Theory of Differential Privacy

Lecture 3:

- Accuracy of: Randomized Response, Laplace Mechanism
- Properties of DP
- Report Noisy Max

EN.601.438/638 Fall 2026

Lydia Zakynthinou

Johns Hopkins University

Sept 9, 2026

Recap

Def. A is ε-DP $\iff$ $\Pr[A(x)\in S] \le e^{\varepsilon} \Pr[A(x')\in S]$

$\forall$ neighboring datasets $x \sim x'$

and $\forall$ measurable event S.

Def. Given function $f: \mathcal{X}^n \to \mathbb{R}^d$ its Global Sensitivity (in ℓ_1)

is $\Delta_1(f) := \sup_{x\sim x'} \|f(x) - f(x')\|_1$

Def. Randomized Response

$$RR_\varepsilon(x) = \begin{cases} x & \text{w.p. } p_\varepsilon = \frac{e^\varepsilon}{1+e^\varepsilon} \\ 1-x & \text{w.p. } q_\varepsilon = \frac{1}{1+e^\varepsilon} \end{cases}$$

$x \in \{0,1\}$

If $x \in \mathcal{X}^n$: $RR_\varepsilon(x) = (RR_\varepsilon(x_1), \dots, RR_\varepsilon(x_n))$

Def. Laplace Mechanism $A_{Lap}(x) = f(x) + Z$, where $Z = (Z_1, \dots, Z_d)$

$f: \mathcal{X}^n \to \mathbb{R}^d$, $A_{Lap}: \mathcal{X}^n \to \mathbb{R}^d$, $Z_i \text{ iid } \mathrm{Lap}\left(\frac{\Delta}{\varepsilon}\right)$.

$\Delta_1(f) \le \Delta$

Examples of Global Sensitivity

- Count $\quad f(x) = \sum_{i=1}^{n} \phi(x_i)$ for $\varphi : \mathcal{X} \to \{0,1\} \Rightarrow \Delta_1 f = 1$

- Proportion $\quad f(x) = \frac{1}{n} \sum_{i=1}^{n} \phi(x_i) \Rightarrow \Delta_1 f = \frac{1}{n}$

- Bounded mean : if $x_i \in [a,b]$ and $f(x) = \frac{1}{n} \sum_{i=1}^{n} x_i \Rightarrow \Delta_1 f = \frac{b-a}{n}$

- Histogram counts

$j \in [d]$

$|\{i \in [n] : x_i \in B_j\}| = c_j$

$$f(x) = \Big(\underbrace{\overset{c_j}{\bigcirc}, \ldots}_{\#\text{buckets}} \Big)$$

$f(x) : \mathcal{X}^n \to \mathbb{R}^d$

$$\Delta_1 f = \sup_{x \sim x'} \| f(x) - f(x') \|_1 = \sum_{j \in [d]} | f(x)_j - f(x')_j | = 2$$

Accuracy: learning the mean

$f(x) = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \quad x_i \in \{0,1\}$.

- **Laplace** $A_L(x) = f(x) + Z, \quad Z \sim Lap\left(\frac{1}{n\varepsilon}\right)$

$$\mathbb{E}[A_L(x)] = \mathbb{E}[\bar{x} + Z] = \bar{x} + \mathbb{E}[Z] = \bar{x} \checkmark$$

$$\sqrt{\mathbb{E}\left((A_L(x) - \bar{x})^2\right)} = \sqrt{\mathbb{E}[Z^2]} = \sqrt{O\left(\frac{1}{n^2\varepsilon^2}\right)} \approx O\left(\frac{1}{n\varepsilon}\right) \checkmark$$

- $RR_\varepsilon(x) = (Y_1, \dots, Y_n)$

$$Y_i = \begin{cases} x_i, & \text{wp } e^\varepsilon/(1+e^\varepsilon) \\ 1-x_i, & \text{wp } 1/(1+e^\varepsilon) \end{cases}$$

$$\bar{Y} = \frac{1}{n}\sum Y_i$$

$$\mathbb{E}[\bar{Y}] = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[Y_i] = \frac{1}{n}\sum_{i=1}^{n}\left(x_i \frac{e^\varepsilon}{1+e^\varepsilon} + (1-x_i)\frac{1}{1+e^\varepsilon}\right)$$

$$= \frac{1}{n}\sum_{i=1}^{n}\left(\frac{e^\varepsilon - 1}{1+e^\varepsilon} x_i + \frac{1}{1+e^\varepsilon}\right) = \frac{e^\varepsilon - 1}{e^\varepsilon + 1}\underbrace{\left(\frac{1}{n}\sum x_i\right)}_{\bar{X}} + \frac{1}{1+e^\varepsilon}$$

