CISC859: Topics in Advanced Networks & Distributed Computing: Network & Distributed System Security

Differential Privacy-2

Review of Laplace mechanism

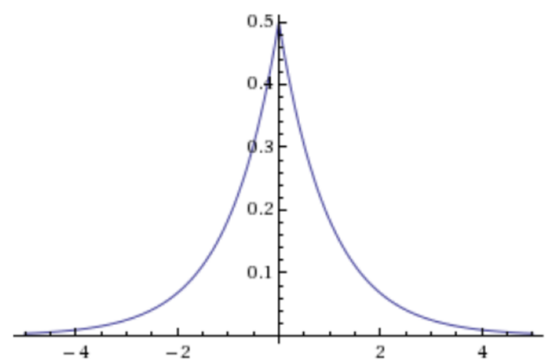
- The Laplace Distribution
- Lap(b) is the probability distribution with p.d.f.:

$$p(x|b) = \frac{1}{2b} \exp(-\frac{|x|}{b})$$

i.e. a symmetric exponential distribution

$Y \sim \text{Lap}(b), \qquad E[|Y|] = b$

$\Pr[|Y| \geq t \cdot b] = e^{-t}$



Answering Numeric Queries: The Laplace Mechanism

Laplace$(D, Q: \mathbb{N}^{|X|} \rightarrow \mathbb{R}^k, \epsilon)$:

1. Let $\Delta = GS(Q)$.
2. For $i = 1$ to k: Let $Y_i \sim \text{Lap}(\frac{\Delta}{\epsilon})$.
3. Output $Q(D) + (Y_1, \ldots, Y_k)$

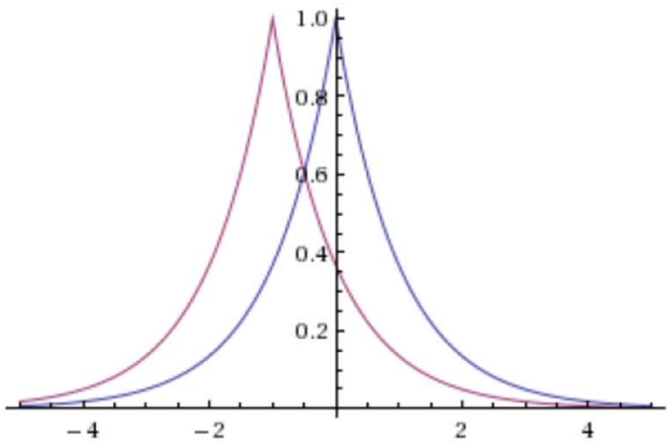
Independently perturb each coordinate of the output with Laplace noise scaled to the sensitivity of the function.

Idea: This should be enough noise to hide the contribution of any single individual, no matter what the database was.

Answering Numeric Queries: The Laplace Mechanism

Laplace$(D, Q\colon \mathbb{N}^{|X|} \rightarrow \mathbb{R}^k, \epsilon)$:

1. Let $\Delta = GS(Q)$.
2. For $i = 1$ to k: Let $Y_i \sim \text{Lap}(\frac{\Delta}{\epsilon})$.
3. Output $Q(D) + (Y_1, \dots, Y_k)$



Why it works

Theorem: The Laplace mechanism is $(\epsilon, 0)$-differentially private.

Proof:

Consider any pair of databases D, D' with $||D - D'||_1 \leq 1$

Consider any event $S \subseteq \mathbb{R}^k$

$$\frac{\Pr[\text{Laplace}(D, Q, \epsilon) \in S]}{\Pr[\text{Laplace}(D', Q, \epsilon) \in S]} = \frac{\int_{x \in S} \Pr[\text{Laplace}(D, Q, \epsilon) = x]}{\int_{x \in S} \Pr[\text{Laplace}(D', Q, \epsilon) = x]}$$

$$\leq \max_{x \in S} \frac{\Pr[\text{Laplace}(D, Q, \epsilon) = x]}{\Pr[\text{Laplace}(D', Q, \epsilon) = x]}$$

Why it works

Theorem: The Laplace mechanism is $(\epsilon, 0)$-differentially private.

