DP Algorithm Primitives

Privacy & Fairness in Data Science
CompSci 590.01 Fall 2018

DUKE
COMPUTER SCIENCE

Outline

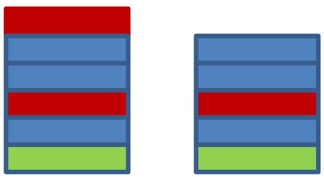
- Recap
- Algorithmic Primitives
  - Randomized Response
  - Laplace Mechanism
  - Exponential Mechanism
- Composition Theorems

Differential Privacy

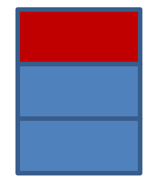
[Dwork ICALP 2006]

For every pair of inputs that differ in one row

For every output …



D_1 ***D_2***



O

Adversary should not be able to distinguish between any D_1 and D_2 based on any O

$$\forall \Omega \in \mathrm{range(A)},\ \ln\left(\frac{\Pr[A(D_1) \in \Omega]}{\Pr[A(D_2) \in \Omega]}\right) \leq \varepsilon, \qquad \varepsilon > 0$$

Outline

- Recap
- Algorithmic Primitives
  - Randomized Response
  - Laplace Mechanism
  - Exponential Mechanism
- Composition Theorems

[W 65]

Randomized Response (a.k.a. local randomization)

D
Disease (Y/N)
Y
Y
N
Y
N
N

With probability p,
Report true value

With probability 1-p,
Report flipped value

O
Disease (Y/N)
Y
N
N
N
Y
N

Differential Privacy Analysis

- Consider 2 databases D, D′ (of size M) that differ in the j^{th} value
  - D[j] ≠ D′[j]. But, D[i] = D′[i], for all i ≠ j

- Consider some output O

$$\frac{P(D \to O)}{P(D' \to O)} \leq e^{\varepsilon} \Leftrightarrow \frac{1}{1+e^{\varepsilon}} < p < \frac{e^{\varepsilon}}{1+e^{\varepsilon}}$$

Utility Analysis

- Suppose y out of N people replied "yes", and rest said "no"
- What is the best estimate for π = fraction of people with disease = Y?

$$\hat{\pi} = \frac{\frac{y}{N} - (1-p)}{2p-1}$$

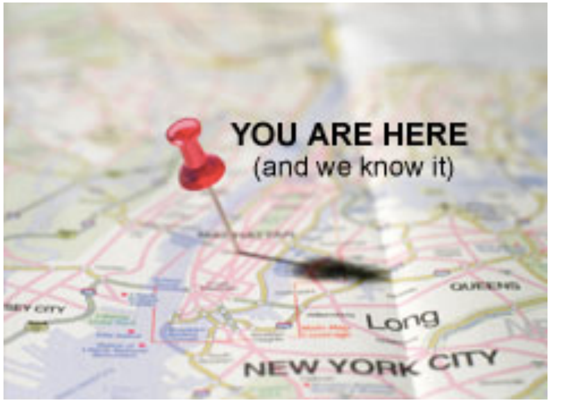
- $E(\widehat{\pi}) = \pi$
- $Var(\hat{\pi}) = \frac{\pi(1-\pi)}{N} + \frac{1}{N\left(16\left(p-\frac{1}{2}\right)^2 - \frac{1}{4}\right)}$

Sampling **Variance due to coin flips**

Randomized response for larger domains

- Suppose area is divided into k x k uniform grid.

- What is the probability of reporting the true location?



- What is the probability of reporting a false location?

Algorithm:

- Report true position: p
- Report any other position: q (< p)

