**CS839: Differential Privacy and Learning** January 27, 2026

# Lecture 3: Privacy Basics II

*Instructor:* Gavin Brown *Scribe:* Hongyi Liu

This lecture continued the "core toolbox" portion of the course: how we *prove* algorithms are differentially private, and how we *design* private mechanisms by calibrating noise to the *sensitivity* of the statistic we want to release. The lecture introduced two lemmas to prove pure and approximate differential privacy, and then defined global sensitivity, reviewed the Laplace mechanism, and introduced the Gaussian mechanism (central for high-dimensional releases and DP-SGD).

## 1 Techniques to show an algorithm is DP

For Pure DP, there are 3 equivalent definitions:

**Claim 1.1.** *Algorithm $\mathcal{A}$ is $\epsilon-$ DP if and only if:*

*1)* $\forall x \sim x', E \subseteq \mathcal{Y}$, $\Pr[\mathcal{A}(x) \in E] \leq e^{\epsilon} \Pr[\mathcal{A}(x') \in E]$

*2)* $\forall x \sim x', y \in \mathcal{Y}$, $\Pr[\mathcal{A}(x) = y] \leq e^{\epsilon} \Pr[\mathcal{A}(x') = y]$

*3)* $\forall x \sim x', y \in \mathcal{Y}$, $L_{\mathcal{A}}^{x \to x'}(y) \leq \epsilon$

*Proof.* Prove left as an exercise. □

However, this does not hold for approximate−DP.

For approximate−DP, here is a main technique to prove that an algorithm is $(\epsilon, \delta) -$ DP using privacy loss as a random variable.

**Claim 1.2.** *Suppose* $\forall x \sim x'$, $\Pr_{y \sim \mathcal{A}(x)}[L_{\mathcal{A}}^{x \to x'}(y) > \epsilon] \leq \delta$, *then* $\mathcal{A}$ *is* $(\epsilon, \delta) -$ DP.

*Proof.* Fix two adjacent datasets $x \sim x'$ and event $E \subseteq \mathcal{Y}$.

Let $\text{BAD} = \{y \in \mathcal{Y} : L_{\mathcal{A}}^{x \to x'}(y) > \epsilon\}$, and by assumption, we have $\Pr[\mathcal{A}(x) \in \text{BAD}] \leq \delta$.

Then, we split the total possibility to bad and good situations:

$$\begin{aligned}\Pr[\mathcal{A}(x) \in E] &= \Pr[\mathcal{A}(x) \in E \wedge \mathcal{A}(x) \notin \text{BAD}] + \Pr[\mathcal{A}(x) \in E \wedge \mathcal{A}(x) \in \text{BAD}] \\ &\leq \Pr[\mathcal{A}(x) \in E \wedge \mathcal{A}(x) \notin \text{BAD}] + \delta\end{aligned}$$

We then justify that the probability of good situations is always bounded since the privacy loss is

bounded by $\epsilon$:

$$\begin{aligned}\Pr[\mathcal{A}(x) \in E \wedge \mathcal{A}(x) \notin \text{BAD}] &= \sum_{y \in E, y \notin \text{BAD}} \Pr[\mathcal{A}(x) = y] \\ &\leq \sum_{y \in E, y \notin \text{BAD}} e^{\epsilon} \Pr[\mathcal{A}(x') = y] \\ &\leq \sum_{y \in E} e^{\epsilon} \Pr[\mathcal{A}(x') = y] \\ &= e^{\epsilon} \Pr[\mathcal{A}(x') \in E]\end{aligned}$$

In conclusion, we have $\Pr[\mathcal{A}(x) \in E] \leq e^{\epsilon} \Pr[\mathcal{A}(x') \in E] + \delta$, i.e. $\mathcal{A}$ is $(\epsilon, \delta) - \mathsf{DP}$. □

## 2 Global sensitivity

Before DP (late 90s/early 2000s), people knew adding noise can help privacy, but lacked a principled answer to: **how much noise is enough?** Differential privacy's starting insight is to calibrate noise to sensitivity.

