Lecture 20

Local Model of Differential Privacy

- Randomized Response
- Mean Estimation
- Frequency Estimator
- Heavy Hitters

Histogram

Connection
Sublinear Algorithms

Next Module : Connections between

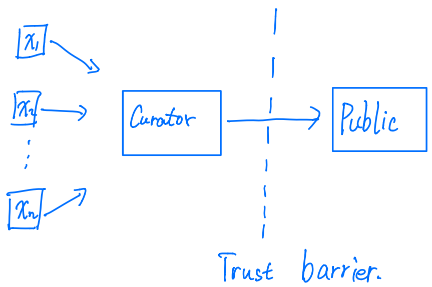
① Adaptive Data Analysis

② "Robustness"
- Mechanism Design
- Adversarial example
- . . .

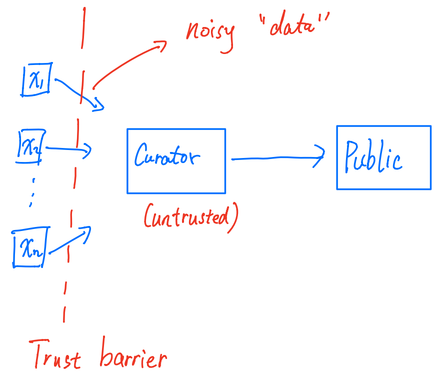
③ Practical Deployments
- local DP.
- US Census. 2020. 2025?

Last Module : 1st week of May : Project Presentation

Central Model



Local Model



Has worse accuracy.

Mean Estimation

$x_1, x_2, \ldots, x_n \in \{0,1\}$

Mean $\mu = \frac{1}{n}\sum_{i=1}^{n} x_i$

Central Model

$$\hat{\mu}_{Lap} = \frac{1}{n}\sum_{i=1}^{n} x_i + Lap\left(\frac{1}{\varepsilon n}\right)$$

$$|\hat{\mu}_{Lap} - \mu| \le O\left(\frac{1}{\varepsilon n}\right) \text{ w.p. } 99\%$$

$$Y_i = x_i + Lap\left(\frac{1}{\varepsilon}\right)$$

Local Model.

Randomized response. (Warner '65)

$$Y_i = \begin{cases} x_i & \text{w.p. } \frac{e^\varepsilon}{1+e^\varepsilon} \\ 1 - x_i & \text{w.p. } \frac{1}{1+e^\varepsilon} \end{cases}$$

$$\hat{\mu}_{RR} = \frac{1}{n}\sum_i \left(\frac{e^\varepsilon+1}{e^\varepsilon-1} Y_i - \frac{1}{e^\varepsilon-1}\right)$$

By Chernoff Bound,

$$|\hat{\mu}_{RR} - \mu| \le O\left(\frac{1}{\varepsilon\sqrt{n}}\right)$$

for $\varepsilon \le 1$.

Local Randomizer $M: X \mapsto Y$

is (ε,δ) - locally differentially private (LDP) if

$\forall \underbrace{x, x'}_{\text{Dataset of size } 1.} \in X, \quad E \subseteq Y$

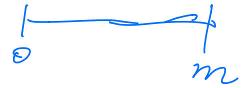
$$\mathbb{P}[M(x) \in E] \le e^\varepsilon \mathbb{P}[M(x') \in E] + \delta.$$

(ε,δ)-DP in the central that makes for size-1 dataset $\Longrightarrow$ (ε,δ)-LDP LR.

Local Model: $(\varepsilon,0)$-LDP v.s. (ε,δ)-LDP ?

Mean Estimation

$x_1, x_2, \dots, x_n \in [0, m]$



$\mu = \frac{1}{n} \sum_i x_i$

Central Model

$\hat{\mu} = \frac{1}{n} \sum_i x_i + Lap\left(\frac{m}{n\varepsilon}\right).$

$|\hat{\mu} - \mu| \le O\left(\frac{m}{n \cdot \varepsilon}\right)$

Sample Complexity for $err \le \alpha$

$n \ge \frac{m}{\alpha \varepsilon}$

Local Model.

$Y_i = x_i + Lap\left(\frac{m}{\varepsilon}\right)$

$\hat{\mu} = \frac{1}{n} \sum_i Y_i$

$|\hat{\mu} - \mu| = \left| \sum_i Lap\left(\frac{m}{\varepsilon}\right) \right|$

$\approx O\left(\frac{m}{\sqrt{n}\, \varepsilon}\right)$

Sample Complexity

$n \ge \frac{m^2}{\varepsilon^2 \alpha^2}$

Communication Issues?

