Lecture 7, Feb 10 2026

Today
- Exponential Mechanism
- Inverse Sensitivity Mechanism

Exponential Mechanism

dataset $x_1, \dots, x_n \in \mathcal{X}$

candidate set $\mathcal{Y}$

utility measure $u(x, y)$

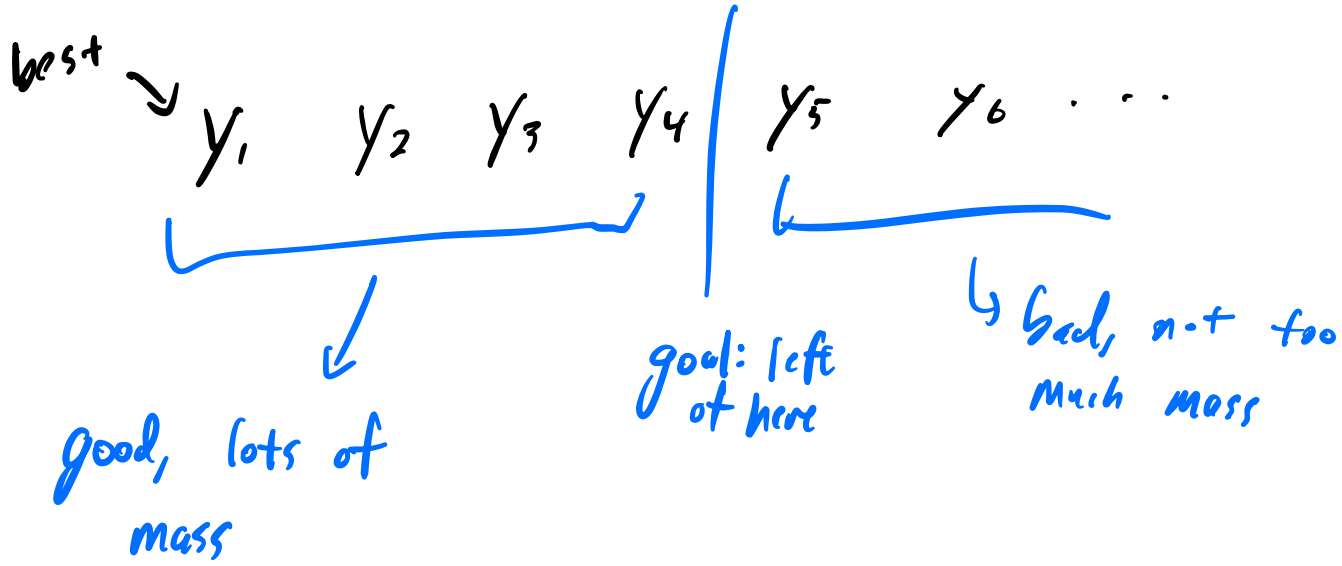
sensitivity bound: $\forall x \sim x', y, \; |u(x, y) - u(x', y)| \le \Delta$

Output $p(y) \propto \exp\left(\frac{\varepsilon}{2\Delta} u(x, y)\right)$

Satisfies ε-DP.

Sketch of Standard utility argument

Fix x, order y by decreasing utility



Claim Suppose $|\mathcal{Y}|$ is finite. Then for any $t > 0$, with $\hat{Y} \sim p(y) \propto \exp\left(\frac{\Sigma}{2} u(x;y)\right)$,

$$\Pr\left[u(\hat{Y};x) < \max_y u(y;x) + \frac{2(t + \log|\mathcal{Y}|)}{\varepsilon}\right] \leq e^{-t}.$$

Proof Assume $\max_y u(y;x) = 0$ (doesn't change anything)

~~$GOOD = \{y \in \mathcal{Y} : u(y;x) > -\frac{2}{\varepsilon}(t + \log|\mathcal{Y}|)\}$~~

$$BAD = \{y \in \mathcal{Y} : u(y;x) \le -\frac{2}{\varepsilon}(t + \log|\mathcal{Y}|)\}.$$

$$\Pr[\hat{Y} \in BAD] = \frac{\sum_{y \in BAD} \exp\left(\frac{\varepsilon}{2} u(y;x)\right)}{\sum_{y \in \mathcal{Y}} \exp\left(\frac{\varepsilon}{2} u(y;x)\right)} \le ???$$

