CSE660
Differential Privacy

September 25, 2017

Marco Gaboardi

Room: 338-B

gaboardi@buffalo.edu

http://www.buffalo.edu/~gaboardi

Outline of the class

Week 1

Introduction, motivation and privacy limitations. Definition of Differential Privacy and the curator model.

Week 2

Basic mechanisms: Randomized Response, Laplace Mechanism,

Week 3

Basic properties following from the definition, Exponential Mechanism and comparison with the other basic mechanisms.

Week 4

The Report Noisy max algorithm. The Sparse Vector technique.

Week 5

Formalizing privacy proofs using an approximate probabilistic coupling argument.

Releasing Many Counting Queries with Correlated Noise. The smallDB algorithm.

Week 6

The MWEM algorithm. The DualQuery algorithm.

Outline of the class

Week 8
Studying the experimental accuracy. Adaptivity and adaptive MWEM.
Week 7
Variations on differential privacy: Renyi DP, zero-concentrated DP
Week 9
PAC learning and private PAC learning
Week 10
The local model for differential privacy.
Week 11
More algorithms for the local model.
Week 12
Differentially Private Hypothesis Testing
Week 13
Differential Privacy and Generalization in Adaptive Data Analysis
Week 14
Project presentations

(ε,δ)-Differential Privacy

Definition

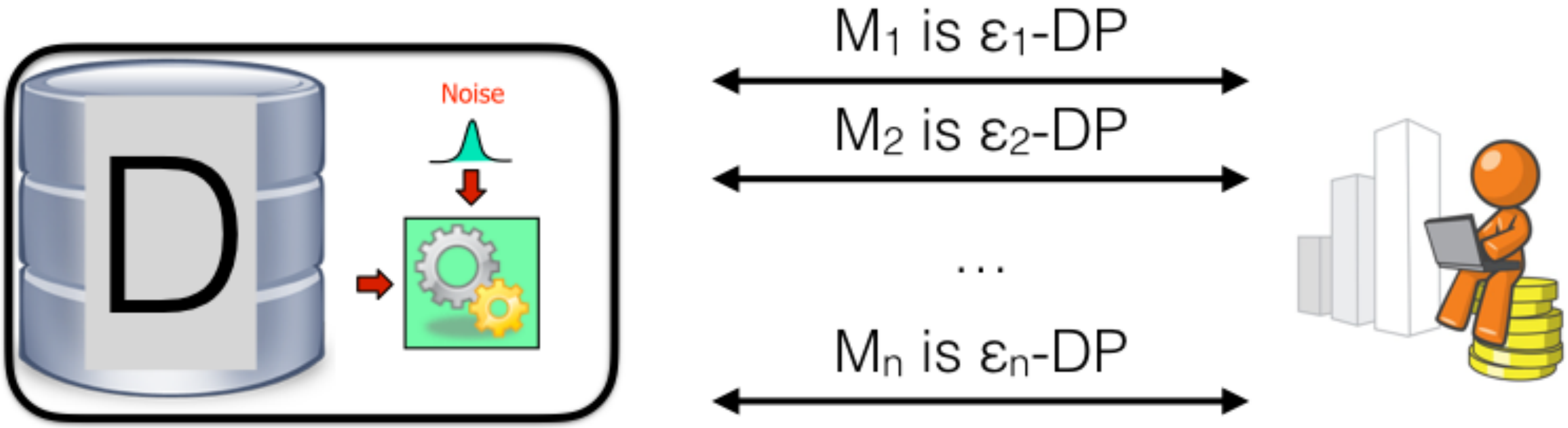
Given $\varepsilon,\delta \geq 0$, a probabilistic query $Q: X^n \rightarrow R$ is (ε,δ)-differentially private iff for all adjacent database b_1, b_2 and for every $S \subseteq R$:

$$\Pr[Q(b_1) \in S] \leq \exp(\varepsilon)\Pr[Q(b_2) \in S] + \delta$$

Some important properties

- Resilience to post-processing
- Group privacy
- Composition

Composition



M_1 is ε_1-DP

M_2 is ε_2-DP

…

M_n is ε_n-DP

The overall process is $(\varepsilon_1+\varepsilon_2+\ldots+\varepsilon_n)$-DP

Composition

Question: How about histograms?

Example III

Let's consider an arbitrary universe domain $\mathcal{X}$ and let's consider the following predicate for $y \in \mathcal{X}$

$$q_y(x) = \begin{cases} 1 & \text{if } y = x \\ 0 & \text{otherwise} \end{cases}$$

we call a point function the associated counting query

$$q_y : \mathcal{X}^n \to [0,1]$$

Question: What is the sensitivity?

