

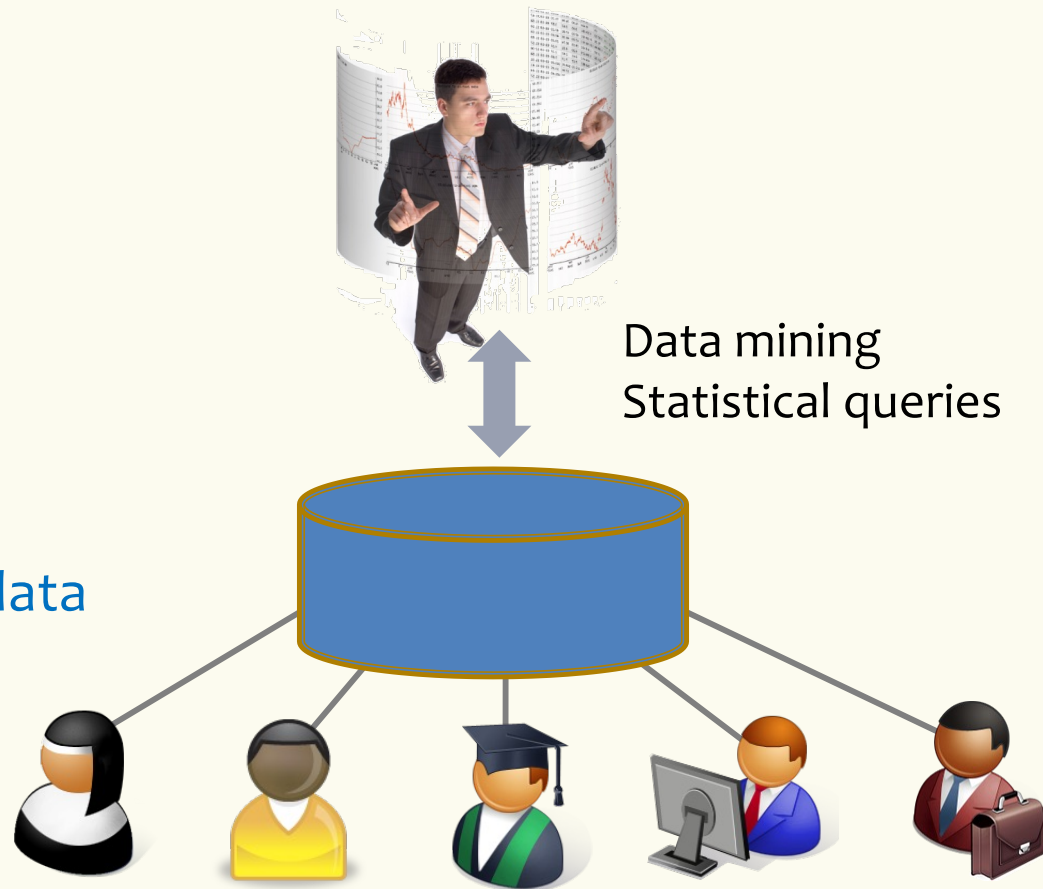
CSE 312

Foundations of Computing II

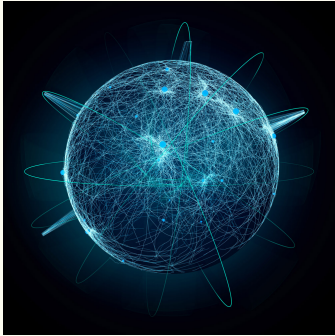
Lecture 26: Differential Privacy

Setting

Medical data
Query logs
Social network data
...



