- Scribes?
- lecture recording

- Secure Computation (how?): given a functionality $\mathcal{F}$ to compute, how to do it securely?
- privacy (what?): what functionalities $\mathcal{F}$ are "safe" to compute in the first place?

Given database D, want to answer query q on D

will give an approximate/noisy answer $q'(D)$.

Informally: releasing $q'(D)$ is private if the answer would be "roughly the same" whether or not a particular user's data was in D

Differential Privacy

let $D = (x_1, \ldots, x_n) \in X^n$ (x_i is data of user i)

D, D' are neighboring if they differ in data of one user.

Mechanism $M: X^n \to Y$ is ε-diff. private if for all neighboring D, D' and all $T \subseteq Y$,

$$\Pr[M(D) \in T] \le e^{\varepsilon} \cdot \Pr[M(D') \in T]$$

M is (ε, δ)-diff. private if for D, D', T as above

$$\Pr[M(D) \in T] \leq e^{\varepsilon} \cdot \Pr[M(D') \in T] + \delta$$

Note: $\varepsilon = \Omega(1/n)$, δ can be cryptographically small

need to look at privacy/utility tradeoff

Composition

- If M is ε-diff private & D, D' differ in data of k users, then for any $T \subseteq Y$,

$$\Pr[M(D) \in T] \leq e^{k \cdot \varepsilon} \cdot \Pr[M(D') \in T]$$

- If $M_1, \ldots, M_\ell$ are ε-diff. private,

  then $(M_1 \times \cdots \times M_\ell)$ is $\ell \cdot \varepsilon$-diff. private

  In fact, if $\ell \leq 1/\varepsilon^2$, then for any $\delta > 0$

  $(M_1 \times \cdots \times M_\ell)$ is $O((\ell \log 1/\delta)^{1/2} \cdot \varepsilon, \delta)$-diff. private

---

Laplace mechanism

For a query $q : X^n \to \mathbb{R}$, define global sensitivity

$$GS_q = \max_{D \sim D'} |q(D) - q(D')|$$

Idea: answer query q by returning $q(D)$ + noise, where noise depends on GS_q & ε.

what distribution to use?

$\mathrm{Lap}(\sigma)$: $\Pr[z] \propto e^{-|z|/\sigma}$

mean 0, std. dev $\sigma \cdot \sqrt{2}$

$\Pr[\mathrm{Lap}(\sigma) > \sigma \cdot t] \leq e^{-t}$

Laplace mechanism: return $q(D) + \mathrm{Lap}\left(\frac{GS_q}{\varepsilon}\right)$

Theorem: This is ε-diff. private

Proof: Fix $D \sim D'$, t.

$$\frac{\Pr[M(D) = t]}{\Pr[M(D') = t]} = \frac{\Pr[\mathrm{Lap}(GS/\varepsilon) = t - q(D)]}{\Pr[\mathrm{Lap}(GS/\varepsilon) = t - q(D')]}$$

$$= \frac{e^{-\varepsilon \cdot |t - q(D)|/GS}}{e^{-\varepsilon \cdot |t - q(D')|/GS}} \leq e^{\varepsilon} \quad \square$$

Utility? $\Pr\left[|M(D) - q(D)| > \frac{GS}{\varepsilon} \cdot \log 1/\beta\right] \leq \beta$

Exponential mechanism

Abstract mechanism based on scoring function $\mathrm{score}(D, y)$

Let $GS = \max_y \max_{D \sim D'} |\mathrm{score}(D, y) - \mathrm{score}(D', y)|$.

Mechanism: On input D, output y w/ prob. proportional to $e^{\varepsilon \cdot \mathrm{score}(D,y)/2 \cdot GS}$.

This is ε-diff. private for any scoring function

Utility: w/ prob. $O(1)$, the output y satisfies

$$\mathrm{score}(D, y) \geq \max_{y^*} \mathrm{score}(D, y^*) - O\left(\frac{GS \log |Y|}{\varepsilon}\right)$$