PRIVACY:

RANDOMIZED RESPONSE:

No trusted party.

Each user $y_i = \begin{cases} x_i & \text{with probability } 1/2 + \gamma \\ 1 - x_i & \text{with probability } 1/2 - \gamma \end{cases}$

$x_i : \{0, 1\}$

$0 < \gamma < 1/2$

$$E[y_i] = (1/2 + \gamma)\, x_i + (1/2 - \gamma)(1 - x_i)$$

$$= 1/2 - \gamma + 2\gamma x_i$$

$$x_i = \frac{E[y_i] - (1/2 - \gamma)}{2\gamma}$$

$$\bar{p} = \frac{\left(\frac{y_1 + \dots + y_n}{n}\right) - \left(\frac{1}{2} - \gamma\right)}{2\gamma}$$

$$\boxed{\Pr\left[|\bar{p} - p| > \delta\right] \leq \frac{\delta}{4\gamma\sqrt{n}}}$$

With 75% chance $|\hat{p} - p| \leq \frac{1}{\gamma\sqrt{n}} = \alpha$

$\uparrow \gamma \rightarrow \downarrow$ Noise $\rightarrow \uparrow$ accuracy

$\uparrow$ people $\nearrow$

$n \geq \frac{1}{\gamma^2 \alpha^2} \longrightarrow$ To get α error, $(1/2 - \gamma)$ privacy.

Differential Privacy:

$M : X^n \rightarrow Y$

is ε - Differentially private if $\forall$ neighboring databases D, D' $\forall\ y \in Y$

$$e^{-\varepsilon} \Pr[M(D') = y] \leq \Pr[M(D) = y] \leq e^{\varepsilon} \Pr[M(D') = y]$$

Small $\varepsilon \rightarrow \uparrow$ Privacy.

Computing the Mean Differentially Privately:

Randomized Response: [Input Perturbation]

$x_i \in \{0, 1\}$

Each user $y_i = \begin{cases} x_i & \text{with probability } 1/2 + \gamma \\ 1 - x_i & \text{with probability } 1/2 - \gamma \end{cases}$

$0 \leq \gamma \leq 1/2$

$$\tilde{f}(x) = \frac{\sum y_i}{n} \qquad f(x) = \frac{\sum x_i}{n}$$

$$\boxed{Pr\left[\left|\tilde{f}(x) - f(x)\right| \geq \frac{1}{\gamma\sqrt{n}}\right] \leq 1/4}$$

Randomized Response is $O(\gamma)$ Differentially private.

$$\frac{Pr[RR_\gamma(x) = a]}{Pr[RR_\gamma(x') = a]} \leq \frac{1/2 + \gamma}{1/2 - \gamma} = O(\gamma)$$

Summary: Randomised Response with noise rate γ is $O(\gamma)$- DP and achieves error $\frac{1}{\gamma\sqrt{n}}$ for the mean.

$\uparrow \gamma \rightarrow \downarrow$ Noise

Noise Rate $\rightarrow \gamma$ [Privacy Parameter]

$O(\gamma)$ DP

Error $1/\gamma\sqrt{n}$

ε - Differential Privacy + error $\leq \alpha$

$$\boxed{n \geq \frac{c}{\varepsilon^2\alpha^2}}$$

$$\frac{1}{\gamma\sqrt{n}} \leq \alpha$$

$$n \geq \frac{1}{\gamma^2 c^2}$$

PRIVACY - ε ERROR - α SIZE - n

OUTPUT PERTURBATION:

→ LAPLACIAN MECHANISM:

