Today: Beating Global Sensitivity

- Local notions of sensitivity
- Propose-Test-Release
- Gaussian Mean Estimation

Local Notions of Sensitivity

Def Global sensitivity of $f: \mathcal{X}^n \to \mathbb{R}$ is $\Delta_f = \max_{x \sim x'} |f(x) - f(x')|$

Privacy guarantees must be worst-case

In many cases, global sensitivity is large because of pathological or atypical instances.

Example Median of data in $[0, R]$,

$\Delta_{med} = R$, only happens when half/half 0 & R

Example Mean of $x_1, \dots, x_n \in \mathbb{R}$.

$\Delta_{mean} = \infty$ b/c outliers

But if $x_1, \dots, x_n, x_{n+1} \overset{iid}{\sim} \mathcal{N}(\mu, \sigma^2)$ then

$$\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) - \left(\frac{1}{n}\sum_{i=2}^{n+1} x_i\right) = \frac{1}{n}(x_1 - x_{n+1})$$

$$\sim \mathcal{N}\left(0, \frac{2\sigma^2}{n}\right)$$

Want to take advantage of when data is not worst case.

Def The **local sensitivity of f at x** is

$$LS_f(x) \triangleq \max_{\substack{x' \\ \text{s.t. } x' \sim x}} |f(x) - f(x')|.$$

Q: Does $f(x) + Lap\left(\frac{LS_f(x)}{\varepsilon}\right)$ give DP?

A: No! (median, 2 points on $\frac{R}{2}$)

Def The **down sensitivity of f at x** is

$$DS_f(x) \triangleq \max_{\substack{x' \\ \text{s.t. } x' \subseteq x \\ |x'| = n-1}} |f(x) - f(x')|$$

Ex Mean of Gaussian data has small down-sensitivity, $LS = \infty$

Instance optimality, b/c for any pair $x \sim x'$ definitely must hide difference.

I know four techniques for beating global sensitivity

1. Inverse sensitivity mechanism
2. Privately estimate $LS_f(x)$
3. Smooth sensitivity

Skipped today

4. Propose-test-Release

If someone asks

$$SS_f^\varepsilon(x) = \max_{x'} LS_f(x') \cdot e^{-\varepsilon d(x,x')}$$

Can do: $f(x) + \frac{SS_f^\varepsilon(x)}{\varepsilon} \cdot Z$ for ε-DP ← Cauchy

$f(x) + \frac{SS_f^\varepsilon(x)}{\varepsilon} \cdot ZeY$ gives (ε, δ)-DP

σ scale

Z: Std Laplace

Y: Std normal

Propose-Test-Release

Simplest form: "guess" that local sensitivity is small on x & all its neighbors.

We'll see more general views as a wrapper around a "goal" algorithm.

Not obviously implementable on a computer, but many efficient examples.

Algorithm Propose-Test-Release

Input: $\varepsilon, \delta \in (0,1)$, $x \in \mathcal{X}^n$, $A : \mathcal{X}^n \to \mathcal{Y}$

Output: $\tilde{y} \in \mathcal{Y}$ or "FAIL"

1) $SAFE = \{x' \in \mathcal{X}^n : \forall x'' \sim x',\ A(x') \approx_{(\varepsilon,\delta)} A(x'')\}$
2) $UNSAFE = \mathcal{X}^n \setminus SAFE$
3) $k \leftarrow \min_{x' \in UNSAFE} d(x, x')$
4) If $\tilde{k} = k + Lap(1/\varepsilon) \leq \frac{\log(1/\delta)}{\varepsilon} + 1$
5) Return "FAIL"
6) Else return $\tilde{y} = A(x)$

(Privacy proof)

Gaussian mean in d dimensions

Assume bound on R

Iteratively refine it

For $x \sim \mathcal{N}(a, I)$, wp $1-\beta$

$$\|x - u\|_2 \leq \sqrt{d} + \sqrt{\log 2/\beta}$$