BU CS599 S1
Foundations of Private Data Analysis
Spring 2023

Lecture 09: Approximate DP

Adam Smith

BU

Today

- Gaussian mechanism: motivation
- (ϵ, δ)-differential privacy
- Gaussian analysis
- Truncated Laplace mechanism
- Stable histograms

Gaussian noise

- Suppose we have $f: \mathcal{U}^n \to \mathbb{R}^d$
- Consider the Gaussian mechanism
  $A(x) = f(x) + (Z_1, \ldots, Z_d)$ where each $Z_i \sim N(0, \sigma^2)$

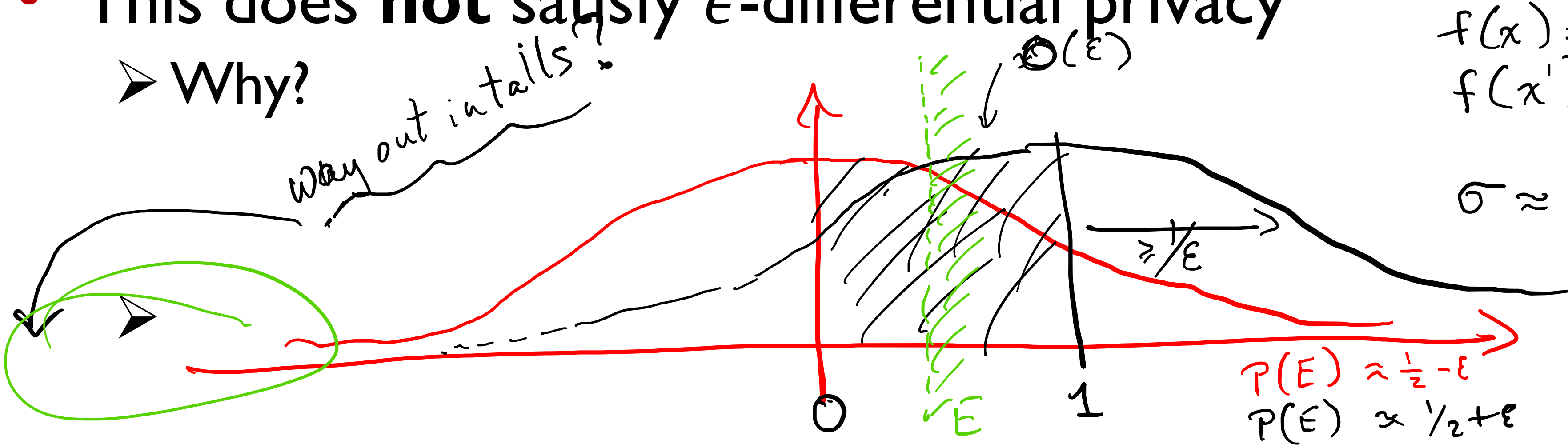
$$x \longrightarrow \boxed{A} \xrightarrow{f(x) + \vec{Z}}$$

- Density of $N(0, \sigma^2)$ is $h(z) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{z^2}{2\sigma^2}\right)$

- This does **not** satisfy ϵ-differential privacy
  - ➢ Why?

1-dimension
$f(x) = 0$
$f(x') = 1$

$\sigma \approx 1/\varepsilon$



(ϵ, δ)-*Differential Privacy*

- A randomized algorithm $A: \mathcal{U}^* \to \mathcal{Y}$ is (ϵ, δ)-differentially private if
  for all neighboring data sets x, x':
  for all events E:

$$\Pr(A(x) \in E) \leq e^{\varepsilon} \Pr(A(x') \in E) + \delta.$$

[Meaningful when $\delta \ll 1/n$ where n is data set size.]