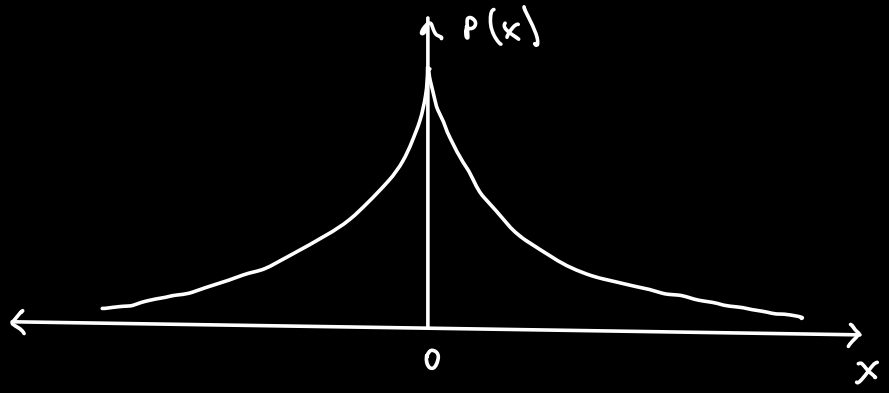
RELEASING MEAN:

$$\tilde{f}(x) = f(x) + z \qquad x = \{0,1\}$$

↓

Laplacian $(0, b)$

Pdf: $$P(x) = \frac{1}{2b} e^{-|x|/b}$$



→ $b = 1/\varepsilon n$ => ε-DP

PROOF: $|f(x) - f(x')| \leq 1/n$

$$Pr[M(x) = a] = Pr[f(x) + z = a]$$

$$= Pr[z = a - f(x)]$$

$$= \frac{1}{2b} \cdot e^{-|a - f(x)|/b}$$

$$\Pr[M(x') = a] = \Pr[f(x') + z = a]$$

$$= \Pr[z = a - f(x')]$$

$$= \frac{1}{2b} \cdot e^{-|a - f'(x)|/b}$$

$$\frac{\Pr[M(x) = a]}{\Pr[M(x') = a]} = e^{-[|a - f(x)| - |a - f'(x)|]/b}$$

$$= e^{(|a - f'(x)| - |a - f(x)|)/b}$$

$$\leq e^{1/nb}$$

$$\varepsilon = \frac{1}{nb}$$

$$\boxed{b = \frac{1}{n\varepsilon}}$$

Steps:

1. Take x, x' → neighboring
2. find $|f(x) - f(x')|$
3. $\frac{Pr[M(x) = a]}{Pr[M(x') = a]}$
   i. $M(x) = f(x) + z = a$
   ii. $Pr(z = a - f(x))$
   ↳ Use distributing PDF.

LAPLACIAN HAS DECAYING TAILS:

$$\boxed{Pr[|z| > t \cdot b] \leq e^{-t}}$$

$t = z$

$z \rightarrow a - f(x)$

$$Pr[|z| > 2b] \leq e^{-2} \leq 1/4$$

$$\boxed{Pr\left[|Answer - f(x)| \geq 2 \cdot \frac{1}{n\varepsilon}\right] \leq 1/4}$$

$$Error = 2 \cdot \frac{1}{n\varepsilon} = \alpha$$

α error, ε-privacy $\Rightarrow$ $\boxed{n \geq \frac{2}{\alpha\varepsilon}}$

Any Query: [Scalar Output]

Sensitivity: $f: X^n \to \mathbb{R}$

$$S_1(f) = \max_{x, x'} |f(x) - f(x')|$$

$$S_1(\text{mean}) = 1/n$$

Laplacian Mechanism:

$S_1(f) = S$, $b = S/\varepsilon$ (noise) $\Rightarrow$ ε-DP.

k-Output Queries:

$$f: X^n \to \mathbb{R}^k$$

$$S_1(f) = \max_{x, x'} \sum_{i=1}^{k} |f(x)_i - f(x')_i|$$

$$\to M(x) = f(x) + (z_1, z_2, \dots, z_k)$$

If $S_1(f) \le S$ + $b = S/\varepsilon$ $\Rightarrow$ ε-DP !

