Lecture 5

- Recap

  Definition of Differential Privacy

  Randomized Response

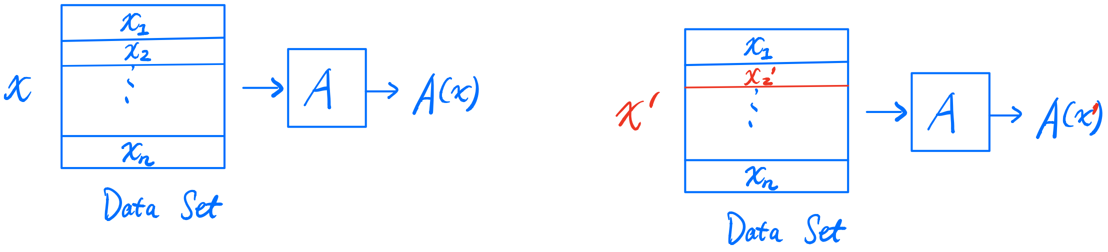
- Laplace Mechanism.

- HW1 is posted; due on Sep 26 Sunday

- Recitation on Friday.

  No Problem set; DP Review
  + Randomized Response
  + Laplace Mech.

Neighboring datasets



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

Definition. (Differential Privacy).

A is ε-differentially private if

for all neighbors x and x'

for all subsets $\underbrace{E}_{\text{event}}$ of outputs

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

Definition. (Differential Privacy).

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

$$\mathbb{P}[A(x) \in E] \leq \underset{\approx(1+\varepsilon)}{e^{\varepsilon}} \mathbb{P}[A(x') \in E]$$

ε: Privacy (Loss) parameter

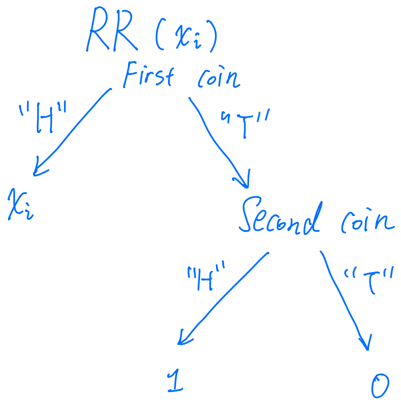
Small constant : $\frac{1}{10}$, 1, but not $\frac{1}{2^{80}}$ or 100

Example : Randomized Response (In lecture 1)

Each person has a secret bit $x_i = 0$ or $x_i = 1$
(Have you ever done XYZ ?)

Input : $x_1, \dots, x_n$

Output : $y_1, \dots, y_n$



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

Basic Proof Strategy :

for all neighbors x and x'
for all subsets E of outputs $(E \subseteq Y)$.

$$\begin{tikzcd} \mathbb{P}[A(x) \in E] \le e^{\varepsilon}\, \mathbb{P}[A(x') \in E] \arrow[d, Leftrightarrow] \\ \mathbb{P}[A(x) = y] \le e^{\varepsilon}\, \mathbb{P}[A(x') = y] \qquad (*) \end{tikzcd}$$

for all y in $\underbrace{Y}_{\text{output space.}}$

$\uparrow$ we prove.

How to construct an estimate?

Input: $(x_1, x_2, \dots, x_n) \in \{0,1\}^n$

For $i = 1, \dots, n$:

$$Y_i = \begin{cases} x_i & \text{with probability } \frac{3}{4} \\ 1 - x_i & \text{with prob. } \frac{1}{4}. \end{cases}$$

Return: $(Y_1, \dots, Y_n)$.

Goal: want to estimate $p = \frac{1}{n} \sum_{i=1}^{n} x_i$

Observe: $\hat{y} = \frac{1}{n} \sum_{i=1}^{n} Y_i$.

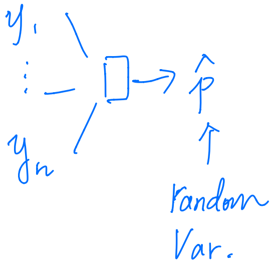
Estimation: $\hat{p} = a\, \hat{y} + b$

Question: What should a & b be?

