Probabilistic Methods for Privacy and Secure Information Flow

Catuscia Palamidessi

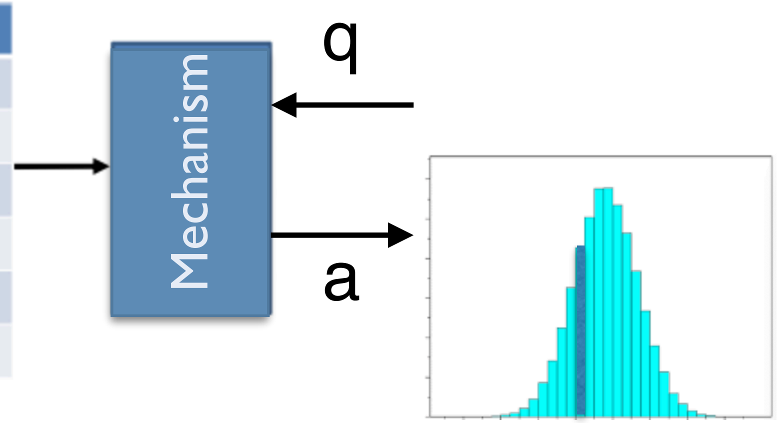
Lecture 3

Resume of previous lecture

Privacy via randomization

Centralized model

Sex	Blood	...	HIV
F	B	...	Y
M	A	...	N
M	O	...	N
M	O	...	Y
F	A	...	N
M	B	...	Y



Differential privacy

- Compositionality: robustness to combination attacks
- Bayesian Interpretation of differential privacy: strong adversary model

Compositionality

Differential privacy is compositional:

Definition Let $\mathcal{K}_1$ and $\mathcal{K}_2$ be two mechanisms on $\mathcal{X}$. Their composition $\mathcal{K}_1 \times \mathcal{K}_2$ is defined as follows:

if $\mathcal{K}_1(x)$ reports z_1 and $\mathcal{K}_2(x)$ reports z_2, then $(\mathcal{K}_1 \times \mathcal{K}_2)(x)$ reports (z_1, z_2)

Theorem (Compositionality) If $\mathcal{K}_1$ and $\mathcal{K}_2$ are respectively ε_1 and ε_2-differentially-private, then their composition $\mathcal{K}_1 \times \mathcal{K}_2$ is $(\varepsilon_1 + \varepsilon_2)$-differentially private.

Proof: Let x and x' be two adjacent DB. Then:

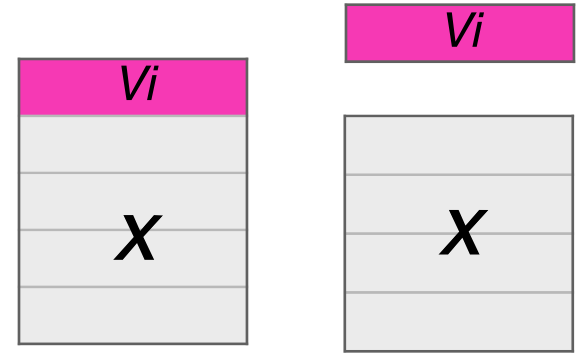
$$\begin{array}{rcl} p((\mathcal{K}_1 \times \mathcal{K}_2)(x) = (z_1, z_2)) & = & p(\mathcal{K}_1(x) = z_1) \quad p(\mathcal{K}_2(x) = z_2) \\ & \leq & e^{\varepsilon_1}\, p(\mathcal{K}_1(x') = z_1) \quad e^{\varepsilon_2}\, p(\mathcal{K}_2(x') = z_2) \\ & = & e^{\varepsilon_1 + \varepsilon_2}\, p((\mathcal{K}_1 \times \mathcal{K}_2)(x') = (z_1, z_2)) \end{array}$$

Bayesian interpretation of DP

Consider an individual i whose value is represented by the random variable *Vi* with the same distribution as *V*

The individual *i* may or may not be present in the DB

The rest of the elements of the DB (or the whole DB) is represented by the random variable *X*



Theorem $\mathcal{K}$ is ε-differentially-private iff $\forall v \in \mathcal{V}, \forall x \in \mathcal{X}, \forall z \in \mathcal{Z}$

$$e^{-\varepsilon}\, p(V_i = v | X = x) \ \leq \ p(V_i = v | X = x, Z = z) \ \leq \ e^{\varepsilon}\, p(V_i = v | X = x)$$

where Z represents the reported answer of $\mathcal{K}$.