$$M(x) = f(x) + (z_1, z_2 \dots z_k)$$

$$\frac{\Pr[M(x) = (a_1, a_2 \dots a_k)]}{\Pr[M(x') = (a_1, a_2, \dots a_k)]}$$

$$\Pr[M(x) = A] = \Pr[f(x) + (z_1, z_2 \dots z_k) = (a_1, a_2 \dots a_k)]$$

$$= \Pr[(z_1, z_2 \dots z_k) = (a_1 - f(x)_1, a_2 - f(x)_2, \dots a_k - f(x)_k)]$$

$$= \frac{1}{2b} \cdot e^{-|a_1 - f(x)_1|/b} \cdot \frac{1}{2b} \cdot e^{-|a_2 - f(x)_2|/b} \dots \frac{1}{2b} e^{-|a_k - f(x)_k|/b}$$

$$\frac{\Pr[M(x) = A]}{\Pr[M(x') = A]} = e^{\sum_i^k (|a_i - f(x')_i| - |a_i - f(x)_i|)/b}$$

$$\leq e^{s/b}$$

POST-PROCESSING DP:

If $M: X^n \to Y$ ε-DP, $F: Y \to Z$ $F \circ M: X^n \to Z$ also ε-DP

GROUP PRIVACY:

If x, x' can differ in k rows, $M: X^n \to Y$ is ε-DP if:

$$\Pr[M(x) = y] \leq e^{k\varepsilon} \cdot \Pr[M(x') = y]$$

COUNTING QUERIES:

$S_1 = 1$

k: Queries:

$S_1 = k$

HISTOGRAM QUERIES:

$S_1 = 2$

$f: x^n \rightarrow \mathbb{R}^k$

$\hookrightarrow y = \mathbb{R}^k$ (k buckets)

$Z\left(2/\varepsilon\right)$ noise added to each bucket.

So Accuracy is affected:

$$Pr\left[|z| > t \cdot b\right] \leq e^{-t}$$

So Probability of a bucket having large error is low.

$$Pr\left[|M(y_i) - f(y_i)| > t/\varepsilon\right] \leq e^{-t}$$

But Probability of atleast one bucket having large error depends on k:

$$Pr\left[\exists i \; |M(y_i) - f(y_i)| > t/\varepsilon\right] \leq k \cdot e^{-t}$$

So $t = 10 + \log k$

$$Pr\left[\exists i \; |M(y_i) - f(y_i)| > (10+\log k)/\varepsilon\right] \leq e^{-10}.$$

So, it only depends "logarithmically" on k.

So we can guarantee that the error will atmost be $(10 + \log k)/\varepsilon$ with $(1 - e^{-10})$ confidence.

QUERY: MOST COMMON FIRST NAME:

Histogram Query.

DIFFERENTIALLY PRIVATE SELECTION

QUERY: MOST COMMON DISEASE

One person can have >1 disease

Not histogram!

Option 1: COUNTING QUERY

$S_1 = k$ → Noise → k/ε on each disease's count.

→ Release all counts.

Option 2: "Noisy Max"

→ Noise $1/\varepsilon$ to each

→ Release disease with max noisy count.

→ ε-DP

EXPONENTIAL MECHANISM:

Data:

	Price Quoted
User 1	1
2	1
3	1
4	3.01

For a given Price set by auctioneer → We have a Set revenue.

$P = 1 \rightarrow R = 4$

$P = 3.01 \rightarrow R = 3.01$ } very sensitive.

$P = 3.02 \rightarrow R = 0$ }

"Range" $R \rightarrow$ Price | Diseases

$x^n \rightarrow$ User quoted prices | User Medical Record

"Utility" Function $u : x^n \times R \rightarrow \mathbb{R}$ | Count of Disease
↳ Revenue

GOAL: $\arg\max_{r \in R} u(x, r)$

Which Price to set for max revenue?

Which Disease has highest count?

$M_\varepsilon(x, u, R)$ selects and outputs $r \in R$ with Probability $\propto e^{(\varepsilon \cdot u(x,r)/\Delta u)}$

$(2\varepsilon - DP)$

$$\Delta u = \max_{r \in R} \max_{x, x'} |u(x, r) - u(x', r)|$$

PROOF:

$$\frac{Pr[M_\varepsilon(x) = r]}{Pr[M_\varepsilon(x') = r]} = \frac{e^{(\varepsilon \cdot u(x,r)/\Delta u)}}{\sum_{s \in R} e^{(\varepsilon u(x,s)/\Delta u)}} \cdot \frac{\sum_{s \in R} e^{\frac{\varepsilon \cdot u(x',s)}{\Delta u}}}{e^{(\varepsilon \cdot u(x',r)/\Delta u)}}$$

