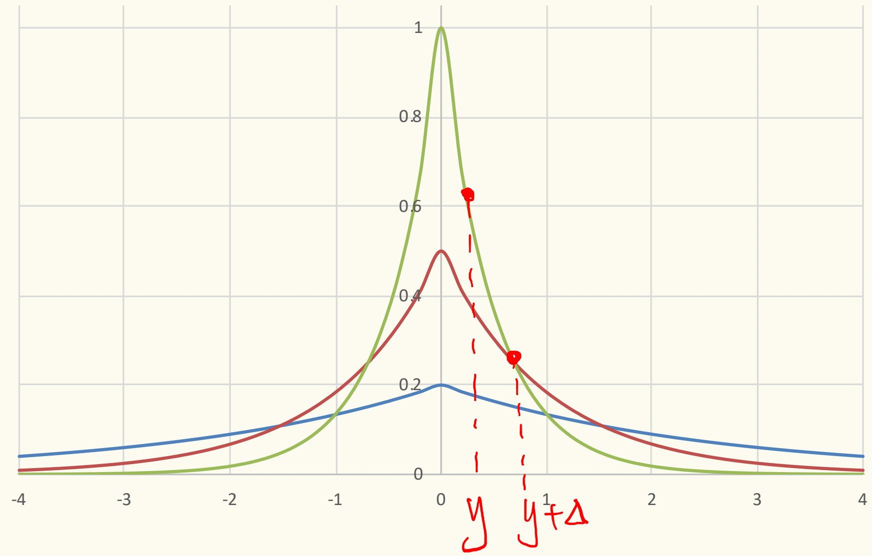
Better Solution – Laplacian Noise

Mechanism M taking input $\vec{\boldsymbol{x}} = (x_1, \ldots, x_n)$:

- Return $M(\vec{\boldsymbol{x}}) = \sum_{i=1}^{n} x_i + Y$

"Laplacian mechanism with parameter ϵ"

Here, Y follows a **Laplace distribution** with parameter ϵ



$$f_Y(y) = \frac{\epsilon}{2} e^{-\epsilon|y|}$$

Key property: For all y, Δ

$$\boxed{\frac{f_Y(y)}{f_Y(y+\Delta)}} \leq \boxed{e^{\epsilon\Delta}}$$

Laplacian Mechanism – Privacy

Theorem. The Laplacian Mechanism with parameter ϵ satisfies ϵ-differential privacy

Show: $\forall\, \vec{\boldsymbol{x}}, \vec{\boldsymbol{x}}'$ differ at one entry, $[a, b]$

$$\mathbb{P}(M(\vec{\boldsymbol{x}}) \in [a,b]) \le e^{\epsilon}\, \boxed{\mathbb{P}(M(\vec{\boldsymbol{x}}') \in [a,b])}$$

$y = y' - s$

$$\Delta = \underbrace{\sum_{i=1}^{n} x'_i}_{= s'} - \underbrace{\sum_{i=1}^{n} x_i}_{= s} \qquad |\Delta| \le 1$$

$y' = y + s$

$$\boxed{\mathbb{P}(M(\vec{\boldsymbol{x}}) \in [a,b])} = \mathbb{P}(s + Y \in [a,b]) = \int_{a-s}^{b-s} f_Y(y)\,\mathrm{d}y = \boxed{\int_a^b f_Y(y' - s)\,\mathrm{d}y'} \quad (1)$$

$$= \int_a^b f_Y(y - s' + \Delta)\,\mathrm{d}y \le e^{\epsilon\Delta} \int_a^b f_Y(y - s')\,\mathrm{d}y \le e^{\epsilon} \boxed{\int_a^b f_Y(y - s')\,\mathrm{d}y} \quad (2)$$

$$= e^{\epsilon}\, \boxed{\mathbb{P}(M(\vec{\boldsymbol{x}}') \in [a,b])}$$

How Accurate is Laplacian Mechanism?

Let's look at $\sum_{i=1}^{n} x_i + Y$

- $\mathbb{E}(\sum_{i=1}^{n} x_i + Y) = \sum_{i=1}^{n} x_i + \mathbb{E}(Y) = \boxed{\sum_{i=1}^{n} x_i}$
- $\text{Var}(\sum_{i=1}^{n} x_i + Y) = \text{Var}(Y) = \frac{2}{\epsilon^2}$

$\varepsilon = \frac{1}{100}$

This is accurate enough for large enough n!

$\frac{\sqrt{2}}{\varepsilon} \approx 100 \cdot \sqrt{2}$

Differential Privacy – What else can we compute?

- **Statistics:** counts, mean, median, histograms, boxplots, etc.
- **Machine learning:** classification, regression, clustering, distribution learning, etc.
- **…**

Differential Privacy – Nice Properties

- **Group privacy:** If M is ϵ-differentially private, then for all $T \subseteq \mathbb{R}$, and <u>for all</u> databases $\vec{\boldsymbol{x}}, \vec{\boldsymbol{x}}'$ which differ at (at most) k entries,

