Differential Privacy

27 April 2021 09:05

Lots of data, lots of "analysis" one can do.

<u>Eg</u> :

{ Understand causality b/w smoking and lung disease. }

↓

Population wide studies.

<u>Q</u> :

{ How to conduct useful population-wide studies / "data analysis" without compromising individual <u>privacy</u> }

↓

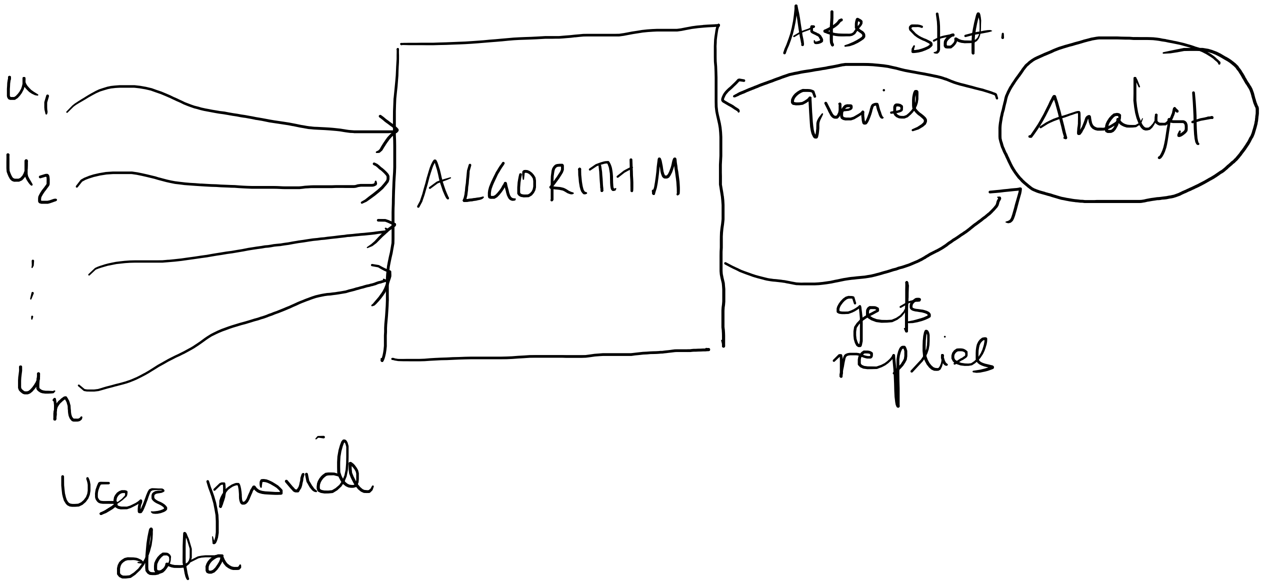
Means users are incentivized to join the study

[Collectively we can learn something, but individually don't lose anything]

{ different from "CRYPTOGRAPHY" } Which is like a vault + key.

{ Here we want to contribute data but the <u>analyst</u> learns but the analyst learns nothing "private"

INITIAL ATTEMPT AT MODELING PRIVACY



Attempt ①

Want the analyst to not learn anything new about any individual.

Natural Issue

You will learn something new from the answer of the query.

Example

Example

Imagine each u_i = (# left feet, # right feet)

- Analyst Asks Avg(# left ft - # right ft)
- Reply (likely) = 0

⇒ Analyst "learns" that u_1 has equal # of left & right ft.

(Maybe need few more queries to get min(# left ft)
Min (# right ft), etc.

- Why is this definition vacuous ?

A: what we learnt is something "global".
Nothing specific about the particular individual

Q: How do we capture "Individual Privacy" ?

[The earlier definition of privacy is broken]

DMNS - Dwork, McSherry, Nissim, Smith

[Differential Privacy]

[What if the Algorithm guarantees that the answer to the query is "almost" the same regardless of whether u_1 (or any fixed individual) is present in the input or not?]

The analyst cant even distinguish if u_1 was part of the study or not, so how can he/she learn anything about u_1's data?

The Differential Privacy Model

Input: Database X of 'n' rows, each corr to a user, fn f.

Algo takes X and outputs $\tilde{f}(X)$

[it can be scalar output, vector output, etc]

f is the statistical query

$\tilde{f}$ is the output.

