Lecture 4

- How to define "Privacy" ?
  → Differential Privacy
- Revisit Randomized Response
- Laplace Mechanism

How to define "privacy"?

Approaches

1. Think of possible attacks; Defenses against these attacks

   Examples: K-anonymity

2. Formulate general criteria

K - anonymity.

- Input Table ⟼ Output Table
- Generalization :

  Replace a single value with a set of possible values

  - 2 ⟼ [1,3]
  - Male ⟼ {Male, Female}

- Table is k-anonymous if every row matches at least (k-1) others in the non-sensitive attributes

	Non-Sensitive			Sensitive
	Zip code	**Age**	**Nationality**	**Condition**
1	130**	<30	*	AIDS
2	130**	<30	*	Heart Disease
3	130**	<30	*	Viral Infection
4	130**	<30	*	Viral Infection
5	130**	≥ 40	*	Cancer
6	130**	≥ 40	*	Heart Disease
7	130**	≥ 40	*	Viral Infection
8	130**	≥ 40	*	Viral Infection
9	130**	3*	*	Cancer
10	130**	3*	*	Cancer
11	130**	3*	*	Cancer
12	130**	3*	*	Cancer

Figure 1: A 4-anonymous table.

- Seems to resist "Linkage attacks"
  → Can't identify a record uniquely
  → Seems hard to link to other info sources

- What can go wrong?
  → Everyone in their 30's has cancer
  → Rule out other info.

	Non-Sensitive			Sensitive
	Zip code	Age	Nationality	Condition
1	130**	<30	*	AIDS
2	130**	<30	*	Heart Disease
3	130**	<30	*	Viral Infection
4	130**	<30	*	Viral Infection
5	130**	≥40	*	Cancer
6	130**	≥40	*	Heart Disease
7	130**	≥40	*	Viral Infection
8	130**	≥40	*	Viral Infection
9	130**	3*	*	Cancer
10	130**	3*	*	Cancer
11	130**	3*	*	Cancer
12	130**	3*	*	Cancer

Figure 1: A 4-anonymous table.

Composition.

Cross referencing:

28 years old
Zipcode 13012
In both data sets

Overlap datasets

	Non-Sensitive			Sensitive
	Zip code	Age	Nationality	Condition
1	130**	<30	*	AIDS
2	130**	<30	*	Heart Disease
3	130**	<30	*	Viral Infection
4	130**	<30	*	Viral Infection
5	130**	≥40	*	Cancer
6	130**	≥40	*	Heart Disease
7	130**	≥40	*	Viral Infection
8	130**	≥40	*	Viral Infection
9	130**	3*	*	Cancer
10	130**	3*	*	Cancer
11	130**	3*	*	Cancer
12	130**	3*	*	Cancer

	Non-Sensitive			Sensitive
	Zip code	Age	Nationality	Condition
1	130**	<35	*	AIDS
2	130**	<35	*	Tuberculosis
3	130**	<35	*	Flu
4	130**	<35	*	Tuberculosis
5	130**	<35	*	Cancer
6	130**	<35	*	Cancer
7	130**	≥35	*	Cancer
8	130**	≥35	*	Cancer
9	130**	≥35	*	Cancer
10	130**	≥35	*	Tuberculosis
11	130**	≥35	*	Viral Infection
12	130**	≥35	*	Viral Infection

- K-anonymity issues
  - → Specifies a set of acceptable output (k-anonymous tables)
  - → Does not specify the "algorithmic" process
  - → "Flexibility" may leak info.

Meaningful definitions

Consider the algorithms

	Non-Sensitive			Sensitive
	Zip code	Age	Nationality	Condition
1	130**	<30	*	AIDS
2	130**	<30	*	Heart Disease
3	130**	<30	*	Viral Infection
4	130**	<30	*	Viral Infection
5	130**	≥40	*	Cancer
6	130**	≥40	*	Heart Disease
7	130**	≥40	*	Viral Infection
8	130**	≥40	*	Viral Infection
9	130**	3*	*	Cancer
10	130**	3*	*	Cancer
11	130**	3*	*	Cancer
12	130**	3*	*	Cancer

Figure 1: A 4-anonymous table.