$$p + q(k^2 - 1) = 1$$
$$p \le e^{\varepsilon} q$$

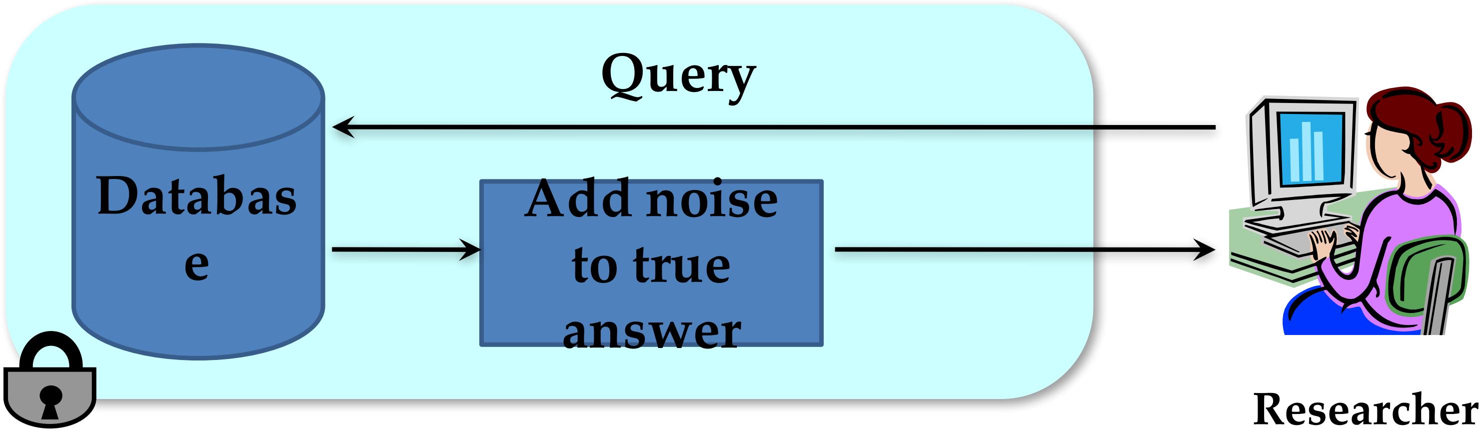
$$q = \frac{1}{e^{\varepsilon} + (k^2 - 1)}$$

- For ε = ln(3), k = 10: $p = \frac{3}{102}$

Outline

- Recap
- Algorithmic Primitives
  - Randomized Response
  - Laplace Mechanism
  - Exponential Mechanism
- Composition Theorems