Example III

$\text{Budget}=\varepsilon_{\text{global}} - \varepsilon \;\; - \varepsilon \;\; - \varepsilon \; - \varepsilon$
$- \varepsilon \;\; - \varepsilon \;\; - \varepsilon \;\; - \varepsilon$

Can we do better?

$D \in X^{10} =$

	D1	D2	D3
I1	0	0	0
I2	1	0	1
I3	0	1	0
I4	1	0	1
I5	0	0	0
I6	0	0	1
I7	1	1	0
I8	0	0	0
I9	0	1	0
I10	1	0	1

$q^*_{000}(D) = .3+L(1/n\varepsilon)$

$q^*_{001}(D) = .1+L(1/n\varepsilon)$

$q^*_{010}(D) = .2+L(1/n\varepsilon)$

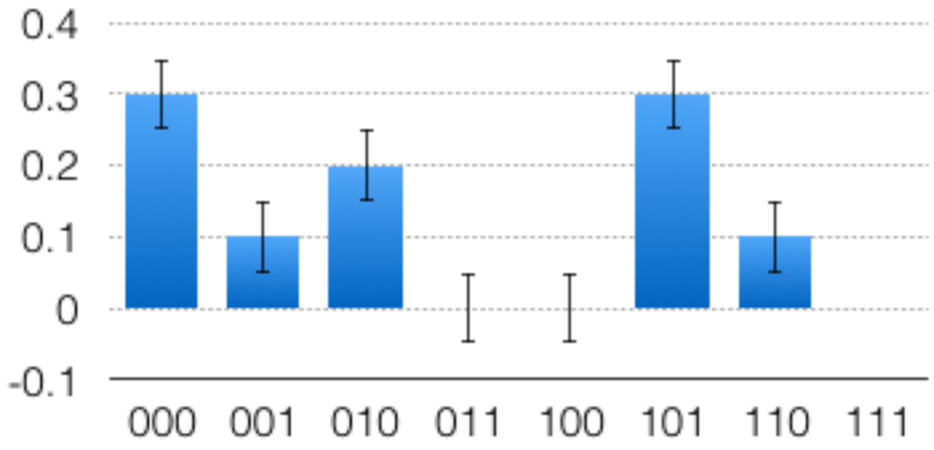
$q^*_{011}(D) = 0+L(1/n\varepsilon)$

$q^*_{100}(D) = 0+L(1/n\varepsilon)$

$q^*_{101}(D) = .3+L(1/n\varepsilon)$

$q^*_{110}(D) = .1+L(1/n\varepsilon)$

$q^*_{111}(D) = 0+L(1/n\varepsilon)$



Example III

$D \in X^{10} =$

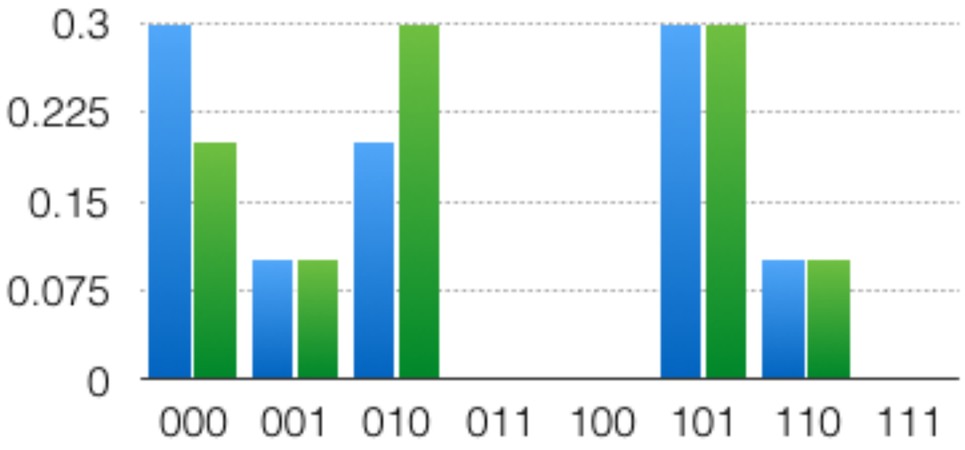
	D1	D2	D3
I1	0	0	0
I2	1	0	1
I3	0	1	0
I4	1	0	1
I5	0	0	0
I6	0	0	1
I7	1	1	0
I8	0	0	0
I9	0	1	0
I10	1	0	1

$D' \in X^{10} =$

	D1	D2	D3
I1	0	0	0
I2	1	0	1
I3	0	1	0
I4	1	0	1
I5	0	1	0
I6	0	0	1
I7	1	1	0
I8	0	0	0
I9	0	1	0
I10	1	0	1

$q_{000}(D) = .3$
$q_{001}(D) = .1$
$q_{010}(D) = .2$
$q_{011}(D) = 0$
$q_{100}(D) = 0$
$q_{101}(D) = .3$
$q_{110}(D) = .1$
$q_{111}(D) = 0$

$q_{000}(D') = .2$
$q_{001}(D') = .1$
$q_{010}(D') = .3$
$q_{011}(D') = 0$
$q_{100}(D') = 0$
$q_{101}(D') = .3$
$q_{110}(D') = .1$
$q_{111}(D') = 0$



Example III

$\text{Budget}=\varepsilon_{\text{global}} - 2\varepsilon$

Can we do better?

