Module 2: Semantic Privacy

Privacy for Data Analysis and ML
CS848 Fall 2024

UNIVERSITY OF WATERLOO | DSg Data Systems Group

Logistics

- Project
  - Project ideas will be posted on Learn (next Tue noon)
  - Start brainstorm your project
  - Choose project due is Sep 24
  - Project proposal due is Oct 3

- Paper reading and presentation
  - Site: https://uauw-fall2024privacy.hotcrp.com/
  - Link and more instructions will be sent to your email (by next Thur class)

Recap: Empirical Privacy

1. De-anonymizing Data:
   A case study on de-anonymizing Netflix data

2. Measures of Anonymity/Privacy:
   k-Anonymity, l-Diversity, t-Closeness

3. Privacy Attacks Practicum:
   Privacy desiderata

4. Privacy Risks in ML:
   Membership inference attacks

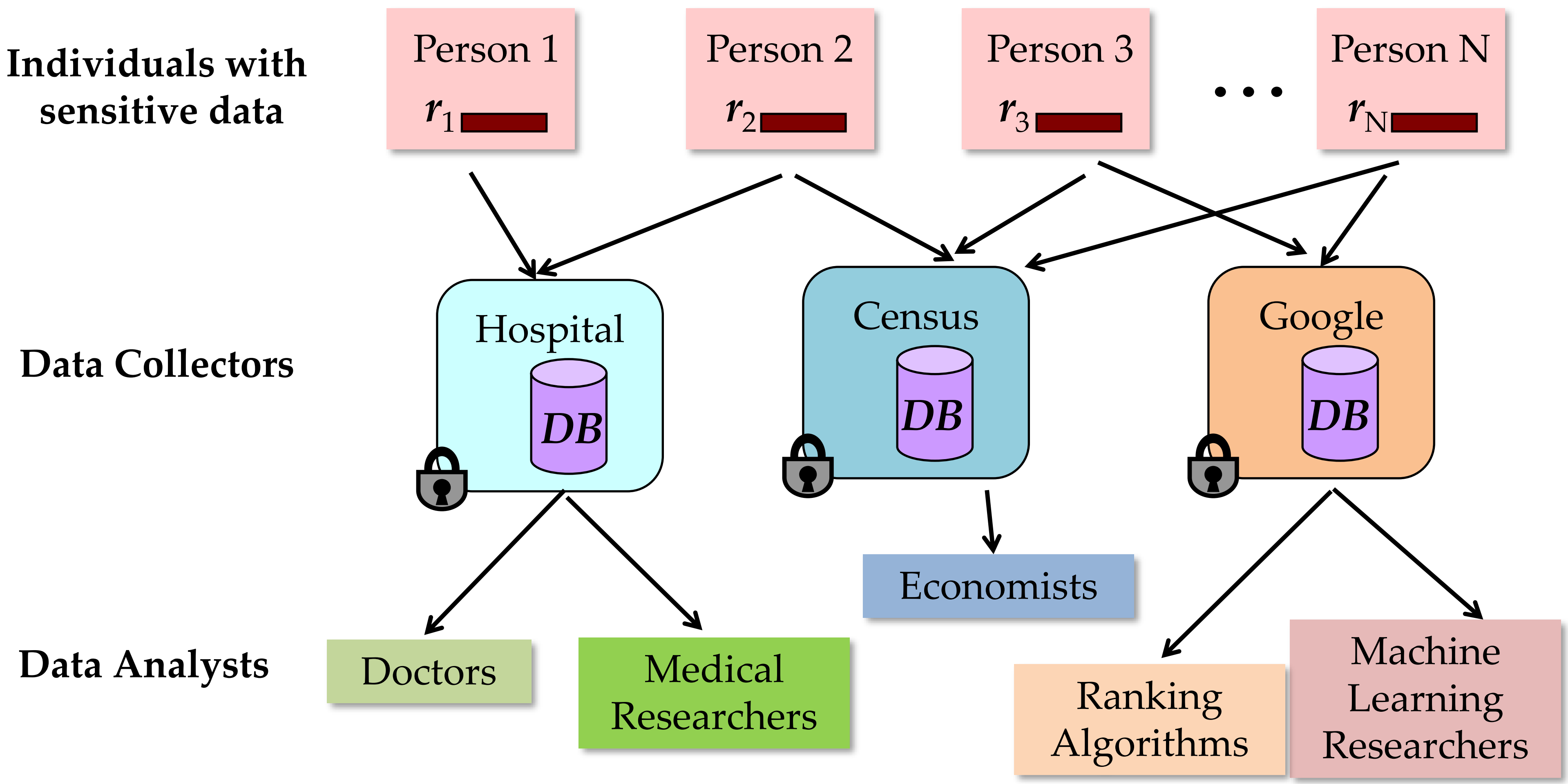
Module 2: Semantic Privacy

- Problem (30 mins)
  - Why Differential Privacy (DP)?
- Basic DP Algorithms (45 mins)
  - Building blocks for DP
- Designing Complex DP Algorithms (60 mins)
  - Composition and in-class exercises

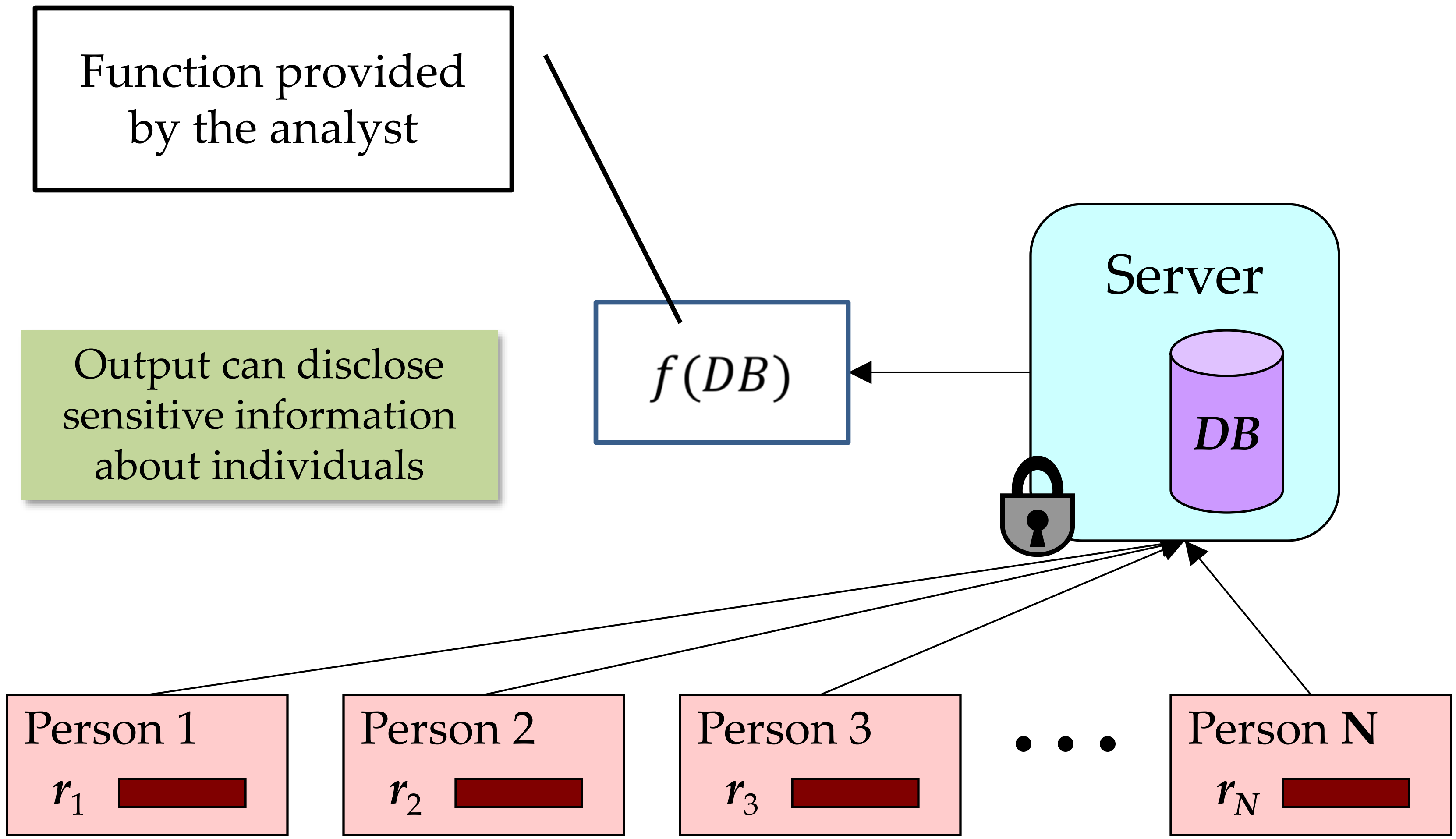
Why Differential Privacy (DP)?