Proof: Let $y = \text{Laplace}(D, Q, \epsilon), y' = \text{Laplace}(D', Q, \epsilon)$

$$\frac{\Pr[y = x]}{\Pr[y' = x]} = \prod_{i=1}^{k} \frac{\Pr[y_i = x_i]}{\Pr[y_i' = x_i]} = \prod_{i=1}^{k} \frac{\Pr[Q(D)_i + Y_i = x_i]}{\Pr[Q(D')_i + Y_i = x_i]}$$

$$= \prod_{i=1}^{k} \frac{\Pr[Y_i = x_i - Q(D)_i]}{\Pr[Y_i = x_i - Q(D')_i]} = \prod_{i=1}^{k} \frac{\exp(-\epsilon \frac{|x_i - Q(D)_i|}{\Delta})}{\exp(-\epsilon \frac{|x_i - Q(D')_i|}{\Delta})}$$

$$= \prod_{i=1}^{k} \exp\left(\epsilon \frac{|x_i - Q(D')_i| - |x_i - Q(D)_i|}{\Delta}\right) \le \prod_{i=1}^{k} \exp\left(\epsilon \frac{|Q(D)_i - Q(D')_i|}{\Delta}\right)$$

$$= \exp\left(\frac{\epsilon}{\Delta} \sum_{i=1}^{k} |Q(D)_i - Q(D')_i|\right) \le \exp\left(\frac{\epsilon}{\Delta}\Delta\right) = \exp(\epsilon)\,.$$

Take away message

- Low sensitivity queries can be answered with very little noise! $$\mathsf{E}[\mathrm{Lap}(\frac{1}{\epsilon})] = \frac{1}{\epsilon}$$
- A subset-sum query $Q : \{0,1\}^{|X|} \to \mathbb{R}$ has sensitivity $GS(Q) = 1$
- Any k of them jointly have sensitivity k. So Laplace Mechanism lets you answer any k subset-sum queries with error $o(k/\epsilon)$

Privacy for Non-Numeric Queries

The Exponential Mechanism

Output Perturbation

- We know how to handle (a single) numeric query.
  - "How many people in this room have blue eyes?"
  - Perturb the answer by an amount proportional to the sensitivity of the query.
  - Noise of magnitude $O(\triangle/\epsilon)$ drawn from the Laplace distribution suffices for $(\epsilon, 0)$ differential privacy

When Output Perturbation Doesn't Make Sense

- What about if we have a non-numeric valued query?
  - "What is the most common eye color in this room?"
- What if the perturbed answer isn't almost as good as the exact answer?
  - "Which price would bring the most money from a set of buyers?

Example: Items for sale

Could set the price of apples at $1.00 for profit: $4.00

Could set the price of apples at $4.01 for profit $4.01

Best price: $4.01
2^{nd} best price: $1.00
Profit if you set the price at $4.02: $0
Profit if you set the price at $1.01: $1.01

The Exponential Mechanism

- A mechanism $M : \mathbb{N}^{|X|} \rightarrow R$ for some abstract range R
  - e.g., $R = \{\text{Red,Blue,Green,Brown,Purple}\}$
  - $R = \$1.00, \$1.01, \$1.02, \ldots,$

- Paired with a *quality score:*

$$q : \mathbb{N}^{|X|} \times R \rightarrow \mathbb{R},$$

$q(D, r)$ represents how good output r is for database D

The Exponential Mechanism

- Relative parameters for privacy, solution quality:
  - Sensitivity of q

$$GS(q) = \max_{r \in R, D, D' : ||D - D'||_1 \leq 1} |q(D, r) - q(D', r)|$$

  - Size and structure of R
    - ➔ How many elements of R are high quality? How many are low quality?