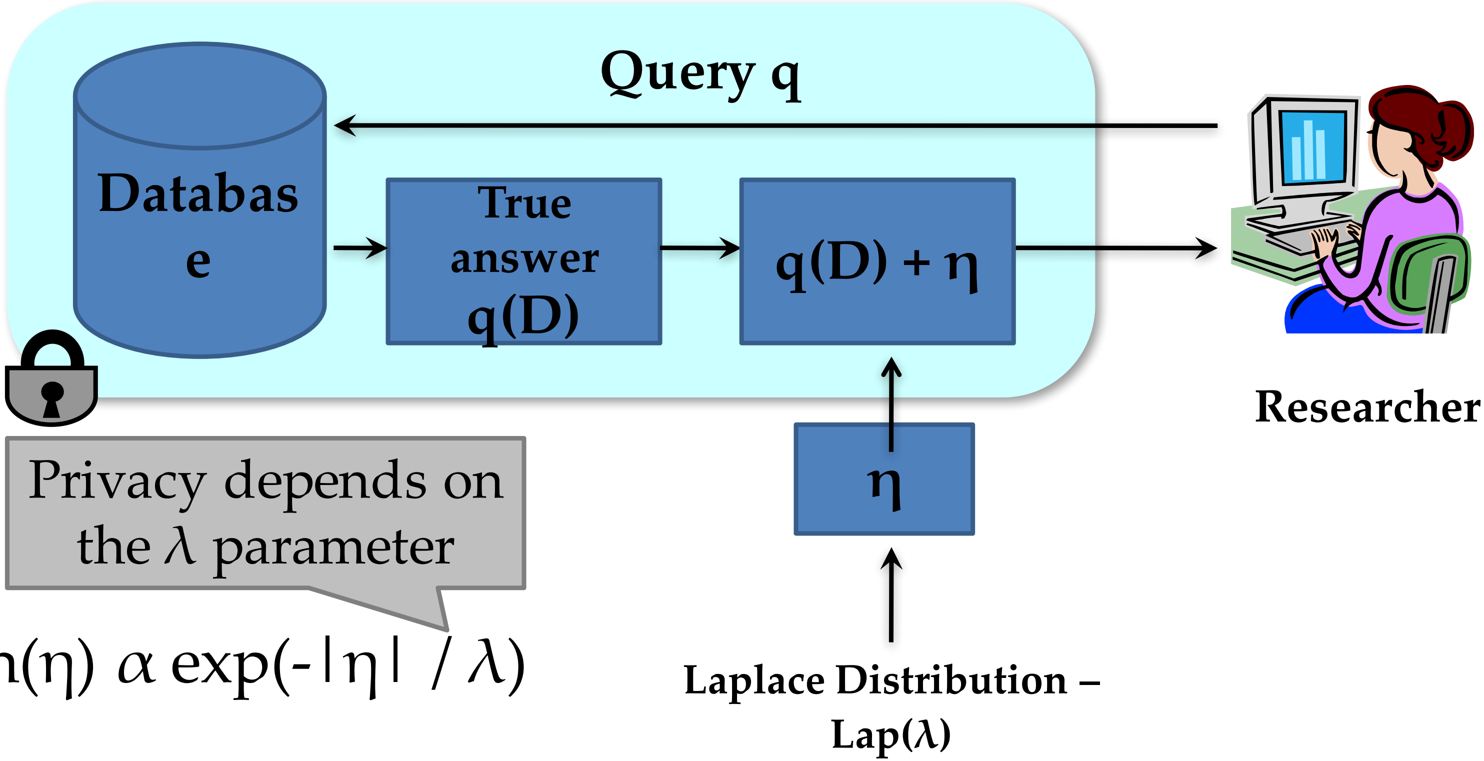
Output Randomization



- Add noise to answers such that:
  - Each answer does not leak too much information about the database.
  - Noisy answers are close to the original answers.

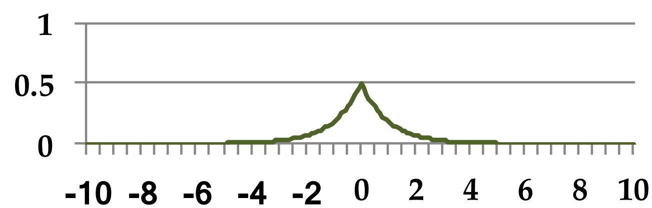
[DMNS 06]

Laplace Mechanism



$h(\eta) \,\alpha\, \exp(-|\eta| / \lambda)$

Mean: 0,
Variance: $2\lambda^2$



How much noise for privacy?

Sensitivity: Consider a query q: $I \rightarrow R$. S(q) is the smallest number s.t. for any neighboring tables D, D',

$$| q(D) - q(D') | \leq S(q)$$

Thm: If **sensitivity** of the query is **S**, then the following guarantees ε-differential privacy.

$$\lambda = S/\varepsilon$$

Sensitivity: COUNT query

- Number of people having disease
- Sensitivity = 1

- Solution: $3 + \eta$,
  where η is drawn from Lap($1/\varepsilon$)
  - Mean = 0
  - Variance = $2/\varepsilon^2$

D

Disease (Y/N)
Y
Y
N
Y
N
N

Sensitivity: SUM query

- Suppose all values x are in [a,b]

- Sensitivity = b

Privacy of Laplace Mechanism

- Consider neighboring databases D and D′
- Consider some output O

$$\frac{\Pr[A(D)=O]}{\Pr[A(D')=O]} = \frac{\Pr[q(D)+\eta=O]}{\Pr[q(D')+\eta=O]}$$

$$= \frac{e^{-|O-q(D)|/\lambda}}{e^{-|O-q(D')|/\lambda}}$$

$$\le e^{|q(D)-q(D')|/\lambda} \le e^{S(q)/\lambda} = e^{\varepsilon}$$

Utility of Laplace Mechanism

- Laplace mechanism works for **any function** that returns a real number

- Error: E(true answer – noisy answer)^2

  = Var(Lap(S(q)/ε))

  = 2*S(q)^2 / ε^2

Utility Theorem

Thm: $P[|A(D) - q(D)| > t \cdot \lambda] = e^{-t}$

$$P[|A(D) - q(D)| > t \cdot \lambda] = \int_{-\infty}^{-t} \frac{e^{-\frac{|x|}{\lambda}} dx}{2\lambda} + \int_{t}^{\infty} \frac{e^{-\frac{|x|}{\lambda}} dx}{2\lambda}$$

$$= 2\int_{t}^{\infty} \frac{e^{-\frac{|x|}{\lambda}} dx}{2\lambda} = e^{-t}$$

Cor: $P\left[|A(D) - q(D)| > \frac{S(q)}{\varepsilon} \ln\left(\frac{1}{\delta}\right)\right] \leq \delta$

Laplace Mechanism vs Randomized Response

Privacy

- Provide the same ε-differential privacy guarantee

- Laplace mechanism assumes data collected is trusted
- Randomized Response does not require data collected to be trusted
  – Also called a *Local* Algorithm, since each record is perturbed

Laplace Mechanism vs Randomized Response

Utility

- Suppose a database with N records where μN records have disease = Y.
- Query: # rows with Disease=Y

