CS-GY 6763: Lecture 8
A Brief Introduction to Differential Privacy

Lucas Rosenblatt

Congrats on finishing your midterm!

As a reward (haha), you get to hear from me about a new topic: *Differential Privacy* or *DP*.

In this short lecture, we'll:

- Motivate the problem of data privacy
- Introduce basic DP
- Demonstrate that we can use it to protect Kanye's* house in Wyoming from robbers

*(Note: I made these slides before Kanye went off the **deeper** deep end.)

Data privacy problem (informal):

Data utility will eventually be consumed!

As the "*Fundamental Law of Information Recovery*" states:

> *...overly accurate answers to too many questions will destroy privacy in a spectacular way... [Dwork et. al 2014]*

Goal: postpone this inevitability as long as possible

FAMOUS EXAMPLE: GOVERNOR WELD'S PRIVACY BREACH

Mid-1990s: Massachusetts Group Insurance Commission (GIC) releases data w/ every single hospital visit of state employees.

William Weld, then Governor of Massachusetts, says data **protected by deleting patient identifiers.** (name and SSN)

Graduate student Latanya Sweeney (now Dr. Sweeney and famous Prof): *"You're full of it!"*

Figure 1: As it happened.

FAMOUS EXAMPLE: GOVERNOR WELD'S PRIVACY BREACH

To prove him wrong: uses info on Governor Weld's city of residence (Cambridge) which has seven ZIP codes, 54,000 residents.

For 20 bucks, Sweeney buys voter rolls from the city. This database contains the name, address, ZIP code, birth date, and sex of every voter.

In GIC data: only six people in Cambridge shared Weld's birth date, and only three of them men, and of them, only he lived in his ZIP code.

In baller move, Sweeney sends the Governor's health records (which included diagnoses and prescriptions) to his office.

There are many more examples of data privacy breaches.

Solving this problem has generated lots of work over the past two decades, and spawned a “data privacy community”

Examples of influential approaches:

- PII scrubbing (BROKEN)
- k-anonymity [Sweeney 2002]
- l-diversity [Machanavajjhala 2007]
- **Differential Privacy [Dwork 2006]**

In many ways, differential privacy has “won,” with widest adoption (and almost always the strongest guarantees). But there are many open questions even still...

Differentially Private Formalism

NEIGHBORING DATASETS

Consider datasets x and x' as being from the data universe X

We will refer to two datasets as *adjacent* or *neighboring* if they "differ by at most one element."

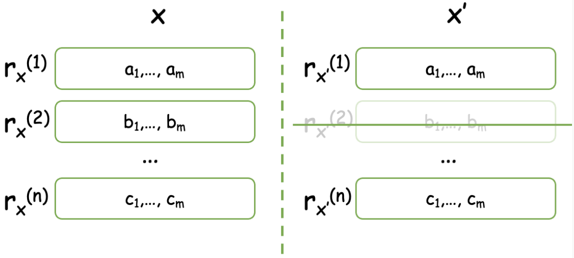


Figure 2: Intuitively, consider two identical datasets where a single row has been removed or added.

(Note: we could be more formal with this definition, but for our purposes this will be fine.)

Differential Privacy [Dwork et. al 2006]

A randomized algorithm $M : \mathbb{N}^{|X|} \rightarrow \mathbb{Y}$ (mapping domain $\mathbb{N}^{|X|}$ to $\mathbb{Y}$) is ϵ-differentially private if for all $S \in \mathbb{Y}$ and every pair of adjacent databases x and x' ($||x - x'||_1 \leq 1$) in $\mathbb{N}^{|X|}$, the following holds:

$$Pr[M(x) \in S] \leq e^{\epsilon} \times Pr[M(x') \in S]$$

Note: Equivalently, we can have this hold over individual output $y \in Y$, and not over sets $S \in Y$:

$$Pr[M(x) = y] \leq e^{\epsilon} \times Pr[M(x') = y]$$

Differential Privacy [Dwork et. al 2006]

$\varepsilon = 0.01$

$$Pr[M(x) \in S] \leq e^{\epsilon} \times Pr[M(x') \in S]$$

$e^{\varepsilon} = (1+\varepsilon)$

The *multiplicative* error approximation e^{ϵ} here is a really strong guarantee

We might have first tried additive error!

If you thought this, that's great - you're halfway to conceptualizing $(\epsilon, \delta) - DP$.

If you're confused as to why we want multiplicative, I'd be happy to discuss after lecture, and I can point you to resources that present more formal motivation.

Our first DP mechanism

Let's consider one of the simplest statistics over a database we might want to answer in a private manner: a counting query.