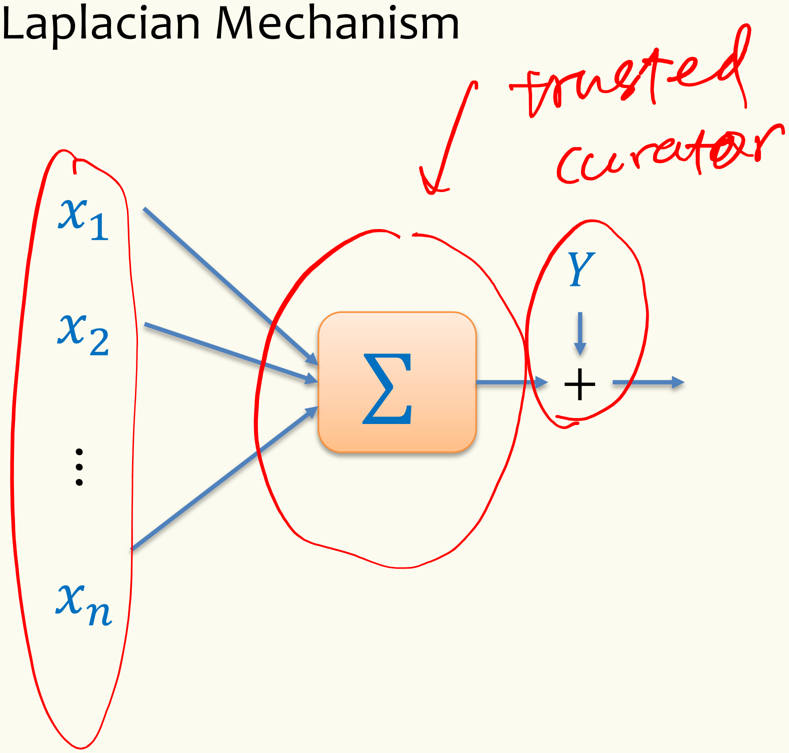
$$\mathbb{P}(M(\vec{\boldsymbol{x}}) \in T) \leq e^{k\epsilon}\, \mathbb{P}(M(\vec{\boldsymbol{x}}') \in T)$$

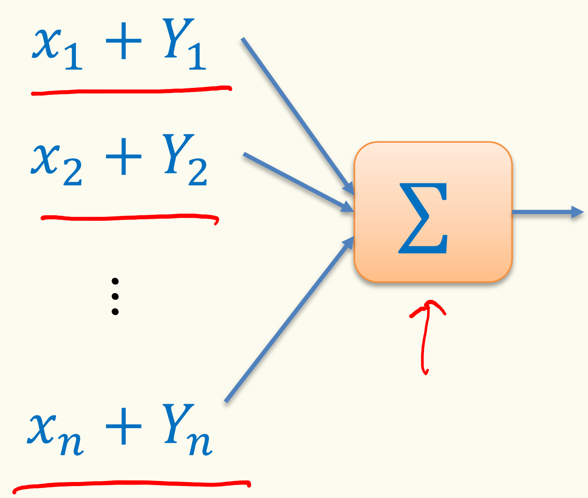
- **Composition:** If we apply two ϵ-DP mechanisms to data, combined output is 2ϵ-DP.
  - How much can we allow ϵ to grow? (So-called “privacy budget.”)
- **Post-processing:** Postprocessing does not decrease privacy.

Local Differential Privacy

What if we don't trust aggregator?

Laplacian Mechanism





Solution: Add noise <u>locally</u>!

Example – Randomize Response

For a given parameter α

Mechanism M taking input $\vec{\boldsymbol{x}} = (x_1, \ldots, x_n)$:

- For all $i = 1, \ldots, n$:
  - $y_i = x_i$ w/ probability $\frac{1}{2} + \alpha$, and $y_i = 1 - x_i$ w/ probability $\frac{1}{2} - \alpha$.
  - $\hat{x}_i = \frac{y_i - \frac{1}{2} + \alpha}{2\alpha}$
- Return $M(\vec{\boldsymbol{x}}) = \sum_{i=1}^{n} \hat{x}_i$

S. L. Warner. Randomized response: A survey technique for eliminating evasive answer bias. Journal of the American Statistical Association, 60(309):63–69, 1965

Example – Randomize Response

For a given parameter α

Mechanism M taking input $\vec{\boldsymbol{x}} = (x_1, \dots, x_n)$:

- For all $i = 1, \dots, n$:
  - $y_i = x_i$ w/ probability $\frac{1}{2} + \alpha$, and $y_i = 1 - x_i$ w/ probability $\frac{1}{2} - \alpha$.
  - $\hat{x}_i = \frac{y_i - \frac{1}{2} + \alpha}{2\alpha}$
- Return $M(\vec{\boldsymbol{x}}) = \sum_{i=1}^{n} \hat{x}_i$

Theorem. Randomized Response with parameter α satisfies ϵ-differential privacy, if $\alpha = \boxed{\frac{e^{\epsilon}-1}{e^{\epsilon}+1}}$.

Fact 1. $\mathbb{E}\big(M(\vec{\boldsymbol{x}})\big) = \sum_{i=1}^{n} x_i$

Fact 2. $\mathrm{Var}\big(M(\vec{\boldsymbol{x}})\big) \approx \frac{n}{\epsilon^2}$

$\frac{\sqrt{n}}{\varepsilon}$

Differential Privacy – Challenges

- **Accuracy vs. privacy:** How do we choose ϵ?
  - Practical applications tend to err in favor of accuracy.
  - See e.g. https://arxiv.org/abs/1709.02753
- **Fairness:** Differential privacy hides contribution of small groups, by design
  - How do we avoid excluding minorities?
  - Very hard problem!

Literature

- Cynthia Dwork and Aaron Roth. *"The Algorithmic Foundations of Differential Privacy"*.
  - https://www.cis.upenn.edu/~aaroth/Papers/privacybook.pdf
- https://privacytools.seas.harvard.edu/