**Definition 2.1.** The gobal sensitivity of $f : X^n \to \mathbb{R}$, denoted by $\Delta_f$, is

$$\Delta_f \triangleq \max_{x \sim x'} \|f(x) - f(x')\|$$

**Note**: This generalizes to functions of higher dimensions $f : X^n \to \mathbb{R}^d$ using $L_1$ or $L_2$ norms.

## 3 Laplace mechanism

If you can bound global sensitivity, you can get a DP algorithm by adding Laplace noise scaled to sensitivity as follows:

**Claim 3.1.** *For any function $f$, the algorithm that computes $f$ and adds Laplace noise with scale $\Delta_f$, i.e. $f(x) + \text{Lap}(\Delta_f/\epsilon)$, preserves $\epsilon - \mathsf{DP}$.*

*Proof.* Proof is similar to the proof of Claim 3.2. Laplace mechanism is $\epsilon - \mathsf{DP}$ in the last lecture. □

**Notes**:

- If you have some function to compute, and you can bound its global sensitivity. This leads to some DP algorithm.
- A lot of DP researches is about doing better than just adding noise calibrated to the global sensitivity.

## 4 Gaussian mechanism

The density of Gaussian distribution $\mathcal{N}(\mu, \sigma^2)$ is

$$p(z) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \exp\left(\frac{-(z-\mu)^2}{2\sigma^2}\right)$$

**Lemma 4.1** (Gaussian tail bound)**.** *For $z \sim \mathcal{N}(\mu, \sigma^2)$,* $\Pr[|z - \mu| > t\sigma] \leq \frac{2}{t\sqrt{2\pi}} \exp\left(\frac{-t^2}{2}\right)$.

**Claim 4.2.** *Fix any function $f : X^n \to \mathbb{R}$, set $\sigma^2 = \frac{2\Delta_f^2 \log(2/\delta)}{\epsilon^2}$ assuming $0 \leq \epsilon \leq 1, 0 \leq \delta \leq \frac{1}{2}$, and then the algorithm $\mathcal{A}(x) = f(x) + \mathcal{N}(0, \sigma^2)$ is $(\epsilon, \delta) - \mathsf{DP}$.*

*Proof.* Fix two adjacent datasets $x \sim x'$. By assumption, $|f(x) - f(x')| \leq \Delta_f$, and it is sufficient to assume that $f(x) = 0$, thus $|f(x')| \leq \Delta_f$.

For $y \in \mathcal{Y}$, the privacy loss is

$$\begin{aligned}
L_{\mathcal{A}}^{x \to x'}(y) &= \log\left(\frac{\Pr[\mathcal{A}(x) = y]}{\Pr[\mathcal{A}(x') = y]}\right) \\
&= \log\left(\frac{\Pr[\mathcal{N}(0, \sigma^2) = y - f(x)]}{\Pr[\mathcal{N}(0, \sigma^2) = y - f(x')]}\right) \\
&= \log\left(\frac{\exp\left(\frac{-y^2}{2\sigma^2}\right)}{\exp\left(\frac{-(y - f(x'))^2}{2\sigma^2}\right)}\right) \\
&= \frac{(y - f(x'))^2}{2\sigma^2} - \frac{y^2}{2\sigma^2} \\
&= \frac{f(x')^2 - 2yf(x')}{2\sigma^2}
\end{aligned}$$

Then we upper-bound the privacy loss function,

$$\begin{aligned}
L_{\mathcal{A}}^{x \to x'}(y) &= \frac{f(x')^2 - 2yf(x')}{2\sigma^2} \\
&\leq \frac{f(x')^2}{2\sigma^2} + \frac{|y||f(x')|}{\sigma^2} \\
&\leq \frac{\Delta_f^2}{2\sigma^2} + \frac{|y|\Delta_f}{\sigma^2}
\end{aligned}$$

We want to show that $\Pr[L_{\mathcal{A}}^{x \to x'}(y) > \epsilon] \leq \delta$. Since $\frac{\Delta_f^2}{2\sigma^2} + \frac{|y|\Delta_f}{\sigma^2} \leq \epsilon \iff |y| \leq \frac{\epsilon\sigma^2}{\Delta_f} - \frac{\Delta_f}{2}$, we have,

$$\begin{aligned}
\Pr[L_{\mathcal{A}}^{x \to x'}(y) > \epsilon] &= 1 - \Pr[L_{\mathcal{A}}^{x \to x'}(y) \leq \epsilon] \\
&\leq 1 - \Pr\left[|y| \leq \frac{\epsilon\sigma^2}{\Delta_f} - \frac{\Delta_f}{2}\right] \\
&= 1 - \left(1 - \Pr\left[|y| > \frac{\epsilon\sigma^2}{\Delta_f} - \frac{\Delta_f}{2}\right]\right) \\
&= \Pr\left[|y| > \frac{\epsilon\sigma^2}{\Delta_f} - \frac{\Delta_f}{2}\right]
\end{aligned}$$