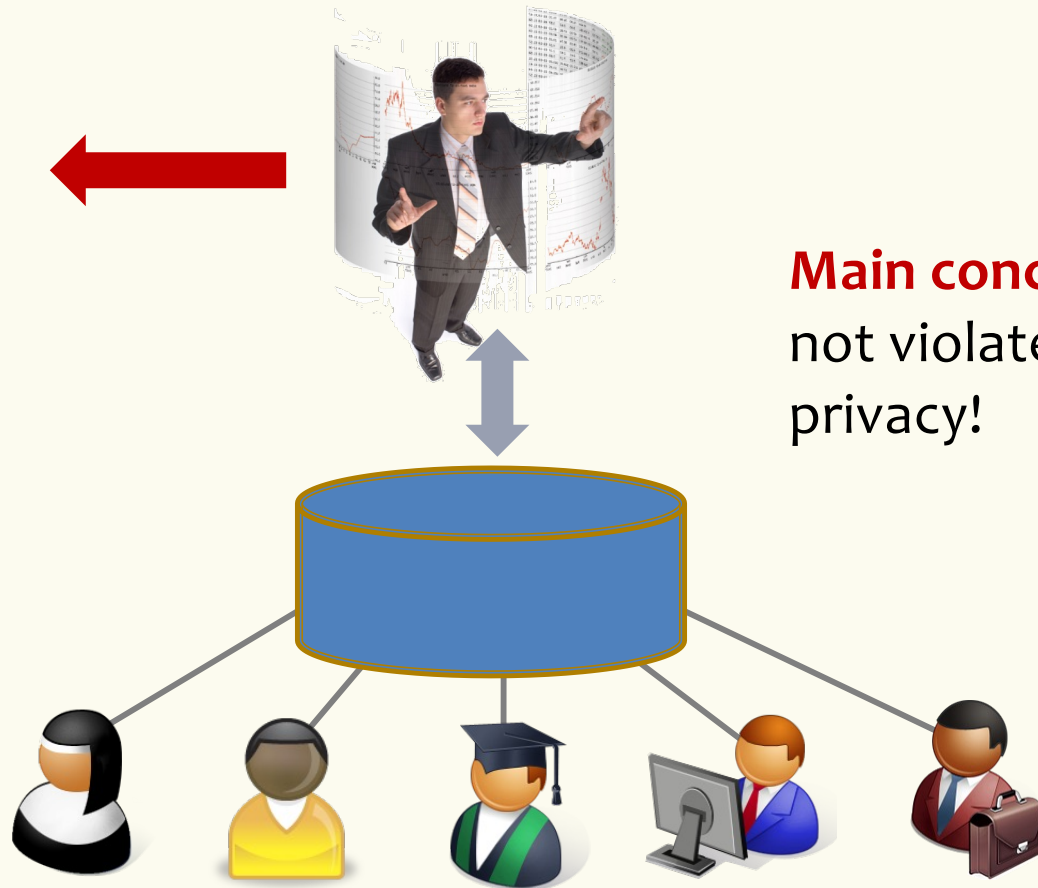
Setting – Data Release



Internet

Publish:

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



Main concern: Do
not violate user
privacy!

Example – Linkage Attack

- 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
 - 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!)

“Linkage”

