differential privacy

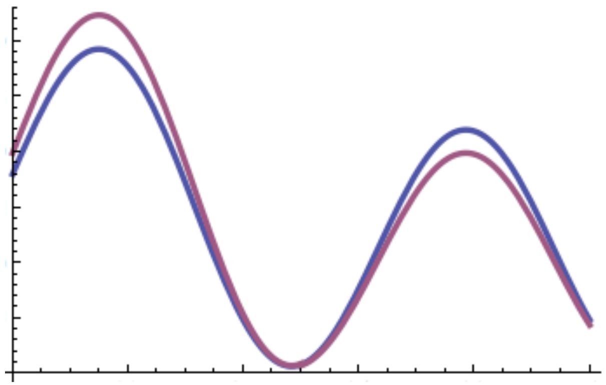
[DinurNissim03, DworkNissimMcSherrySmith06, Dwork06]

ε-Differential Privacy for algorithm M:

for any two neighboring data sets x_1, x_2, differing by the addition or removal of a single row

any $S \subseteq$ range(M),

$$\Pr[M(x_1) \in S] \leq e^{\varepsilon} \Pr[M(x_2) \in S]$$



sensitivity of a function f

$$\Delta f = \max_{x1,\ x2} \ |f(x_1) - f(x_2)|_1$$

for neighboring data sets $x_1,\ x_2$

measures how much one person can affect output

sensitivity is $1/|x|$ for queries returning the average value of count queries mapping X to $\{0,1\}$

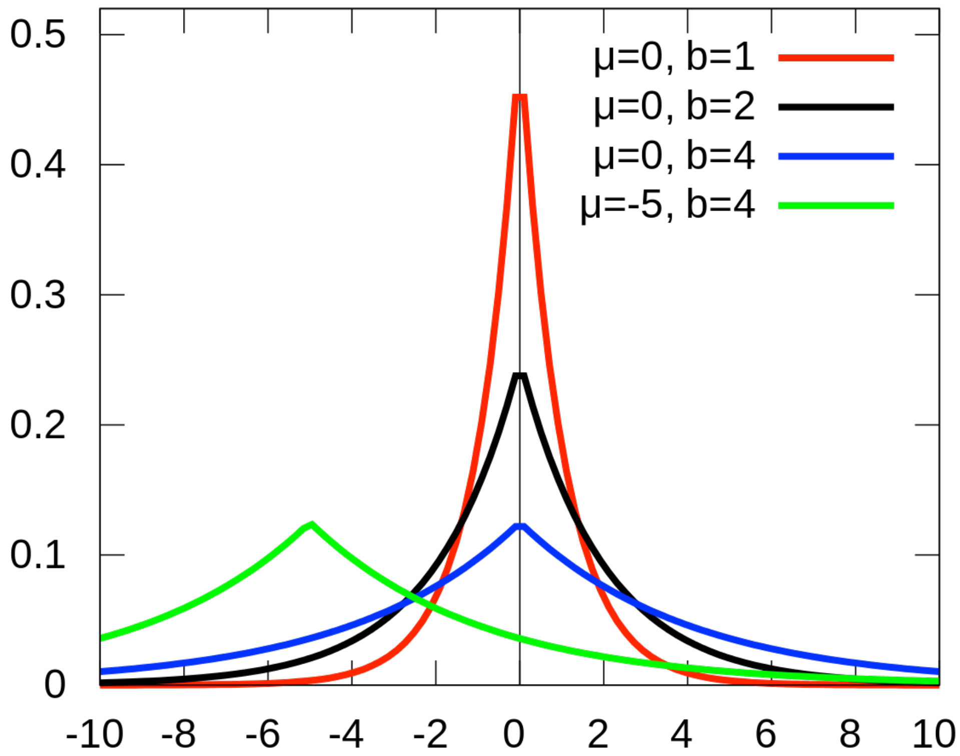
linear queries : $X \to [0,1]$ over the dataset (think statistical queries)

scale noise with sensitivity [DworkNissimMcSherrySmith06]

$$\Delta f = \max_{x1,\ x2} \ |f(x_1) - f(x_2)|_1$$

for neighboring data sets $x_1,\ x_2$

Laplace distribution $\mathrm{Lap}(b)$



$\mathrm{pdf}(z)$ = exp(-$|z|/b)/2b$

variance = $2b^2$

For $\mathrm{Y}\sim\mathrm{Lap}(b)$, $\mathrm{Pr}[|\mathrm{Y}| \geq bt] =$ exp(-t)

Laplace mechanism

Def. Given $f : \mathbb{N}^{|X|} \rightarrow R^k$ the Laplace Mechanism is defined as

$$M_{\mathrm{L}}(x, f(.), \varepsilon) = f(x) + (\mathrm{Y}_1, \ldots, \mathrm{Y}_\mathrm{k})$$

where the Y_i are iid random draws from $\mathrm{Lap}(b)$ with $b = \Delta f/\varepsilon$.

