CS839: Differential Privacy and Learning January 20, 2026

Lecture 1: Introduction to Differential Privacy

Instructor: Gavin Brown *Scribe:* Sarthak Choudhary

Overview

This lecture introduces the fundamental definitions of differential privacy (DP). We begin by defining what it means for two probability distributions to be indistinguishable, then define the notion of adjacent datasets, and finally arrive at the formal definition of (ε, δ)-differential privacy. We also discuss the intuition behind the privacy parameters ε and δ, and distinguish between "pure" and "approximate" differential privacy.

1 Differential Privacy: First Definition

We begin with the core definition of differential privacy. Consider an algorithm $\mathcal{A}$ that takes a dataset as input and produces some output. Intuitively, we want the algorithm's output to be insensitive to the presence or absence of any single individual's data.

Definition 1.1 (Differential Privacy — Probabilistic Form)**.** An algorithm $\mathcal{A} : \mathcal{X}^n \to \mathcal{Y}$ is (ε, δ)-*differentially private* (or (ε, δ)-DP) for $\varepsilon \geq 0$ and $\delta \geq 0$ if, for all $x, x' \in \mathcal{X}^n$ such that $x \sim x'$ (i.e., x and x' are adjacent), and for all measurable events $O \subseteq \mathcal{Y}$,

$$\Pr[\mathcal{A}(x) \in O] \leq e^{\varepsilon} \Pr[\mathcal{A}(x') \in O] + \delta.$$

Here, $\mathcal{X}^n$ denotes the space of datasets (each dataset consists of n records from some data universe $\mathcal{X}$), and $\mathcal{Y}$ denotes the output space. The relation $x \sim x'$ indicates that datasets x and x' are *adjacent*, which we define precisely below.

2 Indistinguishability of Distributions

Before giving an equivalent formulation of differential privacy, we introduce the notion of indistinguishability between probability distributions.

Definition 2.1 ((ε, δ)-Indistinguishability)**.** Two probability distributions p and q over a common measurable space are (ε, δ)-*indistinguishable*, written $p \approx_{(\varepsilon,\delta)} q$, if for all measurable events E:

$$p(E) \leq e^{\varepsilon}\, q(E) + \delta,$$
$$q(E) \leq e^{\varepsilon}\, p(E) + \delta.$$

Note that this definition is symmetric: $p \approx_{(\varepsilon,\delta)} q$ if and only if $q \approx_{(\varepsilon,\delta)} p$.

3 Adjacent Datasets

The notion of adjacency captures the idea that two datasets differ in one individual's data.

Definition 3.1 (Adjacent Datasets)**.** Two datasets $x = (x_1, \ldots, x_n) \in \mathcal{X}^n$ and $x' = (x'_1, \ldots, x'_n) \in \mathcal{X}^n$ with $x \neq x'$ are *adjacent*, written $x \sim x'$, if there exists an index $i^* \in \{1, \ldots, n\}$ such that for all $i \neq i^*$,

$$x_i = x'_i.$$

In other words, x and x' are adjacent if they differ in one coordinate. This models the scenario where one individual's data is changed (or replaced) while all other individuals' data remain the same.

4 Equivalent Definition via Indistinguishability

Using the notation of indistinguishability, we can give an equivalent and often more convenient definition of differential privacy.

Definition 4.1 (Differential Privacy — Indistinguishability Form)**.** An algorithm $\mathcal{A}$ is (ε, δ)-*differentially private* if, for all adjacent datasets $x \sim x'$,

$$\mathcal{A}(x) \approx_{(\varepsilon,\delta)} \mathcal{A}(x').$$

This definition states that the output distributions of $\mathcal{A}$ on any two adjacent datasets must be (ε, δ)-indistinguishable. We can get different "flavors" of DP by changing the notions of adjacency or indistinguishability. The central example of this is the difference between pure ($\delta = 0$) DP and approximate ($\delta > 0$) DP. Another example is *user-level* or *person-level* DP, where an individual may be associated with multiple records and we wish to protect against changing all the individual's data.

5 Interpreting the Privacy Parameters

The parameters ε and δ quantify the privacy guarantee.

The parameter ε. Think of ε as being on the order of 1 or 0.1. For small ε, we have the approximation

$$e^{\varepsilon} \approx 1 + \varepsilon,$$

so smaller values of ε mean that the output distributions on adjacent datasets are closer to identical. The multiplicative factor e^{ε} bounds the "privacy loss" in a worst-case sense.

Composition of Exponentials. For any $\varepsilon_1, \varepsilon_2 \geq 0$,

$$e^{\varepsilon_1} \cdot e^{\varepsilon_2} = e^{\varepsilon_1 + \varepsilon_2}.$$

This basic fact underlies the composition properties of differential privacy: running two ε-DP algorithms sequentially yields a 2ε-DP algorithm (in the pure DP case).

The parameter δ. The parameter δ should be thought of as very small—typically $\delta \approx 1/n^2$ or even cryptographically negligible, such as $\delta = 2^{-\Omega(n)}$. The δ term allows for a small probability of "catastrophic" privacy failure, so we want it to be negligible.

6 Pure vs. Approximate Differential Privacy

- **Approximate DP:** When $\delta > 0$, we call the guarantee (ε, δ)-*differential privacy* or *approximate differential privacy*.
- **Pure DP:** When $\delta = 0$, we obtain ε-*differential privacy* or *pure differential privacy*. In this case, the definition simplifies to: for all adjacent $x \sim x'$ and all events O,

$$\Pr[\mathcal{A}(x) \in O] \leq e^{\varepsilon} \Pr[\mathcal{A}(x') \in O].$$

Pure DP provides a stronger guarantee since there is no additive slack term.

7 Exercises

Exercise 7.1 (TV Distance Privacy)**.** Show that "TV distance privacy" corresponds to the case $\varepsilon = 0$. That is, characterize what it means for an algorithm to be $(0, \delta)$-differentially private in terms of the total variation distance between output distributions.

Exercise 7.2 (Transitivity of Indistinguishability)**.** Suppose $\mathcal{A}(x) \approx_{(\varepsilon,\delta)} \mathcal{A}(x')$ and $\mathcal{A}(x') \approx_{(\varepsilon,\delta)} \mathcal{A}(x'')$. Determine values ε' and δ' such that

$$\mathcal{A}(x) \approx_{(\varepsilon',\delta')} \mathcal{A}(x'').$$