$D \in X^{10} =$

	D1	D2	D3
I1	0	0	0
I2	1	0	1
I3	0	1	0
I4	1	0	1
I5	0	0	0
I6	0	0	1
I7	1	1	0
I8	0	0	0
I9	0	1	0
I10	1	0	1

$q^*_{000}(D) = .3+L(1/n\varepsilon)$

$q^*_{001}(D) = .1+L(1/n\varepsilon)$

$q^*_{010}(D) = .2+L(1/n\varepsilon)$

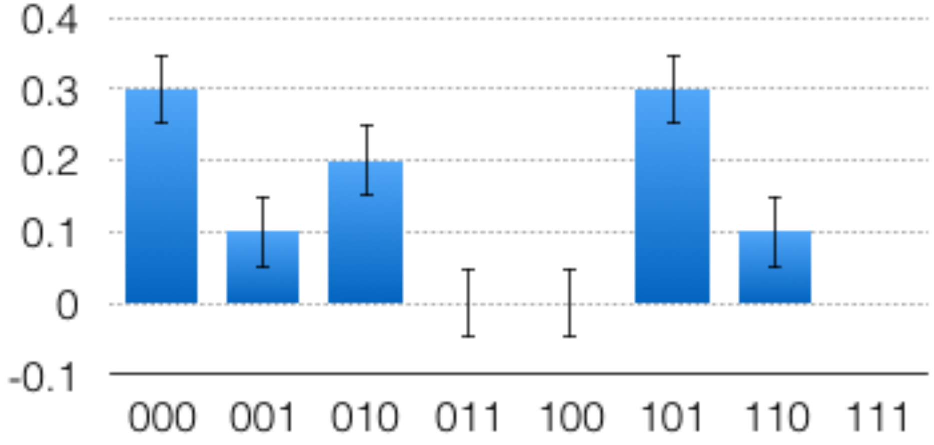
$q^*_{011}(D) = 0+L(1/n\varepsilon)$

$q^*_{100}(D) = 0+L(1/n\varepsilon)$

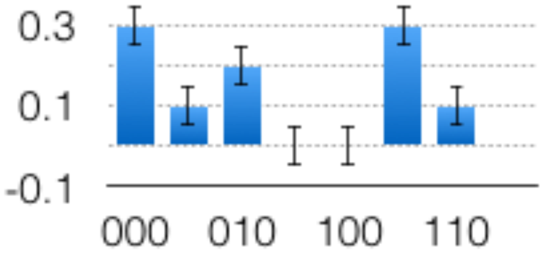
$q^*_{101}(D) = .3+L(1/n\varepsilon)$

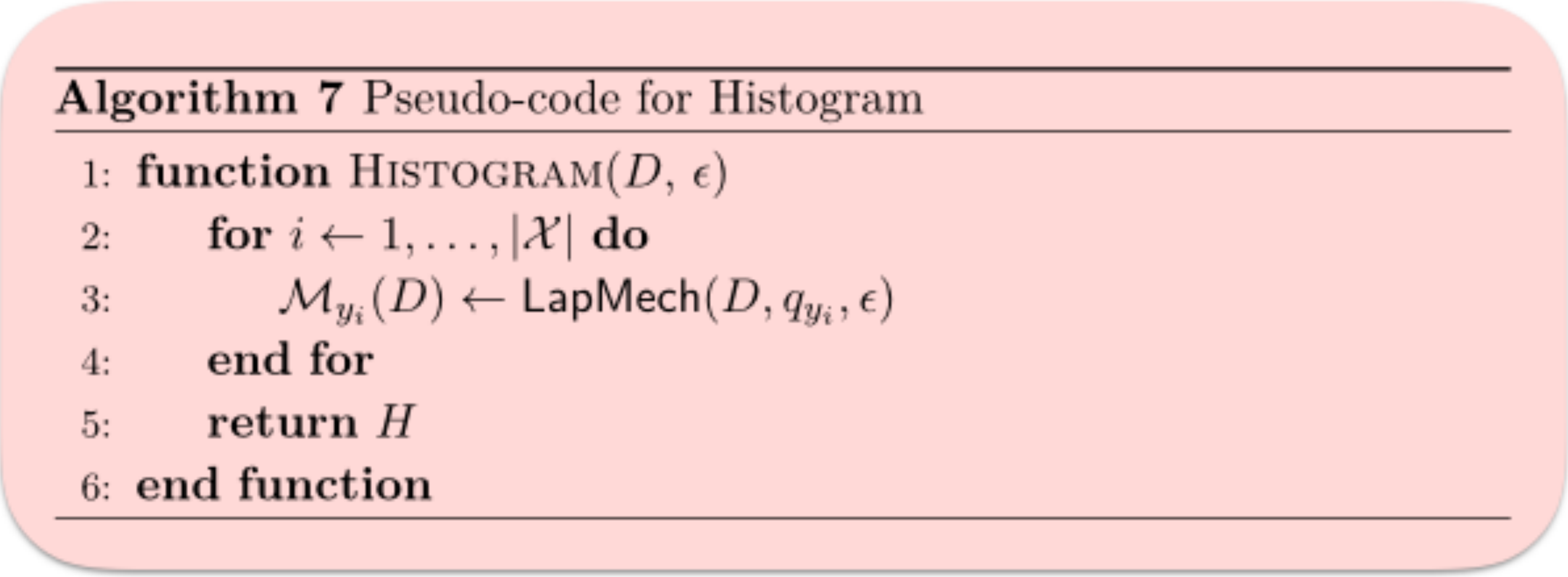
$q^*_{110}(D) = .1+L(1/n\varepsilon)$

$q^*_{111}(D) = 0+L(1/n\varepsilon)$



DP Histograms



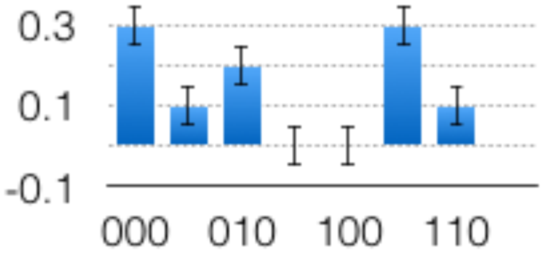


Algorithm 7 Pseudo-code for Histogram

```
function HISTOGRAM($D$, $\epsilon$)
    for $i \leftarrow 1, \ldots, |\mathcal{X}|$ do
        $\mathcal{M}_{y_i}(D) \leftarrow \mathsf{LapMech}(D, q_{y_i}, \epsilon)$
    end for
    return $H$
end function
```

Theorem 1.11. The algorithm Histogram is 2ϵ differentially private.

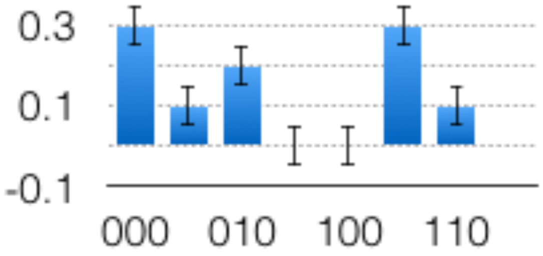
DP Histograms



