Privacy of Numeric Queries Via Simple Value Perturbation

The Laplace Mechanism

Differential Privacy
A Basic Model

- Let X represent an abstract data universe and D be a multi-set of elements from X.
  - i.e. D can contain multiple copies of an element $x \in X$.
- Convenient to represent D as a *histogram:*

$$D \in \mathbb{N}^{|X|}$$

$$D[i] = |\{x \in D : x = x_i\}|$$

Differential Privacy
A Basic Model

- i.e for a database of heights
  - $D = \{5'2, 6'1, 5'8, 5'8, 6'0\} \subset [4-8]$
  - $D = (\ldots, \underbrace{1}_{5'2},0,0,0,0,0,\underbrace{2}_{5'8},0,0,0,\underbrace{1}_{6'0},\underbrace{1}_{6'1},0, \ldots) \in \mathbb{R}^{48}$

Differential Privacy
A Basic Model

- The *size* of a database n:
  - As a set: $n = |D|$.
  - As a histogram: $n = \left||D|\right|_1 = \sum_{i=1}^{|X|} |D[i]|$

Definition: ℓ_1 (Manhattan) Distance.

For $\hat{v} \in \mathbb{R}^d$, $\left||\hat{v}|\right|_1 = \sum_{i=1}^{d} |\hat{v}_i|$.

Differential Privacy
A Basic Model

- The *distance* between two databases:
  - As a set: $|D \Delta D'|$.
  - As a histogram: $\left\|D - D'\right\|_1$

Differential Privacy
A Basic Model

- i.e for a database of heights
  - $D = \{5'2, 6'1, 5'8, 5'8, 6'0\} \subset [4-8]$
  - $D = (\dots, \underbrace{1}_{5'2},0,0,0,0,0,\underbrace{2}_{5'8},0,0,0,\underbrace{1}_{6'0},\underbrace{1}_{6'1},0, \dots) \in \mathbb{R}^{48}$
  - $D' = (\dots, 2,1,0,0,0,0,1,0,0,0,1,1,0, \dots) \in \mathbb{R}^{48}$

$$||D||_1 = |1| + |2| + |1| + |1| = 5$$
$$||D'||_1 = |2| + |1| + |1| + |1| + |1| = 6$$
$$||D - D'||_1 = |-1| + |-1| + |1| = 3$$

Basic Lower Bound: Blatant Non-Privacy

- How much noise is necessary to guarantee privacy?
- A simple model.
  - For simplicity, $D \in \{0,1\}^{|X|}$ (i.e. no repeated elts)
  - A query is a bit vector $Q \in \{0,1\}^{|X|}$
  - $Q(D) = \langle Q, D \rangle = \sum_{i:Q[i]=1} D[i]$
    - A "subset sum query"
    - For $S \subseteq [n]$ write Q_S for the vector:

$$Q_S[i] = \begin{cases} 1, & i \in S \\ 0, & i \notin S \end{cases}$$

Basic Lower Bound: Blatant Non-Privacy

Definition: A mechanism $M: \{0,1\}^n \to R$ is *blatantly non-private* if on any database D, an adversary can use $\mathrm{y} = M(D)$ to reconstruct $D' = A(y)$ that agrees with D on all but $o(n)$ entries:

$$\left||D - D'|\right|_1 \in o(n)$$

Basic Lower Bound: Blatant Non-Privacy

Answering all subset-sum queries requires linear noise.

Definition: A mechanism $M: \{0,1\}^{|X|} \rightarrow R$ adds noise bounded by α if for every $D \in \{0,1\}^{|X|}$ and for every query $S \subseteq [n]$: $M(D) = y$ such that

$$|Q_S(D) - Q_S(y)| \leq \alpha$$

Basic Lower Bound: Blatant Non-Privacy

Theorem: Let M be a mechanism that adds noise bounded by α. Then there exists an adversary that given $M(D)$ can construct a database D' such that $\left|\left|D - D'\right|\right|_0 \leq 4\alpha$

- So adding noise $o(n)$ leads to blatant non-privacy ☹

Basic Lower Bound: Blatant Non-Privacy

Proof: Consider the following adversary.

- Let $r = M(D)$
- For each $D' \in \{0,1\}^{|X|}$
  - If $|Q_S(D') - Q_S(r)| \leq \alpha$ for all $S \subseteq X$ then:
    - Output D'

Claim 1: *The algorithm always outputs some D'.*

Yes: $D' = D$ passes all tests.

Claim 2: $\left|\left|\mathrm{D}' - \mathrm{D}\right|\right|_0 \leq 4\alpha$

Let $S0 = \{x \in X : x \in D', x \notin D\}$

Let $S1 = \{x \in X : x \in D, x \notin D'\}$

Observe $\left|\left|D' - D\right|\right|_1 = |S0| + |S1|$

So: If $\left|\left|\mathrm{D}' - \mathrm{D}\right|\right|_1 > 4\alpha$ then $\max(|S0|, |S1|) > 2\alpha$. WLOG assume $|S0| > 2\alpha$.

We know $Q_{S0}(D) = 0$, so by accuracy: $Q_{S0}(r) \leq \alpha$.

But $Q_{S0}(D') > 2\alpha$, so it must be: $|Q_{S0}(D') - Q_{S0}(r)| > |2\alpha - \alpha| = \alpha$

So it would have failed one of the tests…

Basic Lower Bound: Blatant Non-Privacy

- Bad news!
  - Accuracy $n/2$ is trivial.
  - Accuracy $n/40$ already lets an adversary reconstruct 9/10ths of the database entries!
- But that attack required answering all possible queries…
  - Guiding lower bound: Going forward, we will only try to be accurate for restricted classes of queries.