Output $\hat{p} = \dfrac{(e^{\varepsilon}+1)\bar{Y} - 1}{e^{\varepsilon} - 1} \Rightarrow \mathbb{E}[\hat{p}] = \bar{x}$

$$\mathbb{E}[(\hat{p} - \bar{x})^2] \approx \left(\frac{e^{\varepsilon/2}\sqrt{n}}{(e^{\varepsilon}-1)n}\right)^2 \approx \left(\frac{e^{\varepsilon/2}}{(e^{\varepsilon}-1)\sqrt{n}}\right)^2 \underset{\text{for small } \varepsilon}{\approx} \left(\frac{1}{\varepsilon\sqrt{n}}\right)^2$$

★ Hint to statistical estimation: $x \sim \mathrm{Ber}(p)$. Error of Laplace if I want to estimate p:

$$\mathbb{E}[(\hat{p} - p)^2] = \mathbb{E}[((\bar{x} - p) + (\hat{p} - \bar{x}))^2]$$

$$= \mathbb{E}[((\bar{x} - p) + z)^2]$$

$$= \underbrace{\mathbb{E}[(\bar{x} - p)^2]}_{} + 2\,\mathbb{E}[\overset{\text{ind.}}{z}(\bar{x} - p)] + \mathbb{E}[z^2]$$

$$= \frac{p(1-p)}{n} + 0 + O\left(\frac{1}{n^2\varepsilon^2}\right)$$

"sampling error" or "statistical error"

"cost of privacy" or " " error"

Properties of DP

- **Post-processing:** If $A : \mathcal{X}^n \to \mathcal{Y}$ is ε-DP and $B : \mathcal{Y} \to \mathcal{Z}$ (does not access x), then $B(A(x))$ is also ε-DP.

Proof sketch: For $T \subseteq \mathcal{Z}$: (assume B is deterministic)

$$\Pr[B(A(x)) \in T] = \Pr[A(x) \in B^{-1}(T)]$$
$$\leq \Pr[A(x') \in B^{-1}(T)] \cdot e^{\varepsilon}$$
$$= \Pr[B(A(x')) \in T]$$

For B randomized: define $B(R, A(x))$ where R are the random bits of B.

$$\Pr[B(R, A(x)) \in T] = \mathop{\mathbb{E}}_{r \sim R}\Big[\Pr[\underbrace{B(r, A(x))}_{\text{for fixed } r}) \in T]\Big] \ldots\ldots$$

$B(r, A(x))$ deterministic

$\Rightarrow$ proceed as above.

Basic Adaptive Composition.

$$x \to A_1 \to w_1$$

$$(x, w_1) \to A_2 \to w_2$$

$$A(x) = (w_1, w_2)$$

If A_1 is ε_1-DP and $A_2(\cdot, w_1)$ is also ε_2-DP $\forall w_1$

then $A(x) = (A_1(x), A_2(x, A_1(x)))$ is $(\varepsilon_1 + \varepsilon_2)$-DP.

Proof sketch:

$$\Pr[A(x) = (w_1, w_2)] = \Pr[A_1(x) = w_1] \cdot \Pr[A_2(w_1, x) = w_2]$$

$$\leq e^{\varepsilon_1} \Pr[A_1(x') = w_1] \cdot e^{\varepsilon_2} \Pr[A_2(w_1, x') = w_2]$$

$$= e^{\varepsilon_1 + \varepsilon_2} \cdot \Pr[A(x') = (w_1, w_2)].$$

... for K adaptive composition:

$$\left(\sum_{i=1}^{K} \varepsilon_i\right) - DP$$

Group Privacy

If $d_H(x, x') = k$

If A is ε-DP then $\Pr[A(x) \in S] \le e^{k\varepsilon} \cdot \Pr[A(x') \in S]$.