Theorem 1.11. The algorithm Histogram is 2ϵ differentially private.

Proof. Fix any pair of adjacent datasets $D \sim_1 D'$ and consider the answer to the point function queries on the two datasets: $q_{y_1}(D), \dots, q_{y_{|\mathcal{X}|}}(D)$ and $q_{y_1}(D'), \dots, q_{y_{|\mathcal{X}|}}(D')$, respectively. It is immediate to see that since D and D' differ only in one position, there will be two elements y and y' for which $q_y(D) \neq q_y(D')$ and $q_{y'}(D) \neq q_{y'}(D')$ while for every other elements the results of the point function queries will be the same.

DP Histograms



Theorem 1.11. The algorithm Histogram is 2ϵ differentially private.

Now, let's fix a possible output $(r_1, \ldots, r_{|\mathcal{X}|})$. We have:

$$\frac{\Pr[\mathcal{M}(D) = (r_1, \ldots, r_{|\mathcal{X}|})]}{\Pr[\mathcal{M}(D') = (r_1, \ldots, r_{|\mathcal{X}|})]} = \frac{\prod_{i=1}^{|\mathcal{X}|} \Pr[\mathcal{M}_{y_i}(D) = r_i]}{\prod_{i=1}^{|\mathcal{X}|} \Pr[\mathcal{M}_{y_i}(D') = r_i]}$$