Proof

Only if) By the Bayes law, we have

$$p(V_i = v | X = x, Z = z) = \frac{p(Z = z | X = x, V_i = v)\; p(V_i = v | X = x)}{p(Z = z | X = x)}$$

And now, just observe that, since $\mathcal{K}$ is ε-DP, we have

$$e^{-\varepsilon}\, p(Z = z | X = x) \ \leq \ p(Z = z | X = x, V_i = v) \ \leq \ e^{\varepsilon}\, p(Z = z | X = x)$$

Note that the above inequalities holds independently from whether the individual i is in the DB or not.

If) Analogous, just reverse the reasoning.

Strong adversary

In the Bayesian interpretation of DP, the conditioning on $X=x$ represents the fact that the adversary knows the rest of the DB. This scenario is called *strong adversary hypothesis* (SAH).

Is this hypothesis necessary for the boundaries expressed by the Bayesian interpretation of DP ?

Strong adversary

In the Bayesian interpretation of DP, the conditioning on $X = x$ represents the fact that the adversary knows the rest of the DB. This scenario is called *strong adversary hypothesis* (SAH).

Is this hypothesis necessary for the boundaries expressed by the Bayesian interpretation of DP ?

Yes. But we can have a similar result without this hypothesis, only with weaker bounds.

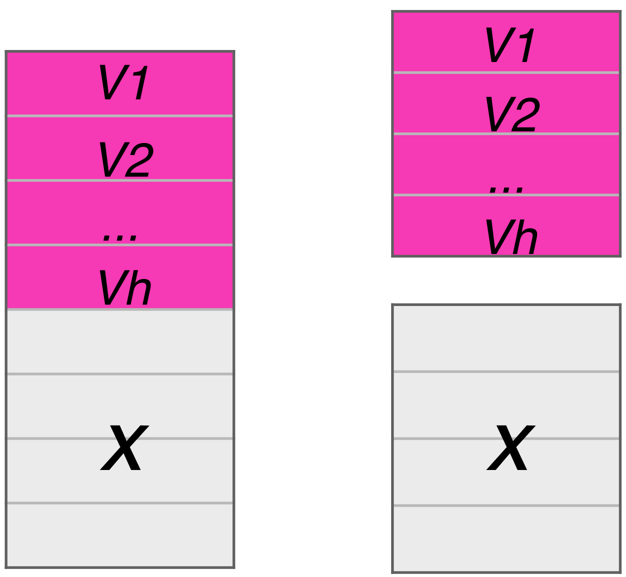
Strong adversary

In the Bayesian interpretation of DP, the conditioning on $X = x$ represents the fact that the adversary knows the rest of the DB. This scenario is called *strong adversary hypothesis* (SAH).

Is this hypothesis necessary for the boundaries expressed by the Bayesian interpretation of DP ?

Yes. But we can have a similar result without this hypothesis, only with weaker bounds.

Consider individuals 1,2, ... h whose value is represented by the RV $\boldsymbol{V} = V1\ V2\ ...Vh$



Bayesian interpretations of DP w/o the SAH

Theorem The following statements are equivalent

1. $\mathcal{K}$ is ε-DP

2. $e^{-h\varepsilon}\, p(\mathbf{V} = \mathbf{v} | X = x) \;\leq\; p(\mathbf{V} = \mathbf{v} | X = x, Z = z) \;\leq\; e^{h\varepsilon}\, p(\mathbf{V} = \mathbf{v} | X = x)$

3. $e^{-h\varepsilon}\, p(V_i = v | X = x) \;\leq\; p(V_i = v | X = x, Z = z) \;\leq\; e^{h\varepsilon}\, p(V_i = v | X = x)$

Furthermore, we can drop the conditioning on $X = x$ if we know that there is no correlation between the V_i's and X (given the result of $\mathcal{K}$, i.e., Z).

Proof

(1) $\leftrightarrow$ (2)) This part can be proved in a way analogous to the previous theorem

(2) $\leftrightarrow$ (3)) Observe that (2) holds for every tuple of values of $\mathbf{V}$ and then marginalize w.r.t. V_i

(3) $\leftrightarrow$ (1)) For $h = 1$, (3) coincides with (1).

Note: The same results hold if we replace the value of V_i with the presence/absence of i in the DB.

Differential Privacy: continuous case

We now consider the continuous case. Namely, $\mathcal{K}(x)$ determines a probability density function on $\mathcal{Z}$. The only thing that change is that we consider measurable subsets $\mathcal{S}$ of $\mathcal{Z}$ rather than single z.