- Two probability distributions P, Q on the set same $\mathcal{Y}$ are (ϵ, δ)-close if for all events $E \subseteq \mathcal{Y}$, (Σ-algebra)

$$P(E) \leq e^{\varepsilon} Q(E) + \delta \quad \underline{\text{and}} \quad Q(E) \leq e^{\varepsilon} P(E) + \delta.$$

  We write $P \approx_{\epsilon,\delta} Q$

- (Same term and notation for random variables whose distributions are close) e.g. $\forall x, x'$ neighbors, $A(x) \approx_{\varepsilon,\delta} A(x')$

Privacy Loss Random Variable $\quad L_{x,x'}(A(x))$

- Given mechanism A and neighboring data sets x, x',

$$L_{x,x'}(\cancel{z}\,y) = \ln\left(\frac{\Pr(A(x) = \cancel{z}\,y)}{\Pr(A(x') = \cancel{z}\,y)}\right) \leftarrow \text{can be densities.}$$

- For 1-dimensional Gaussians: and function f.

$$L_{x,x'}(y) = \ln \frac{\cancel{\frac{1}{\sigma\sqrt{2\pi}}}}{\cancel{\frac{1}{\sigma\sqrt{2\pi}}}} \exp\left(\frac{-1}{2\sigma^2}\left((y-f(x))^2 - (\quad -f(x'))^2\right)\right)$$

$\Delta = f(x') - f(x)$

$$= \frac{-1}{2\sigma^2}\left(2\underbrace{(y-f(x))}_{\sim N(0,\sigma^2) \text{ where } y = A(x)}\Delta + \Delta^2\right)$$

$$L_{x,x'}(A(x)) \sim N\left(\frac{-\Delta^2}{2\sigma^2}, \left(\frac{\Delta\sigma}{\sigma^2}\right)^2\right)$$

$$= N\left(\frac{-\Delta^2}{2\sigma^2}, \left(\frac{\Delta}{\sigma}\right)^2\right)$$

From Privacy Loss to DP

$$L_{xx'}(y) = \ln \frac{\Pr(A(x)=y)}{\Pr(A(x')=y)}$$

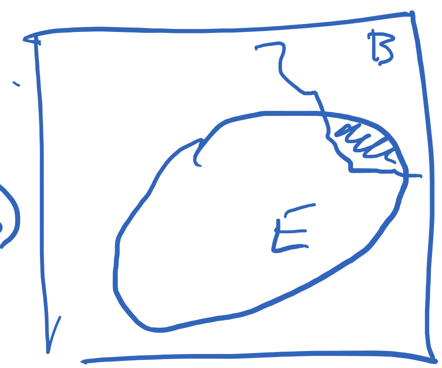
- Lemma: Suppose that for all neighboring x, x',

$$\Pr_{y\sim A(x)}\big(\underbrace{L_{x,x'}(y) \le \epsilon}\big) \ge 1-\delta.$$

Then A is (ϵ, δ)-differentially private.

Proof: Fix x, x' neighboring, event E.

Define $B = \{y : L_{x,x'}(y) > \varepsilon\}$.



$$\Pr(A(x)\in E) \le \Pr(A(x)\in E\setminus B) + \Pr(A(x)\in B)$$

$$= \sum_{y\in E\setminus B} \Pr(A(x)=y) + \delta$$

$$\le \sum_{y\in E\setminus B} \left(\Pr(A(x')=y)\cdot e^{\varepsilon}\right) + \delta$$

(by definition of B)

$$= e^{\varepsilon} \Pr(A(x')\in E\setminus B) + \delta$$

$$\le e^{\varepsilon} \Pr(A(x')\in E) + \delta.$$

1-d Gaussian mechanism

- **Proposition:** If $f: \mathcal{U}^* \to \mathbb{R}$ and $GS_f \leq \Delta$, then the 1-d Gaussian mechanism with $\sigma = \frac{\Delta\sqrt{2\ln(2/\delta)}}{\epsilon}$ is (ϵ, δ)-differentially private

Setting $\delta = e^{-100}$ and ~~increase~~ get noise $\approx \frac{\Delta}{\varepsilon} \cdot 15$.