$\tilde{f}$ is the output.

$\tilde{f}(x)$ can be some "noisy/approximate" response to $f(x)$.

If x and x' are two databases which differ in a single row, then

want $\boxed{\tilde{f}(x) \text{ "}\approx\text{" } \tilde{f}(x')}$ ①

↑ guarantees privacy.

<u>and</u>

② $|f(x) - \tilde{f}(x)|$ is "small" for all x

↑ guarantees utility of study.

Need to formalize "$\approx$" meaning and "small"

{ Just ① is easy to satisfy:
Output $\tilde{f}(x) = 0$ always.
Full privacy, No utility }

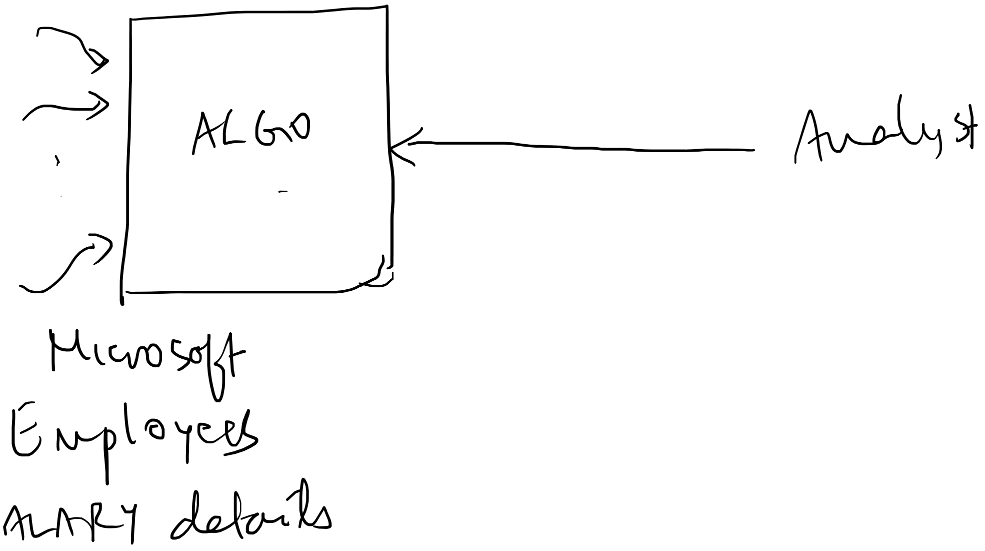
{ How to get both together? }

[illegible] DP many simple queries for

L

There are many simple queries for which we need to introduce noise, else we break privacy.

Ex ①



Analyst asks " count # people with Salary $\geq$ 2M $ "

Sps Algo replies ①.

{ Then analyst can learn that CEO's salary $\geq$ 2M $. }

Output will be 0 if CEO doesn't take part in survey.

take part in survey.

$\Rightarrow$ CEO's privacy is compromised acc. our definition

---

Is this just a 1 vs 0 issue?

Perhaps not.

Sps Answer = 1000.

and tomorrow, the answer is 1001.

Maybe Microsoft hired a new employee,

$\Rightarrow$ { likely that this person has salary $\geq$ 2M \$. }

we may learn something about the new individual.

So our Model / definition captures those issues well

$\Rightarrow$ we must add some noise to $f(x)$ so that

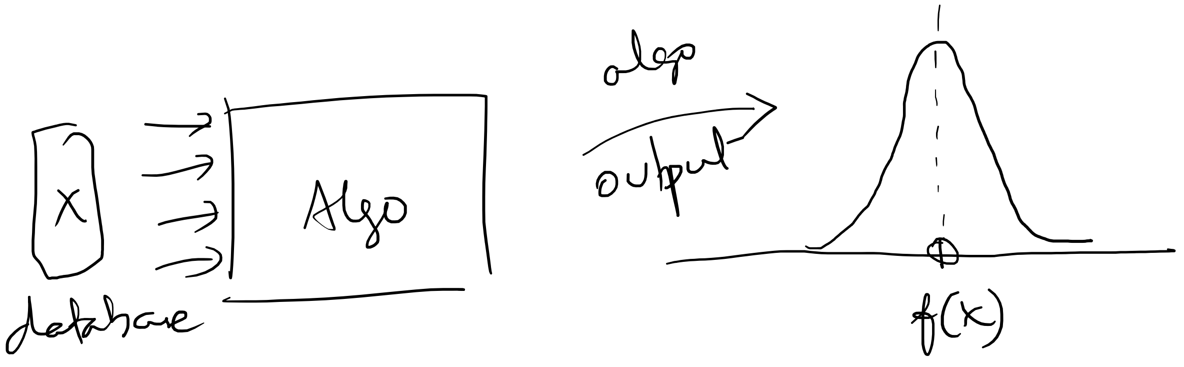
$\tilde{f}(x) \approx \tilde{f}(x')$ for all

d having .