Definition (Differential Privacy) $\mathcal{K}$ is ε-differentially-private iff for every pair of databases $x_1, x_2 \in \mathcal{X}$ s.t. $x_1 \sim x_2$ and for every measurable $\mathcal{S} \subseteq \mathcal{Z}$ we have

$$p(\mathcal{K}(x_1) \in \mathcal{S}) \leq e^{\varepsilon}\, p(\mathcal{K}(x_2) \in \mathcal{S})$$

where $p(\mathcal{K}(x) \in \mathcal{S})$ represents the probability that $\mathcal{K}$ applied to x report an answer in $\mathcal{S}$

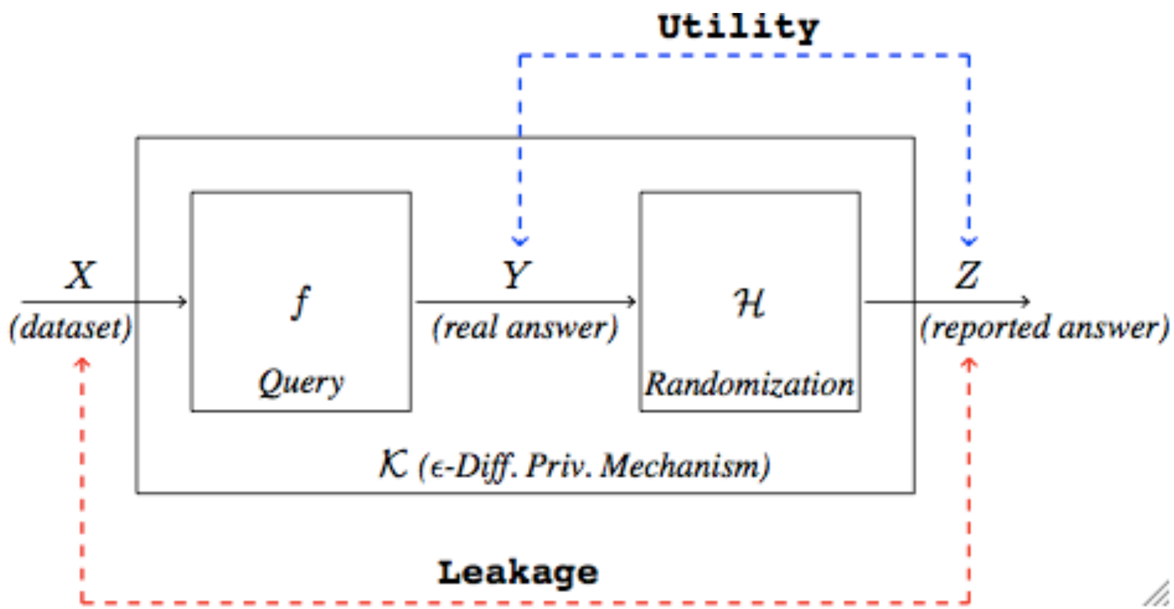
Note: $p(\mathcal{K}(x) \in \mathcal{S})$ represents a conditional probability. We will write it as $p(Z \in \mathcal{S} | X = x)$ when we need to make this fact more explicit.

Some "real" DP mechanisms

Oblivious Mechanisms

- Given $f: \mathcal{X} \rightarrow \mathcal{Y}$ and $\mathcal{K}: \mathcal{X} \rightarrow \mathcal{Z}$, we say that $\mathcal{K}$ is oblivious if it depends only on $\mathcal{Y}$ (not on $\mathcal{X}$)

- If $\mathcal{K}$ is oblivious, it can be seen as the composition of f and a randomized mechanism $\mathcal{H}$ (noise) defined on the exact answers $\mathcal{K} = \mathcal{H} \circ f$



- Privacy concerns the information flow between the databases and the reported answers, while utility concerns the information flow between the correct answer and the reported answer

A typical oblivious DP mechanism: Laplace noise

- Randomized mechanism for a query $f : \mathcal{X} \to \mathcal{Y}$.
- A typical randomized method: **add Laplace noise to** $y = f(x)$**.**
  Namely, report z with a probability density function defined as:

$$dP_y(z) = c\, e^{-\frac{|z-y|}{\Delta f}\varepsilon}$$

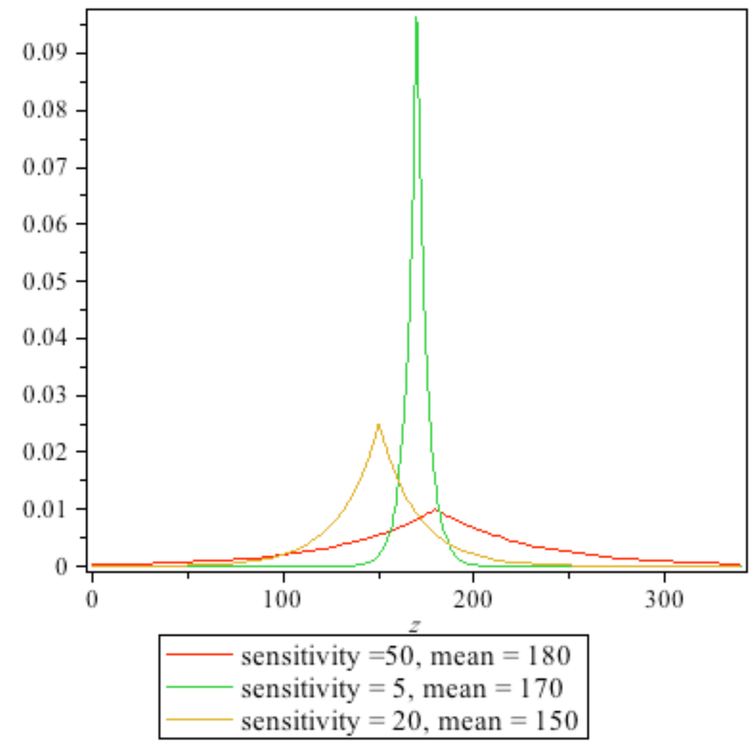
where Δf is the *sensitivity* of f:

$$\Delta f = \max_{x \sim x' \in \mathcal{X}} |f(x) - f(x')|$$

($x \sim x'$ means x and x' are adjacent, i.e., they differ only for one record)

and c is a normalization factor:

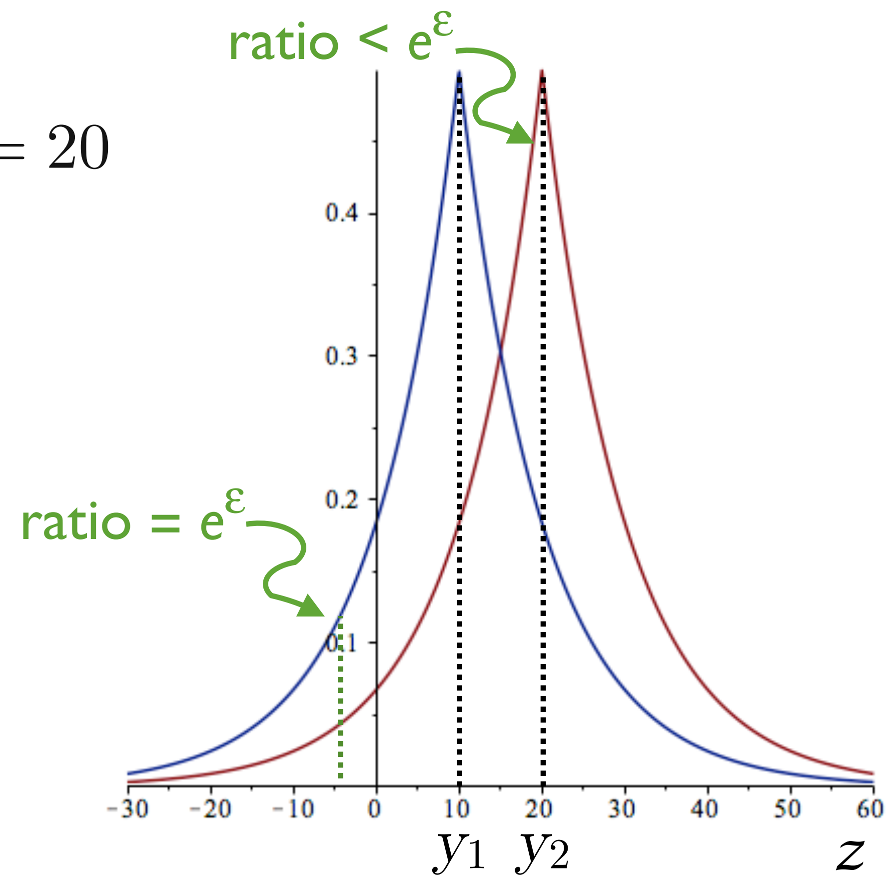
$$c = \frac{\varepsilon}{2\,\Delta f}$$



Example of Laplace Mechanism

- $\varepsilon = 1$
- $\Delta_f = |f(x_1) - f(x_2)| = 10$
- $y_1 = f(x_1) = 10,\ y_1 = f(x_2) = 20$

