

Theory of Differential Privacy

Lecture 6:

Exponential Mechanism

EN.601.438/638 Fall 2026

Lydia Zakynthinou

Johns Hopkins University

Sept 21, 2026

Setup

$x \in \mathcal{X}^n$, candidates $y \in \mathcal{Y}$, score $q(y; x) : \mathcal{Y} \times \mathcal{X}^n \rightarrow \mathbb{R}$.

Goal: $\hat{y}$ s.t. $q(\hat{y}; x) \approx q^*(x) := \sup_y q(y; x)$.

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

Def.: Exp. Mech.

For finite output set, let μ be a counting / base measure.

"normalizing constant", $Z_x = \int_{\mathcal{Y}} \exp\left(\frac{\varepsilon}{2\Delta q} q(z; x)\right) d\mu(z) < \infty$

The Exp. Mech. has density $p_x(y) = \frac{\exp\left(\frac{\varepsilon}{2\Delta q} q(y; x)\right)}{Z_x}$ with respect to μ .

If $\mathcal{Y}$ finite:

$$\Pr[M_{EM}(x) = y] = \frac{\exp(\varepsilon \cdot q(y; x) / 2\Delta q)}{\sum_{z \in \mathcal{Y}} \exp(\varepsilon \cdot q(z; x) / 2\Delta q)}$$

Privacy

The Exp. Mech. is ε -DP.

Proof sketch: Fix $x \sim x'$. for fixed point y :

$$\begin{aligned} \frac{p_x(y)}{p_{x'}(y)} &= \frac{\exp(\varepsilon \cdot q(y; x) / 2\Delta q)}{\exp(\varepsilon \cdot q(y; x') / 2\Delta q)} \cdot \frac{Z_{x'}}{Z_x} \\ &= \exp\left(\left(\frac{\varepsilon}{2\Delta q}\right) \cdot \overbrace{(q(y; x) - q(y; x'))}^{\leq \Delta q}\right) \cdot \frac{Z_{x'}}{Z_x} \\ &\leq e^{\varepsilon/2} \cdot \frac{Z_{x'}}{Z_x} \quad (1) \end{aligned}$$

$$Z_{x'} = \int_Y \exp\left(\left(\frac{\varepsilon}{2\Delta q}\right) \cdot q(y; x')\right) d\mu(z)$$

$$\leq \int_Y \exp\left(\left(\frac{\varepsilon}{2\Delta q}\right) \cdot (q(y; x) + \Delta q)\right) d\mu(z)$$

$$= Z_x \cdot e^{\varepsilon \Delta q / 2\Delta q} = Z_x \cdot e^{\varepsilon/2} \quad (2)$$

$$(1) \stackrel{(2)}{\implies} \frac{P_x(y)}{P_{x'}(y)} \leq e^\varepsilon.$$



Equivalently, E.M. can be written as $P_x^{(a)}(y) \propto e^{a \cdot q(y; x)}$

Then privacy budget is $\varepsilon = 2a\Delta q$.

Example 1:

Learn $f: \mathcal{X}^n \rightarrow \mathbb{R}^d$, with ℓ_1 -sensitivity Δ_1 .

Let $q(y; x) = -\|y - f(x)\|_1$. Then:

$$\Delta q \leq \Delta_1.$$

$$\text{E.M.: } P_x^{(a)}(y) \propto \exp\left(-\frac{\epsilon}{2\Delta_1} \|y - f(x)\|_1\right).$$

$$\begin{aligned} Z_x &= \int_{\mathbb{R}^d} \exp(-\alpha \|y - f(x)\|_1) dy \xrightarrow{z = y - f(x)} \int_{\mathbb{R}^d} \exp(-\alpha \|z\|_1) dz \\ &= \prod_{i=1}^d \int_{\mathbb{R}} e(-\alpha \cdot |z_i|) dz_i = \left(\frac{2}{\alpha}\right)^d. \end{aligned}$$

Independent of x so $Z_x = Z_x!$

$\Rightarrow$ Set $\alpha = \frac{\epsilon}{\Delta_1}$. The E.M. $\nearrow$ with $\alpha = \frac{\epsilon}{\Delta_1}$ is ϵ -DP.*

! Laplace Mech. : $f(x) + \text{Lap}\left(\frac{\Delta_1}{\epsilon}\right) \approx \text{laplace}(f(x), \frac{\Delta_1}{\epsilon}) \equiv \text{E.M. } P_x^{(\epsilon/\Delta_1)}(y)$ for

* confirm by going over the privacy proof sketch above.

above q .

Example 2:

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

$$q(y; x) = \mathbf{1}_{\{y=x\}} \quad \Delta q = 1.$$

$$p_x^{(a)}(y) \propto \exp(a \cdot q(y; x)).$$

$$\Pr[Y=x] \propto \exp(a) \quad \Pr[Y \neq x] \propto 1$$

$$Z_x = Z_{x'} = \frac{1}{1+e^a}$$

$$\Pr[Y=x] = \frac{e^a}{1+e^a}$$

⚡ E.M. with $a = \varepsilon$ and the above choice of $q \equiv \mathbb{P} \mathbb{R}_\varepsilon$.