Summary

Post-processing $\varepsilon \to \varepsilon$

Basic adaptive comp. $\varepsilon \to k \cdot \varepsilon$ $\longrightarrow$ example. k Laplace

$\varepsilon_{total} \to$ each $A_{Lap,j}$ $\varepsilon_j = \frac{\varepsilon_{total}}{k}$

$$\mathrm{Lap}\left(\frac{\Delta}{\varepsilon_j}\right) = \mathrm{Lap}\left(\frac{k \cdot \Delta}{\varepsilon_{total}}\right).$$

Group privacy $\varepsilon \to k \cdot \varepsilon$

Selection

Candidates $\mathcal{Y} = \{1, \ldots, d\}$.

$x \in \mathcal{X}^n$

Score function $q(y;x)$

Assume: $|q(y;x) - q(y;x')| \leq \Delta \quad \forall y \in \mathcal{Y}$.

Goal:

$$y^* = \underset{y \in \mathcal{Y}}{\operatorname{argmax}}\ q(y;x)$$

want $\hat{y}$ s.t. $q(\hat{y};x) \geq q(y^*;x) - \text{"small"}$

Naive approach: release $\forall y \in \mathcal{Y}$: $q(y;x) + Lap\left(\frac{d\Delta}{\varepsilon}\right)$

choose $y = \underset{y \in \mathcal{Y}}{\operatorname{argmax}} (q(y,x) + Z_y) \Rightarrow$ error $\mathbb{E}[|Z_y|] \approx O\left(\frac{d\Delta}{\varepsilon}\right)$.

Report Noisy Max

Step 1:

$\forall y \in \mathcal{Y}: q(y;x) + Z_y = \tilde{q}(y;x)$, $\quad Z_y \sim \mathrm{Exp}\left(\frac{2\Delta}{\varepsilon}\right)$.

Step 2:

Output: $\hat{y} = \underset{y \in \mathcal{Y}}{\operatorname{argmax}}\ \tilde{q}(y;x)$

$\Pr[Z_y \geq t] = e^{-t/b}$, $t \geq 0$.

Proof sketch:

Privacy: fix $Z_u = z_u \quad \forall u \neq y$.

Candidate y "wins" if $q(y;x) + Z_y \geq \max_{u \neq y} (q(u;x) + z_u)$.

Threshold: $T_x = \max_{u \neq y} \{q(u;x) + z_u\} - q(y;x)$.

Equiv. if $Z_y \geq T_x$

- sensitivity of T_x: $|T_x - T_{x'}|$

$\rightarrow \left| (q(u;x) + \cancel{z_u} - q(y;x)) - (q(u;x') + \cancel{z_u} - q(y;x')) \right|$

$$= \left| (q(u;x) - q(u;x')) + (q(y;x') - q(y;x)) \right| \leq 2\Delta$$

$$\Rightarrow |T_x - T_{x'}| \leq 2\Delta \quad \text{(Eq. 1)}$$

$$\frac{\Pr[\hat{y} = y \mid z_{-y}, x]}{\Pr[\hat{y} = y \mid z_{-y}, x']} = \frac{\Pr[z_y \geq T_x \mid z_{-y}, x]}{\Pr[z_y \geq T_{x'} \mid z_{-y}, x']} = \frac{e^{-\frac{[T_x]_+}{(2\Delta/\varepsilon)}}}{e^{-\frac{[T_{x'}]_+}{(2\Delta/\varepsilon)}}}$$

$$= \exp\left(\frac{[T_{x'}]_+ - [T_x]_+}{(2\Delta/\varepsilon)}\right) \leq \exp\left(\frac{|T_{x'} - T_x|}{(2\Delta/\varepsilon)}\right) \overset{\text{(Eq. 1)}}{\leq} e^{\varepsilon}$$

Then "average" over z_{-y}:

$$\Pr[\hat{y} = y \mid x] = \mathop{\mathbb{E}}_{\substack{z_u \sim \mathrm{Exp}(2\Delta/\varepsilon) \\ \forall u \neq y}} \left[\Pr[\hat{y} = y \mid z_{-y}, x]\right]$$

Then "sum" over set S:

$$\Pr[\hat{y} \in S \mid x] = \sum_{y \in S} \Pr[\hat{y} = y \mid x] \ldots \Rightarrow \boxed{\text{Report Noisy Max is } \varepsilon\text{-DP}}$$