Intro to DP Stats

- Overview of Goals & Techniques
- Subsample & Aggregate
- Friendly Core

Statistics with DP

As always, privacy is worst-case

Accuracy/utility with assumptions

- eg, $\|x_i\|_2 \leq B$
- eg, $x_i \sim \mathcal{N}(\mu, \Sigma)$
- assumptions of some sort are necessary

Types of goals

- Parameter estimation

  ex/ given $x_1, \dots, x_n \sim N(\mu, \Pi)$,
  find $\tilde{\mu}$ with $\|\tilde{\mu} - \mu\|_2 \leq \alpha$

- distribution learning

  ex/ given $x_1, \dots, x_n \sim P$, find $\tilde{P}$
  such that $TV(P, \tilde{P}) \leq \alpha$

- Hypothesis testing

  ex/ decide if $x_1, \dots, x_n$ come from
  P or Q

- prediction

  ex/ Linear regression, excess risk
  find $\tilde{\theta}$ such that

$$\mathop{\mathbb{E}}_{x,y} \left(\langle \tilde{\theta}, x \rangle - y \right)^2 - \min_{\theta'} \mathop{\mathbb{E}}_{x,y} \left(\langle \theta', x \rangle - y \right)^2 \leq \alpha$$

Some Important Tools

- Bound global sensitivity
- DP optimization
- subsample-and-aggregate (today)
- Robustness-to-privacy transformation
- "Stable, data-dependent noise"
  - ↳ not a real name

We'll also see lower bounds techniques

- fingerprinting
- Packing
- group privacy

These often lead to lower bounds for optimization as well.

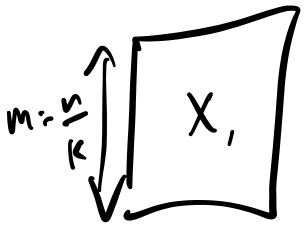
Subsample & Aggregate

Have statistic we want to privatize

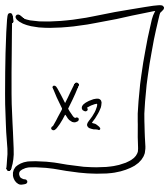
$$f: X^+ \to \mathcal{Y}$$

Dataset $X = x_1, \ldots, x_n$

Partition into k buckets

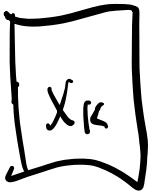


$f(X_1) = y_1$



$f(X_2) = y_2$

⋮



$f(X_k) = y_k$

Changing one point can only change one entry

Example Hypothesis Testing

Null $H_0: x_1, \dots, x_n \sim P$

Alternate $H_1: x_1, \dots, x_n \sim Q$

Non-private $f: X^* \to \{0, 1\}$ (null, alt.)

Get bucket estimates $y_1, \dots, y_K$

Aggregate with exponential mechanism:

$$\Pr[A(x) = 0] = \frac{\exp\left(\frac{\varepsilon}{2} \#\{y_i = 0\}\right)}{\exp\left(\frac{\varepsilon}{2} \#\{y_i = 0\}\right) + \exp\left(\frac{\varepsilon}{2} \#\{y_i = 1\}\right)}$$

Privacy: ε-DP because count is 1-sensitive

Utility: Assume null hypothesis.

$$\Pr[A(x) \text{ is wrong}] = \Pr[A(x) = 1]$$

$$\leq \frac{\Pr[A(x) = 1]}{\Pr[A(x) = 0]}$$

$$= \exp\left(\frac{\varepsilon}{2}\left(\#\{y_i = 1\} - \#\{y_i = 0\}\right)\right)$$

$$= \exp\left(\frac{\varepsilon}{2}\left(k - 2 \cdot \#\{y_i = 0\}\right)\right)$$

$$= \exp\left(\frac{k\varepsilon}{2}\left(1 - \frac{2 \cdot \#\{y_i = 0\}}{k}\right)\right)$$

small whenever $k \geq \frac{1}{\varepsilon}$ and

$$\#\{y_i = 0\} \gg \frac{k}{2}.$$

ie, whenever base test is reliable on inputs of size $\frac{n}{k} \approx n\varepsilon$

Example For statistics $f(X) \in \mathbb{R}$, can apply a DP median algorithm similar to "median of means."

Q: in high dimensions, can use mean estimation.

Friendly Core

Tsfadia, Cohen, Kaplan, Mansour, Stemmer 2022

"general purpose" DP aggregator, conceptually useful

Idea: check that most bucket estimates are similar/agree with each other, if so can do private aggregation.

Given: predicate $f: X \times X \to \{0,1\}$

Def (1) Two points x, y are **f-friends** if $f(x,y) = 1$

(2) A dataset $X = (x_1, \ldots, x_n)$ is **f-friendly** if $\forall\, x_i, x_j$. exists z (not necessarily in X) such that $f(x_i, z) = f(x_j, z) = 1$

(3) An algorithm A satisfies **f-friendly (ε,δ)-DP** if $\forall\, X \sim X'$ such that $X \cup X'$ is f-friendly, we have $A(X) \approx_{(\varepsilon,\delta)} A(X')$

Example Predicate $f(x,y) = \mathbb{1}\{\|x-y\|_2 \le R\}$

Useful for mean estimation.

Example Clustering: non-private algorithm produces m-tuple $y = (y_1, \dots, y_m)$ of cluster centers. Then y, y' are friends if we can reorder y' so that all pairs of centers are close.

Alg Basic Filter

Input: dataset $x_1, \dots, x_n$, predicate f

Return: weights $w_1, \dots, w_n \in [0,1]$

(1) For $i = 1, \dots, n$

(2) Count friends $c_i \leftarrow \sum_{j=1}^{n} f(x_i, x_j)$

(3) w_i gets $\begin{cases} 0 & \text{if } c_i \le n/2 \\ 1 & \text{if } c_i = n \\ \frac{c_i - n/2}{n} & \text{otherwise} \end{cases}$

(4) End For

Output is a friendly dataset!

Changing one input doesn't change output weights.

Claim Suppose distributions P & Q are supported in an ℓ_2 ball of radius R. Then

$$\|u(P) - u(Q)\|_2 \le 2R \cdot TV(P,Q)$$

Proof Rescale so $\text{supp}(P), \text{supp}(Q) \subseteq B(0,R)$.

Then $$\left\| \mathop{\mathbb{E}}_{x\sim P}[x] - \mathop{\mathbb{E}}_{y\sim Q}[y] \right\|_2 = \left\| \sum_x (x p(x) - x q(x)) \right\|_2$$

$$\le \sum_x \|x\|_2 \cdot |p(x) - q(x)|$$

$$\le R \cdot \|p - Q\|_1 = 2R\, TV(P,Q) \quad \boxtimes$$

(Analysis for $\mathcal{U}(u, \Pi)$, if time)