By Δu definition, $\forall s \in R$

$$|u(x, s) - u(x', s)| \leq \Delta u$$

Numerator: $e^{\varepsilon u(x', s)/\Delta u} \leq e^{\varepsilon \cdot (u(x,s) + \Delta u)/\Delta u}$

$$= e^{\varepsilon} \cdot e^{\varepsilon u(x,s)/\Delta u}$$

Denominator: $e^{\varepsilon u(x,s)} \geq e^{\varepsilon} \cdot e^{\varepsilon u(x', s)/\Delta u}$

$$\frac{Pr\left[M_{\varepsilon}(x) = r\right]}{Pr\left[M_{\varepsilon}(x') = r\right]} \leq e^{\varepsilon} \cdot e^{\varepsilon} \leq e^{2\varepsilon}$$

ACCURACY:

$$Pr\left[u\left[M_{\varepsilon}(x, u, R)\right) \leq OPT_u(x) - \frac{\Delta u}{\varepsilon} \cdot (\ln(|R|) + t)\right] \leq e^{-t}$$

↓

Logarithmic on number of ranges (like histogram)

For Laplacian

$$Pr\left[|M(x) - f(x)| \geq \underbrace{\frac{S_1(x) \cdot t}{\varepsilon}}_{\text{extra } \ln|R| \text{ here.}}\right] \leq e^{-t}$$

Approximate Differential Privacy:

(ε, δ) - Differential privacy

$$Pr[M(x) = a] \leq e^{\varepsilon} Pr[M(x') = a] + \delta$$

So for unlikely scenarios

$$Pr[M(x) = a] \leq \delta$$

need not be ε-DP.

$$\delta << 1/n$$

L_2 Sensitivity:

$$S_2(f) = \max_{x, x'} \| f(x) - f(x') \|_2$$

Gaussian Mechanism:

$$Pr(z = y) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-y^2/2\sigma^2}$$

$$N(0, \sigma^2)$$

$f: X^n \to \mathbb{R}^k$

$$\sigma = \sqrt{2 \ln\left(\frac{1.25}{\delta}\right)} \cdot \frac{S_2(f)}{\varepsilon}$$ ensures

(ε, δ) - DP

Group Privacy:

$$\Pr[M[x] = y] \leq e^{k\varepsilon} \Pr[M[x'] = y] + k e^{k\varepsilon} \delta.$$

Adaptive Queries:

$M_1 \circ M_2 \circ \cdots M_k \rightarrow k\varepsilon\text{-DP} \quad | \quad (k\varepsilon, k\delta)\text{-DP}$

if each ε-DP. $\quad | \quad (\varepsilon, \delta)\text{-DP}$

Advanced Composition Theorem:
If $M_1, \ldots M_k$ are ε-DP mechanisms for answering adaptive queries. Then, the composition is $\left(\varepsilon\sqrt{2k\ln\left(\frac{1}{\delta}\right)} + k\varepsilon^2, \delta\right)$-DP
(as long as $\varepsilon < 1$).

Good when $\varepsilon << 1/\sqrt{k}$

Private ERM:

→ Output Perturbation

→ Objective Perturbation

→ Gradient Perturbation

↓ Adaptive Queries!

SGD

$\left(\widetilde{\nabla_\theta L}\right) = \nabla_\theta L + \text{noise}$

↳ Gaussian

THEOREM:

We can use Gaussian Mechanism + DP-SGD to get (ε, δ)-DP and error α if $n > \frac{\sqrt{d}\sqrt{\log(1/\delta)}}{\varepsilon \cdot \alpha^2}$.

