Lecture 2 (Draft)

Today

- Two private algorithms
- Properties of DP

Rewrite def, plus: composition, postprocessing, group privacy → will make quantitative later

Randomized Response

Simplest DP algorithm.
Operates on a single bit.

Alg 1: $RR_\varepsilon(x)$

Input: $x \in \{0,1\}$, $\varepsilon \geq 0$

Returns: $\tilde{x}$

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

Claim RR_ε is ε-DP.

Proof

$$\Pr[RR_\varepsilon(0)=0] = \frac{e^\varepsilon}{1+e^\varepsilon}$$

$$= e^\varepsilon \cdot \frac{1}{1+e^\varepsilon}$$

$$= e^\varepsilon \cdot \Pr[RR_\varepsilon(1)=0]$$

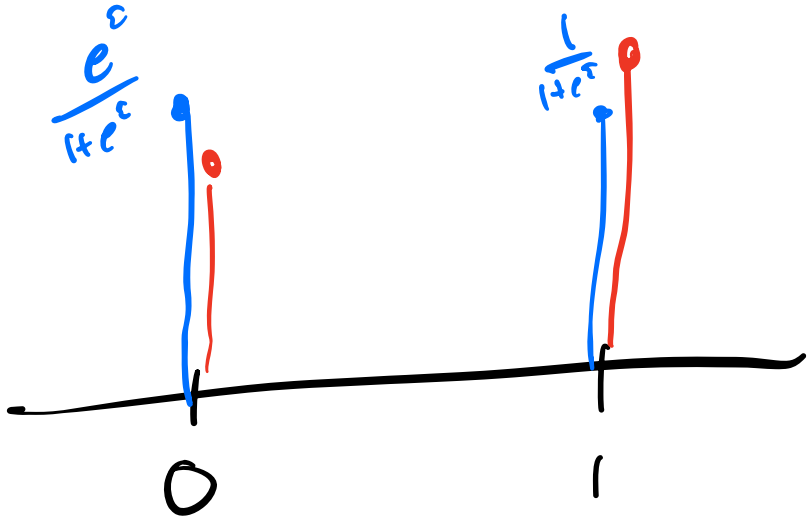
$$\Pr[RR_\varepsilon(0)=1] = \frac{1}{1+e^\varepsilon}$$

$$\leq \frac{e^\varepsilon}{1+e^\varepsilon}$$

$$= \Pr[RR_\varepsilon(1)=1]$$

$$\leq e^\varepsilon \Pr[RR_\varepsilon(1)=1]$$

$\boxtimes$



Laplace Noise & Mean Estimation

Given: n iid samples from $\text{Bern}(p)$

Goal: estimate p

$X \sim \text{Bern}(p)$, $p \in [0,1]$

$\Pr[X = 1] = p$

$\Pr[X = 0] = 1 - p$

Alg 2 Laplace Mechanism

Input: $x_1, \ldots, x_n \in \{0,1\}$, $\varepsilon \geq 0$

Returns: $\hat{p}$

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

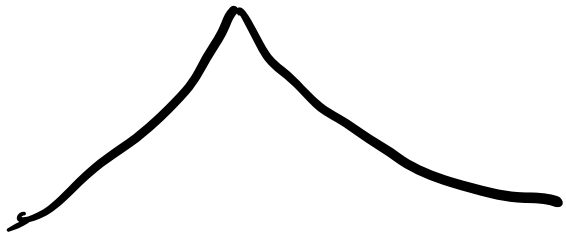
② $\tilde{s} \leftarrow \bar{s} + Z$, $Z \sim \text{Lap}(1/\varepsilon)$

③ Return $\frac{1}{n} \tilde{s}$

Two separate analyses.

$\text{Lap}(\mu, b)$

$p(z) = \frac{1}{2b} e^{-\frac{|x-\mu|}{b}}$



Claim Alg 2 is ε-DP.

Pf Suffices to show releasing $\tilde{s}$ is ε-DP.

Fix adjacent x, x' & event $E \subseteq \mathbb{R}$.