Counting Query

Consider row x_i in database X, and predicate function $pred(x_i) : x_i \to \{0, 1\}$ (think, abusing notation, of $pred(x_i) = \mathbb{1}[\text{condition over } x_i]$).

We define a counting query $Q_{pred} : X \to \{0, n\}$ as:

$$Q_{pred} = \sum_{x_i \in X} pred(x_i)$$

Example *predicates*?

x_i in zip:

$x_{i\,zip} == 15217$ AND $x_{i\,naught} < 6$

What's so interesting about counting queries?

They have low sensitivity!

(Global) Sensitivity

Consider *any* two neighboring datasets x and x' in X.

The sensitivity ΔQ of a query Q is then,

$$\Delta Q = \max_{x,x'} |Q(x) - Q(x')|$$

This can be stated more generally: for some function $f : x \to \mathbb{R}$ which maps dataset x to a real number,

$$\Delta f = \max_{x,x'} |f(x) - f(x')|$$

Question for you: *What's the sensitivity of a counting query?* =1

Now that we've established a counting query Q, let's consider a method of answering Q in an ϵ-differentially private manner.

Here's a sketch of the algorithm for answering Q privately on dataset x:

- Calculate $Q(x)$.
- Return $Q(x)$ + *noise*

...Is that it?

DP Procedure for Counting Queries

- Calculate $Q(x)$.
- Return $Q(x) + \textit{noise}$

It's not quite it!

We need to add just enough *noise* to get the guarantee from our ϵ-DP definition.

Turns out, the procedure as stated *is* ϵ-differentially private *if* "*noise*" is drawn from the *Laplace distribution*.

Algorithm 1: The Laplace Mechanism for Counting Queries

- Calculate $Q(x)$.
- Return $Q(x) + s \sim \text{Lap}(\frac{\Delta Q}{\epsilon})$

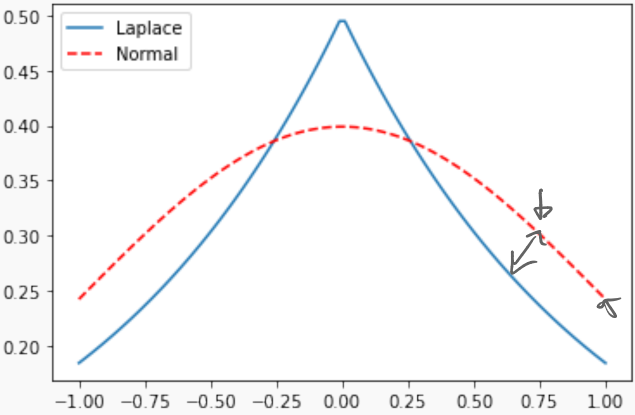
$\frac{\Delta Q}{\varepsilon} \leftarrow 0.01$

We will prove this, but first we'll need a couple tools.

Laplace distribution

The Laplace distribution, centered at 0 with parameter b has variance $2b^2$ and the density function

$$f(x|b) = \frac{e^{\frac{-|x|}{b}}}{2b}$$



The density of the Laplace vs. Normal.

Laplace concentration bound

$$\Pr_{x \sim Lap(b)} [|x| > \alpha b] \leq e^{-\alpha}$$

Useful note: ϵ-DP is: $\frac{Pr[M(x)=y]}{Pr[M(x')=y]} \leq e^{\epsilon}$ ← ε-DP

Algorithm 1: The Laplace DP Mechanism (M_L)

- Calculate $Q(x)$.
- Return $Q(x)$ + $s \sim \text{Lap}(\frac{\Delta Q}{\epsilon})$ ←

Proof. *Algorithm 1, M_L, is ϵ-DP*:

Let x and x' be neighboring datasets in X^n, and let $y \in \mathbb{R}$

Note $b = \frac{\Delta Q}{\epsilon}$.

$Pr[M_L(x, Q) = y] = Pr[Q(x) + s = y]$

$= Pr[s = y - Q(x)]$

$= \frac{1}{2b} e^{\frac{-|y-Q(x)|}{b}}$

Proof.
(cont):

Similarly, $Pr[M_L(x', Q) = y] = \frac{1}{2b} e^{\frac{-|y-Q(x')|}{b}}$ ← Triangle Inequality

Plugging into the definition of DP, we have:

$$\frac{Pr[M_L(x,Q)=y]}{Pr[M_L(x',Q)=y]} = \frac{\frac{1}{2b} e^{\frac{-|y-Q(x)|}{b}}}{\frac{1}{2b} e^{\frac{-|y-Q(x')|}{b}}}$$

