Privacy Enhancing Technologies

Differential Privacy

What was the "issue" so far?

If we release too many answers with bounded noise we leak a lot of information (namely all inconsistent DBs)

Relax this condition: noise "small" w.h.p

Randomized Response (Warner '65)

Setting: $D = \{x_1, \ldots, x_n\}$

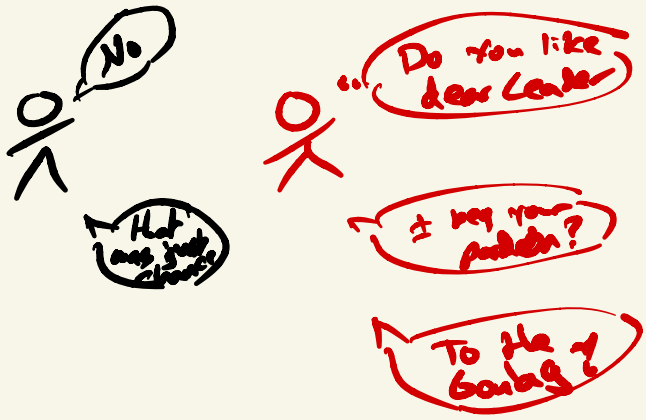
we want to learn $\sum_{i=1}^{n} x_i$ (we don't even trust a curator)

Protocol: before reporting your bit, flip a biased coin and answer

$$y_i = \begin{cases} x_i & \text{w.p } 1/2+\gamma \\ 1-x_i & \text{w.p } 1/2-\gamma \end{cases}$$

for $\gamma \in [0, 1/2]$

"Privacy argument": plausible deniability



Utility

- What we want: $p = \frac{1}{n}\sum x_i$

- The noisy answer is biased:

$$\mathbb{E}\left[\frac{1}{n}\sum y_i\right] = \frac{1}{n}\sum \mathbb{E}[y_i]$$

$$= \frac{1}{n}\sum\left(\left(\tfrac{1}{2}+\gamma\right)x_i + \left(\tfrac{1}{2}-\gamma\right)(1-x_i)\right)$$

$$= \frac{1}{n}\sum\left[2\gamma x_i + (1/2-\gamma)\right]$$

$$= 2\gamma \cdot p + 1/2 - \gamma$$

$$= C_1 \cdot p + C_2$$

$\Rightarrow \hat{p} = \left(\frac{1}{n}\sum y_i - 1/2 + \gamma\right) \cdot \frac{1}{2\gamma}$

is an unbiased estimate of p!

$$\mathbb{E}[\hat{p}] = p$$

$\Rightarrow$ what is $\mathrm{Var}[\hat{p}]$?

$n\hat{p}$ is a sum of "rescaled" Bernoulli R.V

Asymptotically: $|p - \hat{p}| \leq \tilde{O}\left(\frac{1}{\gamma\sqrt{n}}\right)$

with high probability $(1 - o(1))$

$\Rightarrow$ if we want to estimate p with additive error α, then we need $n = \tilde{O}\left(\frac{1}{\gamma^2\alpha^2}\right)$



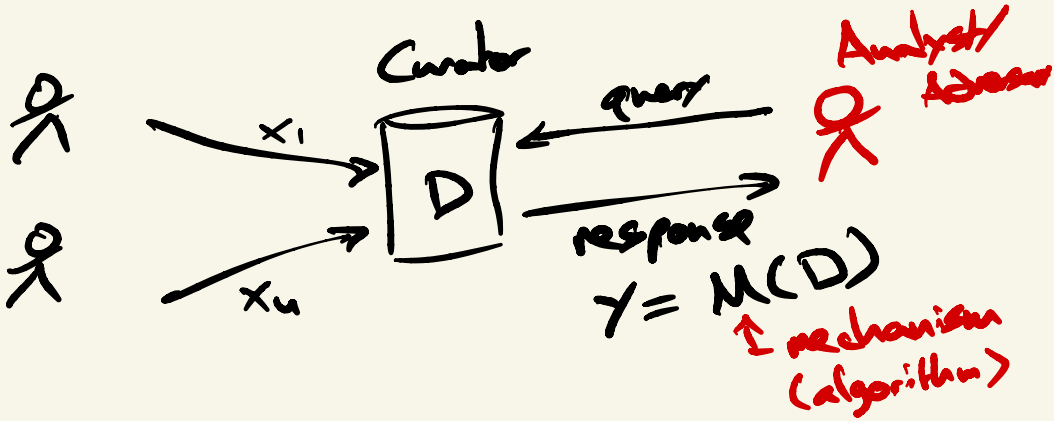
assume some "population" has property p

(formally, there is a distribution D, $\mathbb{E}[x]$ for $x \sim D$ is p)