The Exponential Mechanism

Exponential$(D, R, q: \mathbb{N}^{|X|} \rightarrow R, \epsilon)$:

1. Let $\Delta = GS(q)$.
2. Output $r \sim R$ with probability proportional to:

$$\Pr[r] \sim \exp\left(\frac{\epsilon q(D,r)}{2\Delta}\right)$$

$$\Pr[r] = \frac{\exp(\frac{\epsilon q(D,r)}{2\Delta})}{\sum_{r' \in R} \exp(\frac{\epsilon q(D,r')}{2\Delta})}$$

The Exponential Mechanism

Exponential$(D, R, q: \mathbb{N}^{|X|} \to R, \epsilon)$:

1. Let $\Delta = GS(q)$.
2. Output $r \sim R$ with probability proportional to:

$$\Pr[r] \sim \exp\left(\frac{\epsilon q(D, r)}{2\Delta}\right)$$

- Idea: Make high quality outputs exponentially more likely at a rate that depends on the sensitivity of the quality score (and the privacy parameter)

Why it works

Theorem: The Exponential Mechanism preserves $(\epsilon, 0)$-differential privacy.

Proof: Fix any $D, D' \in \mathbb{N}^{|X|}$ with $\left||D, D'|\right|_1 \leq 1$ and any $r \in R$...

$$\frac{\Pr[\text{Exponential}(D, R, q, \epsilon) = r]}{\Pr[\text{Exponential}(D', R, q, \epsilon) = r]} =$$

$$\frac{\left(\dfrac{\exp(\frac{\epsilon q(D,r)}{2\Delta})}{\sum \exp(\frac{\epsilon q(D,r')}{2\Delta})}\right)}{\left(\dfrac{\exp(\frac{\epsilon q(D',r)}{2\Delta})}{\sum \exp(\frac{\epsilon q(D',r')}{2\Delta})}\right)} = \left(\frac{\exp(\frac{\epsilon q(D,r)}{2\Delta})}{\exp(\frac{\epsilon q(D',r)}{2\Delta})}\right)\left(\frac{\sum_{r'} \exp(\frac{\epsilon q(D',r')}{2\Delta})}{\sum_{r'} \exp(\frac{\epsilon q(D,r')}{2\Delta})}\right)$$

Why it works

$$\bigstar = \left(\frac{\exp(\frac{\epsilon q(D,r)}{2\Delta})}{\exp(\frac{\epsilon q(D',r)}{2\Delta})} \right) =$$

$$\exp\left(\frac{\epsilon(q(D,r) - q(D',r))}{2\Delta} \right) \leq$$

$$\exp\left(\frac{\epsilon\Delta}{2\Delta} \right) = \exp\left(\frac{\epsilon}{2} \right)$$

Why it works

$$= \left(\frac{\sum_{r'} \exp(\frac{\epsilon q(D',r')}{2\Delta})}{\sum_{r'} \exp(\frac{\epsilon q(D,r')}{2\Delta})} \right) \le$$

$$\left(\frac{\sum_{r'} \exp(\frac{\epsilon (q(D,r') + \Delta)}{2\Delta})}{\sum_{r'} \exp(\frac{\epsilon q(D,r')}{2\Delta})} \right) =$$

$$= \left(\frac{\exp(\frac{\epsilon}{2}) \sum_{r'} \exp(\frac{\epsilon q(D,r')}{2\Delta})}{\sum_{r'} \exp(\frac{\epsilon q(D,r')}{2\Delta})} \right) = \exp(\frac{\epsilon}{2})$$

Why it works

- Recall:

$$\frac{\Pr[\text{Exponential}(D,R,q,\epsilon)=r]}{\Pr[\text{Exponential}(D',R,q,\epsilon)=r]} =$$

$$\leq \exp\left(\frac{\epsilon}{2}\right)\exp\left(\frac{\epsilon}{2}\right)$$

$$= \exp(\epsilon)$$

But is the answer any good?

- It depends…

How good the answer is?