Proof: $N\left(\frac{\Delta^2}{2\sigma^2}, \left(\frac{\Delta}{\sigma}\right)^2\right)$

$\leq \varepsilon$ w.p. $\geq 1-\delta$.

$$\Pr(N(0,1) > \underbrace{t}_{\approx \varepsilon}) \sim \underbrace{e^{-t^2/2}}_{\approx \delta}$$

Multivariate Gaussian: ℓ_2 sensitivity

- Given $f: \mathcal{U}^n \to \mathbb{R}^d$, let

$$\Delta_2 = \sup_{x,x' \text{ neighboring}} \|f(x) - f(x')\|_2 = \sqrt{\sum_j (f(x)_j - f(x')_j)^2}$$

- For example, say we have a list of k counting queries
  - How many people are men, how many are Asian, how many are diabetic, …
  - Laplace mechanism tells us to add noise k/ϵ per entry
  - But $\Delta_2 = \sqrt{k}$

With Gaussian noise, can get error $\propto \frac{\sqrt{k \ln(1/\delta)}}{\varepsilon}$ per entry.

Multivariate Gaussian Analysis

- **Proposition:** If $f : \mathcal{U}^* \to \mathbb{R}^d$ ~~and $GS_f \le$~~ with ℓ_2 sensitivity Δ_2, then the ~~1-d~~ Gaussian mechanism with $\sigma = \frac{\Delta_2 \sqrt{2\ln(2/\delta)}}{\epsilon}$ is (ϵ, δ)-differentially private

Density of noise: $h(z_1, \ldots, z_d) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left(\frac{-z_1^2}{2\sigma^2}\right) \times \cdots \times \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-z_d^2}{2\sigma^2}\right)$

$$= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^d \exp\left(\frac{-1}{2\sigma^2} \|\vec{z}\|_2^2\right)$$

$$L_{x,x'}(y) = \ln \frac{h(\vec{y} - \vec{f}(x))}{h(\vec{y} - f(x'))} = \ln\left(\frac{\cancel{\left(\frac{1}{\sigma\sqrt{2\pi}}\right)^d}}{\cancel{\left(\frac{1}{\sigma\sqrt{2\pi}}\right)^d}} \exp\left(\frac{-1}{2\sigma^2}\left(\|y - f(x)\|_2^2 - \|y - f(x')\|_2^2\right)\right)\right)$$

$$= \left(\frac{-1}{2\sigma^2}\left(\|\vec{z} + \vec{u}\|_2^2 - \|z\|^2\right)\right)$$

$z = y - f(x)$

$u = f(x) - f(x')$

$$= \frac{-1}{2\sigma^2}\left(\langle z+u, z+u \rangle - \langle z, z \rangle\right)$$

$$= \frac{-1}{2\sigma^2}\left(\langle z, z \rangle + 2\langle z, u \rangle + \langle u, u \rangle - \langle z, z \rangle\right) = \frac{-1}{2\sigma^2}\left(2\langle z, u \rangle + \|u\|^2\right)$$

Distribution of $\langle z, u \rangle \sim N(0, \sigma^2 \|u\|^2)$

$$\sim N\left(\frac{\Delta_2^2}{2\sigma^2}, \left(\frac{\Delta}{\sigma}\right)^2\right)$$

Today

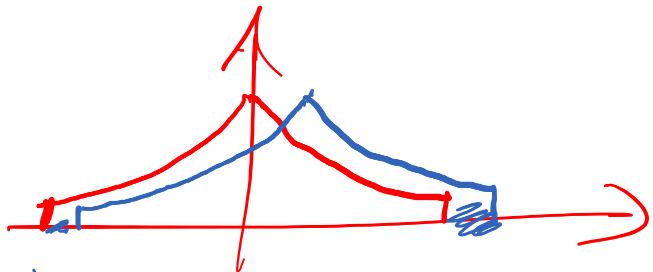
With $(\varepsilon, 0)$-D.P. need $\Omega(k/\varepsilon)$ noise per entry ☹

- Gaussian mechanism: motivation
- (ϵ, δ)-differential privacy
- Gaussian analysis
- Truncated Laplace mechanism
- Stable histograms

Truncated Laplace

- Suppose $Z \sim \mathrm{Lap}(\lambda)$ conditioned on $|Z| \le \lambda \ln(1/\delta)$.

$$\underbrace{\qquad\qquad\qquad\qquad}_{\mathrm{Lap}(\lambda,\ \lambda \ln(1/\delta))}$$



Proposition:

- Adding noise $\mathrm{Lap}(\lambda, \lambda \ln(1/\delta))$ with $\lambda = \Delta/\varepsilon$ (in 1-D case) satisfies (ε, δ)-DP.