Accuracy of the E.M.

For every $t > 0$: $\Pr [q(M_{EM}(x); x) < q^*(x) - \frac{2\Delta q}{\varepsilon}(\log |Y| + t)] \leq e^{-t}$.

$\Leftrightarrow$ w.p. $1 - \beta$:

$$q(M_{EM}(x); x) \geq q^*(x) - \frac{2\Delta q}{\varepsilon} \log \frac{|Y|}{\beta}.$$

Proof sketch:

$$BAD_t := \{y : q(y; x) < q^*(x) - \frac{2\Delta q}{\varepsilon}(\log |Y| + t)\}.$$

$$\Pr [Y \in BAD_t] = \frac{\sum_{y \in BAD_t} \exp\left(\frac{\varepsilon}{2\Delta q} \cdot q(y; x)\right)}{\sum_{z \in Y} \exp\left(\frac{\varepsilon}{2\Delta q} \cdot q(z; x)\right)} \leq \frac{|BAD_t| \exp\left(\frac{\varepsilon}{2\Delta q} \left(q^*(x) - \frac{2\Delta q}{\varepsilon}(\log |Y| + t)\right)\right)}{\exp\left(\frac{\varepsilon}{2\Delta q} \cdot q^*(x)\right)}.$$

$$= |BAD_t| \cdot \exp(-(\log |Y| + t)) \leq |Y| \cdot \exp(-\log |Y|) \cdot \exp(-t)$$

utility of E.M. $\approx \frac{\Delta q}{\varepsilon} \cdot \log |Y|$.



Continuous median

$x_1, \dots, x_n \in [0, R]$ Assume n is even.

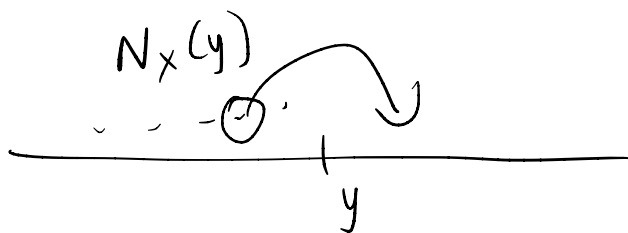
$$y \in [0, R]$$

$$N_x(y) = \# \{i : x_i \leq y\}$$

Goal: approximate median $\equiv y^*$ with $N_x(y^*) = \frac{n}{2}$.

$$q(y; x) = - \left| N_x(y) - \frac{n}{2} \right|$$

$$\Delta q = 1.$$



E.M. has density:
$$p_x(y) = \frac{\exp\left(-\frac{\varepsilon}{2} \left| N_x(y) - \frac{n}{2} \right| \right)}{Z_x}, \text{ for } y \in [0, R]$$

Sort x :

$$0 = x_{(0)} \leq x_{(1)} \leq \dots \leq x_{(n)} \leq x_{(n+1)} = \mathcal{R}.$$

for $j=0, \dots, n$: $I_j = [x_{(j)}, x_{(j+1)})$, $\ell_j = |x_{(j+1)} - x_{(j)}|$.

$$\forall y \in I_j : N_x(y) = j$$

Define $w_j = \ell_j \cdot \exp\left(-\frac{\varepsilon}{2} \left|j - \frac{n}{2}\right|\right)$.

Then E.M. is equivalent to:

① Pick $j \in \{0, \dots, n\}$ w.p. $\frac{w_j}{\sum w_j}$.

② Sample uniformly from I_j .

can run this!

• utility

$$\mathcal{G} := \left\{ y : \left| N_x(y) - \frac{n}{2} \right| \leq \frac{\alpha n}{2} \right\} \leftarrow N_x(y) \approx (1 \pm \alpha) \frac{n}{2},$$

$$0 < \alpha < \frac{1}{2}.$$

$$\mathcal{B} := \{y : |N_X(y) - \frac{n}{2}| > \alpha n\} \leftarrow (1+2\alpha)\frac{n}{2} \leq N_X(y)$$

$$\text{or } N_X(y) \leq (1-2\alpha)\frac{n}{2}$$

$$\Pr[Y \in \mathcal{B}] \leq \frac{R \cdot \exp(-\frac{\varepsilon}{2} \cdot \alpha n)}{\text{len}(G) \cdot \exp(-\frac{\varepsilon}{2} \frac{\alpha n}{2})} = \frac{R}{\text{len}(G)} \cdot \exp\left(-\frac{\varepsilon \alpha n}{4}\right) = \beta$$

$$\Leftrightarrow \text{w.p. } 1-\beta, \text{ if } n \geq \frac{1}{\varepsilon \alpha} \left(\log\left(\frac{R}{\text{len}(G)}\right) + \log\frac{1}{\beta} \right)$$

then E.M. returns y with rank $(1 \pm \alpha) \frac{n}{2}$.