Differential Privacy (Dwork, McSherry, Nissim, Smith) 2006

- Algorithmic Property.
  - → Rigorous guarantees against arbitrary external info.
  - Resists known attacks.

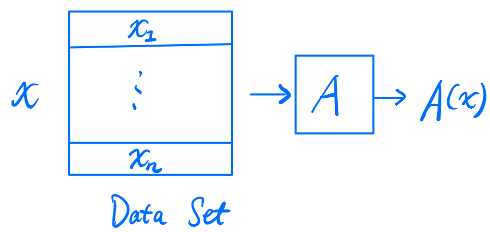
Data domain $\mathcal{X}$ (e.g. $\{0,1\}^d$, $\mathbb{R}^d$).

Data set $x = (x_1, x_2, \dots, x_n) \in \mathcal{X}^n$

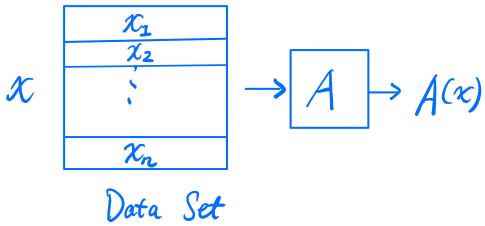
(Think of x as fixed, not random)

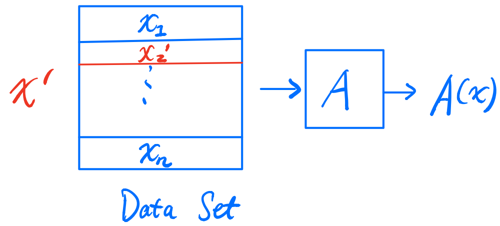
Randomized Algorithm $\mathcal{A}$

$\Rightarrow$ $A(x)$ is a random variable.



Thought Experiment.





x' is a neighbor of x
if they differ in one data point.

Idea of DP: Neighboring data sets induce close output distributions

Definition. (Differential Privacy).

A is ε-differentially private if

for all neighbors x and x'

for all subsets E of outputs

$$\mathbb{P}[A(x) \in E] \leq e^{\varepsilon} \, \mathbb{P}[A(x') \in E]$$

↑

This is an algorithmic property.

Definition. (Differential Privacy)

A is ε-differentially private if
for all neighbors x and x'
for all subsets E of outputs

$$\mathbb{P}[A(x) \in E] \leq e^{\varepsilon}\, \mathbb{P}[A(x') \in E]$$

What is ε?

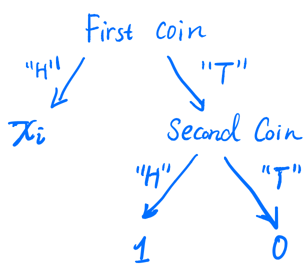
- Measure of info leakage (called max divergence)
- Small constant : $\frac{1}{10}$, 1, but not $\frac{1}{2^{80}}$ or 100

Example : Randomized Response (In lecture 1)

Each person has a secret bit $x_i = 0$ or $x_i = 1$

(Have you ever done XYZ?)