PROBLEM

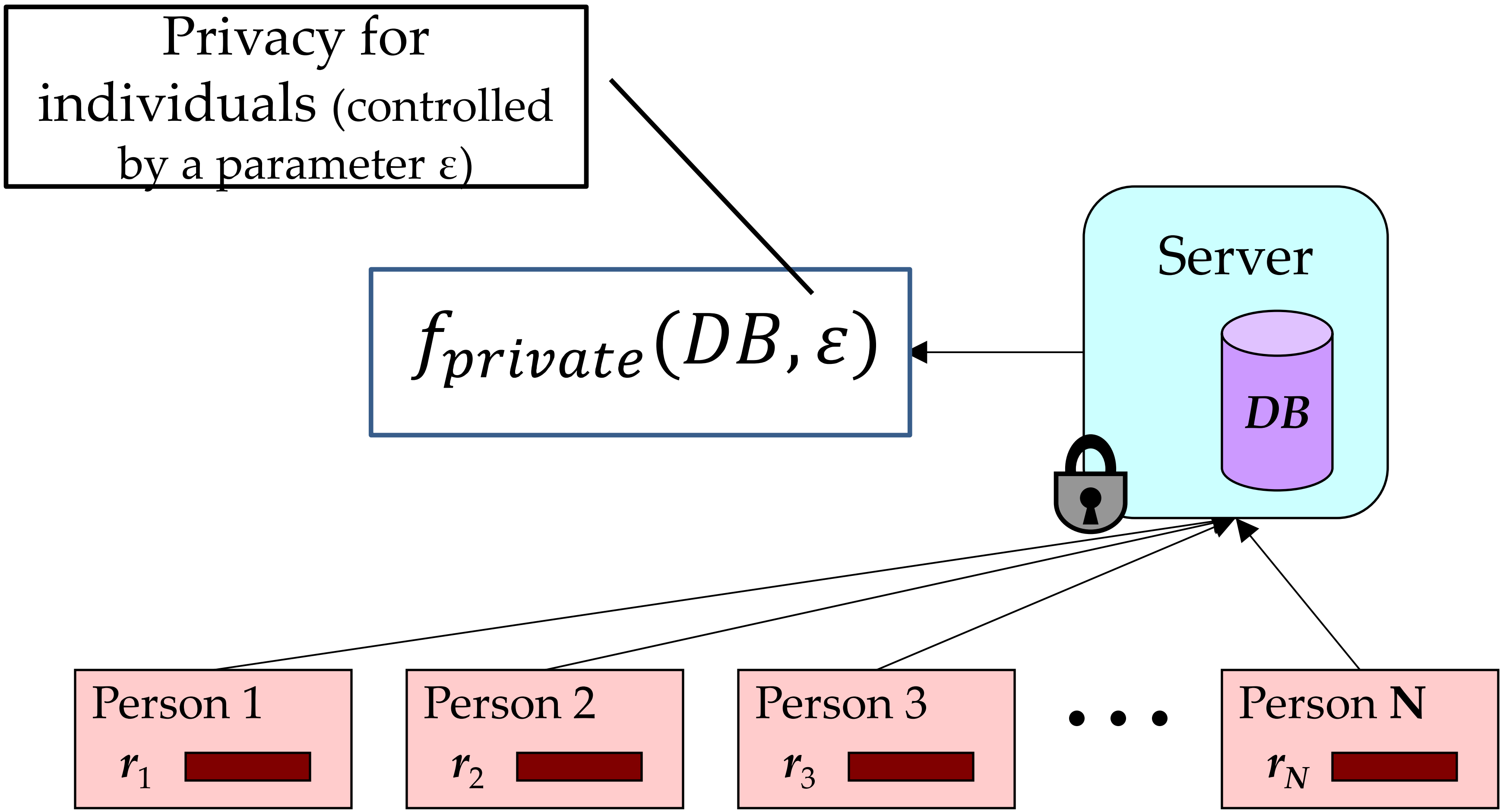
Statistical Databases



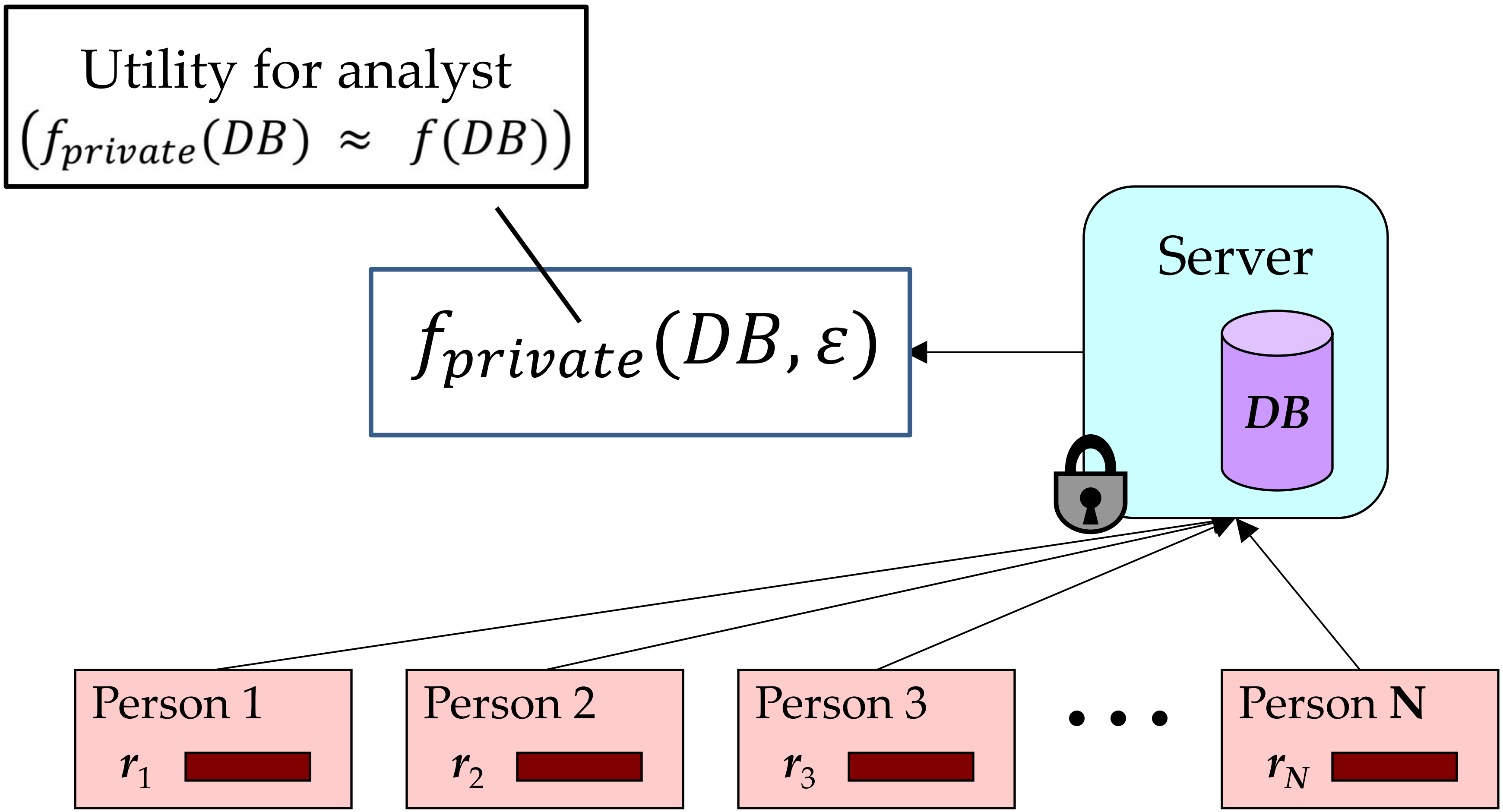
Statistical Database Privacy



Statistical Database Privacy

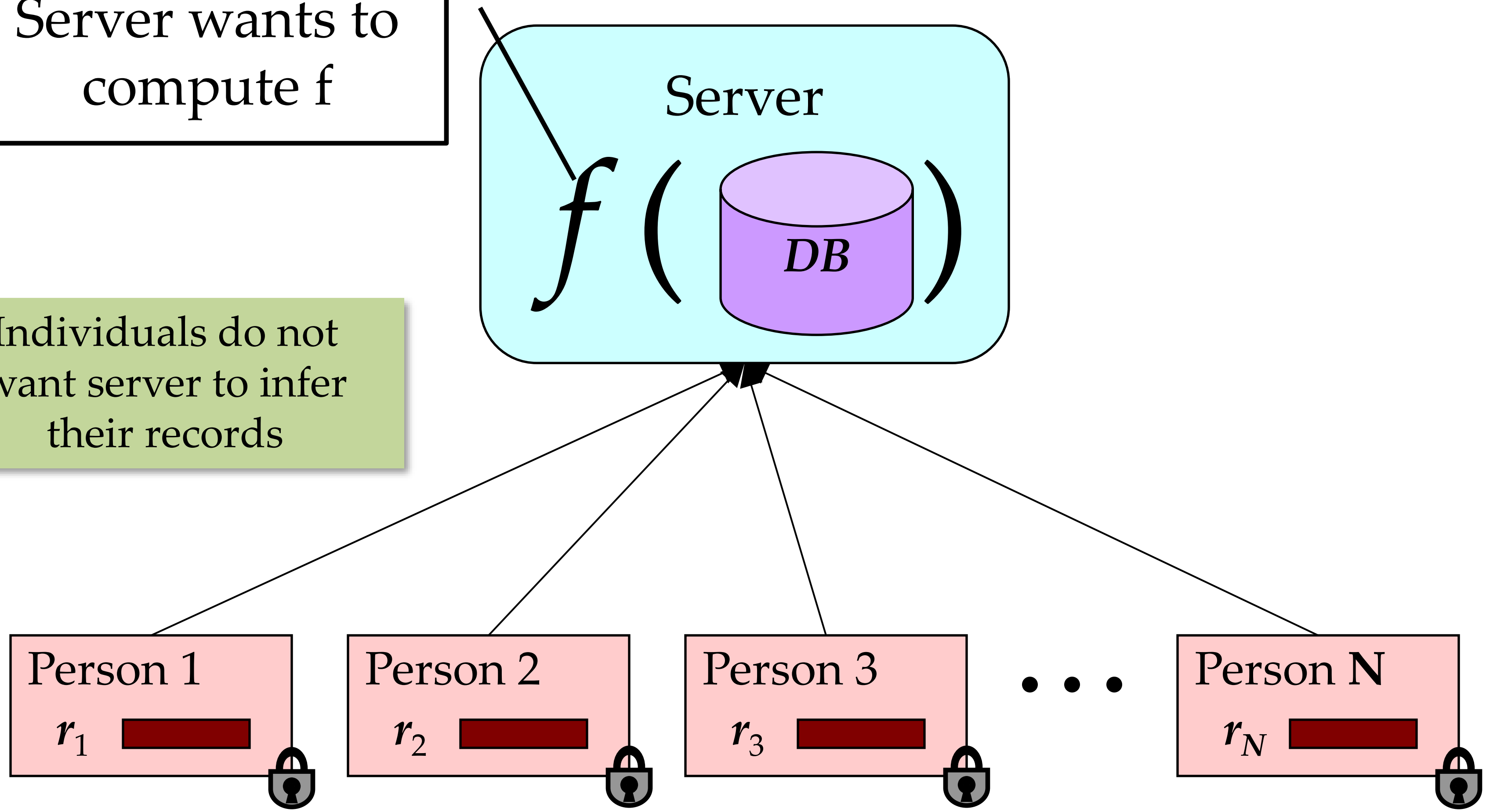


Statistical Database Privacy



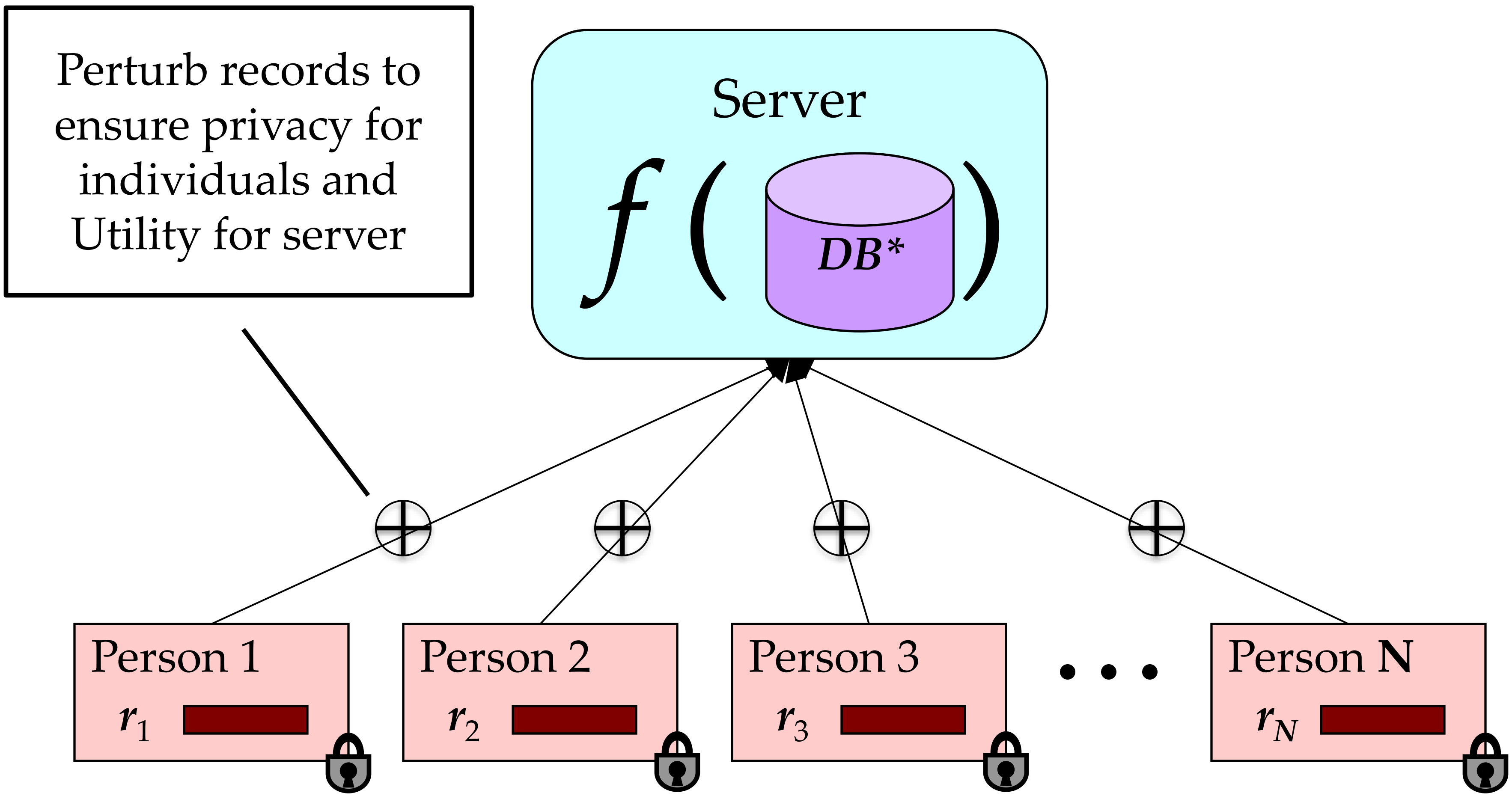
Statistical Database Privacy (untrusted collector)

Server wants to compute f



Individuals do not want server to infer their records

Statistical Database Privacy (untrusted collector)



Statistical Databases in real-world applications

Application	Data Collector	Private Information	Analyst	Function (utility)
Medical	Hospital	Disease	Epidemiologist	Correlation between disease and geography
Genome analysis	Hospital	Genome	Statistician/ Researcher	Correlation between genome and disease
Advertising	Google/FB	Clicks/Browsing	Advertiser	Number of clicks on an ad by age/region/gender …
Social Recommen-dations	Facebook	Friend links / profile	Another user	Recommend other users or ads to users based on social network

Statistical Databases in real-world applications

- Settings where data collector may not be trusted (or may not want the liability …)

Application	Data Collector	Private Information	Function (utility)
Location Services	Verizon/AT&T	Location	Traffic prediction
Recommen-dations	Amazon/Google	Purchase history	Recommendation model
Traffic Shaping	Internet Service Provider	Browsing history	Traffic pattern of groups of users

Privacy is *not* …

Statistical Database Privacy is not …

- Encryption:

Statistical Database Privacy is not …

- Encryption:
  Alice sends a message to Bob such that Trudy (attacker) does not learn the message. Bob should get the correct message …

- Statistical Database Privacy:
  Bob (attacker) can access a database
  - Bob must learn aggregate statistics, but
  - Bob must not learn new information about individuals in database.

