Lecture 3

Intro to Differential Privacy

1. Randomized Response [Warner '65] –

"How many students cheated?"

n people, individual i, $X_i \in \{0,1\}$

Y_i (depends on X_i, randomness)

Send Y_i to analyst

Goal: estimate $p = \frac{1}{n}\sum X_i$

Attempt 1

$Y_i = X_i$

$$Y_i = \begin{cases} X_i & \text{w.p. } 1 \\ 1-X_i & \text{w.p. } 0 \end{cases}$$

$\hat{p} = \frac{1}{n}\sum Y_i \Rightarrow \hat{p} = p$. Perfect acc.

NO Privacy

Attempt 2

$$Y_i = \begin{cases} X_i & \text{w.p. } 1/2 \\ 1-X_i & \text{w.p. } 1/2 \end{cases}$$

Y_i is perfectly private

$\tilde{p} = \frac{1}{n}\sum Y_i$. $\tilde{p} \sim Bin(n, \frac{1}{2})$

↖ Doesn't depend on p

No acc!

Randomized Response ($\gamma \in (0, \frac{1}{2})$)

$Y_i = \begin{cases} X_i & \text{w.p. } 1/2 + \gamma \\ 1 - X_i & \text{w.p. } 1/2 - \gamma \end{cases}$

$\gamma = \frac{1}{4}$

$\gamma = \frac{1}{2} \to$ Attempt 1

$\gamma = 0 \to$ Attempt 2

Analysis

$$E[Y_i] = (\tfrac{1}{2}+\gamma)X_i + (\tfrac{1}{2}-\gamma)(1-X_i)$$
$$= 2\gamma X_i + \tfrac{1}{2} - \gamma$$
$$E\left[\frac{1}{2\gamma}(Y_i - 1/2 + \gamma)\right] = X_i$$

$$\widetilde{p} = \frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{2\gamma}\left[Y_i - \frac{1}{2} + \gamma\right]\right]$$

$$\mathrm{Var}[\mathrm{Ber}(p)] = p(1-p) \le \frac{1}{4}.$$

$$E[\widetilde{p}] = \frac{1}{n}\sum X_i \triangleq p$$

$$\mathrm{Var}[\widetilde{p}] = \mathrm{Var}\left[\frac{1}{n}\sum\left(\frac{1}{2\gamma}\left(Y_i - \frac{1}{2} + \gamma\right)\right)\right] = \frac{1}{4\gamma^2 n^2}\sum \mathrm{Var}\left(Y_i - \tfrac{1}{2} + \gamma\right)$$

$$\le \frac{1}{4\gamma^2 n^2} \cdot n \cdot \frac{1}{4} = \frac{1}{16\gamma^2 n}$$

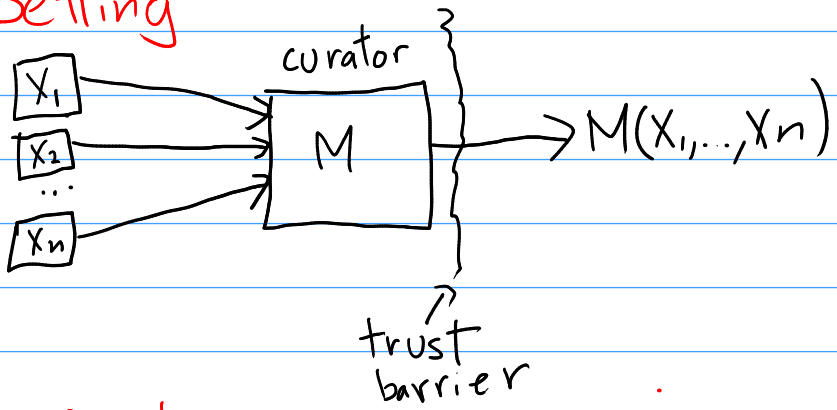
Chebyshev's inequality: $|\widetilde{p} - p| \le O\left(\frac{1}{\gamma\sqrt{n}}\right) \to 0$ as $n \to \infty$

Chernoff $\to$ high prob.