if we sample $x_1, \ldots, x_n \overset{iid}{\sim} D$ and report the sample mean $\hat{p} = \frac{1}{n}\sum x_i$

then $|p - \hat{p}| \leq \tilde{O}(1/\sqrt{n})$ w.h.p

Central model of differential privacy

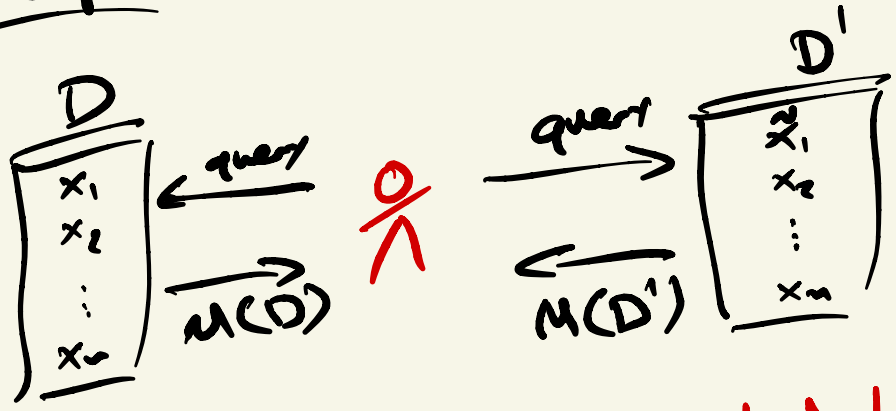


What would "plausible deniability" mean in this setting?

Attempt 1: "the Adv learns nothing from interacting with the DB" (the query results can be simulated without D)

⚠ : impossible if we want any utility

Attempt 2: (DP, DNNS06)



DP: the adv. cannot distinguish between D and D'

The guarantee (informally): whether you are in the DB or not doesn't change the query answer "too much"

Def: Neighboring datasets: D and D' are neighboring, $D \sim D'$, if $|D| = |D'|$ and they differ in at most one row

Def (DP)

A randomized algorithm (mechanism) $M: X^n \to Y$ is ε-differentially private if

$$\forall D \sim D' \in X^n$$

and all events $S \subseteq Y$:

$$\Pr[M(D) \in S] \leq e^{\varepsilon} \cdot \Pr[M(D') \in S]$$

} check Carol's notes

Randomized Response is (locally) DP

$$M\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \to \begin{matrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{matrix}$$

M is the "mechanism" that produces all local random bits

We show: M is ε-DP

As a consequence (post-processing)
any function applied to $Y_1, \ldots, Y_n$
is ε-DP

In particular, our estimate $\hat{p}$ is ε-DP

Let D, D' to be databases that differ only
in the value of x_j for some j

Let $\vec{b} \in \{0,1\}^n$ be an arbitrary bit vector

we want: $\Pr[M(D) = \vec{b}] \leq e^{\varepsilon} \Pr[M(D') = \vec{b}]$

$$\frac{\Pr[M(D) = \vec{b}]}{\Pr[M(D') = \vec{b}]} = \frac{\prod_{i=1}^{n} \Pr[Y_i = b_i]}{\prod_{i=1}^{n} \Pr[Y_i' = b_i]} = \frac{\Pr[Y_j = b_j]}{\Pr[Y_j' = b_j]}$$

$$\leq \frac{1/2 + \gamma}{1/2 - \gamma} = e^{\varepsilon}$$

For example, if we set $\gamma = 1/n$

(keep your bit w.p $3/n$ and flip w.p $1/n$)

