Lecture 12

Approximate Differential Privacy.

- (ε, δ) - DP
- Gaussian Mechanism.

Announcements:

→ Guest lecture (Jon Ullman)
Will post Zoom / Youtube Link.
This Weds.

→ Recitation / Make-up Class on Friday.

Definition. (Differential Privacy)

A is ε-differentially private if

for all neighbors x and x'

for all subsets E of outputs

$$\mathbb{P}[A(x) \in E] \leq e^{\varepsilon}\, \mathbb{P}[A(x') \in E]$$

Is this too stringent?

Suppose there is some $E \subseteq Y$ such that

$$\mathbb{P}[A(x) \in E] \leq \frac{1}{2^{100}}$$

and $\mathbb{P}[A(x') \in E] = 0.$

(ε, δ) - DP

A is (ε, δ)-differentially private if
for all neighbors x and x'
for all subsets E of outputs

$$\mathbb{P}[A(x) \in E] \leq e^{\varepsilon} \mathbb{P}[A(x') \in E] + \delta$$

e^{ε}: Multiplicative Approximation

δ: Additive Approximation

Naming Convention:

ε-DP, $(\varepsilon, 0)$-DP, "pure" DP

(ε, δ)-DP, "approximate" DP

Interpretation of δ:
probability of "Privacy Failure"
$<< \frac{1}{n}$.
(think "Name & Shame")

Preserve Nice Properties.

① Post-Processing

(ε, δ)-DP $A: \mathcal{X}^n \to \mathcal{Y}$, $f: \mathcal{Y} \to \mathcal{Y}'$

$f(A(\cdot))$ is (ε, δ)-DP

② Adaptive Composition

$A_1 : \mathcal{X}^n \to \mathcal{Y}_1$ $(\varepsilon_1, \delta_1)$-DP

$A_2 : \mathcal{X}^n \times \mathcal{Y}_1 \to \mathcal{Y}_2$ $(\varepsilon_2, \delta_2)$-DP

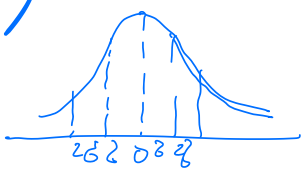
"Basic" Composition: $(A_1(x), A_2(x, A_1(x)))$ is $(\varepsilon_1+\varepsilon_2, \delta_1+\delta_2)$-DP.

Gaussian Mechanism

1-dim case $f: \mathcal{X}^n \mapsto \mathbb{R}$, w/ $GS_f = \Delta$

$$A(x) = f(x) + N\left(0, \frac{2\Delta^2 \log(2/\delta)}{\varepsilon^2}\right)$$

σ^2 ↑ Variance



Gaussian dist $= N(\mu, \sigma^2)$

$$p_{\mu,\sigma^2}(y) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

Gaussian Mechanism

1-dim case $f: \mathcal{X}^n \to \mathbb{R}$, w/ $GS_f = \Delta$ ← Sensitivity

$$A(x) = f(x) + N\left(0, \frac{2\Delta^2 \log(2/\delta)}{\varepsilon^2}\right)$$

Lap Mech.
$+ Lap\left(\frac{\Delta}{\varepsilon}\right)$

↑ Variane or the error

δ : delta

Theorem. For any $\varepsilon \leq 1$ $\delta > 0$, 1-d Gaussian Mech. satisfies (ε, δ)-DP.

Refs: CDP : Concentrated DP.
RDP : Renyi DP.

How to prove (ε, δ)-DP

Recap: How to prove $(\varepsilon, 0)$-DP

For $(\varepsilon, 0)$-DP, suffices to show.

$$\forall y \in \mathcal{Y}, \quad \mathbb{P}[A(x)=y] \le e^{\varepsilon} \mathbb{P}[A(x')=y]$$

(y: output)

$$\Leftrightarrow \forall E \subseteq \mathcal{Y} \quad \mathbb{P}[A(x) \in E] \le e^{\varepsilon} \mathbb{P}[A(x') \in E]$$

($\Leftrightarrow$: Equivalent)

Strategy: for all y in "Good Set"

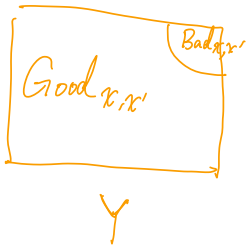
Ratio $\dfrac{\mathbb{P}[A(x)=y]}{\mathbb{P}[A(x')=y]} \le e^{\varepsilon}$

$$\mathbb{P}[\text{"Good set"}] \ge 1-\delta.$$

Lemma. Let $A: \mathcal{X}^n \to Y$. For any $x, x' \in \mathcal{X}^n$, and any $\varepsilon > 0$.