RR



Input : $x_1, \dots, x_n$

Output : $y_1, \dots, y_n$

RR is $\ln(3)$- diffentially private

Proof. • Fix two neighboring data sets

$x = (x_1, \ldots, x_i, \ldots, x_n)$, $x' = (x_1, \ldots, x_i', \ldots, x_n)$

• To start, fix some output $y = (y_1, \ldots, y_n) \in \{0,1\}^n$

$$\frac{\mathbb{P}[RR(x) = y]}{\mathbb{P}[RR(x') = y]} = \frac{\mathbb{P}[Y_i = y_i \mid x_i]}{\mathbb{P}[Y_i = y_i \mid x_i']} \qquad 3 \text{ or } \frac{1}{3}$$

$$\Rightarrow \mathbb{P}[RR(x) = y] \leq e^{\ln(3)} \mathbb{P}[RR(x') = y]$$

◦ To Complete, For any $E \subseteq \{0,1\}^n$

$$\mathbb{P}[RR(x) \in E] = \sum_{y \in E} \mathbb{P}[RR(x) = y]$$
$$\leq e^{\varepsilon} \sum_{y \in E} \mathbb{P}[RR(x') = y] = \mathbb{P}[RR(x') \in E]$$

Basic Proof Strategy :

for all neighbors x and x'
for all subsets E of outputs

$$\mathbb{P}[A(x) \in E] \le e^{\varepsilon} \mathbb{P}[A(x') \in E]$$

$$\mathbb{P}[A(x) = y] \le e^{\varepsilon} \mathbb{P}[A(x') = y]$$

Noise addition:

function f

Input $x = (x_1, \dots, x_n) \longrightarrow \boxed{A} \longrightarrow$ Output $A(x) = f(x) + \text{noise}$

Randomized

- Goal: Release approximation to $f(x) \in \mathbb{R}^d$
  e.g., # ppl wearing socks,
- Intuition: $f(x)$ can be released accurately if f is insensitive to the change of individual examples $x_1, \dots, x_n$

Sensitivity.

- Intuition: $f(x)$ can be released accurately if f is insensitive to the change of individual examples $x_1, \ldots, x_n$

Global Sensitivity:

$$GS_f = \max_{x, x' \text{ neighbors}} \| f(x) - f(x') \|_1$$

Example: $f(x) \equiv$ fraction of people wearing socks

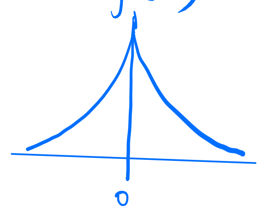
$$GS_f = \frac{1}{n}$$

Laplace Mechanism.

$$A_L(x) = f(x) + (z_1, \dots, z_d)$$

where each z_i drawn i.i.d. from $\mathrm{Lap}\left(\frac{GS_f}{\varepsilon}\right)$

Laplace Distribution : $\mathrm{Lap}(b)$



$$\mathrm{PDF}(x) = \frac{1}{2b} \exp\left(\frac{-|x|}{b}\right)$$

$$\mathop{\mathbb{E}}_{X \sim \mathrm{Lap}(b)} [|X|] = b.$$

Theorem. A_L is ε-differentially private.

Examples.

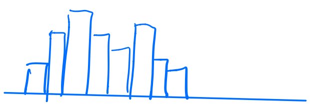
- Proportion. $f(x) = \frac{1}{n} \sum_{i=1}^{n} x_i$

  "fraction of people wearing socks"

  $GS_f = \frac{1}{n}$.

- Histogram. Data domain $\mathcal{X} = B_1 \cup B_2 \cup \cdots \cup B_d$

  $f(x) = (n_1, \dots, n_d)$, $n_j = \#\{i : x_i \in B_j\}$



Examples

- Sequence of d $\underbrace{\text{statistical queries}}_{\text{averages}}$

  properties $\phi_1, \dots, \phi_d$ with each $\phi_j : \mathcal{X} \mapsto [0,1]$

  For each j, $f_j(x) = \frac{1}{n} \sum_{i=1}^{n} \phi_j(x_i)$

  $$GS_{f_j} \le \frac{1}{n}$$

  $$f(x) = \left(f_1(x), \dots, f_d(x) \right), \quad f(x) - f(x') \in \left[-\frac{1}{n}, \frac{1}{n}\right]^d$$

  $$GS_f \le \frac{d}{n}$$