For any $y \in E$,

$$\Pr_{\tilde{s} \leftarrow A(x)}[\tilde{s} = y] = \frac{1}{2b} \exp\left(-\varepsilon \left| y - \sum_i x_i \right|\right)$$

$$= \frac{1}{2b} \exp\left(-\varepsilon \left| y - \left(\sum_i x_i' - x_i' + x_i\right)\right|\right)$$

$$\leq \frac{1}{2b} \exp\left(-\varepsilon \left| y - \sum_i x_i' \right| + \varepsilon \left| x_i' - x_i \right|\right)$$

$$\leq e^{\varepsilon} \Pr_{\tilde{s} \leftarrow A(x')}[\tilde{s} = y]$$

$$\Pr_{\tilde{s} \leftarrow A(x)}[\tilde{s} \in E] = \int_E \Pr_x[\tilde{s} = y]\, dy \leq e^{\varepsilon} \int_E \Pr_{x'}[\tilde{s} = y]\, dy$$

$$= e^{\varepsilon} \Pr_{x'}[\tilde{s} \in E].$$

$\square$

Claim When $x_1, \dots, x_n \overset{iid}{\sim} \text{Bern}(p)$, Alg 2 satisfies $\mathbb{E}\left[(p - \tilde{p})^2\right] \le \frac{p(1-p)}{n} + \frac{2}{\varepsilon^2 n^2}$.

Proof $\mathbb{E}\left[(p - \tilde{p})^2\right] = \mathbb{E}\left[\left(p - \frac{1}{n}\bar{s} + \frac{1}{n}\bar{s} - \frac{1}{n}\tilde{s}\right)^2\right]$

$= \mathbb{E}\left[\left(p - \frac{1}{n}\bar{s} + \frac{1}{n}\bar{s} - \left(\frac{1}{n}\bar{s} + \frac{1}{n}z\right)\right)^2\right]$

$= \mathbb{E}\left[\left(p - \frac{1}{n}\bar{s}\right)^2 + \frac{1}{n^2}(z)^2\right]$

$= \frac{1}{n^2}\mathbb{E}\left[(np - s)^2\right] + \frac{1}{n^2}\mathbb{E}\left[z^2\right]$

↙ Variance of $\text{Bin}(n, p)$ $= np(1-p)$

↓ Variance of $\text{Lap}(1/\varepsilon)$ $= \frac{2}{\varepsilon^2}$

$= \underbrace{\frac{p(1-p)}{n}}_{\text{Sampling error}} + \underbrace{\frac{2}{\varepsilon^2 n^2}}_{\text{"cost of privacy"}}$ ☒

Properties

Claim (Immunity to Postprocessing) If $A: \mathcal{X}^n \to \mathcal{Y}$ is (ε, δ)-DP and $f: \mathcal{Y} \to \mathcal{Z}$ is a randomized fn, then $f \circ A$ is (ε, δ)-DP.

Proof

f randomized, uses random coins, write

$$B: \mathcal{Y} \times \mathcal{R} \to \mathcal{Z}$$

Fix adjacent datasets X, X'.

Fix event $E \subseteq \mathcal{Z}$.

$$\Pr_{\substack{A(x), \\ R}}[f(A(x), R) \in E] = \sum_r \Pr[R = r] \Pr[f(A(x), R) \in E \mid R = r]$$

$$= \sum_r \Pr[R = r] \Pr_{A(x)}[f(A(x), r) \in E]$$

$$= \mathbb{E}_R\left[\Pr_{A(x)}[f(A(x), R) \in E]\right]$$

For each r, $\exists$ event $E'_r \subseteq \mathcal{Y}$ s.t.
$A(x) \in E'_r$ then $B(A(x), r) \in E$.

$$\cdots = \mathop{\mathbb{E}}_{R}\left[\Pr_{A}\left[A(x) \in E'_r\right]\right]$$

$$\leq \mathop{\mathbb{E}}_{R}\left[e^{\varepsilon}\Pr_{A}\left[A(x') \in E'_r\right] + \delta\right]$$

$$= \mathop{\mathbb{E}}_{R}\left[e^{\varepsilon}\Pr_{A}\left[B(A(x'), R) \in E\right] + \delta\right]$$

$$= e^{\varepsilon}\Pr_{A,R}\left[B(A(x'), R) \in E\right] + \delta.$$

□

Claim (Group Privacy) Suppose A is (ε, δ)-DP and let x, x' differ in at most k entries. Then $A(x) \approx_{(\varepsilon', \delta')} A(x')$ for

$\varepsilon' = k\varepsilon, \quad \delta' = k \cdot e^{\varepsilon k}\delta.$

Proof for $k=2$

. . .

□

Claim (Basic Composition) (Actually, $M_1, \ldots, M_k$) Let algs A, B be (ε, δ)-DP. Then $(A(x), B(x))$ is $(2\varepsilon, 2\delta)$-DP.

Note: even holds adaptively, if $B(x)$ gets A's output as input.

Proof for $\delta = 0$

Let $M(x) = (A(x), B(x))$

$$L_M^{x \to x'}(y) \triangleq \ln\left(\frac{\Pr[M(x)=y]}{\Pr[M(x')=y]}\right)$$

Privacy loss

Pure DP equiv to, $\forall y, x, x' \quad |L_M^{x \to x'}(y)| \leq \varepsilon.$