Let $y^* \in \operatorname{argmax} u(y;x)$

$$\Pr[\hat{Y} \in BAD] \le \frac{\Pr[\hat{Y} \in BAD]}{\Pr[\hat{Y} = y^*]} = \frac{\sum_{y \in BAD} \exp\left(\frac{\varepsilon}{2} u(y;x)\right)}{\exp\left(\frac{\varepsilon}{2} \cdot u(y^*;x)\right)}$$

$$\le |BAD| \cdot \exp\left(\frac{\varepsilon}{2} \cdot \left[-\frac{2}{\varepsilon}(t + \log|\mathcal{Y}|)\right)\right)$$

$$= |BAD| \cdot e^{-t} \cdot \frac{1}{|\mathcal{Y}|} \le e^{-t}$$

$\boxtimes$

Task: median estimation (continuous)

$x_1, \dots, x_n \in [0, R]$ — known a priori, ie pulled out of thin air

empirical CDF: $\hat{F}_x(y) = \frac{1}{n} \#\{i : x_i \le y\}$

Use $u(x, y) = -n\left|\frac{1}{2} - \hat{F}_x(y)\right|$

sample $p(y) \propto \exp\left(\frac{\varepsilon}{2} u(x, y)\right)$.

$\Pr\left[\left|\frac{1}{2} - F_x(\hat{y})\right| > \alpha\right]$, need to set good vs bad.

try: $\le \dfrac{\Pr\left[\left|\frac{1}{2} - F_x(\hat{y})\right| > \alpha\right]}{\Pr\left[\left|\frac{1}{2} - F_x(\hat{y})\right| \le \alpha\right]} \le \dfrac{Vol(BAD)}{Vol(not\ BAD)} \cdot \dfrac{\exp\left(-\frac{\varepsilon}{2} \cdot n\alpha\right)}{\exp\left(-\frac{\varepsilon}{2} n\alpha\right)}$

$\to \le R$

$\hookrightarrow = \left|x_{n(\frac{1}{2}+\alpha)} - x_{n(\frac{1}{2}-\alpha)}\right|$

Uh-oh!

$GOOD = \left\{y : \left|\frac{1}{2} - F_x(y)\right| \le \frac{\alpha}{2}\right\}$

$$\Pr[\hat{y} \in BAD] \le \frac{R}{\left|x_{n(\frac{1}{2}-\frac{\alpha}{2})} - x_{n(\frac{1}{2}+\frac{\alpha}{2})}\right|} \cdot \frac{\exp\left(-\frac{\Sigma}{2} n\alpha\right)}{\exp\left(-\frac{\Sigma}{4} n\alpha\right)}$$

$$= \frac{1}{w_{x,\frac{\alpha}{2}}} \cdot \exp\left(\frac{\log R}{w_{x,\frac{\alpha}{2}}} - \frac{\Sigma n\alpha}{4}\right) \le \beta$$

$$-\frac{\log R}{w_{x,\frac{\alpha}{2}}} + \frac{\Sigma n\alpha}{4} \ge \log 1/\beta$$

$$n \ge \frac{4}{\Sigma\alpha}\left(\frac{\log R}{w_{x,\frac{\alpha}{2}}} + \log 1/\beta\right)$$

For Gaussian $\mathcal{N}(\mu, \sigma^2)$, whp

$$\hat{F}_X(z) \approx \frac{1}{2} + \frac{x-\mu}{\sqrt{2\pi\sigma^2}} \pm O\left(\frac{1}{\sqrt{n}}\right)$$

Inverse Sensitivity Mechanism

Want to privately release $f: \mathcal{X}^n \to \mathcal{Y}$.

Run exponential mechanism with

$$u(y; x) \triangleq -\text{len}_f(y; x) = -\min_{x'} \{ d(x, x') : f(x') = y \}$$

median is basically an example.

History

- Folklore mechanism
- Asi & Duchi 2021: investigation of its instance-optimality
- Concurrent 2022: Asi, Ullman, Zakynthinou & Hopkins, Kamath, Majid, Narayanan:

  use $f(x) =$ robust estimator

  optimal accuracy for many problems.

Research Q's: when can we do better?

when can we do it efficiently?

→ compute len

→ sample