Statistical Database Privacy is not …

- Computation on Encrypted Data:

Statistical Database Privacy is not …

- Computation on Encrypted Data:
  - Alice stores encrypted data on a server controlled by Bob (attacker).
  - Server returns correct query answers to Alice, without Bob learning *anything* about the data.

- Statistical Database Privacy:
  - Bob is allowed to learn aggregate properties of the database.

Statistical Database Privacy is not …

- The Millionaires Problem:

Statistical Database Privacy is not …

- Secure Multiparty Computation:
  - A set of agents each having a private input xi …
  - … Want to compute a function f(x1, x2, …, xk)
  - Each agent can learn the true answer, but must learn no other information than what can be inferred from their private input and the answer.

- Statistical Database Privacy:
  - Function output *must not disclose* individual inputs.

Statistical Database Privacy is not …

- Access Control:

Statistical Database Privacy is not …

- Access Control:
  - A set of agents want to access a set of resources (could be files or records in a database)
  - Access control rules specify who is allowed to access (*or not access*) certain resources.
  - 'Not access' usually means no information must be disclosed

- Statistical Database:
  - A single database and a single agent
  - Want to release aggregate statistics about a set of records without allowing access to individual records

Privacy Problems

- In today's systems a number of privacy problems arise:
  - Encryption when communicating data across a unsecure channel
  - Secure Multiparty Computation when different parties want to compute on a function on their private data without using a centralized third party
  - Computing on encrypted data when one wants to use an unsecure cloud for computation
  - Access control when different users own different parts of the data

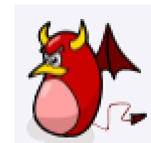
- Statistical Database Privacy:
  Quantifying (and bounding) the amount of information disclosed about individual records by the output of a valid computation.

What *is* privacy?

Privacy Breach: Attempt 1

A privacy mechanism M(D)

that allows

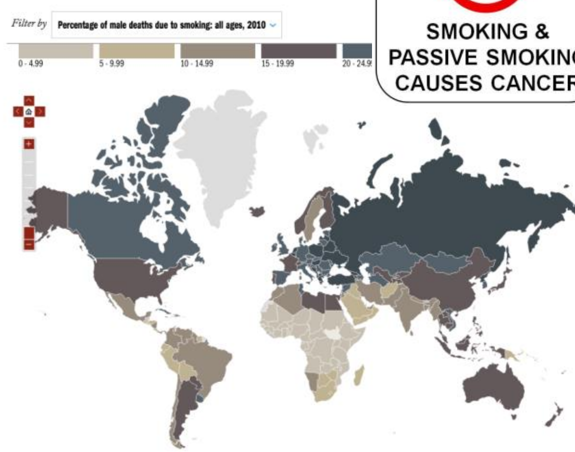
an unauthorized party 

to learn sensitive information about any individual in D,

which  could not have learnt without access to M(D).

Alice

+



=

Alice has Cancer

Is this a privacy breach? NO

Privacy Breach: Attempt 2

A privacy mechanism M(D) that allows
an unauthorized party 
to learn sensitive information about
any individual Alice in D,

which  could not have learnt even with access to M(D)
if Alice was *not in the dataset*.

Differential Privacy

[Dwork ICALP 2006]

For every pair of inputs
that differ in one row

For every output …

D_1 D_2 O

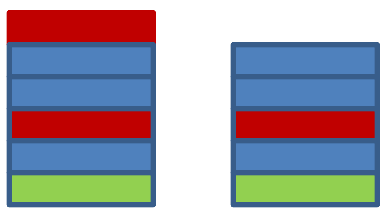
Adversary should not be able to distinguish
between any D_1 and D_2 based on any O

$$\ln\left(\frac{\Pr[A(D_1) = o]}{\Pr[A(D_2) = o]}\right) \leq \varepsilon, \qquad \varepsilon > 0$$

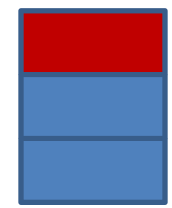
Why pairs of datasets *that differ in one row*?

For every pair of inputs that differ in one row

For every output …



D_1 D_2



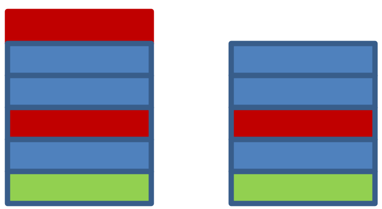
O

Simulate the presence or absence of a single record

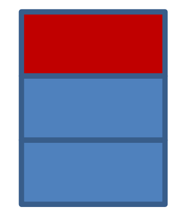
Why *all* pairs of datasets …?

For every pair of inputs that differ in one row

For every output …



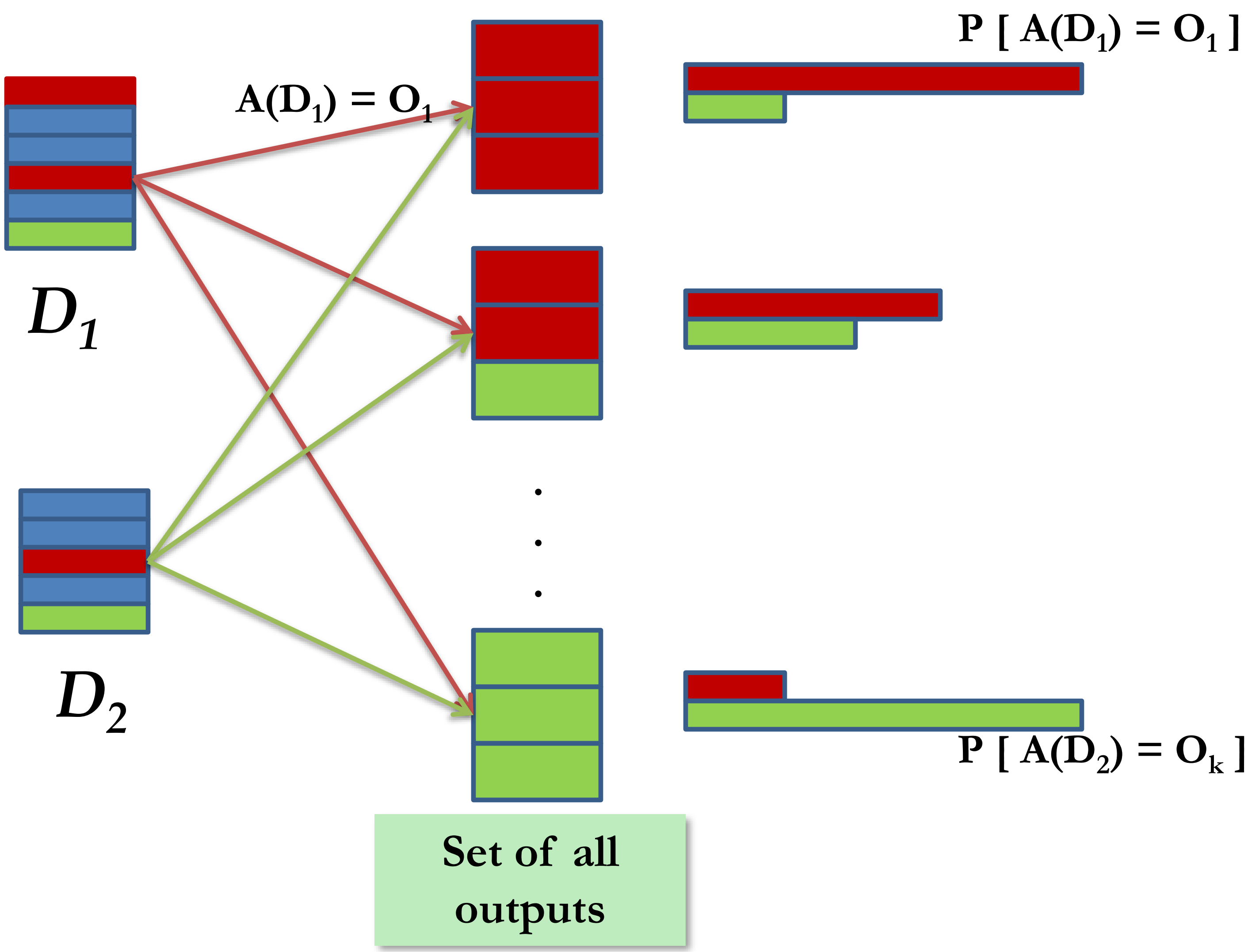
D_1 D_2



O

Guarantee holds no matter what the other records are.

Why *all* outputs?

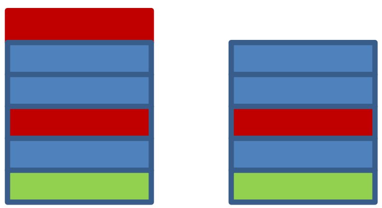


Should not be able to distinguish whether input was D_1 or D_2 no matter what the output

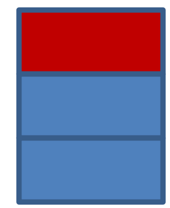
Privacy Parameter ε

For every pair of inputs that differ in one row

For every output …



D_1 D_2



O

$$\Pr[A(D_1) = o] \leq e^{\varepsilon} \Pr[A(D_2) = o]$$

Controls the degree to which D_1 and D_2 can be distinguished.
Smaller the ε more the privacy (and worse the utility)

Desiderata for a Privacy Definition

1. Resilience to background knowledge
   – A privacy mechanism must be able to protect individuals' privacy from attackers who may possess background knowledge

2. Privacy without obscurity
   – Attacker must be assumed to know the algorithm used as well as all parameters [MK15]

3. Post-processing
   – Post-processing the output of a privacy mechanism must not change the privacy guarantee [KL10, MK15]

4. Composition over multiple releases
   – Allow a graceful degradation of privacy with multiple invocations on the same data [DN03, GKS08]