Using the observation above we have

$$\begin{aligned}
\frac{\prod_{i=1}^{|\mathcal{X}|} \Pr[\mathcal{M}_{y_i}(D) = r_i]}{\prod_{i=1}^{|\mathcal{X}|} \Pr[\mathcal{M}_{y_i}(D') = r_i]} &= \frac{\Pr[\mathcal{M}_y(D) = r_y] \Pr[\mathcal{M}_{y'}(D) = r_{y'}]}{\Pr[\mathcal{M}_y(D') = r_y] \Pr[\mathcal{M}_{y'}(D') = r_{y'}]} \\
&= \Big(\frac{\Pr[\mathcal{M}_1(D) = r_1]}{\Pr[\mathcal{M}_1(D') = r_1]}\Big)\Big(\frac{\Pr[\mathcal{M}_2(D) = r_2]}{\Pr[\mathcal{M}_2(D') = r_2]}\Big) \\
&\leq \exp(\epsilon) \exp(\epsilon) = \exp(2\epsilon).
\end{aligned}$$

Composition

We always need to think before applying composition to whether we have other options!

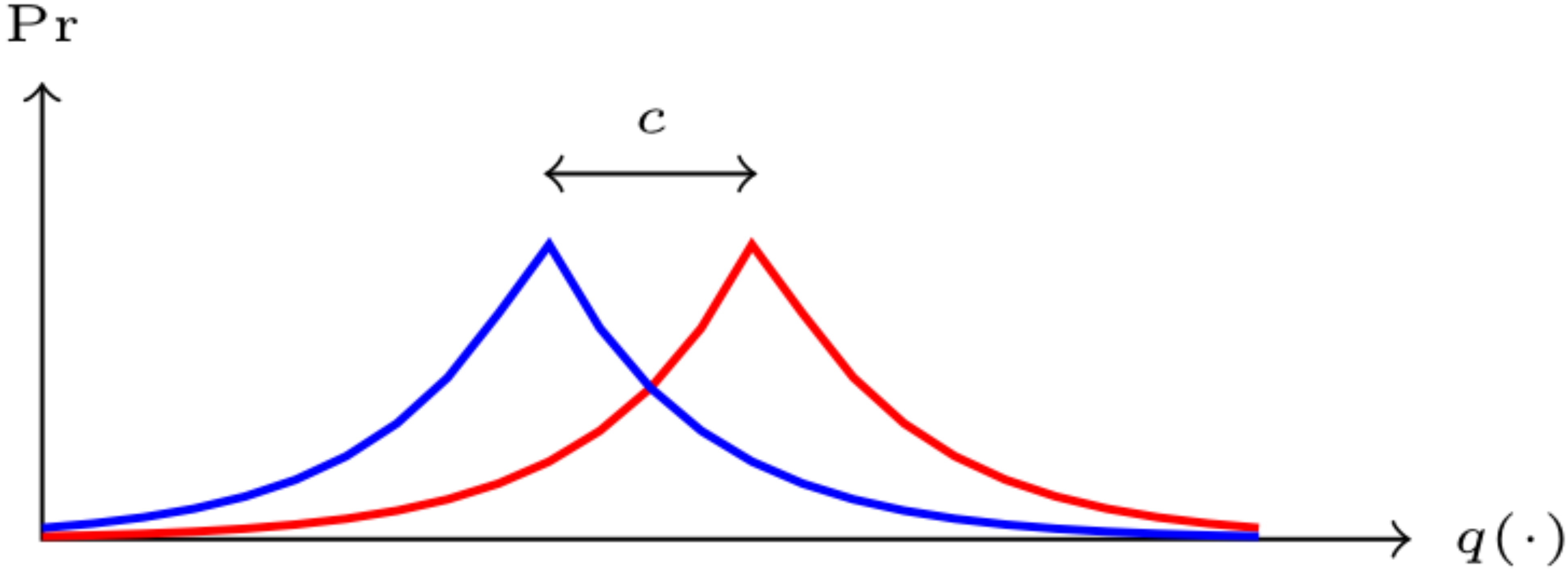
Mechanisms

Question: How about non-numeric data?

Laplace Mechanism

Theorem (Privacy of the Laplace Mechanism)
The Laplace mechanism is ε-differentially private.