Define:

$$OPT_q(D) = \max_{r \in R} q(D, r)$$

$$R_{OPT} = \{r \in R : q(D, r) = OPT_q(D)\}$$

$$r^* = \text{Exponential}(D, R, q, \epsilon)$$

Theorem:

$$\Pr\left[q(r^*) \leq OPT_q(D) - \frac{2\Delta}{\epsilon}\left(\log\left(\frac{|R|}{|R_{OPT}|}\right) + t\right)\right] \leq e^{-t}$$

How good the answer is?

Theorem:

$$\Pr\left[q(r^*) \leq OPT_q(D) - \frac{2\Delta}{\epsilon}\left(\log\left(\frac{|R|}{|R_{OPT}|}\right) + t\right)\right] \leq e^{-t}$$

Corollary:

$$\Pr\left[q(r^*) \leq OPT_q(D) - \frac{2\Delta}{\epsilon}(\log(|R|) + t)\right] \leq e^{-t}$$

Proof:

$|R_{OPT}| \geq 1$ by definition.

How good the answer is?

Theorem:

$$\Pr\left[q(r^*) \le OPT_q(D) - \frac{2\Delta}{\epsilon}\left(\log\left(\frac{|R|}{|R_{OPT}|}\right) + t\right)\right] \le e^{-t}$$

Corollary:

$$\mathrm{E}[q(r^*)] \ge OPT_q(D) \; - \frac{2\Delta}{\epsilon}\Big(\log(|R|) + \log\big(OPT_q(D)\big)\Big) - 1$$

Proof:

$$\Pr\left[q(r^*) \le OPT_q(D) - \frac{2\Delta}{\epsilon}\Big(\log(|R|) + \log(OPT_q(D))\Big)\right] \le \frac{1}{OPT_q(D)}$$

$$\Pr\left[q(r^*) \ge OPT_q(D) - \frac{2\Delta}{\epsilon}\Big(\log(|R|) + \log(OPT_q(D))\Big)\right] \ge 1 - \frac{1}{OPT_q(D)}$$

How good the answer is?

Theorem:

$$\Pr\left[q(r^*) \le OPT_q(D) - \frac{2\Delta}{\epsilon}\left(\log\left(\frac{|R|}{|R_{OPT}|}\right) + t\right)\right] \le e^{-t}$$

Proof:

$$\Pr[q(r^*) \le x] \le \frac{\Pr[q(r^*) \le x]}{\Pr\left[q(r^*) = OPT_q(D)\right]}$$

$$\le \frac{|\mathrm{R}|\exp(\frac{\epsilon x}{2\Delta})}{|\mathrm{R_{OPT}}|\exp(\frac{\epsilon OPT q(D)}{2\Delta})}$$

$$= \frac{|R|}{|R_{OPT}|}\exp\left(\frac{\epsilon\left(x - OPT_q(D)\right)}{2\Delta}\right) = \left(\frac{|R|}{|R_{OPT}|}\right)\exp\left(-\log\left(\frac{|R|}{|R_{OPT}|}\right) - t\right)$$

$$= \left(\frac{|R|}{|R_{OPT}|}\right)\left(\frac{|R_{OPT}|}{|R|}\right)e^{-t} = e^{-t}$$

Example

- So if $R = \{\text{Red,Blue,Green,Brown,Purple}\}$ then we can answer "What is the most common eye color in this room?" with a color that is shared by:

$$OPT - \frac{2}{\epsilon}(\log 5 + 3) < OPT - \frac{7.4}{\epsilon}\text{people}$$

  – Except with probability: $\leq e^{-3} < 0.05$

- *Independent* of the number of people in the room. Very small error if n is large.

Remark

- The exponential mechanism is based on the vector:

$$\hat{q} : \mathbb{N}^{|X|} \to |R| = \left(q(D, r_1), q(D, r_2), \dots, q\left(D, r_{|R|}\right)\right)$$

  - Might have sensitivity $GS(\hat{q}) = |R| \cdot GS(q)$
  - Exponential Mechanism only depends on $GS(q)$

- Error has only logarithmic dependence on $|R|$.
  - Could take exponentially large ranges!
  - But *sampling* from the exponential mechanism efficiently is non-trivial.