- Std dev of Laplace mechanism answer: $O(1/\varepsilon)$
- Std dev of Randomized Response answer: $O(\sqrt{N}/\varepsilon)$

Outline

- Recap
- Algorithmic Primitives
  - Randomized Response
  - Laplace Mechanism
  - Exponential Mechanism
- Composition Theorems

Exponential Mechanism

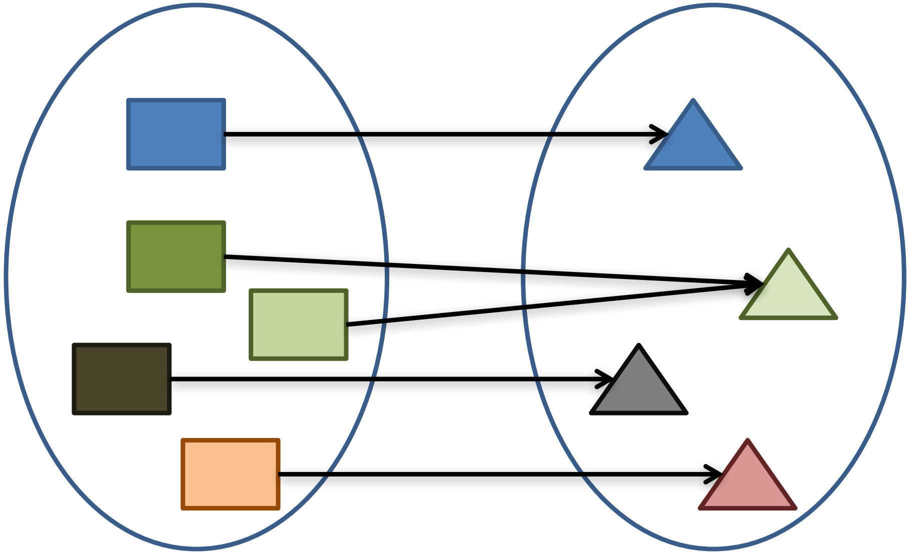
- For functions that do not return a real number …
  - "what is the most common nationality in this room": Chinese/Indian/American…

- When perturbation leads to invalid outputs …
  - To ensure integrality/non-negativity of output

[MT 07]

Exponential Mechanism

Consider some function f (can be deterministic or probabilistic):



How to construct a differentially private version of f?

Exponential Mechanism

- Scoring function *w: Inputs x Outputs* → *R*

- D: nationalities of a set of people
- #(D, O): # people with nationality O
- f(D): most frequent nationality in D
- w(D, O) = #(D, O) - #(D, f(D))

Exponential Mechanism

- Scoring function *w: Inputs x Outputs* → *R*

- Sensitivity of w

$$\Delta_w = \max_{O\&D,D'} |w(D,O) - w(D,O')|$$

where D, D′ differ in one tuple

Exponential Mechanism

Given an input D, and a scoring function w,

Randomly sample an output O from *Outputs* with probability

$$\frac{e^{\frac{\varepsilon}{2\Delta}\cdot w(D,O)}}{\sum_{Q\epsilon Outputs} e^{\frac{\varepsilon}{2\Delta}\cdot w(D,Q)}}$$

- Note that for every output O, probability O is output > 0.

Utility of the Exponential Mechanism

- Depends on the choice of scoring function – weight given to the best output.

- E.g.,
  "What is the most common nationality?"
  w(D,nationality) = # people in D having that nationality

  Sensitivity of w is 1.

- Q: What will the output look like?

Utility of Exponential Mechanism

- Let OPT(D) = nationality with the max score
- Let O_{OPT} = {O ε Outputs : w(D,O) = OPT(D)}

- Let the exponential mechanism return an output O*

Theorem:

$$\Pr\left[w(D,O^*) \leq OPT(D) - \frac{2\Delta}{\varepsilon}\left(\log\frac{|Outputs|}{|O_{OPT}|} + t\right)\right] \leq e^{-t}$$

Utility of Exponential Mechanism

Theorem:

$$\Pr\left[w(D, O^*) \leq OPT(D) - \frac{2\Delta}{\varepsilon}\left(\log\frac{|Outputs|}{|O_{OPT}|} + t\right)\right] \leq e^{-t}$$

Suppose there are 4 nationalities
Outputs = {Chinese, Indian, American, Greek}

Exponential mechanism will output some nationality that is shared by at least K people with probability $1-e^{-3}$(=0.95), where

$$K \geq OPT - 2(\log(4) + 3)/\varepsilon = OPT - 6.8/\varepsilon$$

Laplace versus Exponential Mechanism

- Let f be a function on tables that returns a real number.