Additive err. α: $n \ge \Omega\left(\frac{1}{\alpha^2\gamma^2}\right)$

2. Differential Privacy

Setting



Definition [Dwork, McSherry, Nissim, Smith '06]

TCC '16 T.e.T. award
2017 Gödel prize

Definition 1.1 ((pure) differential privacy [38]). For $\varepsilon \geq 0$, we say that a randomized mechanism $\mathcal{M}\colon \mathcal{X}^n \to \mathcal{Y}$ is *ε-differentially private* if for every pair of neighboring datasets $x \sim x' \in \mathcal{X}^n$ (i.e., x and x' differ in one row), we have:

$$\forall T \subseteq \mathcal{Y},\ \Pr[\mathcal{M}(x) \in T] \leq e^{\varepsilon} \cdot \Pr[\mathcal{M}(x') \in T]\,.$$

Uber
Microsoft
Google
Facebook
Apple
LinkedIn
US. Census Bureau

Technical Comments

- Quantitative
- ε is smallish. $0.1 \leq \varepsilon \leq 5$
- Worst-case definition
  - ↳ Avg case? Steinke-Ullman '20
- Why e^{ε}? $(1 \pm \varepsilon)$?

If ε small $e^{\varepsilon} \approx 1+\varepsilon$ (Taylor expansion)

$e^{\varepsilon_1} \cdot e^{\varepsilon_2} = e^{\varepsilon_1+\varepsilon_2}$

- Symmetric
- Randomized.

- Why multiplicative?
- "Neighbouring"?

Interlude: Hypothesis Testing

Wasserman + Zhou '10

One of the two holds:
H_0: the underlying dataset is X X, X' neighbouring
H_1: " " X'
Based on M on dataset, decide H_0 vs H_1
"If M is DP, guessing is ≈ best strat"
p: prob of adv. guessing H_1 when H_0 (false alarm)
q: " " H_0 when H_1 (missed discovery)
ε-DP $\Leftrightarrow$ $p + e^{\varepsilon} q \geq 1$
" $e^{\varepsilon} p + q \geq 1$
- Operational
- Composition [KOV '15]
- GDP [DRS '19]

Interpreting Differential Privacy

- DP → Prob of any event is comparable whether or not an individual is in the dataset

- Rules out attacks
  - ↳ Linkage
  - ↳ Reconstruction
  - ↳ Arbitrary risks

- Not for "individual" level learning
- Aggregate level ✓

What doesn't DP guarantee?

- Allows stats + ML
- "Smoking causes cancer"

- DP is information theoretic

Randomized Response, Revisited

$M = \gamma$-RR. How DP is M?

Suppose M outputs $(Y_1, \ldots, Y_n)$ $\quad Y_i = \begin{cases} X_i & \text{w.p. } \frac{1}{2}+\gamma \\ 1-X_i & \frac{1}{2}-\gamma \end{cases}$

Consider $a \in \{0,1\}^n$

$\Pr[M(X)=a] = \prod \Pr[Y_i = a_i]$.

Suppose X, X' differ in coord j only

$$\frac{\Pr[M(X)=a]}{\Pr[M(X')=a]} = \frac{\prod \Pr[Y_i=a_i]}{\prod \Pr[Y_i'=a_i]} = \frac{\Pr[Y_j=a_j]}{\Pr[Y_j'=a_j]} \leq \frac{\frac{1}{2}+\gamma}{\frac{1}{2}-\gamma}$$

M is $\ln\left(\frac{\frac{1}{2}+\gamma}{\frac{1}{2}-\gamma}\right)$-DP

$\gamma \leq \frac{1}{4} \Rightarrow \frac{\frac{1}{2}+\gamma}{\frac{1}{2}-\gamma} \leq e^{O(\gamma)} \Rightarrow M$ is $O(\gamma)$-DP.

ε-RR: ε-Local DP

- ε-DP
- Accuracy of $|\tilde{p}-p| \leq O\left(\frac{1}{\varepsilon\sqrt{n}}\right)$

- Laplace Algo