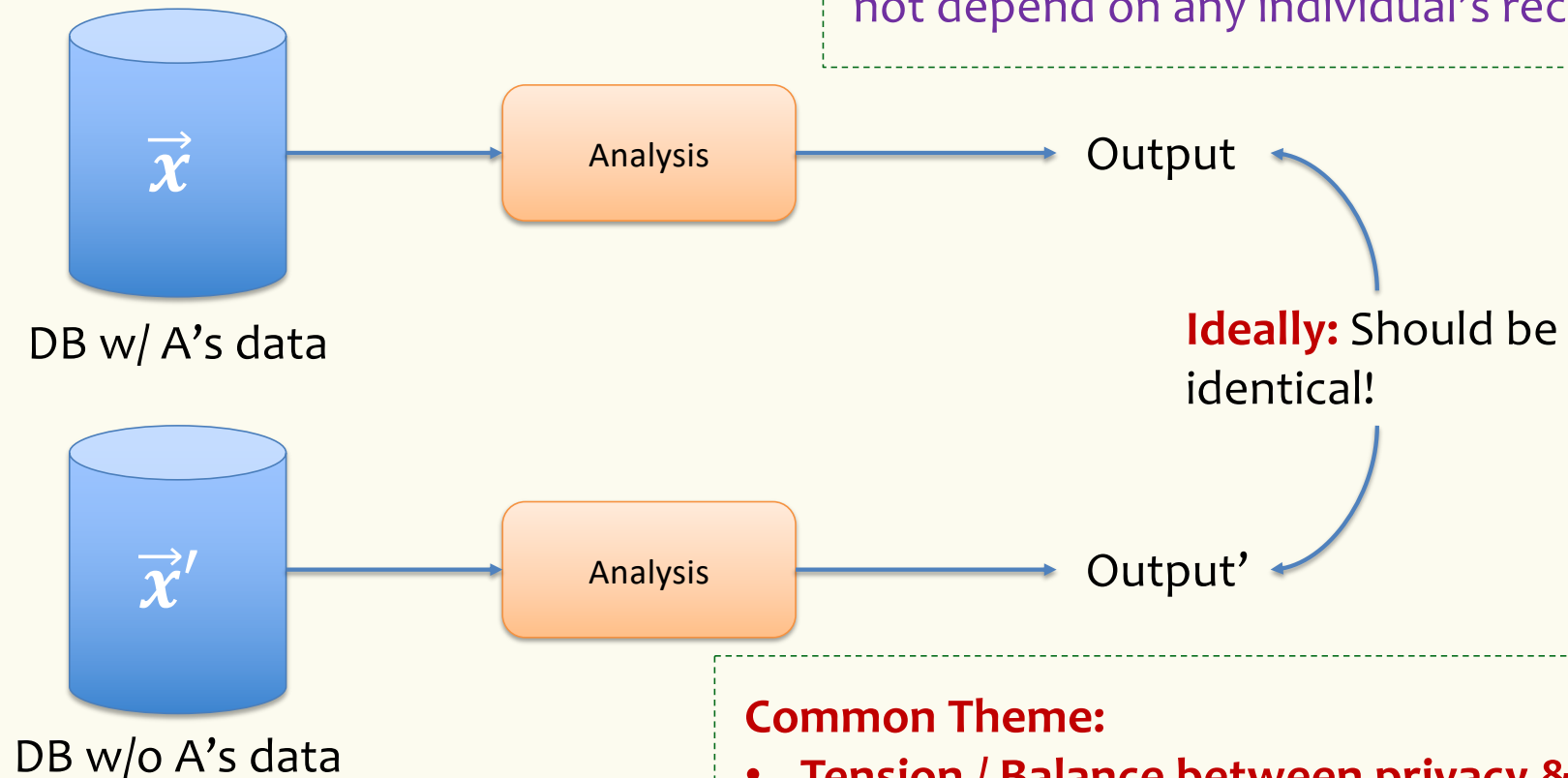


One way out? Differential Privacy

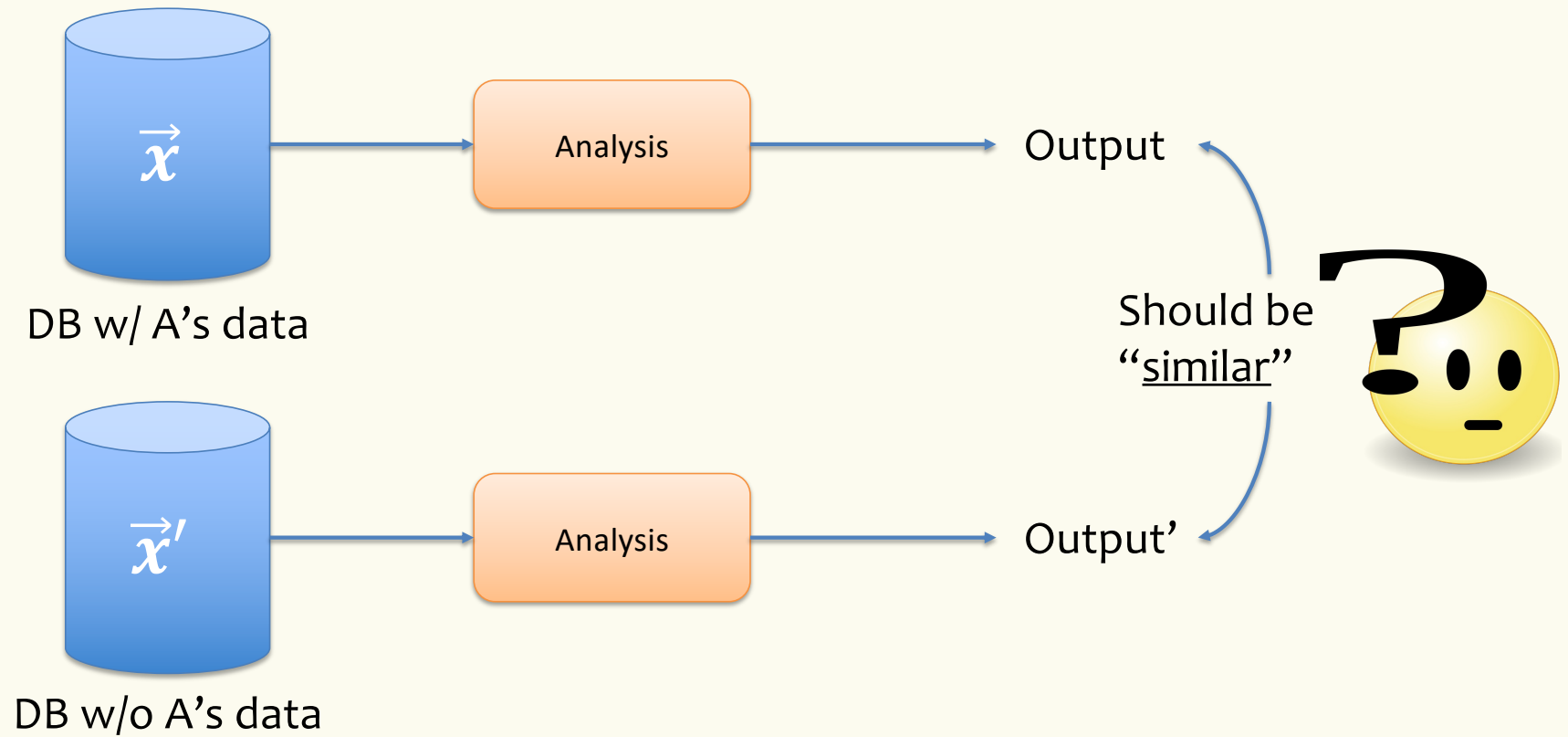
- A **formal definition** of privacy
 - Satisfied in systems deployed by Google, Uber, Apple, ...
- Used by 2020 census
- Idea: *Any information-related risk to a person should not change significantly as a result of that person's information being included, or not, in the analysis.*
 - *Even with side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

