1 Notation and Terminology

The notation presented here will follow closely with [DR14].

- database \( x \)
- data of \( n \) individuals
- data universe \( \mathcal{X} \)

We can think of the data in the following two ways:

1. As a matrix, where each row \( x_i \) of the \( n \) rows is the data of an individual, and \( x \in \mathcal{X}^n \).
   
   For example,
   
   \[
   x = \begin{bmatrix}
   x_1 \\
   x_2 \\
   \vdots \\
   x_n
   \end{bmatrix} = \begin{bmatrix}
   Alice & F & + \\
   Bob & M & - \\
   \vdots & \vdots & \vdots \\
   Zeus & M & +
   \end{bmatrix}
   \]  

2. As a histogram, which records how many of each data type are present in the database, where \( x \in \mathcal{X}^n \), with \( \mathcal{X} = \{+, -\} \). For example,

   \[
   x = \begin{bmatrix}
   + \\
   - \\
   + \\
   - 
   \end{bmatrix}
   \]  

   or equivalently \( x \in \mathbb{N}^{\vert \mathcal{X} \vert} \). For example,

   \[
   x = (2 \ 2)
   \]

   The latter is the notation we will primarily use in the course.

Definition 1. \( l_1 \)-norm
The \( l_1 \)-norm of a database \( x \in \mathbb{N}^{|\mathcal{X}|} \) is defined as

\[
\|x\|_1 = \sum_{i=1}^{|\mathcal{X}|} |x_i|
\]  

This counts the total number of entries in the database by summing the number of entries, or individuals, with data type \( i \) over all \( |\mathcal{X}| \) possible data types. In this course, this sum is typically defined to be \( n \).

**Definition 2. \( l_1 \)-distance (Manhattan Distance)**

The \( l_1 \)-distance between two databases \( x, y \in \mathbb{N}^{|\mathcal{X}|} \) is

\[
\|x - y\|_1
\]

This counts the number of entries on which databases \( x \) and \( y \) differ.

**Definition 3. Neighboring Databases**

We say two databases \( x, y \in \mathbb{N}^{|\mathcal{X}|} \) are neighboring if

\[
\|x - y\|_1 \leq 1
\]

Here we will primarily use the notion of adding or deleting an entry to get to a neighboring database; so one of \( x \) or \( y \) will have \( n \) entries and the other will have \( n - 1 \).

### 1.1 Differential Privacy

**Definition 4. Differential Privacy**

A randomized mechanism \( \mathcal{M} : \mathbb{N}^{|\mathcal{X}|} \rightarrow \text{Range}(\mathcal{M}) \) is \((\varepsilon, \delta)\)-differentially private if \( \forall \mathcal{S} \subseteq \text{Range}(\mathcal{M}) \) and \( \forall \) neighboring \( x, y \in \mathbb{N}^{|\mathcal{X}|} \),

\[
Pr[\mathcal{M}(x) \in \mathcal{S}] \leq e^\varepsilon Pr[\mathcal{M}(y) \in \mathcal{S}] + \delta
\]

The intuition here is that we want to compare the distribution of the output when \( \mathcal{M} \) is run on \( x \) to that when it is run on \( y \). The probability that \( \mathcal{M} \) when run on \( x \) outputs something in \( \mathcal{S} \) should be close, by a factor of \( e^\varepsilon \), to the probability that \( \mathcal{M} \) is run on \( y \) plus the additive \( \delta \) term. When \( \delta = 0 \), we say \( \mathcal{M} \) is \( \varepsilon \)-differentially private.

There are a few things we should note about this definition of differential privacy:

- It is worst case over all possible pairs of neighboring databases \( x, y \in \mathbb{N}^{|\mathcal{X}|} \)
  - We can delete any entry or add any entry
  - We can have any arbitrary combination of the data
  - No assumptions are made about how data look
Figure 1: An example of the significance of $\delta$ in the plots of $M(x)$ vs. $M(y)$

- It is worst case over all possible events $S$
  - Think of $S$ as being some bad outcome
  - Our definition of differential privacy says that the chances of this bad event occurring is about the same regardless of any one individual’s data
  - This is how we ensure we are only able to extract global properties of the data rather than that of individuals