$f(x) \approx f(x)$ for all neighboring datasets.

Model Formally

$\boxed{\text{ALGO}}$ is randomized, adds noise acc. distribution



Distribution of $\tilde{f}(x)$.

Error of Algo

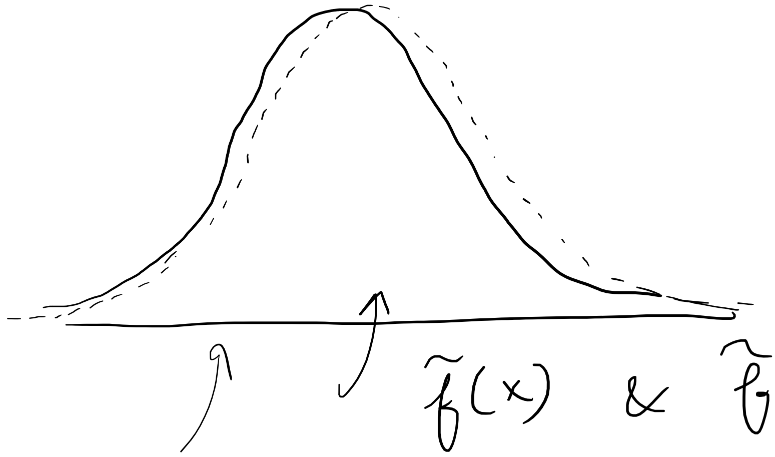
$$\mathop{E}_{\substack{\text{random}\\\text{choices}}}\left[\left(\tilde{f}(x) - f(x)\right)^2\right]$$

Privacy Requirement :-

$\forall$ x, x' differing in one row, want distributions to be nearly identical.

[illegible]

nearly identical.



$\tilde{f}(x)$ & $\tilde{f}(x')$ distribution.

$\forall$ x, x', and for all subsets R of possible outputs of $\tilde{f}$,

$$\boxed{\Pr\left[\tilde{f}(x) \in R\right] \le e^{\varepsilon} \Pr\left[\tilde{f}(x') \in R\right]}$$

like $(1+\varepsilon)$

ϵ - Differential Privacy.

Goal:

[How to answer queries with ϵ - DP, with MIN. ERROR?]

↓

TOMORROW

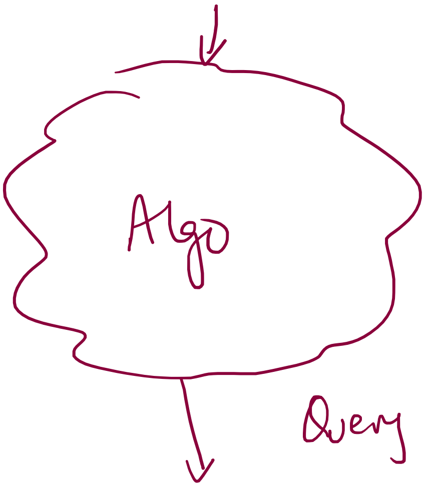
Such a scheme for "counting" queries

Counting Queries

30 April 2021 11:59

Database X

Users	Salary
u_1	s_1
u_2	s_2
⋮	
u_n ⋮	s_n



Query = "Count # users with salary $\geq v$"
$f(x)$

What should a good $\tilde{f}(x)$ be?

Want ① $\tilde{f}(x)$ close to $f(x)$

② $\tilde{f}(x)$ close to $\tilde{f}(x')$ for all neighboring x'.

Idea

- Add noise to real answer.