Lets write $t = \frac{\epsilon\sigma}{\Delta_f} - \frac{\Delta_f}{2\sigma} = \sqrt{2\log(2/\delta)} - \frac{\epsilon}{2\sqrt{2\log(2/\delta)}}$. By Gaussian tail bound, we know that,

$$\begin{aligned}
\Pr[L_{\mathcal{A}}^{x\to x'}(y) > \epsilon] &\le \Pr\left[|y| > \sigma \cdot t\right] \\
&\le \frac{2}{t\sqrt{2\pi}} \exp\left(-\frac{1}{2}t^2\right) \\
&= \frac{2}{t\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(2\log(2/\delta) - \epsilon + \frac{\epsilon^2}{8\log(2/\delta)}\right)\right) \\
&= \frac{\delta}{t\sqrt{2\pi}} \exp\left(\frac{\epsilon}{2} - \frac{\epsilon^2}{16\log(2/\delta)}\right)
\end{aligned}$$

To show that $\Pr[L_{\mathcal{A}}^{x\to x'}(y) > \epsilon] \le \delta$, it suffices to show that $\exp\left(\frac{\epsilon}{2} - \frac{\epsilon^2}{16\log(2/\delta)}\right) \le t\sqrt{2\pi}$.

Since $\epsilon \in [0,1], \delta \in [0,1/2]$, we know that $\log(2/\delta) \ge \log 4$, so $\frac{\epsilon^2}{16\log(2/\delta)} \ge 0$, and thus $\exp\left(\frac{\epsilon}{2} - \frac{\epsilon^2}{16\log(2/\delta)}\right) \le e^{\epsilon/2} \le e^{1/2}$. Besides, $t = \sqrt{2\log(2/\delta)} - \frac{\epsilon}{2\sqrt{2\log(2/\delta)}} \ge \sqrt{2\log 4} - \frac{1}{2\sqrt{2\log 4}}$.

$$\exp\left(\frac{\epsilon}{2} - \frac{\epsilon^2}{16\log(2/\delta)}\right) \le e^{1/2} \approx 1.6487 < 3.4211 \approx \sqrt{2\pi}\left(\sqrt{2\log 4} - \frac{1}{2\sqrt{2\log 4}}\right) \le \sqrt{2\pi} \cdot t$$

With $\Pr[L_{\mathcal{A}}^{x\to x'}(y) > \epsilon] \le \delta$, by Claim 1.2, algorithm $\mathcal{A}$ is $(\epsilon, \delta) - \mathsf{DP}$. □

**Geometric intuition**: We add two Gaussians with different means and overlap. Since Gaussian density function drops dramatically, in the far tails, observing an extreme $y$ gives lots of evidence about which dataset generated it—privacy loss can be large in magnitude. The goal is to show the probability of landing in those tail regions is at most $\delta$.

In Contrast, for Laplace, privacy loss is uniformly bounded (pure DP), so no "catastrophic tail" failure.

**Note**: In higher dimensions, the proof is similar but uses norms instead of squares.

Switching to approximate DP is a weaker privacy guarantee (there's a small probability $\delta$ of large privacy loss). The main benefits of Gaussian mechanism are,

- **High dimensions**: Gaussian works naturally with $L_2$ sensitivity while Laplace typically uses $L_1$ sensitivity. In high dimensions, Gaussian can require much less noise.
- **Composition**: Allowing a small $\delta$ enables stronger composition theorems (get smaller effective $\epsilon$ across many operations. This is the topic of the first question on Homework 1.

When we get into the theory of differentially private statistics, we will formalize a third major advantage of approximate DP algorithms: they are better able to cope with prior uncertainty about the underlying parameters.