Summary

1. Differential Privacy :

$$M : x^n \to y$$

$$Pr[M(x) = y] \leq e^{\varepsilon} Pr[M(x') = y]$$

2. Randomized Response : [Input perturbation]

$$x_i \in \{0, 1\}$$

Method : $$y_i : \begin{cases} x_i & p = 1/2 + \gamma \\ 1 - x_i & p = 1/2 - \gamma \end{cases}$$

Query : Mean

Privacy : $O(\gamma)$ DP

Accuracy : $$Pr\left[|\text{Error}| \geq \frac{1}{\gamma\sqrt{n}} \right] \leq 1/4$$

Data Size : To get ε-DP + error, α

$$n \geq \frac{c}{\varepsilon^2 \alpha^2} \qquad c : \text{constant}$$

3. Laplacian Mechanism [Output Perturbation]:

Laplacian

* PDF: $P(x) = \frac{1}{2b} e^{-|x|/b}$

* $Pr[|z| > t \cdot b] \leq e^{-t}$

Sensitivity $S_1(f) = \max_{x, x'} \sum_{i=1}^{k} |f(x)_i - f(x')_i|$

Privacy: ε-DP if $b = S/\varepsilon$

Accuracy: $Pr\left[|error| \geq t \cdot \frac{S}{\varepsilon}\right] \leq e^{-t}$

Query: Mean $b = 1/n$ $t = 2$

$$Pr\left[|error| \geq \frac{2}{n\varepsilon}\right] \leq 1/4$$

DATA SIZE: $n \geq \frac{2}{\alpha\varepsilon}$

α error, ε-DP

SENSITIVITY: Mean: $(1/n)$

k-Queries: k/n

Counting: 1

Histogram: 2 k buckets

↳ Error $\propto \log(k)$

$P\left(\text{Some bucket error} > \frac{10 + \log k}{\varepsilon}\right) \leq e^{-10}$

3. Noisy Max: ε-DP

Add $1/\varepsilon$ noise to each disease

Release max noisy count

4. Exponential Mechanism: [Return answer like most common disease]

X^n : Data not just number

R : Range

- Utility Function $u : X^n \times R \to \mathbb{R}$
- $\arg\max_{r \in R} u(x, R)$

Method: $M_\varepsilon(x, u, R)$ select $r \in R$ Pr $\propto e^{(\varepsilon \cdot u(x,r)/\Delta u)}$

Privacy : 2ε-DP

$$\Delta u = \max_{r \in R} \max_{x, x'} |u(x, r) - u(x', r)|$$

Accuracy :

$$\Pr[|\text{error}| \geq \frac{\Delta u}{\varepsilon} \cdot (\ln|R| + t)] \leq e^{-t}$$

$$\downarrow$$

$u(x, r) - OPT_u(x)$

(r: answer)

5. Approximate

$$Pr[M(x) = a] \leq e^{\varepsilon} Pr[M(x') = a] + \delta$$

6. Gaussian Mechanism [Approximate]

(ε, δ) - DP if $\sigma = \sqrt{2 \ln\left(\frac{1.25}{\delta}\right)} \cdot \frac{S_2(f)}{\varepsilon}$

7. Adaptive Queries:

$k\varepsilon$ - DP

8. Advanced Composition Theorem:

Adaptive Queries

$\varepsilon << 1/\sqrt{k}$

$$\left(\varepsilon \sqrt{2k \ln(1/\delta)} + k\varepsilon^2, \delta\right) \text{ - DP}$$

$k\delta'$

9. Private ERM

→ Output, Objective, Gradient

↳ SGD

$\tilde{G} = (\nabla_\theta L + \text{gaussian noise})$.

We can use Gaussian Mechanism + DP-SGD to get (ε, δ) - DP and error α if $n > \frac{\sqrt{d}\sqrt{\log(1/\delta)}}{\varepsilon \cdot \alpha^2}$.

HOMEWORK

$$\Pr\left[u(x, M_\varepsilon(x,u,R)) \leq OPT_u(x) - \frac{\Delta u}{\varepsilon}\left((\ln|R| + t)\right)\right] \leq e^{-t}$$

$$\Delta u = \max_{r \in R} \max_{x,x'} |u(x,r) - u(x',r)|$$

$$\Pr \propto e^{(\varepsilon \cdot u(x,r)/\Delta u)}$$

$$\Pr[M_\varepsilon(x,u,r)] = \frac{e^{(\varepsilon \cdot u(x,r)/\Delta u)}}{\sum_{s \in R} e^{(\varepsilon \cdot u(x,s))/\Delta u}}$$