Compute $f(x)$, output $f(x) + R$

where R is some suitable

where R is some suitable noise.

eg $\left\{ \begin{array}{l} R \sim N(0, \sigma^2) \text{ for suitable } \sigma \\ \text{will give privacy with} \\ \text{low error for suitable parameters} \end{array} \right\}$

<u>Our Algo</u> [Technical Difference]

Add noise z w.p $\propto e^{-\frac{|z|}{\sigma}}$

(in contrast gaussian noise has prob $\propto e^{-z^2/\sigma^2}$)

"LAPLACIAN DISTRIBUTION"

$$f_R(z) = \frac{1}{2\sigma} e^{-\frac{|z|}{\sigma}}$$

<u>Check</u>

① $$\int_{-\infty}^{\infty} f_R(z)\, dz = 1$$

② $$\int_{-\infty}^{\infty} z\, f_R(z)\, dz = E[R] = 0$$

③ $$\int_{-\infty}^{\infty} z^2 f_R(z)\, dz = E[R^2] = 2\sigma^2$$

③ $\int_{-\infty} z^2 f_R(z)\,dz = E[R^2]$

$\downarrow$

eg, Integration by parts

If we add noise acc. Laplacian(σ), What is the squared error like?

$$\tilde{f}(x) = f(x) + R$$

$$\Rightarrow \text{error} = E\left[\left(\tilde{f}(x) - f(x)\right)^2\right]$$

$$= E\left[R^2\right] \qquad \text{where } R \sim \text{Lap}(\sigma).$$

$$= 2\sigma^2$$

Want to set σ sufficiently large to get desired privacy.

$\forall$ x, x' differing in a row, and any subset S of output values

$$\left\{ \text{Want} \quad \Pr\left[\tilde{f}(x) \in S\right] \leq e^{\varepsilon} \Pr\left[\tilde{f}(x') \in S\right] \right.$$

Privacy Requirement.

$\Uparrow$

We'll instead set noise such that

We'll infact set noise such that the pdf of Algo output for X and X' are very similar.

Fix an output value `t`.

Let $f_{Alg}(X,t)$ = PDF of Alg outputting t on input X

$$= \frac{1}{2\sigma} \exp\left(-\frac{|f(X)-t|}{\sigma} \right)$$

Similarly,

$$f_{Alg}(X',t) = \frac{1}{2\sigma} \exp\left(-\frac{|f(X')-t|}{\sigma} \right)$$

$$\Rightarrow \quad \frac{f(X,t)}{f(X',t)} \le e^{-\frac{|f(X)-f(X')|}{\sigma}}$$

$$\le e^{-\frac{1}{\sigma}}$$

So we can set $\sigma = \frac{1}{\varepsilon}$. ☺

$$\forall \underbrace{x,x'}_{\substack{\downarrow \\ \text{neighboring} \\ \text{databases}}},t \quad \boxed{\frac{f(x,t)}{f(x',t)} \le e^{\varepsilon}} \leftarrow \text{Privacy}$$

$$\text{And } E\left[\left(\left(\underset{\substack{\uparrow \\ \text{Utility}}}{\tilde{f}}(x)-f(x)\right)\right)^2\right] \le \frac{2}{\varepsilon^2}.$$

Only thing we used in proof is how much f can change from $x \rightsquigarrow x'$.

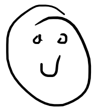
SENSITIVITY of Fn.

$$\Delta_f = \max_{\substack{x,\,x' \\ \text{differing} \\ \text{in one} \\ \text{row}}} |f(x) - f(x')|$$

Noise will simply depend on Δ_f by setting σ appropriately to ensure ϵ-Privacy.

SUMMARY

Simple scheme which works not just for counting queries, but for any low-sensitivity function any low-sensitivity function