Differential Privacy
A Basic Model

Definition: A randomized algorithm with domain $\mathbb{N}^{|X|}$ and range R

$$M: \mathbb{N}^{|X|} \to R$$

is (ϵ, δ)-*differentially private* if:

1) For all pairs of databases $D, D' \in \mathbb{N}^{|X|}$ such that $\left|\left|D - D'\right|\right|_1 \leq 1$ and,

Differing in 1 person's data

2) For all *events* $S \subseteq R$:

$$\Pr[M(D) \in S] \leq e^{\epsilon} \Pr[M(D') \in S] + \delta.$$

Private algorithms *must* be randomized

Resilience to Post Processing

Proposition: Let $M: \mathbb{N}^{|X|} \rightarrow R$ be (ϵ, δ)-differentially private and let $f: R \rightarrow R'$ be an arbitrary function. Then:

$$f \circ M: \mathbb{N}^{|X|} \rightarrow R'$$

is (ϵ, δ)-differentially private.

Thinking about the output of M can't make it less private.

Resilience to Post Processing

Proof:

1) Consider any pair of databases $D, D' \in \mathbb{N}^{|X|}$ with $\left\| D - D' \right\|_1 \leq 1$.

2) Consider any event $S \subseteq R'$.

3) Let $T \subseteq R$ be defined as $T = \{r \in R : f(r) \in S\}$.

Now:

$$\begin{aligned}
\Pr\left[f\left(M(D)\right) \in S\right] &= \Pr[M(D) \in T] \\
&\leq e^{\epsilon} \Pr[M(D') \in T] + \delta \\
&= e^{\epsilon} \Pr\left[f\left(M(D)\right) \in S\right] + \delta
\end{aligned}$$

Randomized mappings f are just convex combinations of functions.

Resilience to Post Processing

Take away message:

1) f as the adversaries analysis: can incorporate arbitrary auxiliary information the adversary may have. Privacy guarantee holds no matter what he does.
2) f as our algorithm: If we access the database in a differentially private way, we don't have to worry about how our algorithm post-processes the result. *We only have to worry about the data access steps*.

Answering Numeric Queries

- Suppose we have some numeric *question* about the private database that we want to know the answer to:

$$Q: \mathbb{N}^{|X|} \rightarrow \mathbb{R}^k. \qquad Q(D) = ?$$

- How do we do it privately?
  - How much noise do we have to add?
  - What are the relevant properties of Q?

Answering Numeric Queries

Definition: The ℓ_1-sensitivity of a query $Q: \mathbb{N}^{|X|} \rightarrow \mathbb{R}^k$ is:

$$GS(Q) = \max_{D,D' : \left|\left|D - D'\right|\right|_1 \leq 1} \left|\left|Q(D) - Q(D')\right|\right|_1$$

i.e. how much can 1 person affect the value of the query?

"*How many people in this room have brown eyes*": Sensitivity 1

"*How many have brown eyes, how many have blue eyes, how many have green eyes, and how many have red eyes*": Sensitivity 1

"*How many have brown eyes and how many are taller than 6'*": Sensitivity 2

Answering Numeric Queries

The Laplace Distribution:

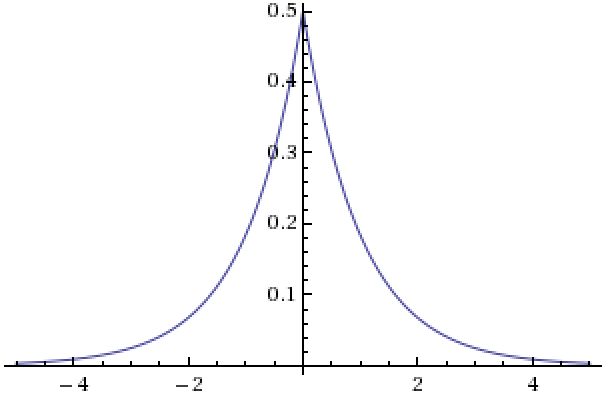
$\mathrm{Lap}(b)$ is the probability distribution with p.d.f.:

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

i.e. a symmetric exponential distribution

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

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



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 $Y_i \sim \mathrm{Lap}(\frac{\Delta}{\epsilon})$.
3. Output $Q(D) + (Y_1, \dots, Y_k)$

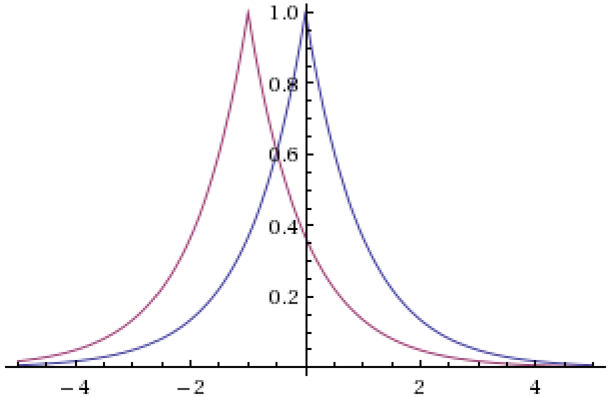
Independently perturb each coordinate of the output with Laplace noise scaled to the sensitivity of the function.

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

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



To Ponder

- Where is there room for improvement?
  - The Laplace mechanism adds *independent* noise to every coordinate…
  - What happens if the user asks (essentially) the same question in every coordinate?
  - Read [Dinur,Nissim03]: a computationally efficient attack that gives blatant non-privacy for a mechanism that adds noise bounded by $o(\sqrt{n})$.