Building blocks for DP

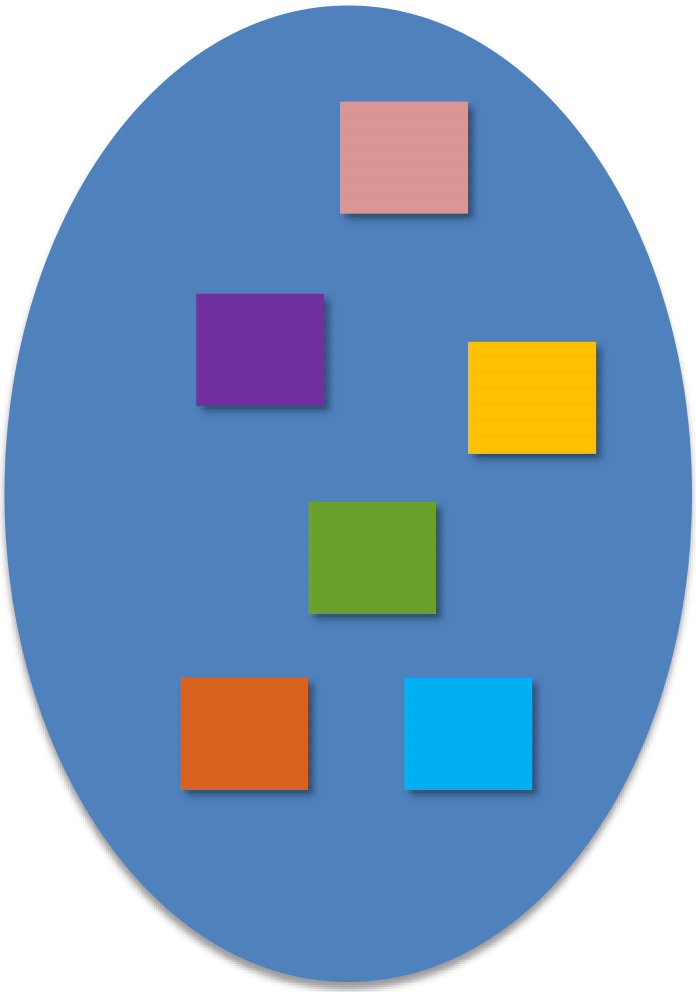
BASIC DP ALGORITHMS

Basic DP Algorithms

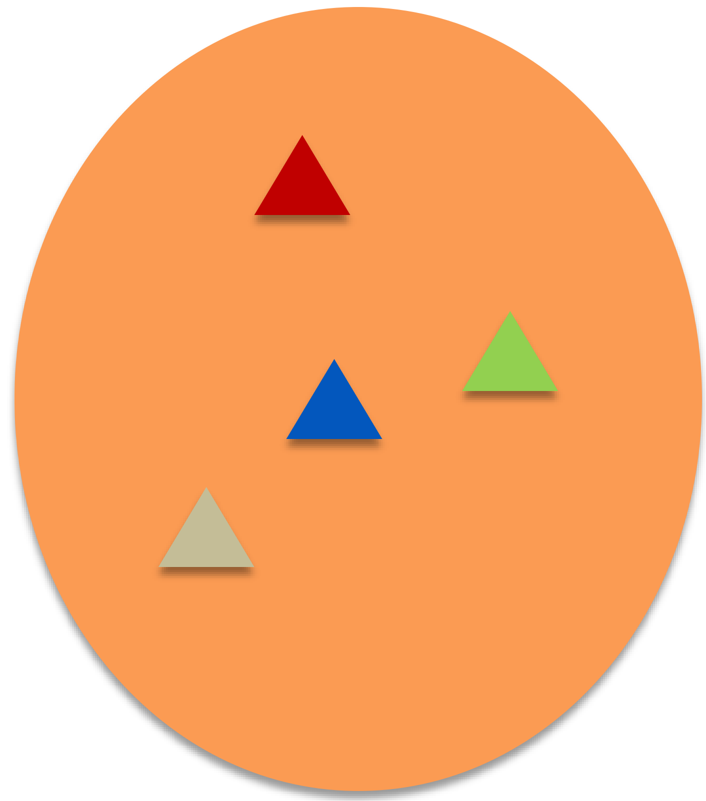
- Randomized Response
- Laplace Mechanism
- Exponential Mechanism
- Gaussian Mechanism
- Noisy Max
- Sparse Vector Technique
- Sample and Aggregate
- ….

Non-trivial deterministic Algorithms do not satisfy differential privacy

Space of all inputs

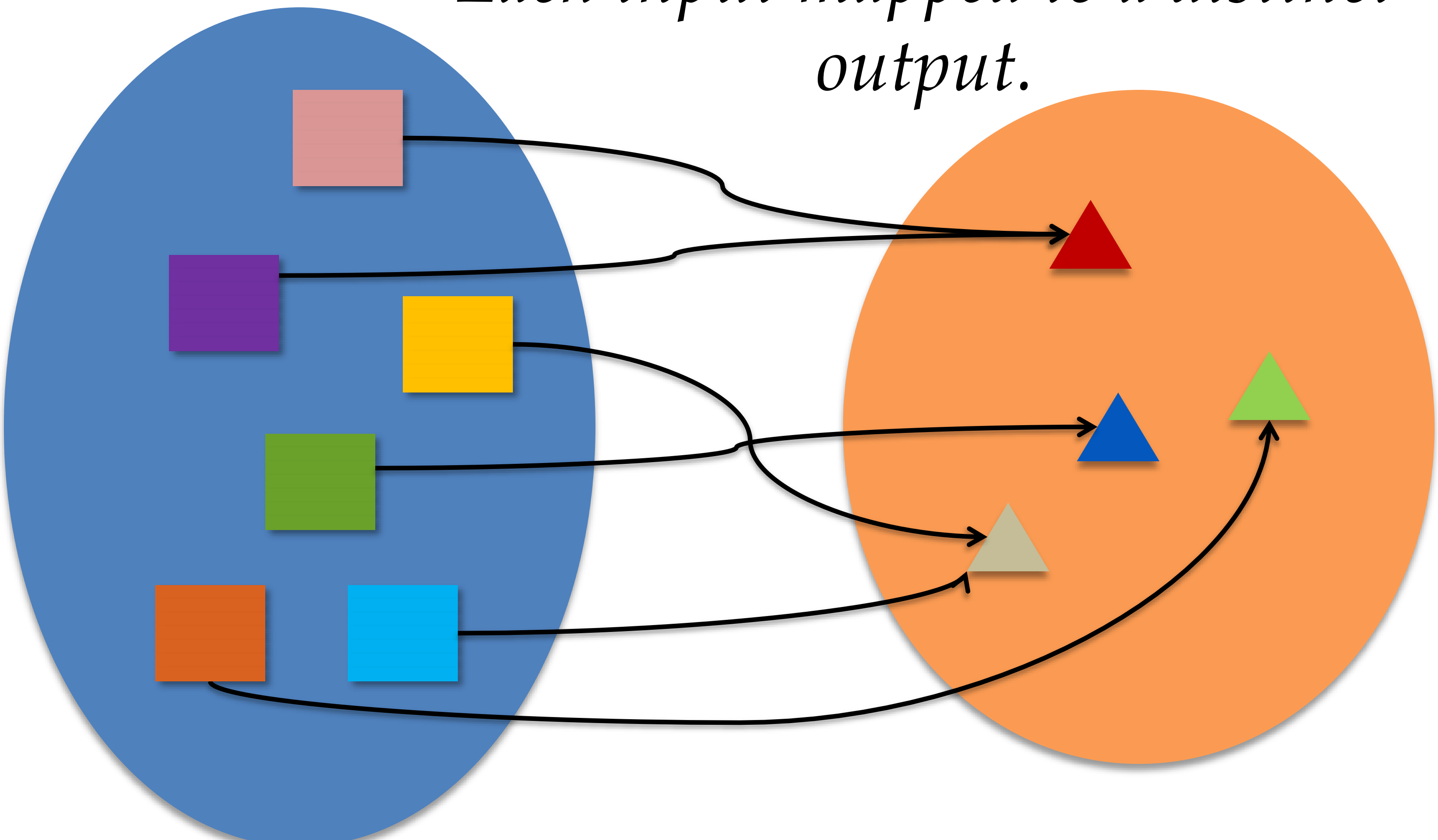


Space of all outputs (at least 2 distinct ouputs)

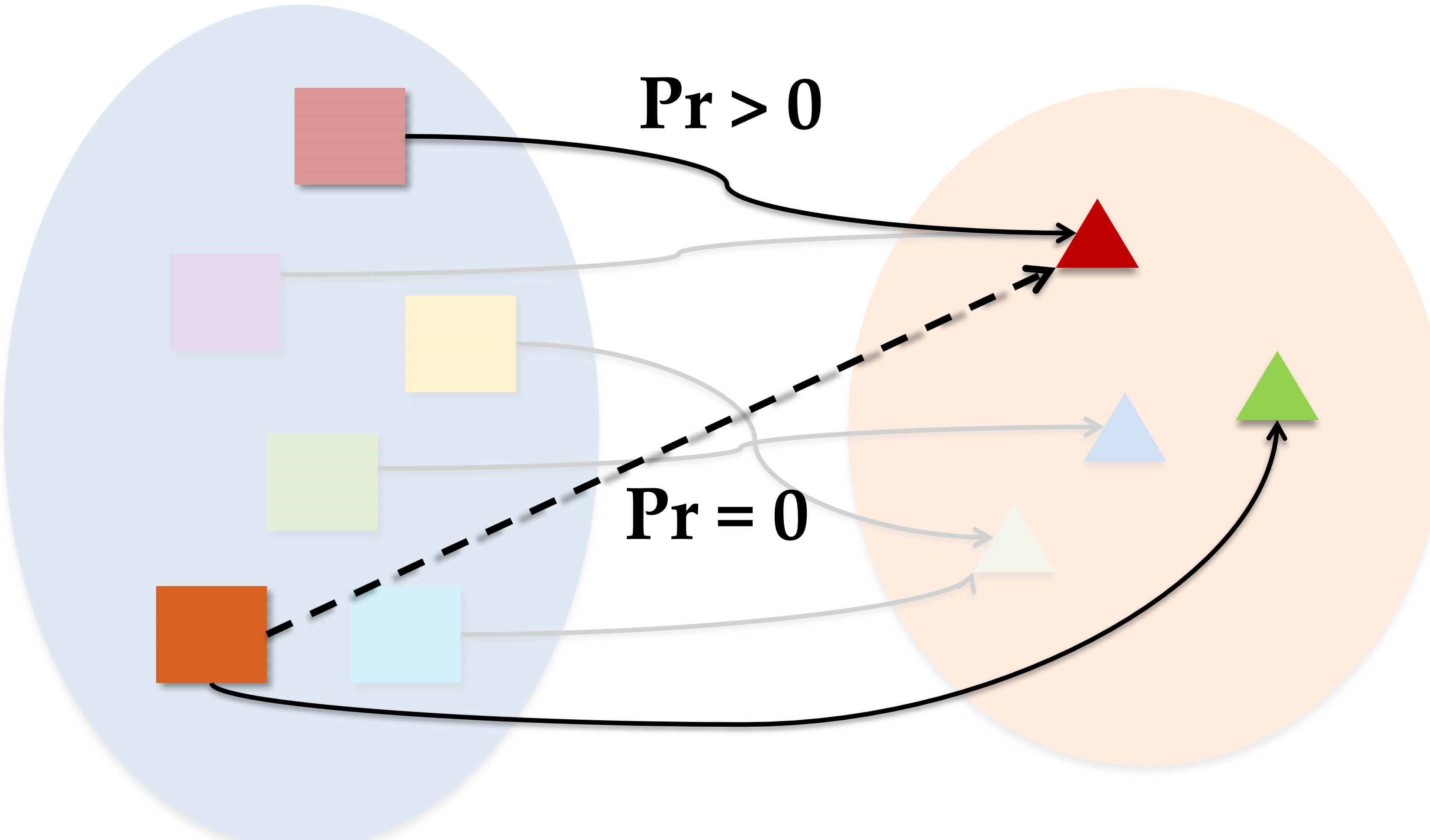


Non-trivial deterministic Algorithms do not satisfy differential privacy

Each input mapped to a distinct output.



There exist two inputs that differ in one entry mapped to different outputs.



Random Sampling …

… also does not satisfy differential privacy

Input

Output

D_1 D_2 O

$$\Pr[D_2 \rightarrow O] = 0 \text{ implies } \log\left(\frac{\Pr[D_1 \rightarrow O]}{\Pr[D_2 \rightarrow O]}\right) = \infty$$

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

[W 65]

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?

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

- $E(\widehat{\pi}) = \pi$

$E(y) = p\pi N + (1-p)(1-\pi)N$

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

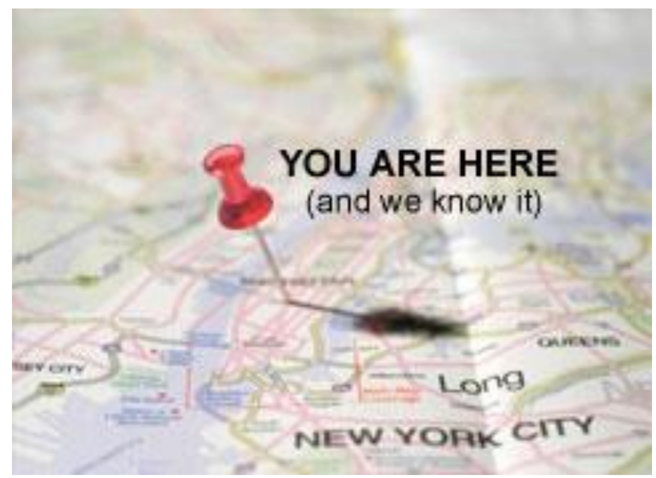
Sampling **Variance due to coin flips**

  - $Std(\hat{\pi}) = \Theta\left(\frac{1}{\sqrt{N}}\right); Std(\hat{\pi}N) = \Theta\left(\sqrt{N}\right)$

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:

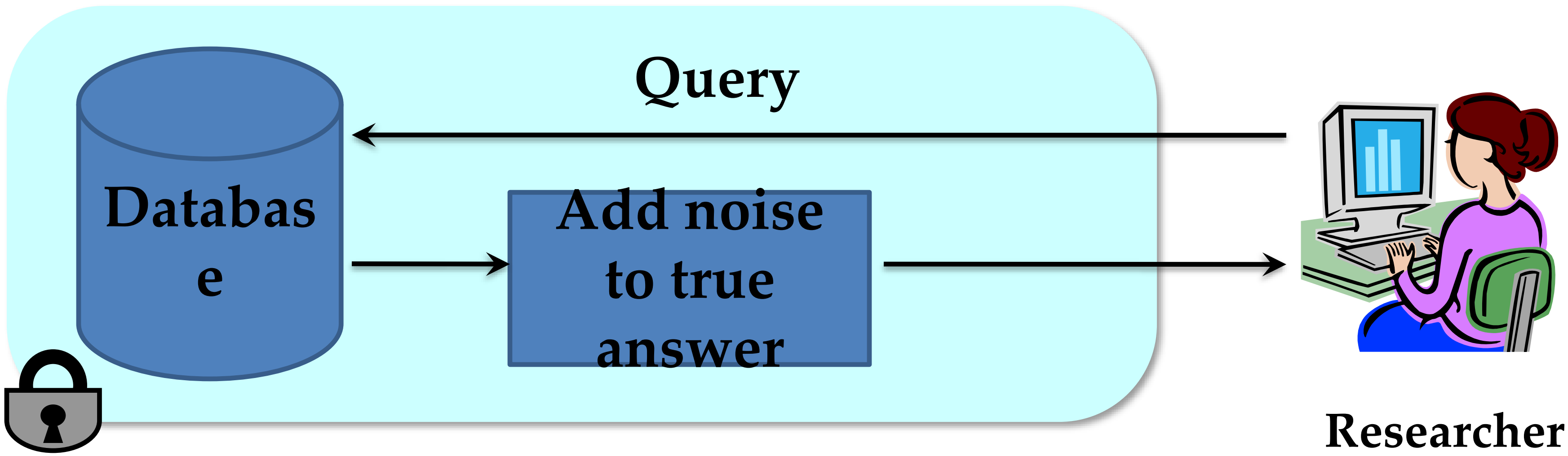
- Report true position: p
- Report any other position: q (< p)

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

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

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

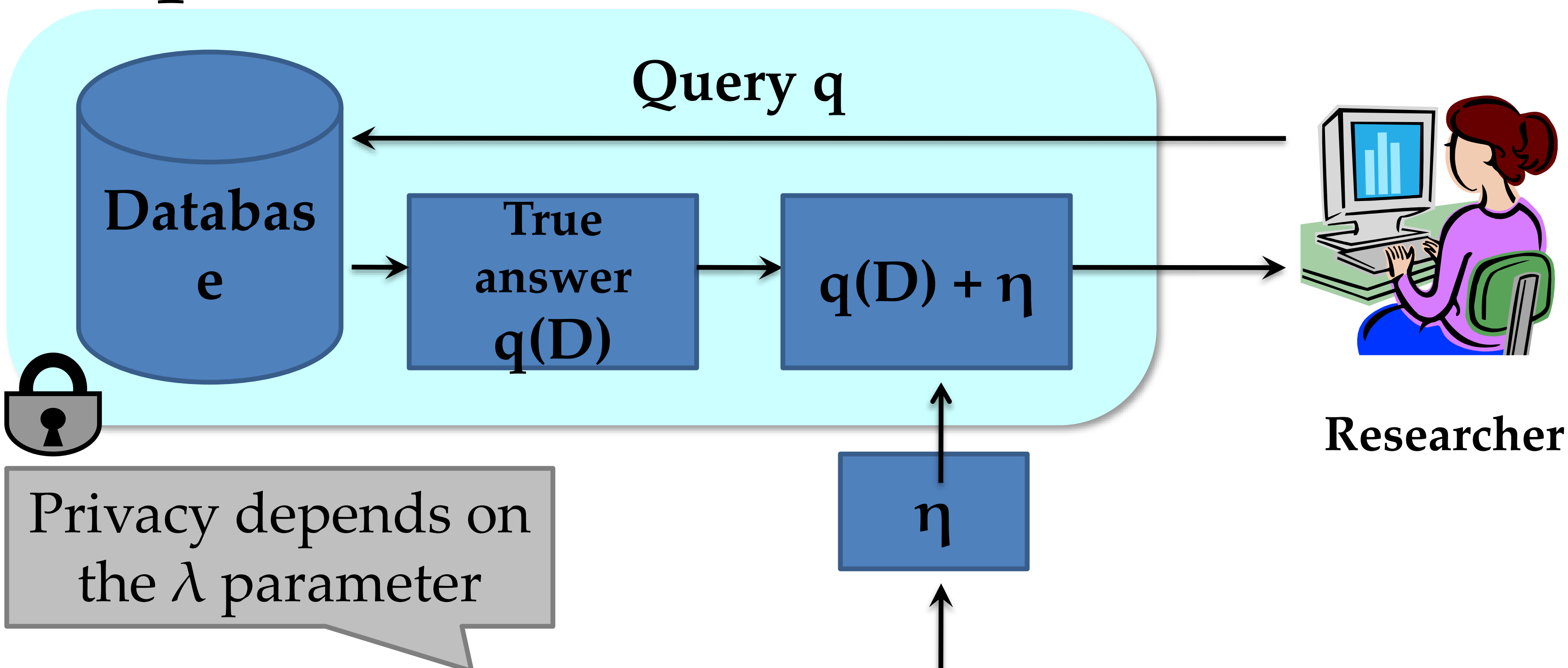
Output Randomization



- 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.

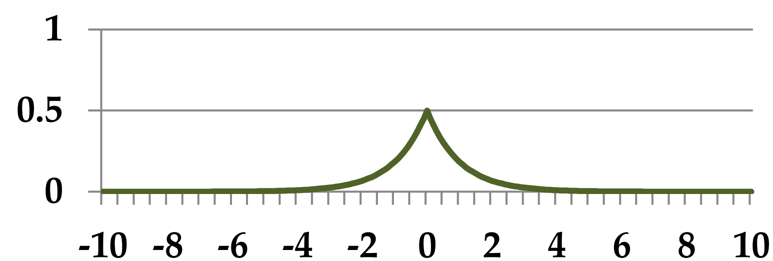
[DMNS 06]

Laplace Mechanism



$h(\eta) \propto \exp(-|\eta| / \lambda)$

Laplace Distribution –
Lap(λ)



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

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

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

h(η) α exp(-|η| / λ)

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

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

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

Utility of Laplace Mechanism

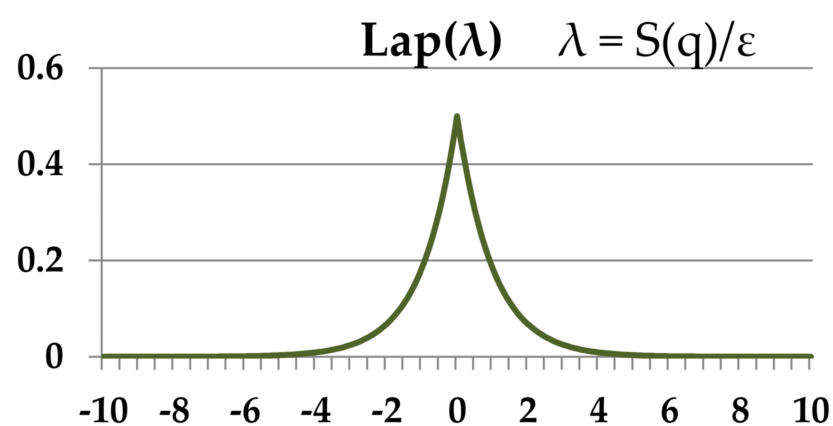
- Laplace mechanism works for **any function** that returns a real number

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

$$= E[(Lap(\lambda))^2]$$

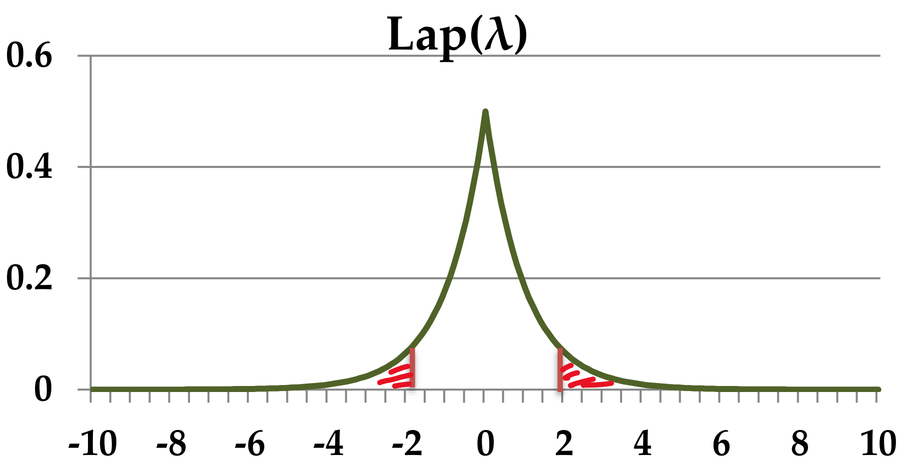
$$= E[(Lap(\lambda))^2] - E[Lap(\lambda)]^2 = Var(Lap(\lambda))$$

$$= 2\lambda^2 = 2 * S(q)^2 / \varepsilon^2$$



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 (RR)

Privacy

- Provide the same ε-DP
- Laplace mechanism assumes data collected is trusted
- RR does not require data collected to be trusted
  - Also called a *Local* Algorithm, since each record is perturbed

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 RR answer: $O(\sqrt{N}/\varepsilon)$

Basic DP Algorithms

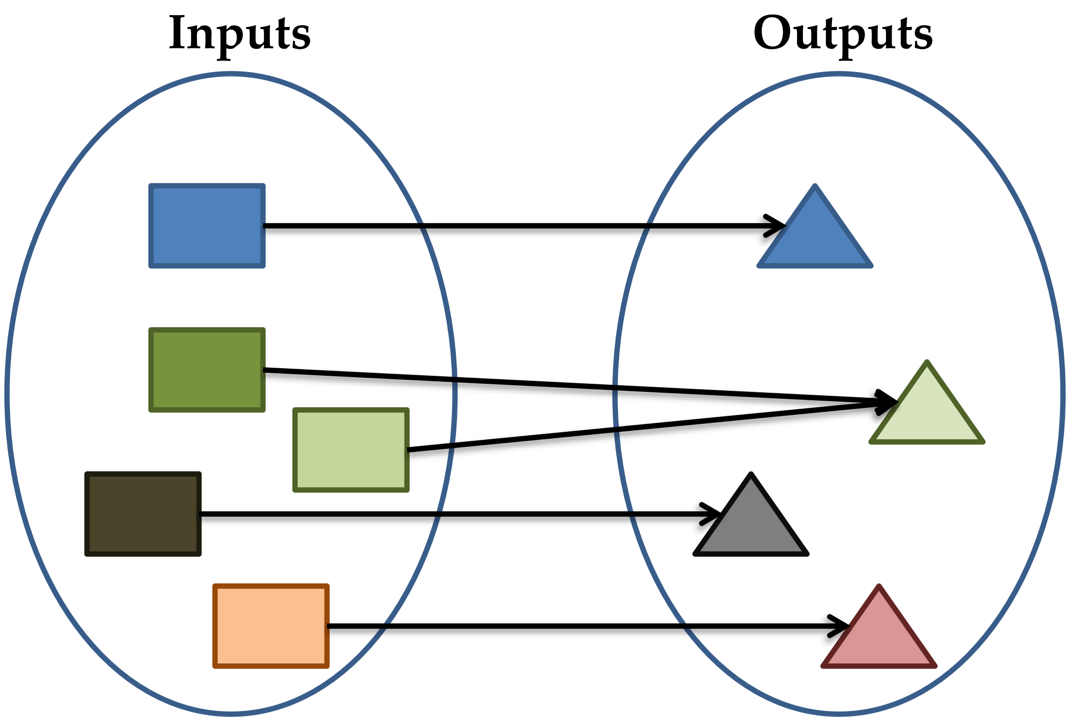
- Randomized Response
- Laplace Mechanism
- **Exponential Mechanism**
- Gaussian Mechanism
- Noisy Max
- Sparse Vector Technique
- Sample and Aggregate
- ….

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

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: $Inputs \times \ Outputs \rightarrow R$
  - D: nationalities of a set of people
  - $\#(D, O)$: # people with nationality O
  - $f(D)$: most frequent nationality in D
  - A possible score function

$$w(D, O) \ = \ \#(D, O) \ - \ \#(D, f(D))$$

- Sensitivity of w:

$$\mathrm{S}_w = \max_{O,D,D':\,|D\Delta D'|=1} |w(D, O) - w(D', O)|$$

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 $1\text{-}e^{-3}$(=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 $f(D) + \eta$ 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

Randomized Response with parameter ε/2

- Exponential mechanism:

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

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.

Basic DP Algorithms

- Randomized Response
- Laplace Mechanism
- Exponential Mechanism
- **Gaussian Mechanism**
- Noisy Max
- Sparse Vector Technique
- Sample and Aggregate
- ….