(If we want discrete output space, subsequently round accordingly.)

Laplace mechanism: Privacy

Thm. The Laplace Mechanism preserves

$(\varepsilon, 0)$-differential privacy.

Laplace Mechanism: Privacy

Thm The Laplace Mechanism preserves $(\varepsilon, 0)$-dp

pf Let x, x' be neighboring databases, so $\|x - x'\|_1 \leq 1$

Let $f : \mathbb{N}^{|X|} \to R$ $(k=1)$

let p_x be prob. density function of $M_L(x, f, \varepsilon)$

$p_{x'}$ " " " " " $M_L(x', f, \varepsilon)$

Let $s \in R$

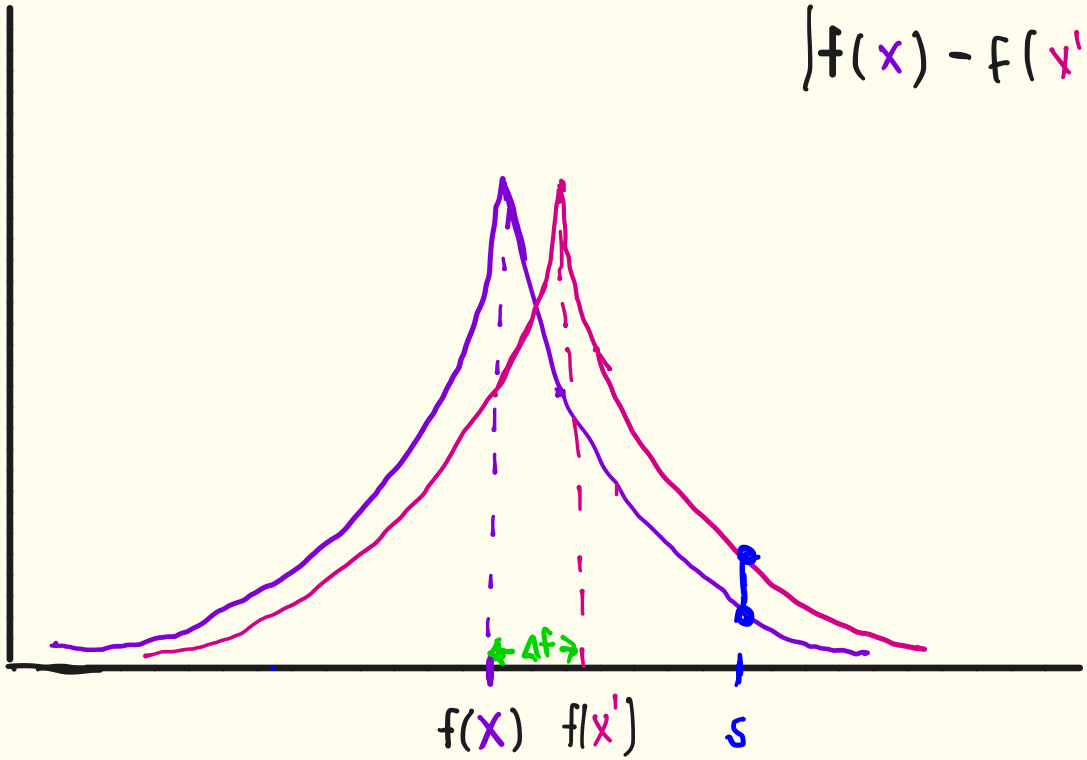
$$\frac{\Pr(s|x)}{\Pr(s|x')} = \left[\frac{\exp\left(-\frac{\varepsilon|f(x)-s|}{\Delta f}\right)}{\exp\left(-\frac{\varepsilon|f(x')-s|}{\Delta f}\right)}\right] = \exp\left[\frac{\varepsilon\left(|f(x')-s| - |f(x)-s|\right)}{\Delta f}\right]$$

$$\leq \exp\left(\frac{\varepsilon|f(x) - f(x')|}{\Delta f}\right) = \exp\left(\frac{\varepsilon\|f(x)-f(x')\|_1}{\Delta f}\right)$$

$$\leq \exp(\varepsilon)$$

$x \qquad x'$

$$|f(x) - f(x')| \leq \Delta f$$



$$\ln\left(\frac{Pr(s|x)}{Pr(s|x')}\right) = \frac{\varepsilon\left(|f(x')-s| - |f(x)-s|\right)}{\Delta f} \leq \frac{\varepsilon \Delta f}{\Delta f} \leq \varepsilon$$