- Define: score function w(D,O) = - |f(D) - O|
- Sensitivity of w = $\max_{D,D'}$ (|f(D) – O| - |f(D') – O|)
  $\leq \max_{D,D'}$ |f(D) – f(D')| = sensitivity of f

- Exponential mechanisms returns an output f(D) + η with probability proportional to

$$e^{-\frac{\varepsilon}{2\Delta}|f(D)+\eta-f(D)|}$$

Laplace noise with parameter 2Δ/ε

Randomized Response vs Exponential Mechanism

- Input: a bit in {0,1}
- Output: a bit in {0,1}

- Score: w(0,0) = w(1,1) = 1; w(0,1) = w(1,0) = 0
- Sensitivity of w = 1

- Exponential mechanism:

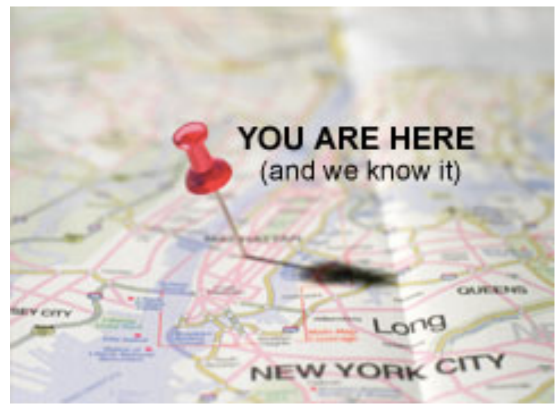
  Output the same value with prob: $\frac{e^{\varepsilon/2}}{1+e^{\varepsilon/2}}$

Randomized Response with parameter ε/2

Randomized response for larger domains

- Suppose area is divided into k x k uniform grid.

- What is the probability of reporting the true location?



- What is the probability of reporting a false location?

Different scoring functions give different algorithms

- Uniform:
  - Report true position: 1
  - Report a false position: 0

- Distance:
  - Report true position (i,j): 0
  - Report false position (x,y): $-(|i-x| + |j-y|)$

- …

Summary of Exponential Mechanism

- Differential privacy for cases when output perturbation does not make sense.
- Idea: Make better outputs exponentially more likely; Sample from the resulting distribution.
- Every differentially private algorithm is captured by exponential mechanism.
  - By choosing the appropriate score function.

Summary of Exponential Mechanism

- Utility of the mechanism only depends on log(|Outputs|)
  - Can work well even if output space is exponential in the input

- However, sampling an output may not be computationally efficient if output space is large.

Outline

- Recap

- Algorithmic Primitives
  - Randomized Response
  - Laplace Mechanism
  - Exponential Mechanism

- Composition Theorems

Sequential Composition

- If $M_1, M_2, \ldots, M_k$ are algorithms that access a private database D such that each M_i satisfies ε_i -differential privacy,

  then the combination of their outputs satisfies ε-differential privacy with

$$\varepsilon=\varepsilon_1+\ldots+\varepsilon_k$$

Privacy as Constrained Optimization

- Three axes
  - Privacy
  - Error
  - Queries that can be answered

- E.g.: Given a fixed set of queries and **privacy budget** ε, what is the minimum error that can be achieved?

Parallel Composition

- If M_1, M_2, ..., M_k are algorithms that access disjoint databases D_1, D_2, …, D_k such that each M_i satisfies ε_i -differential privacy,

  then the combination of their outputs satisfies
  ε-differential privacy with $\varepsilon = \max\{\varepsilon_1,...,\varepsilon_k\}$

Postprocessing

- If M_1 is an εdifferentially private algorithm that accesses a private database D,

  then outputting $M_2(M_1(D))$ also satisfies ε-differential privacy.

Summary

- An algorithm is differentially private if its output is insensitive to the presence or absence of a single row.

- Building blocks
  - Randomized Response
  - Laplace mechanism
  - Exponential Mechanism

- Composition rules help build complex algorithms using building blocks

Next class

- More on Composition

References

[W65] Warner, “Randomized Response” JASA 1965

[DMNS06] Dwork, McSherry, Nissim, Smith, “Calibrating noise to sensitivity in private data analysis”, TCC 2006

[MT07] McSherry, Talwar, “Mechanism Design via Differential Privacy”, FOCS 2007