$\Omega(\log(m))$ bits of communication from each user

1-bit Communication.

$x_i \in [0,m]$, $Y_i = \begin{cases} 1 & \text{w.p. } \underbrace{\frac{1}{e^\varepsilon+1} + \frac{x_i}{m}\cdot\frac{e^\varepsilon-1}{e^\varepsilon+1}}_{\mathbb{E}[Y_i]} \\ 0 & \text{o.w.} \end{cases}$

Estimator : $\hat{\mu} = \frac{m}{n}\sum_{i=1}^{n} \frac{Y_i(e^\varepsilon+1)-1}{e^\varepsilon-1}$

$$\mathbb{E}[\hat{\mu}] = \frac{m}{n}\sum_{i=1}^{n}\frac{\mathbb{E}[Y_i]\,(e^\varepsilon+1)-1}{e^\varepsilon-1}$$

$$= \frac{m}{n}\sum_{i=1}^{n}\frac{\left(\boxed{\frac{1}{e^\varepsilon+1}} + \frac{x_i}{m}\cdot\frac{e^\varepsilon-1}{\cancel{e^\varepsilon+1}}\right)\cancel{(e^\varepsilon+1)}\boxed{-1}}{e^\varepsilon-1} \rightarrow \frac{(e^\varepsilon-1)x_i}{m}$$

$$= \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$|\hat{\mu}-\mu| \le O\left(\frac{m}{\varepsilon\sqrt{n}}\right) \quad \text{w.p. } 99\%.$$

for α-accuracy , $n \ge \frac{m^2}{\alpha^2\varepsilon^2}$.

"Implemented Windows" Ding K Y '17. MSR.

Frequency Estimation

$x_1, \dots, x_n \in [d]$ $\quad \underset{\uparrow \text{ data universe.}}{\{1, \dots, d\}}$ $\quad$ Think d as "Large"

$\forall x \in [d]$

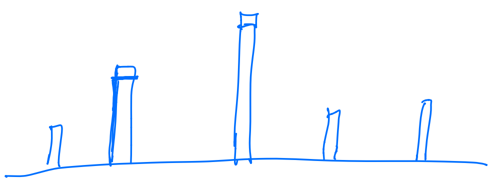
$$f(x) = \sum_{i=1}^{n} \mathbb{1}[x_i = x].$$

$$|f(x) - \hat{f}(x)| \le \alpha.$$

Central Model.

$\forall x$,

$$\hat{f}(x) = f(x) + Lap\left(\frac{2}{\varepsilon}\right)$$



Local Model.

$x_i \in [d]$ $\quad x_i = \underbrace{\boxed{0 \;\; 0\,0 \cdots \overset{x_i\text{-coordinate}}{1} \;\; 0 \cdots 0}}_{\text{length } d.}$

$\downarrow$ Local Randomizer

$$b_i = \boxed{1 \;\; -1 \;\; \cdots\cdots \;\; -1 \;\; 1}$$

$$x_{ij} = 0, \quad b_{ij} = \begin{cases} 1 & w.p. \ \frac{1}{2} \\ -1 & w.p. \ \frac{1}{2} \end{cases}$$

$$x_{ij} = 1, \quad b_{ij} = \begin{cases} 1 & w.p. \ \frac{1}{2} + \frac{\varepsilon}{2} \\ -1 & w.p. \ \frac{1}{2} - \frac{\varepsilon}{2} \end{cases}$$

$$\hat{f} = \left(\sum_i b_i\right) \cdot \frac{1}{\varepsilon}$$

$$\hat{f}(x) = \Big(\underbrace{\sum_{i : x_i = x} b_{ix}}_{\mathbb{E}[\cdot] = \varepsilon} + \underbrace{\sum_{i : x_i \ne x} b_{ix}}_{\mathbb{E}[\cdot] = 0}\Big) \cdot \frac{1}{\varepsilon}$$

$$= \sum_i \mathbb{1}[x_i = x].$$

- Communication $\Omega(d)$ bits
- Runtime for $\left\{ \begin{array}{l} \text{Server} \\ \text{User} \end{array} \right.$ $\Omega(d)$.

Next Lecture:

Communication $\Omega(d) \longrightarrow O(\sqrt{n})$

using idea Count Sketch.

Tree data structure

$\downarrow$

Heavy hitters.

State-of-the-art

$\longrightarrow \tilde{O}(1)$

Count Sketch

Amplification

Heavy Hitters

TreeHist