Define $\text{Good}_{x,x'} = \left\{ y \in Y \;\middle|\; \frac{\mathbb{P}[A(x) = y]}{\mathbb{P}[A(x') = y]} \le e^{\varepsilon} \right\}$

$\text{Bad}_{x,x'}$ = Complement of $\text{Good}_{x,x'}$

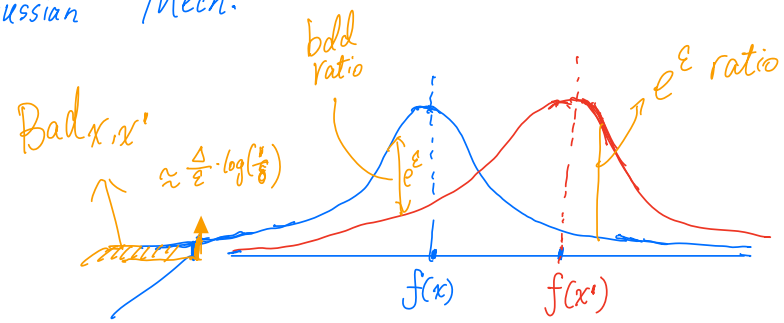


If $\text{Bad}_{x,x'}$ satisfies,

$\mathbb{P}[A(x) \in \text{Bad}_{x,x'}] \le \delta$ for all neighbors x & x',

then A is (ε, δ)-DP.

Gaussian Mech.

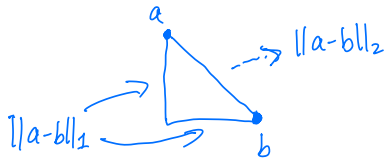


Multivariate Case

$$f : \mathcal{X}^n \longmapsto \mathbb{R}^k$$

$$\Delta_2(f) = \max_{x, x' \text{ neighbors}} \| f(x) - f(x') \|_2$$

$$\|v\|_2 = \sqrt{\sum_{j=1}^{k} v_j^2}$$



$f = (f_1, \ldots, f_k)$ such that

f_j counts

$$f_j(x) = \sum_{i=1}^{n} \varphi_j(x_i), \quad \varphi_i(x) \text{ in } \{0,1\}.$$

↑ Specify property "does x_i smoke or not?"

What is:

① ℓ_1 - sensitivity $\Delta_1 = k$

② ℓ_2 - sensitivity $\Delta_2 = \sqrt{k}$.

Fact: $\Delta_2 \leq \Delta_1 \leq \sqrt{k}\, \Delta_2$

Spherical Multivariate Gaussian

$$p_{\mu,\sigma^2}(y) = \frac{1}{(2\pi\sigma^2)^{k/2}} \exp\left(\frac{-\|y-\mu\|_2^2}{2\sigma^2}\right)$$

Independent 1-dim Gaussian $N(\mu, \sigma^2)$ in each coordinate.

$$N(\mu, \sigma^2 I).$$

↑ identity matrix $\begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$

Multivariate Gaussian Mechanism

$f: \mathcal{X}^n \to \mathbb{R}^k$ w/ ℓ_2-sensitivity Δ_2

$$A(x) = f(x) + N\left(0, \frac{2\Delta_2^2 \log(2/\delta)}{\varepsilon^2} I_{k\times k}\right)$$

same as: Adding per-coordinate independent 1-dim $N\left(0, \frac{2\Delta_2^2 \log(2/\delta)}{\varepsilon^2}\right)$

Theorem: $\forall\ \varepsilon \le 1,\ \delta > 0$

$A(\cdot)$ satisfies (ε, δ)-DP.

Improvement over Laplace Mechanism.

$f = (f_1, \ldots, f_k)$ such that

($f_1, \ldots, f_k$: counts)

$$f_j(x) = \sum_{i=1}^{n} \varphi_j(x_i), \quad \varphi_i(x) \text{ in } \{0,1\}.$$

(φ_j: specify property, "does x_i smoke or not?")

$\Delta_1 = k, \quad \Delta_2 = \sqrt{k}.$

Laplace / Gaussian : $y = f(x) + \text{Noise}.$

Comparison:

Laplace : Privacy $(\varepsilon, 0)$-DP.