then RR is ε-DP for $e^{\varepsilon} = \frac{3/n}{1/n} = 3$

$\varepsilon = \ln(3) = 1.1$

More generally, for $\gamma \leq 1/n$

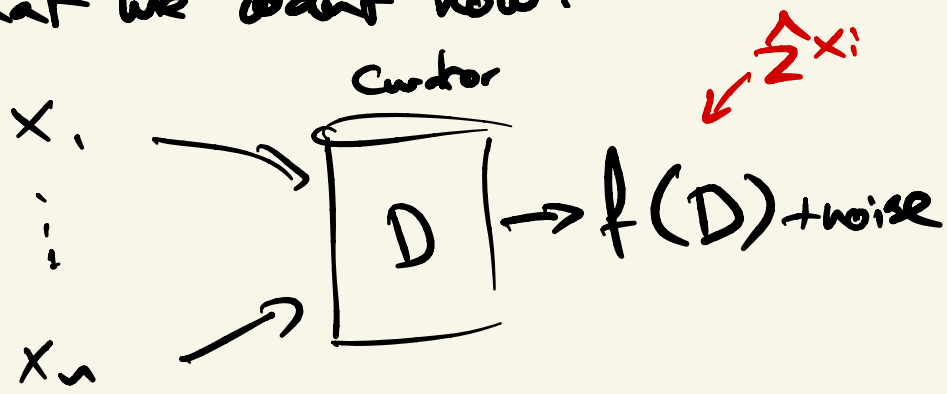
RR is ε-DP for $\varepsilon = O(\gamma)$

The "local" model of DP

- What we showed: outputting all noisy bits $Y_1, \ldots, Y_n$ is DP
- This is ok if you don't trust a curator
- If we have a curator, this seems overkill? Why? The curator only releases

$\hat{p}$ (a function of $Y_1, \ldots, Y_n$)
and not the bits themselves

What we want now:



Def: Sensitivity (ℓ_1): The ℓ_1 sensitivity of a function $f: X^n \to \mathbb{R}^k$

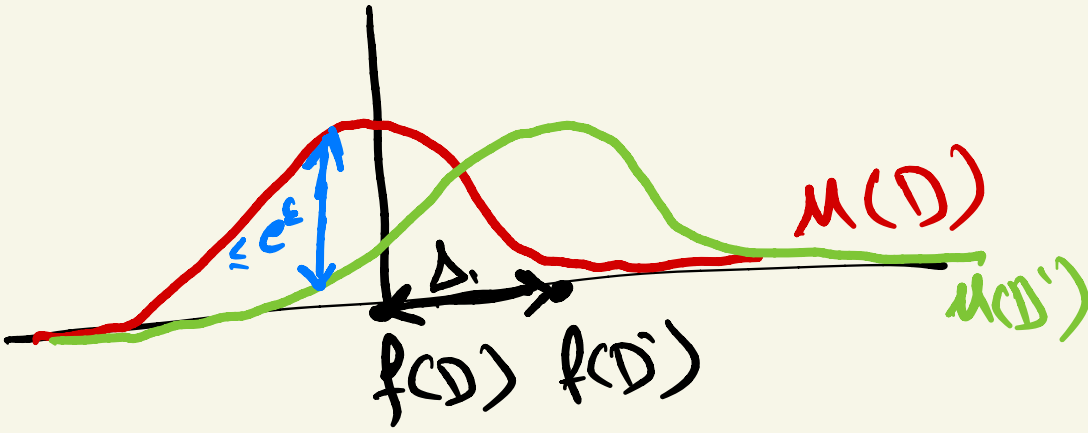
is $\Delta_1 = \max_{D \sim D'} \|f(D) - f(D')\|_1$

$= \max_{D \sim D'} \sum_{i=1}^{k} |f(D)_i - f(D')_i|$

Norms: $\ell_1 : \|x\|_1 = \sum |x_i|$

$\ell_2 : \|x\|_2 = (\sum x_i^2)^{1/2}$

Idea: output $f(D)$ + "noise that grows with the sensitivity of the function"