- $\epsilon$ is the privacy parameter
  - $\epsilon = 0$ implies perfect privacy
  - $\epsilon = \infty$ implies no privacy
  - We choose the appropriate $\epsilon$ based on how much privacy we want to give depending on our application area
  - $e^\epsilon \approx 1 + \epsilon$ for small $\epsilon$
  - Typically $\epsilon$ is a small constant or something that is diminishing in $n$

- $\delta$ serves two major purposes
  - Consider the graphs, in Figure 1 above, of $M(x)$ and $M(y)$ that follow each other closely, but differ at the tails in the following ways
    1. If for some subset $S \subseteq \text{Range}(M)$ there is a "blip" where $M(y) > M(x)$
    2. If for some subset $S \subseteq \text{Range}(M)$, $M(x) = 0$ and $M(y) \neq 0$
  - $\delta$ represents the difference between the outputs

- Randomization is key
  - We cannot have a deterministic differentially private algorithm that does anything nontrivial
- Let $\mathcal{M}(x) = s$. This requires that $Pr[\mathcal{M}(x) = r] = 0 \forall r \neq s$. So, with our definition of differential privacy, we notice that $Pr[\mathcal{M}(y) = r] = 0 \forall r \neq s \forall$ neighboring $y$. This implies that $\mathcal{M}(y) = s$ and that $\forall z \in \mathbb{N}^{|\chi|}, \mathcal{M}(z) = s$. The completion of the proof is left as an exercise for homework.

- What are the implications of $S \subseteq \text{Range}(\mathcal{M})$ vs. $S \in \text{Range}(\mathcal{M})$?
  - When $\delta = 0$, the above statements are equivalent
  - When $\delta > 0$, the above statements are not equivalent
  - The proof of this is left as an exercise for homework

## 2 Laplace Mechanism

### Definition 5. $l_1$-sensitivity (Global Sensitivity)

The $l_1$-sensitivity of a query $f : \mathbb{N}^{|\chi|} \to \mathbb{R}^t$ is

$$\Delta f = \max_{x,y \in \mathbb{N}^{|\chi|}, \|x-y\|_1 \leq 1} \|f(x) - f(y)\|_1$$  \hspace{1cm} (8)

### Definition 6. Laplace Distribution

The Laplace distribution centered at 0 and with scale parameter $b$ has the distribution

$$Lap(z|b) = \frac{1}{2b} \exp\left(-\frac{|z|}{b}\right)$$  \hspace{1cm} (9)

We will abuse the notation $Lap(b)$ to refer to both the distribution $Lap(z|b)$ and random variable $X \sim Lap(b)$. Note that with smaller $b$, $Lap(b)$ is ”pointier”, and with larger $b$, $Lap(b)$ is ”flatter”. This is shown in Figure 2 above.
Definition 7. Laplace Mechanism

Given query $f : \mathbb{N}^{\chi} \rightarrow \mathbb{R}^k$, the Laplace mechanism is

$$ML(x, f, \varepsilon) = f(x) + (Y_1 \ldots Y_k)$$

(10)

where $Y_i \sim \text{Lap}(\frac{\Delta f}{\varepsilon})$ i.i.d. This mechanism was first explored in [DMNS06].

Theorem 8. The Laplace mechanism is $\varepsilon$-differentially private.

Proof. Let $x, y \in \mathbb{N}^{\chi}$ be arbitrary neighboring databases, let $f : \mathbb{N}^{\chi} \rightarrow \mathbb{R}$ be an arbitrary query, and let $s \in \mathbb{R}$.

$$\frac{Pr[ML(x, f, \varepsilon) = s]}{Pr[ML(y, f, \varepsilon) = s]} = \frac{Pr[Lap(\frac{\Delta f}{\varepsilon}) = s - f(x)]}{Pr[Lap(\frac{\Delta f}{\varepsilon}) = s - f(y)]} = \frac{\varepsilon}{\Delta f} \exp(-\frac{|s-f(x)|\varepsilon}{\Delta f})$$

$$= \exp(-\frac{\varepsilon(|s-f(y)|-|s-f(x)|)}{\Delta f}) = \exp(\frac{\varepsilon|f(y)-f(x)|}{\Delta f}) \leq \exp(\varepsilon)$$

(11)

References