Gaussian Mechanism

- The L2–sensitivity of $f: \mathcal{D} \to \mathbb{R}^d$ is:

$$S_2(f) = \max_{D,D' : |D \Delta D'|=1} \|f(D) - f(D')\|_2$$

- **Gaussian mechanism** adds noise scaled to $N(0, \sigma^2)$ to each d component of the output → satisfies (ϵ, δ)-DP if $\sigma \geq cS_2(f)/\epsilon$ for $c^2 > 2\ln\frac{1.25}{\delta}, \epsilon \in (0,1)$

(ϵ, δ)-DP: $\forall S$

$$\Pr[M(D) \in S] \leq e^{\epsilon}\Pr[M(D') \in S] + \delta$$

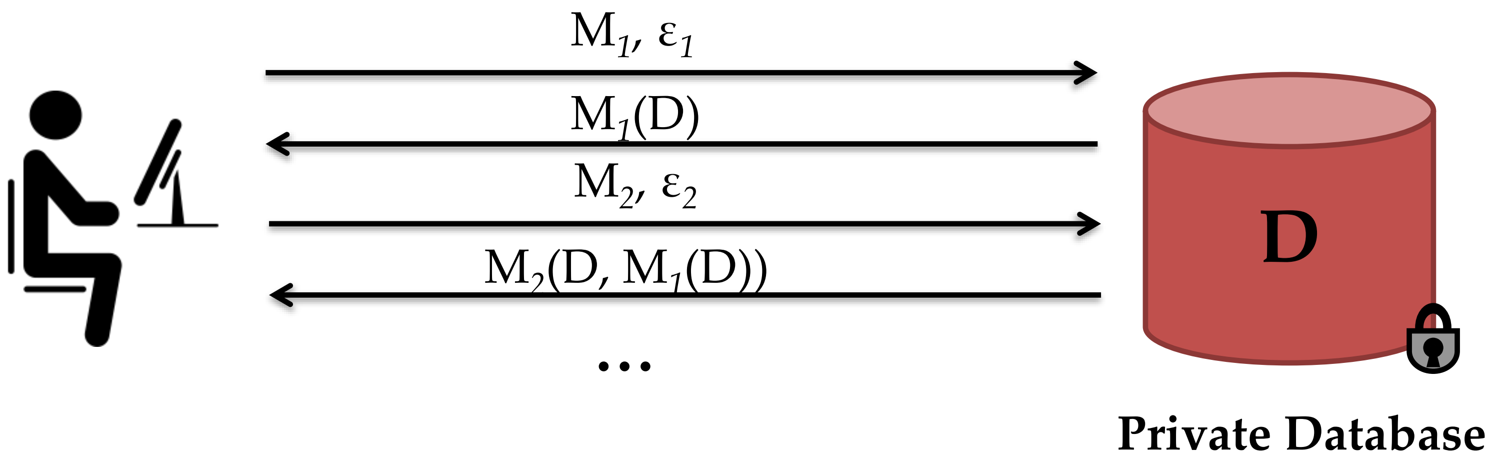
Take a break (5 mins)

- Download the in-class exercise (Jupyter Notebook) and datasets
  - https://cs.uwaterloo.ca/~xihe/cs848_f24/slides/DPExercises/

Composition and in-class exercises

BUILDING COMPLEX DP ALGORITHMS

Sequential Composition

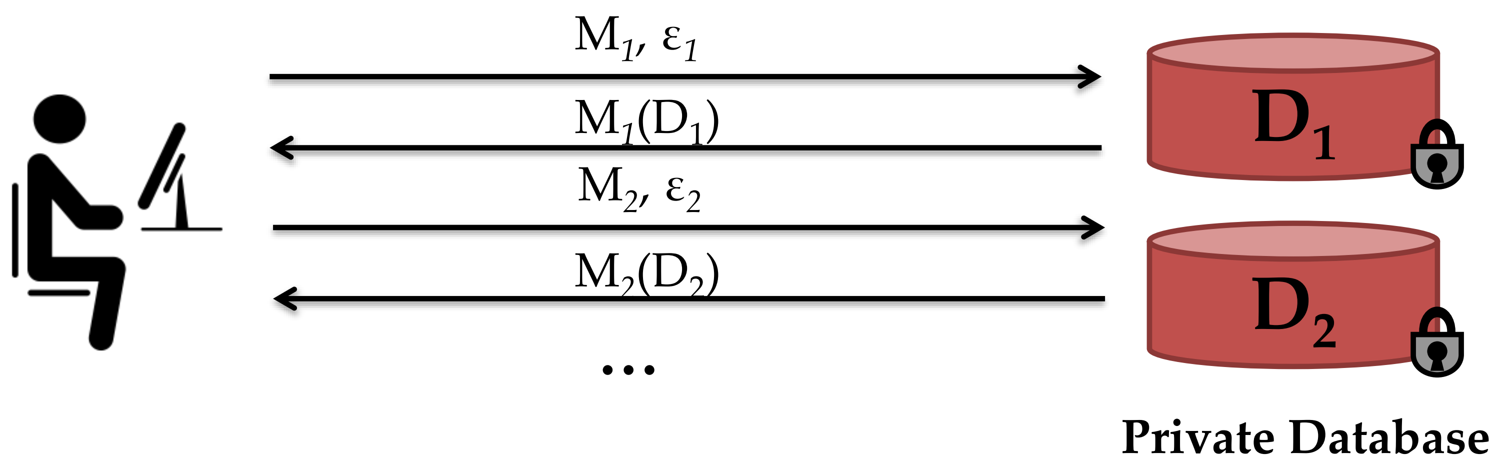


- If $M_1, M_2, \ldots, M_k$ are algorithms that access a private database D such that each M_i satisfies ε_i -differential privacy,

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

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

Parallel Composition

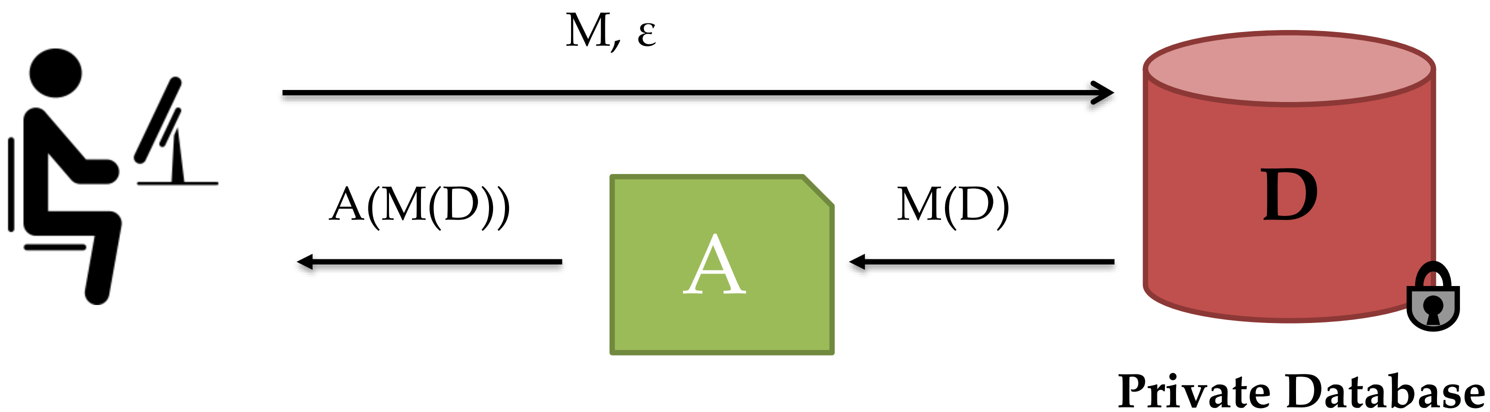


- If M_1, M_2, …, M_k are algorithms that access are algorithms that access disjoint databases D_1, D_2, …, D_k such that each M_i satisfies ε_i -differential privacy,

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

$$\varepsilon = \max(\varepsilon_1, \ldots, \varepsilon_k)$$

Postprocessing



- If M is an ε-differentially private algorithm, any additional post-processing $A \circ M$ also satisfies ε-differential privacy.

Building Complex DP Algorithms

- Composition

- **Problem 1: Answer multiple queries**
  - Examples
  - DP algorithms optimization

- Problem 2: DP Gradient Descent
  - Gradient descent
  - Better composition (RDP)

Problem 1: Answering Multiple Queries

Sex	Height	Weight
M	6′2″	210
F	5′3″	190
F	5′9″	160
M	5′3″	180
M	6′7″	250

Queries:

- # Males with BMI < 25
- # Males
- # Females with BMI < 25
- # Females

- Design an ε-differentially private algorithm that can answer all these questions.
- What is the total error?

Algorithm 1

Return:

- (# Males with BMI < 25) + Lap(4/ε)
- (# Males) + Lap(4/ε)
- (# Females with BMI) < 25 + Lap(4/ε)
- (# Females) + Lap(4/ε)

Privacy

- Sensitivity of count = 1. So each query is answered using a $\varepsilon/4$-DP algorithm.

- By sequential composition, we get ε-DP.

Utility

Error:

$$\sum E\left(\left(\tilde{q}(D) - q(D)\right)^2\right)$$

Total Error:

$Lap(\frac{4}{\epsilon})$ for each query

$$2\left(\frac{4}{\varepsilon}\right)^2 \times 4 = \frac{128}{\varepsilon^2}$$

Algorithm 2

Compute:

- $\widetilde{q_1}$ = (# Males with BMI < 25) + Lap(1/ε)
- $\widetilde{q_2}$ = (# Males with BMI > 25) + Lap(1/ε)
- $\widetilde{q_3}$ = (# Females with BMI < 25) + Lap(1/ε)
- $\widetilde{q_4}$ = (# Females with BMI > 25) + Lap(1/ε)

Return

- $\widetilde{q_1}$, $\widetilde{q_1}+\widetilde{q_2}$, $\widetilde{q_3}$, $\widetilde{q_3}+\widetilde{q_4}$

Privacy

- Sensitivity of count = 1. So each query is answered using a ε-DP algorithm.

- q_1, q_2, q_3, q_4 are counts on disjoint portions of the database. Thus by *parallel composition* releasing $\widetilde{q_1}, \widetilde{q_2}, \widetilde{q_3}, \widetilde{q_4}$ satisfies ε-DP.

- By the *postprocessing theorem,* releasing $\widetilde{q_1}, \widetilde{q_1}+\widetilde{q_2}, \widetilde{q_3}, \widetilde{q_3}+\widetilde{q_4}$ also satisfies ε-DP.

Utility

Error:

$$\sum E\left(\left(\tilde{q}(D) - q(D)\right)^2\right)$$

Tighter privacy analysis gives better accuracy for the same level of privacy

Total Error:

$$2\left(\frac{1}{\varepsilon}\right)^2 + 2 \cdot 2\left(\frac{1}{\varepsilon}\right)^2 + 2\left(\frac{1}{\varepsilon}\right)^2 + 2 \cdot 2\left(\frac{1}{\varepsilon}\right)^2 = \frac{12}{\varepsilon^2}$$

$\widetilde{q_1}$ $\widetilde{q_1} + \widetilde{q_2}$ $\widetilde{q_3}$ $\widetilde{q_3} + \widetilde{q_4}$

Generalized Sensitivity

- Let $f : \mathcal{D} \rightarrow \mathbb{R}^d$ be a function that outputs a vector of d real numbers. The L1-sensitivity of f is given by:

$$S_1(f) = \max_{D,D' : |D \Delta D'|=1} \|f(D) - f(D')\|_1$$

where $\|\mathbf{x} - \mathbf{y}\|_1 = \sum_i |x_i - y_i|$

Generalized Sensitivity

- q_1 = # Males with BMI < 25
- q_2 = # Males with BMI > 25
- q = # Males with BMI

- Let f_1 be a function that answers both q_1, q_2
- Let f_2 be a function that answers both q_1, q

- Sensitivity of $f_1 = 1$
- Sensitivity of $f_2 = 2$

- An alternate privacy proof for Alg 2 is to show that the generalized sensitivity of $\widetilde{q_1}$, $\widetilde{q_2}$, $\widetilde{q_3}$, $\widetilde{q_4}$ is 1.