Ideal Individual's Privacy

For every individual A whose record in DB

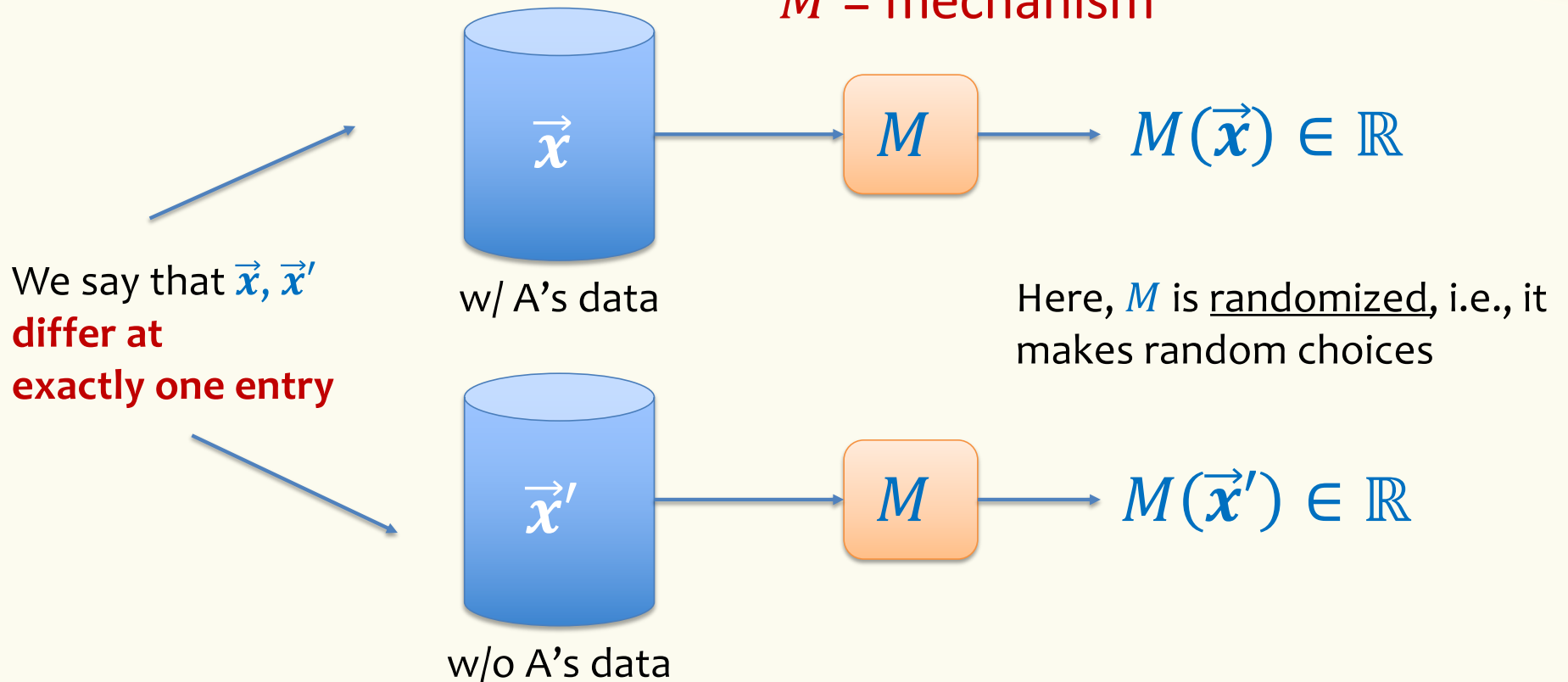


More Realistic Privacy Goal



Setting – Formal

$M = \text{mechanism}$



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,

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

Dwork, McSherry, Nissim, Smith, '06

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

$e^\epsilon \approx 1 + \epsilon$ for small ϵ

Example – Counting Queries

$\forall T \subseteq \mathbb{R}, P(M(\vec{x}) \in T) \leq e^\epsilon P(M(\vec{x}') \in T)$
for all \vec{x}, \vec{x}' that differ on one entry

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

- Query: $q(\vec{x}) = \sum_{i=1}^n x_i$

Poll: pollev.com/rachel312

For what ϵ is $M(\vec{x}) = q(\vec{x})$ ϵ -differentially private?

- a) 0.1
- b) 1
- c) 100
- d) None

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

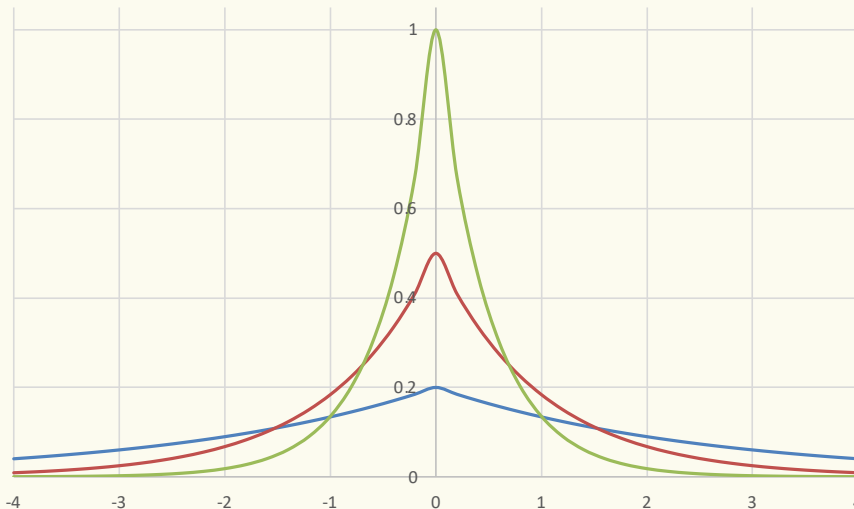
A solution – Laplacian Noise

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

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

where Y follows a **Laplace distribution** with parameter ϵ

“Laplacian mechanism with parameter ϵ ”



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

$$\mathbb{E}[Y] = 0$$

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

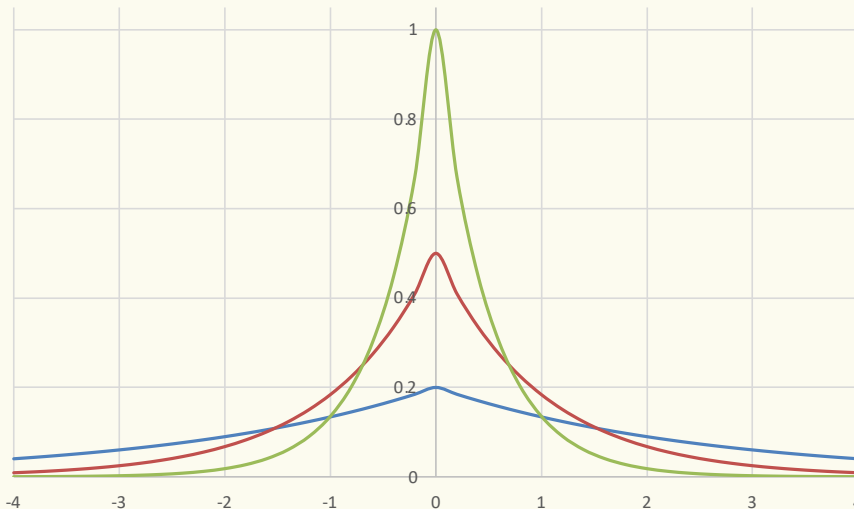
A solution – Laplacian Noise

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

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

where Y follows a **Laplace distribution** with parameter ϵ

“Laplacian mechanism with parameter ϵ ”