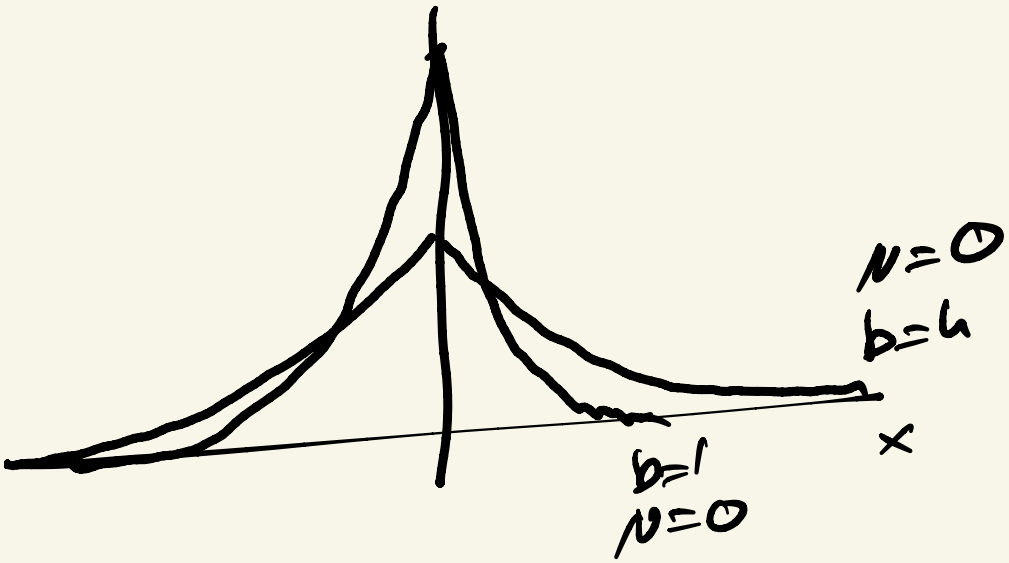


What noise? Laplace noise

Def (Laplace): The Laplace dist. Laplace(μ, b) has density

$$f(x \mid \mu, b) = \frac{1}{2b} e^{-\frac{|x-\mu|}{b}}$$

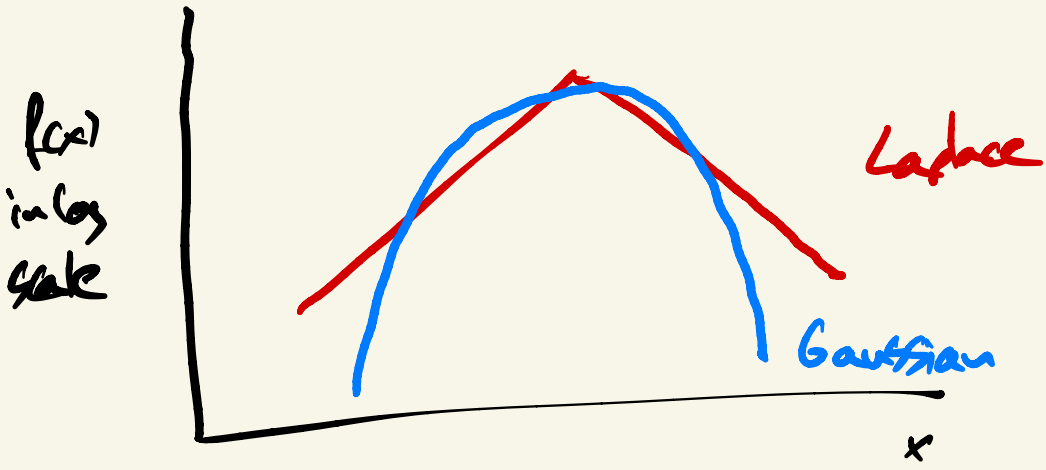
; it has mean ν and variance $2b^2$



Analog: $N(\nu, \sigma^2)$ has

density $f(x \mid \nu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-\nu)^2}{2\sigma^2}}$

Visualization:



Def (Laplace Mechanism): Let $f: X^n \to \mathbb{R}^k$ be a function with l_1 sensitivity Δ_1 and let $\varepsilon > 0$. Then the Laplace Mech:

$$M(D) = f(D) + [Y_1, \ldots, Y_k]$$

where the Y_i are iid samples from

$$\text{Laplace}\left(0, \frac{\Delta_1}{\varepsilon}\right)$$

$$M(D) = [\text{Laplace}(f(D)_1, \tfrac{\Delta_1}{\varepsilon}), \ldots, \text{Laplace}(f(D)_k, \tfrac{\Delta_1}{\varepsilon})]$$

Ex: let $f(D)$ have sensitivity 1

$f(D) = 0$

$f(D') = 1$

Suppose you observe $M(\ldots) = 6'000$

if M is ε-DP you need

$$\frac{Pr[\text{noise} = 6'000]}{Pr[\text{noise} = 9'999]} \leq e^{\varepsilon}$$

Thm: Laplace is ε-DP

Proof: let $p_D(y)$ and $p_{D'}(y)$ be the PDFs of $M(D)$ and $M(D')$ for $D \sim D'$, for $y \in \mathbb{R}^k$ arbitrary

to show: $\frac{p_D(y)}{p_{D'}(y)} \leq e^{\varepsilon}$

$$\frac{p_D(y)}{p_{D'}(y)} = \frac{\prod_{i=1}^{k} \frac{1}{2b} \exp\left(-\frac{|y_i - f(D)_i|}{b}\right)}{\prod_{i=1}^{k} \frac{1}{2b} \exp\left(-\frac{|y_i - f(D')_i|}{b}\right)}$$

$$= \prod_{i=1}^{k} \exp\left(\frac{|y_i - f(D')_i| - |y_i - f(D)_i|}{b}\right)$$

$$\leq \prod_{i=1}^{k} \exp\left(\frac{|f(D')_i - f(D)_i|}{b}\right)$$

$(|a-b| - |a-c| \leq |b-c|)$

$$= \exp\left(\frac{1}{b}\sum |f(D')_i - f(D)_i|\right)$$

$$= \exp\left(\frac{1}{b}\|f(D') - f(D)\|_1\right)$$

$$\leq \exp\left(\frac{\Delta_1}{b}\right) = \exp(\varepsilon)$$

$$b = \frac{\Delta_1}{\varepsilon}$$

Back to counting queries

RR: compute the mean of n bits with ε-DP with error $\tilde{O}\left(\frac{1}{\varepsilon\sqrt{n}}\right)$ This is optimal in the local setting

Laplace: Sensitivity of the mean? $1/n$

For ε-DP, we add Laplace with $b = \frac{\Delta_1}{\varepsilon} = \frac{1}{n\varepsilon}$

The variance is $2b^2 = \frac{1}{\varepsilon^2 n^2}$

And the error, whp, is $\tilde{O}\left(\frac{1}{\varepsilon n}\right)$ This is optimal in the central setting

Tail bound: if $z \sim \text{Laplace}(0, b)$

then $\Pr(|z| \geq t \cdot b) \leq \exp(-t)$

So, if we set $t = \Omega(\log n)$

$$\Pr[|z| \geq \log n \cdot b] = \text{negl}(n)$$

So whp $(= \text{negl}(n))$, the noise is less than

$$\log n \cdot b = \tilde{O}(1/\varepsilon n)$$

What if we want to answer many queries?

Setting: we have k counting functions

$$f = (f_1, \ldots, f_k)$$

↳ each $f_i : X^n \to \mathbb{R}$

$$f_i(D) = \frac{1}{n} \sum_{j=1}^{n} \phi_j(x_j)$$

- What's the sensitivity of f?

  ↳ each f_i has sensitivity $1/n$

  ↳ $\|f(D) - f(D')\|_1 \leq \|[1/n, 1/n, \ldots, 1/n]\|_1$

  $= \frac{k}{n}$

- for ε-DP, we need Laplace with mean 0 and scale $\frac{k}{\varepsilon n}$ in each coordinate

- whp, the maximum noise across all coordinates is $\tilde{O}\left(\frac{k}{\varepsilon n}\right)$ this is optimal ✓

For RR, what's the max. noise for k counting queries? $\tilde{O}\left(\frac{\log k}{\varepsilon n}\right)$ ← ignore log-factors