Note $\frac{1}{b} = \frac{\varepsilon}{\Delta Q}$

$$= e^{-\frac{1}{b}|y-Q(x)| + \frac{1}{b}|y-Q(x')|}$$

$$= e^{\frac{\varepsilon}{\Delta Q}(|y-Q(x')| - |y-Q(x)|)}$$

$$\leq e^{\frac{\varepsilon}{\Delta Q}\Delta Q} = e^{\varepsilon}$$

$|y-Q(x')| - |y-Q(x)| \leq |y-Q(x') - y + Q(x)|$
$\leq |Q(x) - Q(x')|$
$\leq \Delta Q$

□

Recall that the Laplace distribution has variance $2b^2$.

Proposition. M_L has the following bound on its accuracy (directly from the Laplace concentration bound):

$$Pr\left[|M_L(x,Q) - Q(x)| > \frac{\Delta Q}{\epsilon} \times ln\left(\frac{1}{\rho}\right)\right] \leq \rho$$

I won't prove, but convince yourself of this!

Δf

The “Laplace Mechanism,” or M_L, is one of the most fundamental tools of DP.

We framed it for answering counting queries Q in a DP manner, but it’s incredibly versatile.

I’ll briefly demonstrate that **it can be used to privatize a *mean* over a dataset**.

Problem statement: We want to privately estimate the mean of a binary dataset $z_1, \ldots, z_n$, where $z_i \in \{0, 1\}$.

The true *empirical mean* of the dataset is:

$$\tilde{p} = \frac{1}{n} \sum^{n} z_i$$

How might we get an ϵ-differentially private mean estimate, $\hat{p}$?

Question 1: what is the *sensitivity* of the mean, $\Delta\tilde{p}$? 1/n

Question 2: what would be the *sensitivity* of a mean taken over $z_1, \ldots, z_n \in \mathbb{R}$?

unbounded

Claim: ϵ-DP Mean Estimator for binary data ($\hat{p}$): Given a binary dataset $z_1, ..., z_n$, where $z_i \in \{0, 1\}$, an ϵ-DP mean estimator $\hat{p}$ is given by:

$$\hat{p} = \frac{1}{n}\sum^{n} z_i + Lap\left(\frac{1}{\epsilon n}\right)$$

Exercise for you: using Chebyshev's inequality, show that, with enough data and for reasonably small values of ϵ and w.h.p:

$$|\hat{p} - \tilde{p}| \leq O\left(\frac{1}{\epsilon n}\right)$$

In other words, our ϵ-DP mean estimator is accurate, and gets better as n gets bigger!

Let's take a look at a real life potential application of DP: protecting Kanye from robbers!

Kanye West is reportedly building a 10-bedroom mansion and 2 underground garages on his $14 million Wyoming ranch. Take a look at the sprawling property he bought last year.

Katie Warren Updated Jul 23, 2020, 2:00 PM https://www.businessinsider.com/kanye-west-wyoming-ranch-photos-2019-9

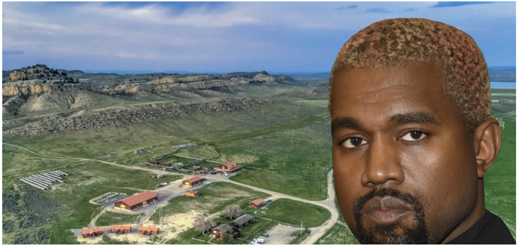


Figure 3: Please don't have said anything crazy today, Kanye.

DEMO: FINDING KANYE'S HOUSE

Note: this is ***real Census data***, but the 2020 U.S. Census actually used DP to privatize results. So we are forced to reinsert Kanye into the zipcodes.

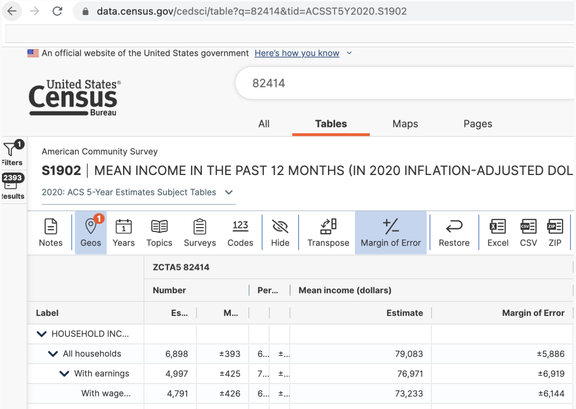


Figure 4: If it were the year 2000, we could find Kanye for sure.