Accuracy $\mathbb{E}\left[\max_{j=1}^{k} |f_j(x) - y_j|\right] \leq O\left(\frac{k \log k}{\varepsilon}\right)$

Gaussian : Privacy (ε, δ)-DP.

Accuracy $\mathbb{E}\left[\max_{j=1}^{k} |f_j(x) - y_j|\right] \leq O\left(\frac{\sqrt{k}\log(k)\log(1/\delta)}{\varepsilon}\right)$

Gaussian →	"Share"	$\sqrt{k}$	Improvement for large k.
	"Pay"	$\sqrt{\log(\frac{1}{\delta})}$	

$\Delta_1 \le k$, $\Delta_2 \le \sqrt{k}$

Laplace Mechanism for $(\varepsilon, 0)$-DP: $\mathbb{E}\left[\max_{j=1}^{k} |f_j(x) - a_j|\right] \le O\left(\frac{k \log k}{\varepsilon}\right)$

Gaussian Mechanism for (ε, δ)-DP: $\mathbb{E}\left[\max_{j=1}^{k} |f_j(x) - a_j|\right] \le O\left(\frac{\sqrt{k \log k \log(1/\delta)}}{\varepsilon}\right)$

Theorem. For any $\varepsilon \le 1$, $\delta > 0$, the (1-dim) Gaussian Mechanism satisfies (ε, δ)-DP.

Better analysis: CDP, Renyi DP.

Proof. Fix any neighbors x, x'. Wlog, $f(x) = 0$, $|f(x')| \le \Delta$

(Hope: Identify Good & Bad y's).

$$\ln\left(\frac{\mathbb{P}[A(x)=y]}{\mathbb{P}[A(x')=y]}\right) \overset{\text{Plug Gaussian}}{=} \ln\left(\frac{\exp\left(-\frac{y^2}{2b^2}\right)}{\exp\left(-\frac{(y-f(x'))^2}{2b^2}\right)}\right)$$

$$= \frac{(y-f(x'))^2}{2b^2} - \frac{y^2}{2b^2}$$

$$= \frac{f(x')^2 - 2f(x')y}{2b^2} \qquad \text{Good} \Rightarrow \le \varepsilon$$

Recall $b^2 = \frac{2\Delta^2 \log(2/\delta)}{\varepsilon^2}$

$$= \frac{\left(f(x')^2 + 2f(x')y\right)}{4\Delta^2 \log(2/\delta)\,/\varepsilon^2}$$

$|f(x')| \le \Delta$, $\varepsilon \le 1$, $\delta > 0$, $\varepsilon^2 \le \varepsilon$

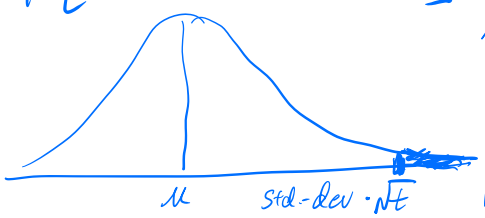
$$\le \frac{\varepsilon}{4} - \frac{y\varepsilon^2}{2\Delta \log(2/\delta)} \qquad \text{want} \Rightarrow \le \varepsilon$$

$$\text{Bad}_{x,x'} \subseteq \left\{ y \;\middle|\; |y| > \frac{\sqrt{2}\,\Delta \log(2/\delta)}{\varepsilon} \right\}$$

↳ Std dev.

$$\mathbb{P}[A(x) \in \text{Bad}_{x,x'}] \le \delta$$

↑ follows from our choice of $b^2 = \frac{2\Delta^2 \log(2/\delta)}{\varepsilon^2}$



$\mathbb{P} \le e^{-t}$

"Name & Shame" Algorithm

$NS_\delta(x_1, x_2, \ldots, x_n)$

For each $i = 1, \ldots, n$

Release $Y_i = \begin{cases} x_i & \text{w.p. } \delta \\ \text{"}\perp\text{"} & \text{w.p. } (1-\delta) \end{cases}$

NS_δ satisfies $(\underset{\downarrow}{\varepsilon}, \delta)$-DP.

If $\delta > \frac{1}{n}$ (e.g. $\frac{20}{n}$), release ≈ 20 in the clear.

$\rightarrow$ $\delta \ll \frac{1}{n}$

$\frac{1}{2^{20}}$, $\frac{1}{2^{100}}$, (Paper writing: $\delta = \frac{1}{n^2}$)