Improving utility of Alg 2

Compute:

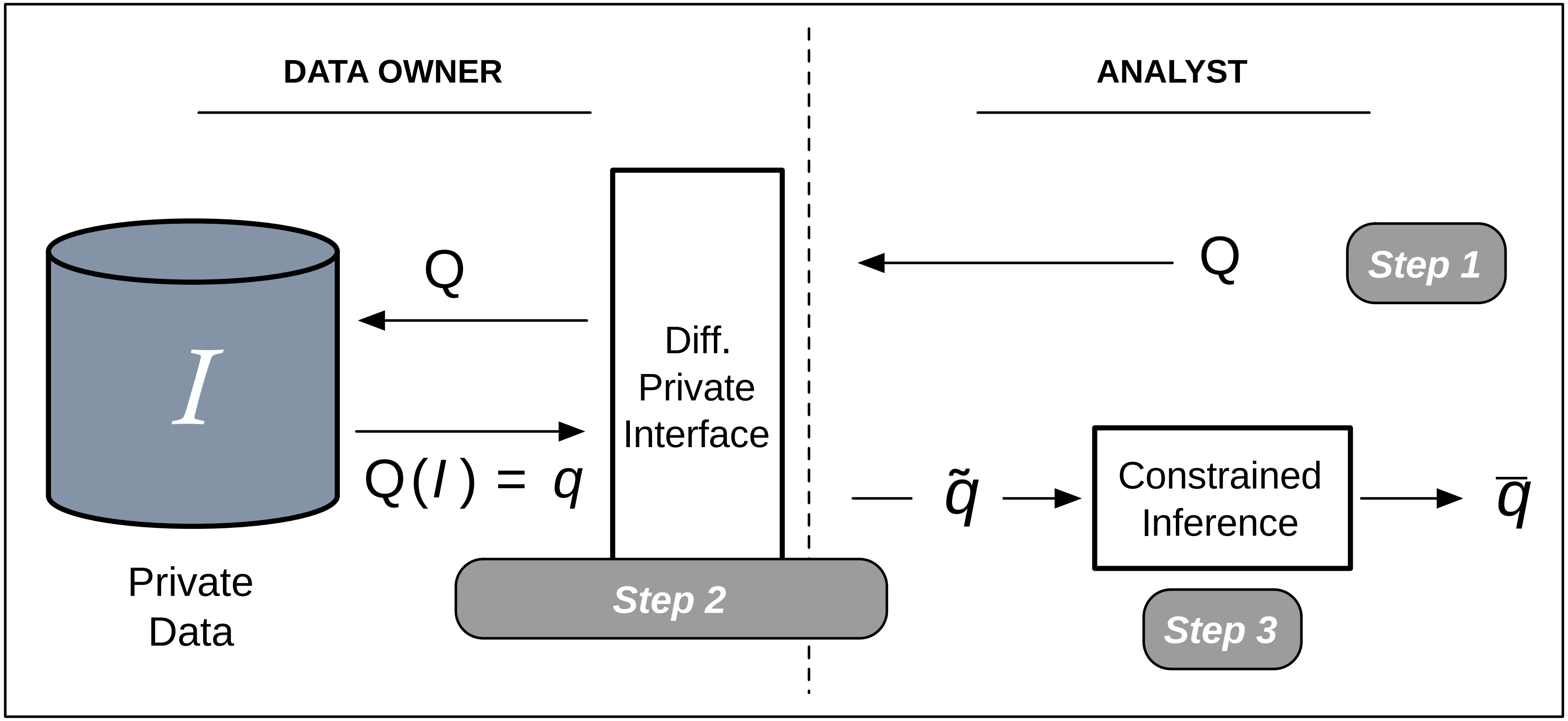
- $\widetilde{q_1}$ = # Males with BMI < 25 + Lap(1/ε)
- $\widetilde{q_2}$ = # Males with BMI > 25 + Lap(1/ε)

Return

- $\widetilde{q_1}$, $\widetilde{q_1}+\widetilde{q_2}$

We know $q_1 \leq q_1 + q_2$,
but $P[\widetilde{q_1} > \widetilde{q_1}+\widetilde{q_2}] > 0$

Constrained Inference



Constrained Inference

- $q_1, q_2, \ldots, q_k$ be a set of queries
- $\widetilde{q_1}, \widetilde{q_2}, \ldots, \widetilde{q_k}$ be the noisy answers
- Constraint C($q_1, q_2, \ldots, q_k$) = 1 holds on true answers (for all typical databases), but does not hold on noisy answers.

- Goal: Find $\overline{q_1}, \overline{q_2}, \ldots, \overline{q_k}$ that are:
  - Close to $\widetilde{q_1}, \widetilde{q_2}, \ldots, \widetilde{q_k}$
  - Satisfy the constraint C($\overline{q_1}, \overline{q_2}, \ldots, \overline{q_k}$)

Least Squares Optimization

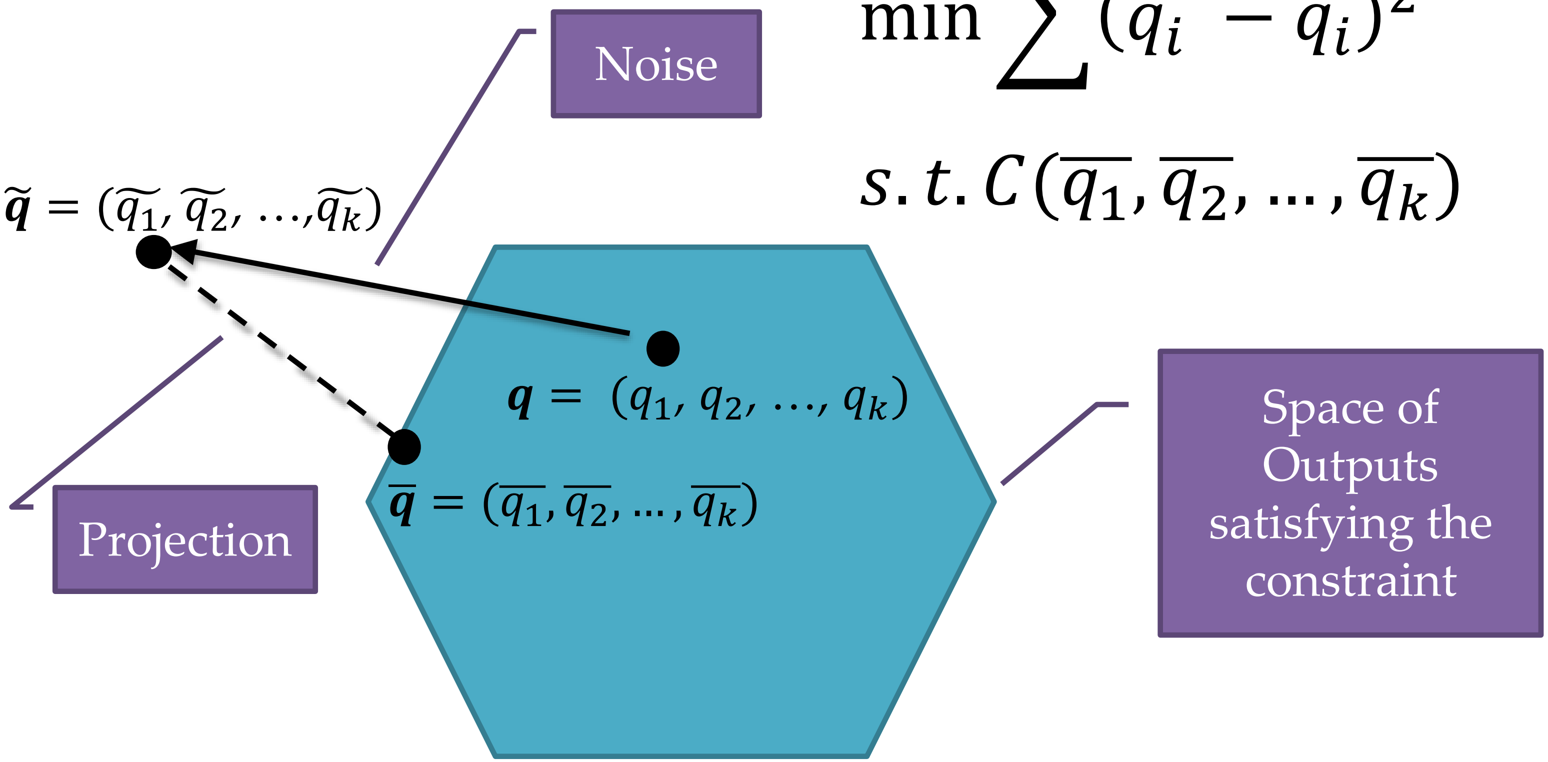
$$\min \sum (\widetilde{q_i} - \overline{q_i})^2$$

$$s.t.\ C(\overline{q_1}, \overline{q_2}, \ldots, \overline{q_k})$$

Geometric Interpretation

$$\min \sum (\tilde{q}_i - \bar{q}_i)^2$$

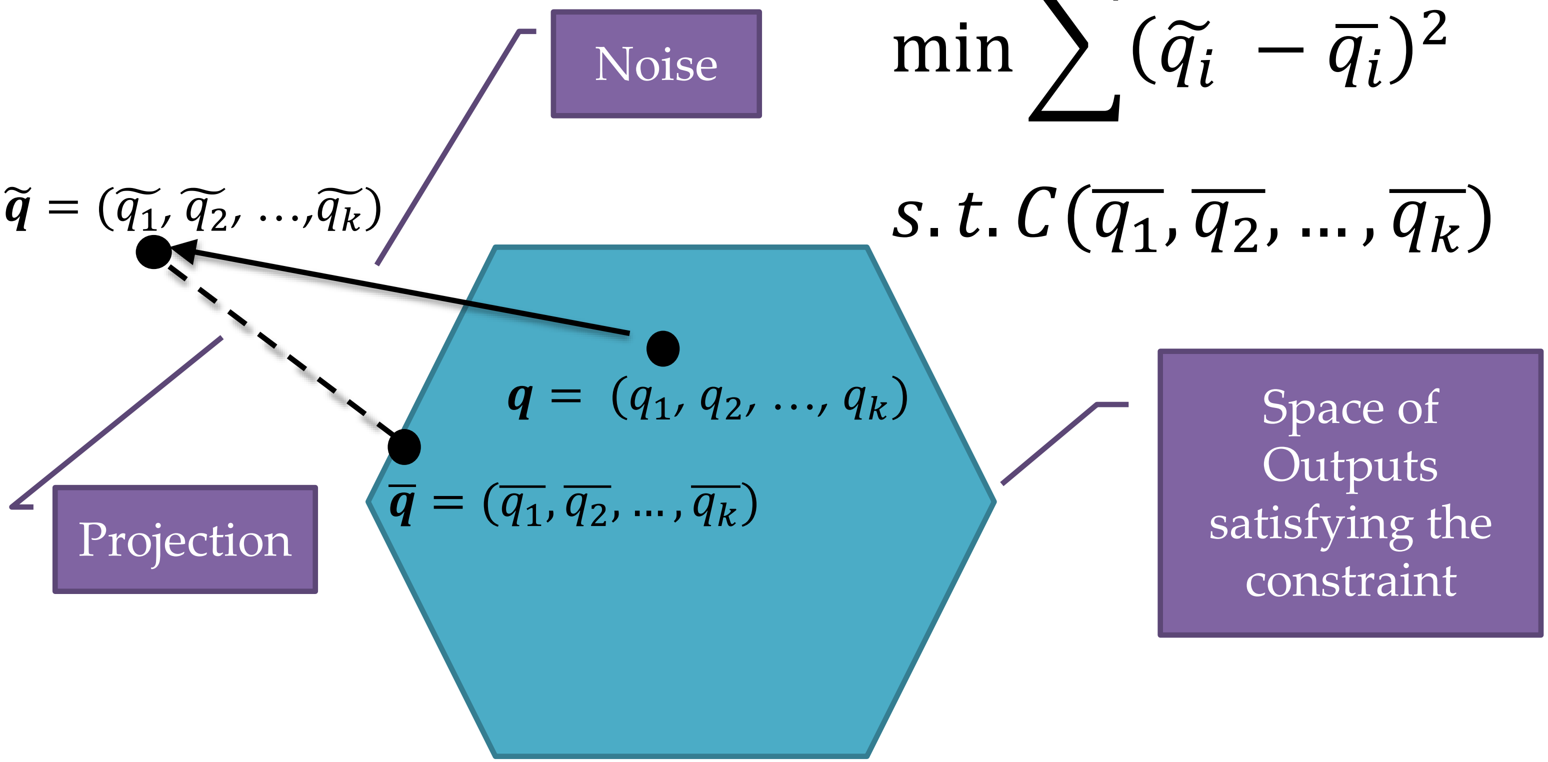
$$s.t.\, C(\overline{q_1}, \overline{q_2}, \ldots, \overline{q_k})$$