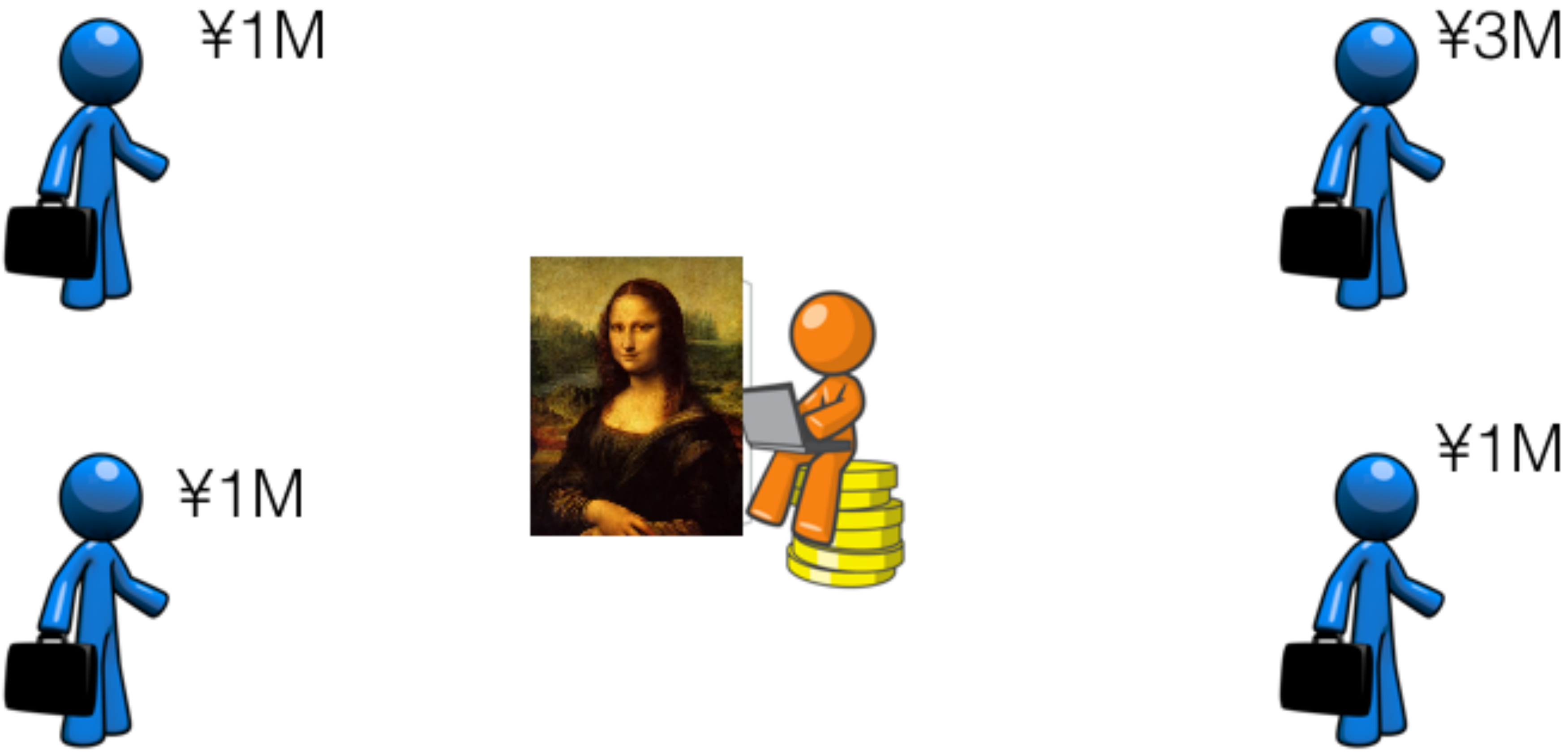
Proof: Intuitively



Another mechanism?

Question: Can we have a similar mechanism for non-numeric queries?

Auctions



How shall we set the price under differential privacy?

Max response

(a)

(b)

(c)

(d)

Suppose that each one of us can vote for one star, and we want to say who is the star that receives most votes.

The answer here is not a number, how can we release it under differential privacy?

Differentially private selection

- We want to select some element that maximize some value,
- The noise added for privacy should not destroy the utility,
- If we cannot return the maximal one, with high probability, we want to return one close to it.