Trials (a,b) $= (2, -\frac{1}{2})$? ✓

"Implicit" Goal: Find a & b,

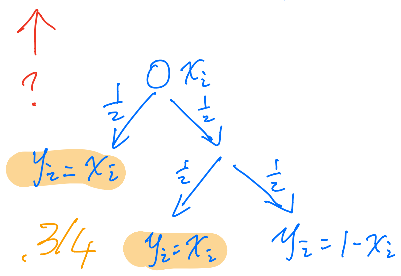
$\mathbb{E}[\hat{p}] = p.$ ← Unbiased estimate



$$\mathbb{E}[\hat{p}] = \mathbb{E}\left[\frac{1}{n} \sum_i (aY_i + b)\right]$$

Linearity of Expectation →

$$= a \cdot \frac{1}{n} \sum_i \mathbb{E}[Y_i] + b.$$



$$\mathbb{E}[Y_i] = \frac{3}{4} x_i + \frac{1}{4}(1 - x_i)$$

$$= \frac{x_i}{2} + \frac{1}{4}.$$

$$\mathbb{E}[\hat{p}] = a \cdot \frac{1}{n} \left(\sum_i \left(\frac{x_i}{2} + \frac{1}{4} \right) \right) + b$$

$$= \frac{a}{2} \Big(\underbrace{\frac{1}{n} \sum_i x_i}_{p} \Big) + \frac{a}{4} + b \underset{\substack{\uparrow \\ \text{unbiased} \\ \text{condition}}}{=} p$$

$$\Longrightarrow \quad a = 2 \;, \quad b = -\frac{1}{2}.$$

In HW1, $\nearrow$ Generalization

Noise addition. (Laplace Mechanism).

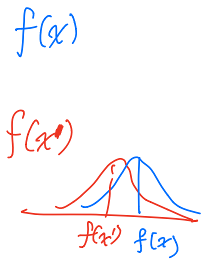
function f

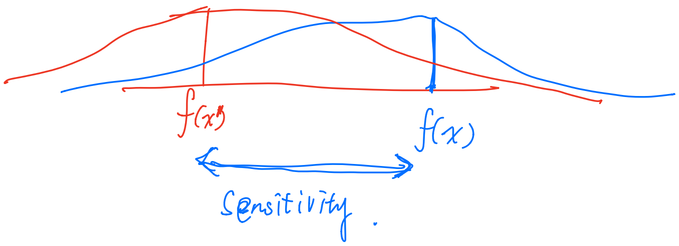
Input $x=(x_1,\ldots,x_n) \longrightarrow \boxed{A} \longrightarrow$ Output $A(x) = f(x) + \text{noise}$

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 $x_1, \ldots, x_n$





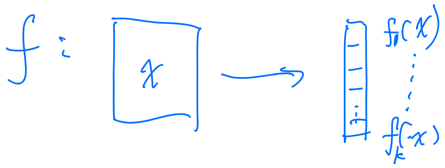
Sensitivity.

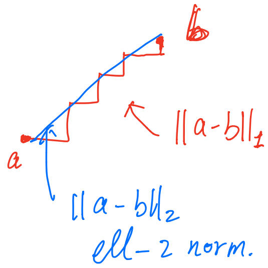
- Intuition: $f(x)$ can be released accurately if f is insensitive to the change of individual examples $x_1, \dots, x_n$

Global Sensitivity:

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

ell_ one norm $\| \cdot \|_1$. $\|v\|_1 = \sum_{j=1}^{d} |v_j|$.



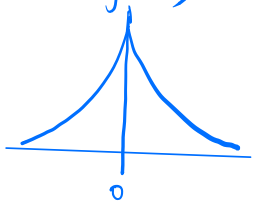


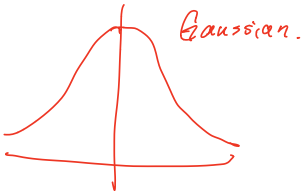
Laplace Mechanism.

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

where each z_i drawn i.i.d. from $\mathrm{Lap}\left(\frac{\overset{\text{Global Sensitivity of } f}{\overset{\downarrow}{GS_f}}}{\varepsilon}\right)$

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





Facts: $z \sim \mathrm{Lap}(b)$, $\mathbb{E}[|z|] = b$

Prob. density function $\longrightarrow$ $PDF(z) = \frac{1}{2b} \exp(-|z|)$.

$\exp(a) = e^a$

Theorem. A_L is ε-differentially private.

Examples.

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

- Proportion.

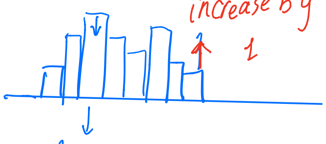
$$f(x) = \frac{1}{n} \sum_{i=1}^{n} x_i$$

" fraction of people wearing socks "

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

$$= \max_{x,x'} \left| \frac{1}{n} x_i - \frac{1}{n} x_i' \right| \le \frac{1}{n}$$

- Histogram.



total difference = 2.

age groups

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

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

count

$$x \longrightarrow x'$$

$$f(x) - f(x')$$