  Then:
- $dP_{y_1} = \frac{1}{2 \cdot 10} e^{\frac{|z-10|}{10}}$
- $dP_{y_2} = \frac{1}{2 \cdot 10} e^{\frac{|z-20|}{10}}$



The ratio between these distribution is

- $= e^{\varepsilon}$ outside the interval $[y_1, y_2]$
- $\leq e^{\varepsilon}$ inside the interval $[y_1, y_2]$

The Laplace mechanism is DP

Remember that the probability density function of the Laplace mechanism is:

$$p(Z=z|X=x) = dP_{f(x)}(z) = c\,e^{-\frac{|z-f(x)|}{\Delta f}\varepsilon}$$

where $c = \frac{\varepsilon}{2\,\Delta f}$

Theorem: The Laplace mechanism is ε–differentially private

Proof: Let $x_1 \sim x_2$ and $y_1 = f(x_1), y_2 = f(x_2)$ We have:

$$\begin{aligned}
\frac{p(Z{=}z|X{=}x_1)}{p(Z{=}z|X{=}x_2)} &= \frac{c\,e^{-\frac{|z-f(x_1)|}{\Delta f}\varepsilon}}{c\,e^{-\frac{|z-f(x_2)|}{\Delta f}\varepsilon}} \\
&= e^{\frac{|z-y_2|}{\Delta f}\varepsilon - \frac{|z-y_1|}{\Delta f}\varepsilon} \\
&\leq e^{\frac{|y_1-y_2|}{\Delta f}\varepsilon} \\
&\leq e^{\varepsilon}
\end{aligned}$$

Sensitivity of the query

- The sensitivity of the query and the level of privacy ε determine the amount of noise of the mechanism:
  - higher sensitivity $\Rightarrow$ more noise
  - smaller ε $\Rightarrow$ more privacy, more noise
- Intuitively, the more the mechanism is noisy, the less useful it is (the reported answer is less precise)
- To reduce the sensitivity, for some queries it may help to assume that the database contains a minimum number of individuals
- **Example:** consider the query "What is the average age of the people in the DB ?". Assume that the age can vary from 0 to 120. Check the sensitivity in the following two cases:
  - the DB contains at least 100 records, or
  - there is no restriction.

The geometric mechanism

- The Laplacian noise is typically used in the case that $\mathcal{Y}$ (the set of true answers of the query) is a **continuous** numerical set, like the Reals.

- If $\mathcal{Y}$ is a **discrete** numerical set, like the Integers, then the typical mechanism used in this case is the **geometric mechanism**, which is a sort of discrete Laplacian.

- In the geometric mechanism, the probability distribution of the noise is:

$$p(z|y) = c\, e^{-\frac{|z-y|}{\Delta f}\varepsilon}$$

  - In this expression, c is a normalization factor, defined so to obtain a probability distribution,

  - Δf is the sensitivity of query f

Normalization constant in a geometric mechanism

- In the geometric mechanism, the probability distribution of the noise is:

$$p(z|y) = c\, e^{-\frac{|z-y|}{\Delta f}\varepsilon}$$

As usual, we can compute c (the normalization factor) by imposing that the sum of the probability on all Z is 1. It turns out that $c = \frac{1-\alpha}{1+\alpha}$ where $\alpha = e^{-\frac{\varepsilon}{\Delta f}}$

$$\text{hence} \quad p(z|y) = \frac{1-\alpha}{1+\alpha}\, \alpha^{|z-y|}$$

- **Exercises:** Compute the geometric mechanism for the following queries:
  - " How many diabetic people weight more than 100 kilos ? "
  - " What is the max weight (in kilos) of a diabetic person ? "

Gaussian noise

The formula for gaussian noise is

$$c\; e^{-\frac{(y-z)^2}{\sigma}\varepsilon}$$

where c is a normalization factor and σ is a suitable constant.

Question: does an oblivious mechanism based on this noise function satisfy **ε**-differential privacy, for some suitable value of σ ?

Gaussian noise

The formula for gaussian noise is

$$c\ e^{-\frac{(y-z)^2}{\sigma}\varepsilon}$$

where c is a normalization factor and σ is a suitable constant.

Question: does an oblivious mechanism based on this noise function satisfy **ε**-differential privacy, for some suitable value of σ ?

A gaussian noise does not satisfy differential privacy.

However it satisfies a more relaxed form of privacy called (ε,δ)-DP