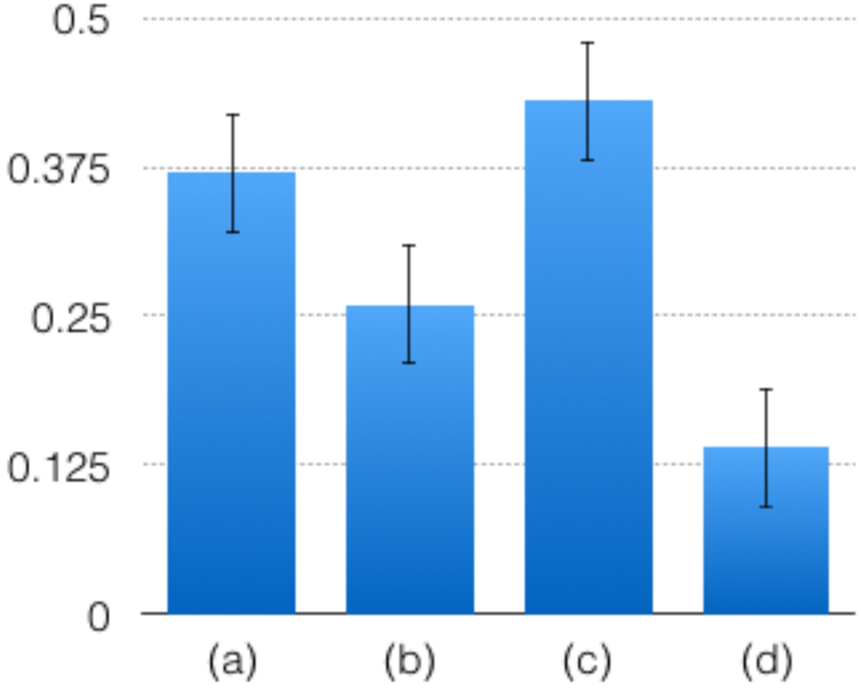
Max response

(a)

(b)

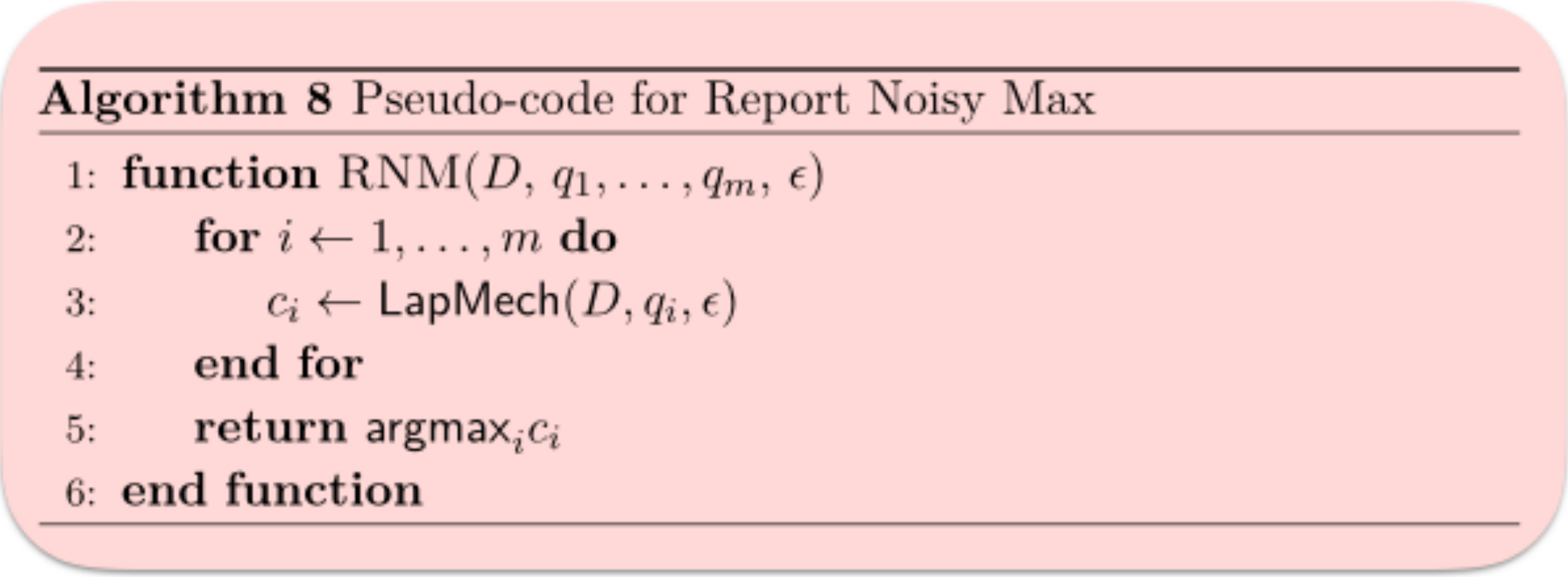
(c)

(d)



Intuition:
We could compute the histogram add Laplace noise to each score and then select the maximal noised score.

Report Noisy Max



Algorithm 8 Pseudo-code for Report Noisy Max

```
**function** RNM($D$, $q_1, \ldots, q_m$, $\epsilon$)
    **for** $i \leftarrow 1, \ldots, m$ **do**
        $c_i \leftarrow \mathsf{LapMech}(D, q_i, \epsilon)$
    **end for**
    **return** $\mathsf{argmax}_i c_i$
**end function**
```

Report Noisy Max

Theorem 1.12. The algorithm RNM is ϵ-differentially private.

Proof. Fix $D \sim D'$, and $c = q(\vec{D})$ and $c' = q(\vec{D}')$ be the results of the counting queries in the two cases. For simplicity we assume that c and c' have the following properties.