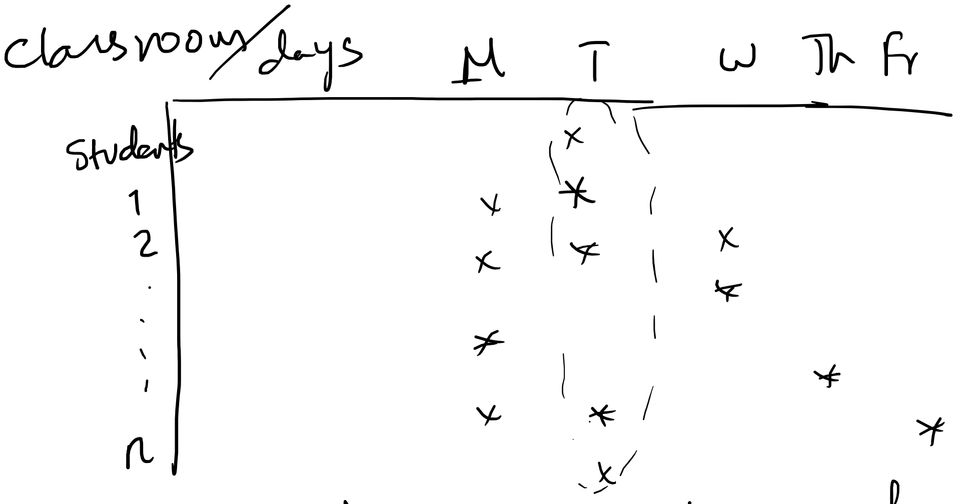


Algo is called "Laplace Mechanism"

Similarly, Gaussian Mechanism also exists (w/ gaussian noise).

In above example, query answer was a numerical value. what if its not?

Example



Each student has a preference for when to conduct exam.

We want to select "Most popular day" in a differentially private manner.

Q1: - Cant simply add noise, (Meaningless)

Q1:- Cant simply add noise, (Meaningless)

Q2:- How to measure utility of a scheme.

Privacy is easy to ~~extend~~

$$\frac{\Pr[\text{Alg selects Monday for } X]}{\Pr[\text{" \quad " \quad " \quad " } X']} \le e^{\varepsilon}$$

Similarly for each other day.

$\boxed{\text{How to get utility?}}$

Want output to be a popular day if not the MOST popular day.

Idea

For each possible output (days in our example)

Day	1	2	3	4	5
compute # people who prefer	n_1	n_2	n_3	n_4	n_5

Output day 'i' as answer with prob $= e^{\varepsilon \cdot n_i}$

$$\overline{\sum e^{\varepsilon n_i}}$$

⇒ Intuitively popular days are more likely to be output.

Privacy + Error Analysis

For any day i and inputs X and X'

$$\frac{\Pr[\text{Alg selects } i \text{ for } X]}{\Pr[\text{Alg selects } i \text{ for } X']}$$

$$= \frac{e^{\varepsilon n_i(x)}}{\sum_j e^{\varepsilon n_j(x)}} \cdot \frac{\sum_j e^{\varepsilon n_j(x')}}{e^{\varepsilon n_i(x)}}$$

$$= \underbrace{e^{\varepsilon\left(n_i(x) - n_j(x')\right)}} \cdot \frac{\sum e^{\varepsilon n_j(x')}}{\sum e^{\varepsilon n_j(x)}}$$

Since X and X' differ in a single row,

each $-1 \le n_i(x) - n_i(x') \le 1$

for all i.

Overall, $$\frac{\Pr[\text{Alg output } i \text{ on } X]}{\Pr[\text{Alg output } i \text{ on } X']} \le e^{\varepsilon} \cdot e^{\varepsilon} = e^{2\varepsilon}.$$

Satisfies 2ε- Differential Privacy.

What about error?

Let database has n people and suppose $n_1 = OPT$ is the day with largest count.

Ideally: Want Alg to output a day with count close to n_1

Let's calculate

$$\Pr[\text{Alg outputs a day with count} \le n_1 - t]$$

Let's fix a day i with count $\le n_1 - t$

$$\Pr[\text{Alg outputs this day}] = \frac{e^{\varepsilon n_i}}{\sum e^{\varepsilon n_j}}$$

Pr [Alg outputs $\qquad\qquad$] $\overline{\qquad\sum e^{\varepsilon n_j}}$

$$\le \frac{e^{\varepsilon(n_1-t)}}{e^{\varepsilon n_1}} \le e^{-\varepsilon t}.$$

So for $t = \frac{\log(n/\delta)}{\varepsilon}$

This probability is $\le \delta/n$

⇒ Union bound over all bad days gives

Pr[outputting a bad day] $\le \delta$

⇒

① with good probability, Algo chooses a day with score $\geq \text{OPT} - \underbrace{\frac{\log(n/\delta)}{\varepsilon}}$

[very useful if OPT is $\gg \frac{\log n}{\varepsilon}$]

AND

② Preserves 2ε- Privacy of users.

EXPONENTIAL MECHANISM.