Laplace mechanism: Accuracy

Thm. The Laplace Mechanism preserves…

Laplace Mechanism – Accuracy

Thm Let $f : \mathbb{N}^{|X|} \to \mathbb{R}^k$, $y = M_L(x, f, \varepsilon)$. Then $\forall \delta \in [0,1]$:

$$\Pr\left[\|f(x) - y\|_\infty \geq \ln\left(\frac{k}{\delta}\right)\left(\frac{\Delta f}{\varepsilon}\right)\right] \leq \delta$$

Pf

$$\Pr\left[\|f(x) - y\|_\infty \geq \ln\left(\frac{k}{\delta}\right)\left(\frac{\Delta f}{\varepsilon}\right)\right] = \Pr\left(\max_{i \in [k]} |Y_i| \geq \ln\left(\frac{k}{\delta}\right)\left(\frac{\Delta f}{\varepsilon}\right)\right)$$

$$\leq k \cdot \Pr\left[|Y_i| \geq \ln\left(\frac{k}{\delta}\right)\left(\frac{\Delta f}{\varepsilon}\right)\right]$$

$$\Pr\left[|Y| \geq t b \right] = \exp(-t)$$

$$b = \Delta f / \varepsilon$$

$$= k\left(\frac{\delta}{k}\right)$$

$$= \delta$$

Notes

1. Could replace Laplacian by gaussian noise

   add noise scaled to $N(0, \sigma^2)$, $\sigma \sim \Delta f \ln(1/\delta)/\varepsilon$

   gives (ε, δ)-dp

2. The simpler randomized response algorithm

   is Local

   However overall its accuracy is worse.

applying the Laplace mechanism

single counting query: how many people in the database satisfy predicate P?

sensitivity = 1

can add noise $\mathrm{Lap}(1/\varepsilon)$

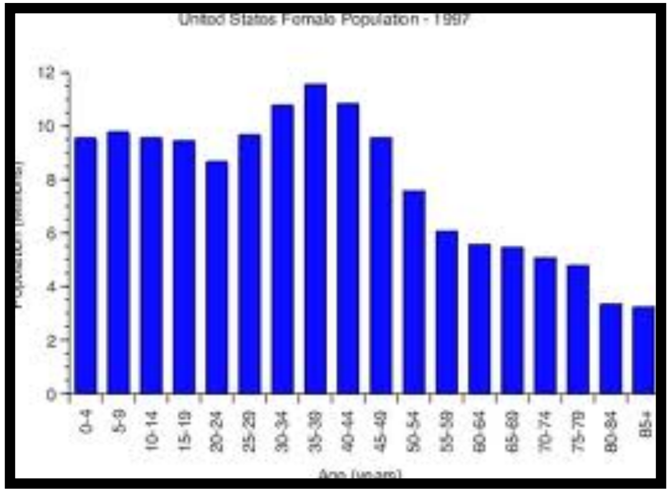
applying the Laplace mechanism

vector-valued queries of dimension d

Can apply composition and add noise $\mathrm{Lap}(d\ \Delta f/\varepsilon)$ in each component of output, where Δf is sensitivity of each component.

(Could also split the noise unevenly. Could also treat the queries separately and apply composition.)

applying the Laplace mechanism



histogram queries

could again use noise $\mathrm{Lap}(d/\varepsilon)$

but actually only need $\mathrm{Lap}(1/\varepsilon)$, since

sensitivity generalizes as max ℓ_1 distance

What is the sensitivity of the query, "how many people in the database are both over age 50 AND smokers?"

example

Suppose we wanted to determine the most commonly-"liked" Facebook page, subject to DP

could give a DP count of the number of likes for each page, but sensitivity would grow with the max number of "likes" a person could give (bad)

but we only want to know the max, not *every* count—could that be easier?

reportNoisyMax

For m count queries add noise $\mathrm{Lap}(1/\varepsilon)$ to each, and report the index of the largest noised query.

Claim: reportNoisyMax is $(\varepsilon,\, 0)$-differentially private.

reportNoisyMax

For m count queries add noise $\mathrm{Lap}(1/\varepsilon)$ to each, and report the index of the largest noised query.

What about accuracy?