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

Key property: For all y, Δ

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

Laplacian Mechanism – Privacy

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

Goal to show: $\forall \vec{x}, \vec{x}'$ differing at one entry, $\forall [a, b]$

$$\Delta = \sum_{i=1}^n x'_i - \sum_{i=1}^n x_i = s' - s \quad |\Delta| \leq 1 \quad P(M(\vec{x}) \in [a, b]) \leq e^\epsilon \cdot P(M(\vec{x}') \in [a, b])$$

$$\begin{aligned} P(M(\vec{x}) \in [a, b]) &= P(s + Y \in [a, b]) = \int_{a-s}^{b-s} f_Y(y) dy = \int_a^b f_Y(y' - s) dy' \\ &= \int_a^b f_Y(y - s' + \Delta) dy \leq e^{\epsilon|\Delta|} \int_a^b f_Y(y - s') dy \leq e^\epsilon \int_a^b f_Y(y - s') dy \\ &= e^\epsilon P(M(\vec{x}') \in [a, b]) \\ P(M(\vec{x}') \in [a, b]) &= P(s' + Y \in [a, b]) = \int_a^b f_Y(y' - s') dy' \end{aligned}$$

How Accurate is Laplacian Mechanism?

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

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

This is accurate enough for large enough ϵ !

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

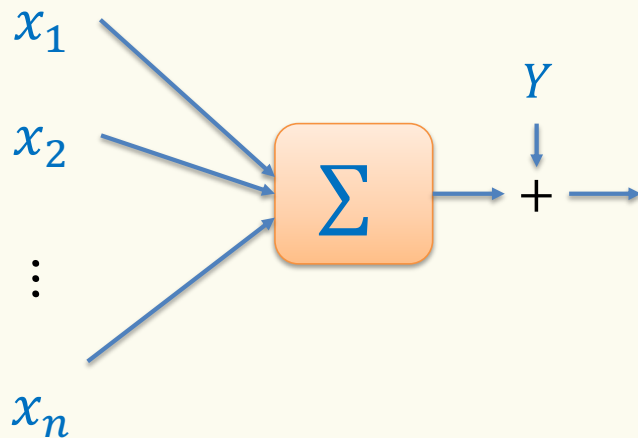
- **Group privacy:** If M is ϵ -differentially private, then for all $T \subseteq \mathbb{R}$, and for all databases \vec{x}, \vec{x}' which differ at (at most) k entries,

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

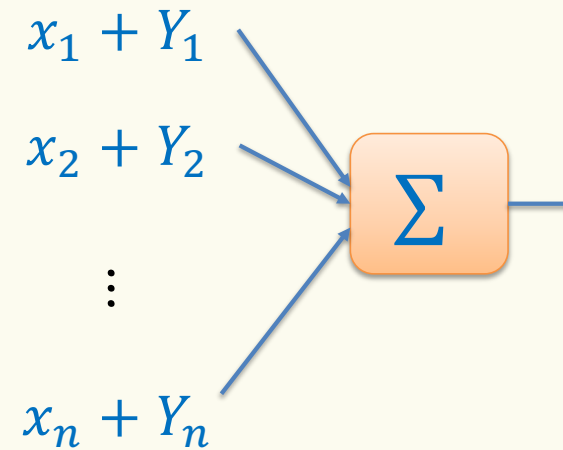
- **Composition:** If we apply two ϵ -DP mechanisms to data, combined output is 2ϵ -DP.
 - How much can we allow ϵ to grow? (So-called “privacy budget.”)
- **Post-processing:** Postprocessing does not decrease privacy.

Local Differential Privacy

Laplacian Mechanism



What if we don't trust aggregator?



Solution: Add noise locally!

Example – Randomized Response

For a given parameter α

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

- 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 $M(\vec{x}) = \sum_{i=1}^n \hat{x}_i$

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 $\vec{x} = (x_1, \dots, x_n)$:

- 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 $M(\vec{x}) = \sum_{i=1}^n \hat{x}_i$

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

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

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

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>
 - E.g. Privacy budgets of 2, 4, 8 per application feature, not tiny as assumed. These exponents add up quickly!
- **Fairness:** Differential privacy hides contribution of small groups, by design
 - How do we avoid excluding minorities?
 - Very hard problem!
- **Ethics:** Does differential privacy incentivize data collection?

Literature

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