- i) $\forall j, c_j \geq c'_j$
- ii) $\forall j, \frac{1}{n} + c'_j \geq c_j$

The general case is a little more involved but overall similar.

Let's now fix i, we want to bound the ratio of the probability that i is selected when running RNM on D and D'. Now let r_j be the noise drawn from Laplace for the round j and r_{-i} be the noise drawn from Laplace for all the rounds except the i-th one.

Report Noisy Max

Theorem 1.12. The algorithm RNM is ϵ-differentially private.

We first argue that $\Pr[i|D, r_{-i}] \leq e^{\varepsilon} \Pr[i|D', r_{-i}]$. Define

$$r^* = \min_{r_i} : c_i + r_i > c_j + r_j \ \forall j \neq i.$$

Note that, having fixed r_{-i}, i will be the output (the argmax noisy count) when the database is D if and only if $r_i \geq r^*$.

We have, for all $1 \leq j \neq i \leq m$:

$$\begin{aligned} c_i + r^* &> c_j + r_j \\ \Rightarrow (1 + c_i') + r^* \geq c_i + r^* &> c_j + r_j \geq c_j' + r_j \\ \Rightarrow c_i' + (r^* + 1) &> c_j' + r_j. \end{aligned}$$

Report Noisy Max

Theorem 1.12. The algorithm RNM is ϵ-differentially private.

Thus, if $r_i \geq r^* + 1$, then the ith count will be the maximum when the database is D' and the noise vector is (r_i, r_{-i}). The probabilities below are over the choice of $r_i \sim \mathrm{Lap}(1/\varepsilon)$.

$$\Pr[r_i \geq 1 + r^*] \geq e^{-\varepsilon} \Pr[r_i \geq r^*] = e^{-\varepsilon} \Pr[i|D, r_{-i}]$$
$$\Rightarrow \Pr[i|D', r_{-i}] \geq \Pr[r_i \geq 1 + r^*] \geq e^{-\varepsilon} \Pr[r_i \geq r^*] = e^{-\varepsilon} \Pr[i|D, r_{-i}],$$

which, after multiplying through by e^{ε}, yields what we wanted to show: $\Pr[i|D, r_{-i}] \leq e^{\varepsilon} \Pr[i|D', r_{-i}]$.

Report Noisy Max

Theorem 1.12. The algorithm RNM is ϵ-differentially private.

We now argue that $\Pr[i|D', r_{-i}] \leq e^{\varepsilon} \Pr[i|D, r_{-i}]$. Define

$$r^* = \min_{r_i} : c'_i + r_i > c'_j + r_j \ \forall j \neq i.$$

Note that, having fixed r_{-i}, i will be the output (argmax noisy count) when the database is D' if and only if $r_i \geq r^*$.

We have, for all $1 \leq j \neq i \leq m$:

$$\begin{aligned} c'_i + r^* &> c'_j + r_j \\ \Rightarrow 1 + c'_i + r^* &> 1 + c'_j + r_j \\ \Rightarrow c'_i + (r^* + 1) &> (1 + c'_j) + r_j \\ \Rightarrow c_i + (r^* + 1) \geq c'_i + (r^* + 1) &> (1 + c'_j) + r_j \geq c_j + r_j. \end{aligned}$$

Report Noisy Max

Theorem 1.12. The algorithm RNM is ϵ-differentially private.

Thus, if $r_i \geq r^* + 1$, then i will be the output (the argmax noisy count) on database D with randomness (r_i, r_{-i}). We therefore have, with probabilities taken over choice of r_i:

$$\Pr[i|D, r_{-i}] \geq \Pr[r_i \geq r^* + 1] \geq e^{-\varepsilon} \Pr[r_i \geq r^*] = e^{-\varepsilon} \Pr[i|D', r_{-i}],$$

which, after multiplying through by e^{ε}, yields what we wanted to show: $\Pr[i|D', r_{-i}] \leq e^{\varepsilon} \Pr[i|D, r_{-i}]$. □

Exponential Mechanism

The Exponential Mechanism generalize this approach.

Suppose that we have a scoring function u(D,o) that to each pair (database, potential output) assign a score (a negative real number).

We want to output approximately the element with the max score.

Exponential Mechanism

Exponential Mechanism:

$\mathcal{M}_E(x, u, \mathcal{R})$

return $r \in \mathcal{R}$ with prob. $\dfrac{\exp(\frac{\varepsilon u(x,r)}{2\Delta u})}{\sum_{r' \in \mathcal{R}} \exp(\frac{\varepsilon u(x,r')}{2\Delta u})}$

where

$$\Delta u = \max_{r \in \mathcal{R}} \max_{x \sim_1 y} \left| u(x,r) - u(y,r) \right|$$