Geometric Interpretation

$$\min \sum (\tilde{q}_i - \overline{q_i})^2$$

$$s.t.\ C(\overline{q_1}, \overline{q_2}, \ldots, \overline{q_k})$$



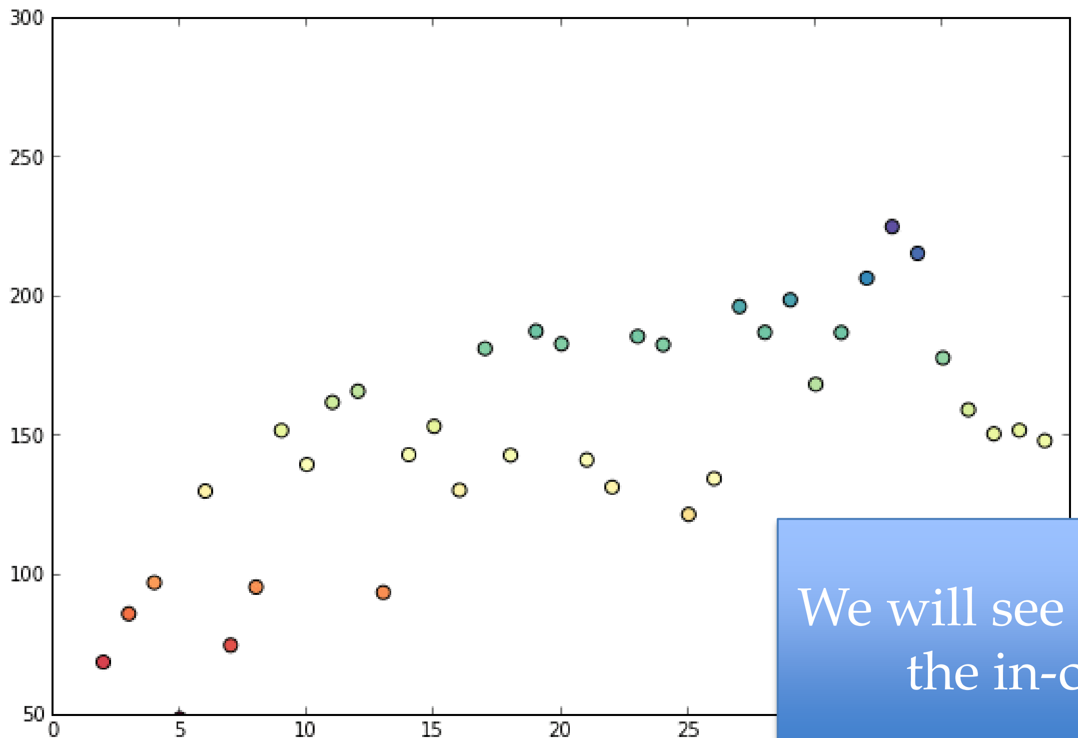
Theorem: $\|\boldsymbol{q} - \overline{\boldsymbol{q}}\|_2 \leq \|\boldsymbol{q} - \tilde{\boldsymbol{q}}\|_2$ when the constraints form a convex space

Ordering Constraint

Isotonic Regression:

$$\min \sum (\widetilde{q_1} - \overline{q_1})^2$$

$$s.t. \overline{q_1} \le \overline{q_1} \le \ldots \le \overline{q_k}$$



We will see such a problem in the in-class exercises

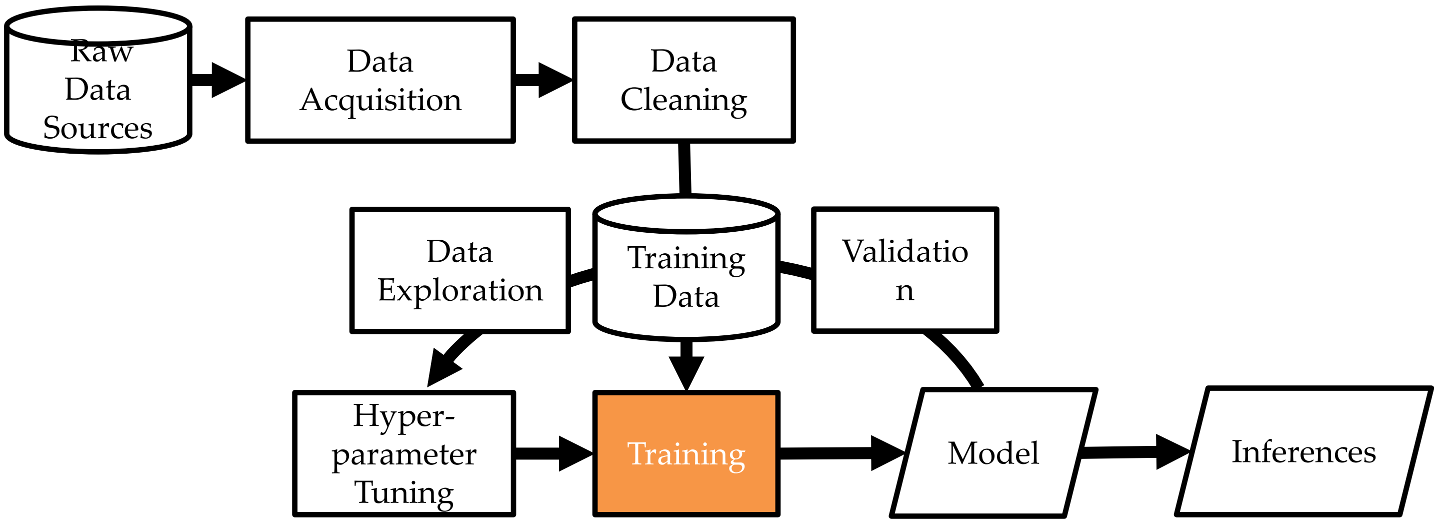
Building Complex DP Algorithms

- Composition

- Problem 1: Answer multiple queries
  - Examples
  - DP algorithms optimization

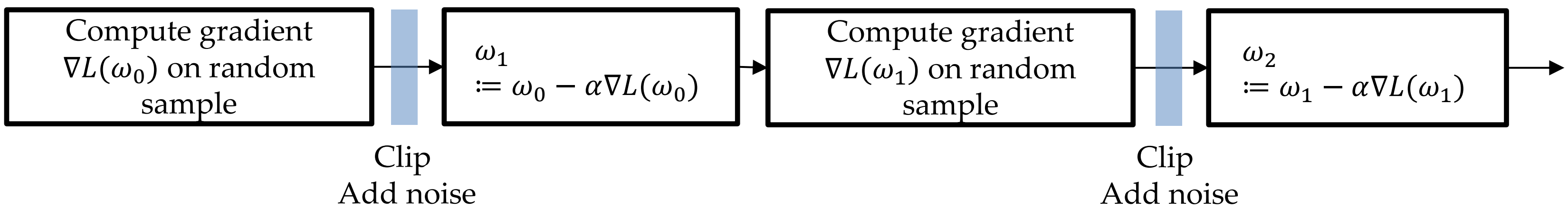
- **Problem 2: DP Learning**
  - DPSGD
  - Better composition (RDP)

DP Training


Raw Data Sources
Data Acquisition
Data Cleaning
Data Exploration
Training Data
Validatio n
Hyper-parameter Tuning
Training
Model
Inferences

DP Training

- DPSGD [ACG+16]



Initialize ω_0 and choose a learning rate α
For $t = 0 \ldots T-1$
    Take a random sample of size L
    Compute gradient per sample and clip gradient to norm bound b
    Add noise $\mathcal{N}(0, b^2\sigma^2)$ to the averaged clipped gradients
    Descent ω_{t+1} from ω_t at learning rate α

We will see DP Gradient Descent in the in-class exercises

Building Complex DP Algorithms

- Composition

- Problem 1: Answer multiple queries
  – Examples
  – DP algorithms optimization

- Problem 2: DP Learning
  – DPSGD
  – How to compose the privacy noise?

In-class exercise Time!!!
(30 mins)

In-class Exercises

Building Complex DP Algorithms

- Composition

- Problem 1: Answer multiple queries
  - Examples
  - DP algorithms optimization

- Problem 2: DP Learning
  - DPSGD
  - **Better composition (RDP)**

Advanced Composition Theorem

- Basic Composition:
  - Compositing (ϵ_1, δ_1)-DP and (ϵ_2, δ_2)-DP is $(\epsilon_1 + \epsilon_2, \delta_1 + \delta_2)$-DP
  - n-fold composition of (ϵ, δ)-DP is $(n\epsilon, n\delta)$-DP

- Advanced Composition:
  - n-fold composition of ϵ-DP is $\left(\sqrt{2n \ln(\frac{1}{\delta})}\, \epsilon, \delta\right)$-DP, for $\delta < 1$
  - Applicable to (ϵ, δ)-DP

Trouble with (ϵ, δ)-DP

- Composing advanced composition

ϵ_1-DP

ϵ_1-DP

...

ϵ_1-DP

(ϵ_1', δ'_1)-DP

(ϵ_2', δ'_2)-DP

...

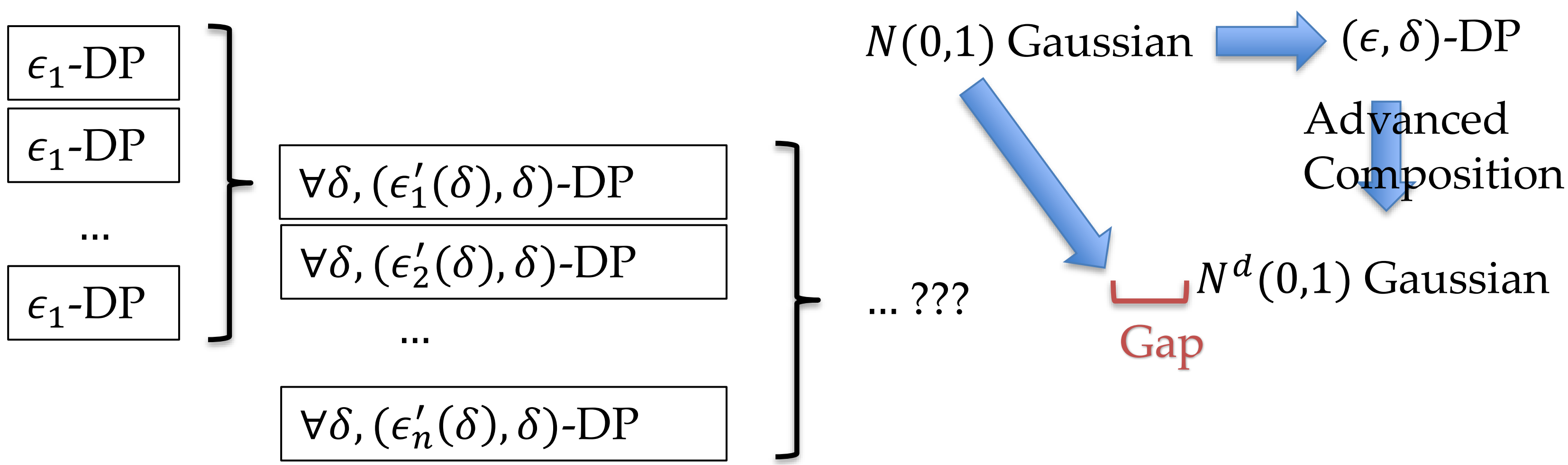
(ϵ_n', δ'_n)-DP

... #P Hard

Murtagh, Vadhan, ``The complexity of computing the optimal composition of differential privacy", TCC 2016-A.

Trouble with (ϵ, δ)-DP

- Composing advanced composition



- Gaussian + Advanced Composition is not tight

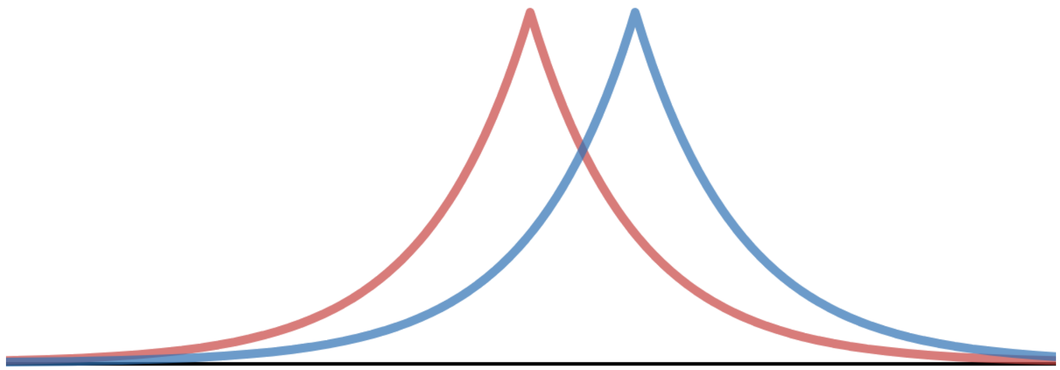
Better Notion of Closeness

- ϵ-DP

$$\max_x P(x)/Q(x) < e^{\varepsilon}$$

- Rényi Divergence at ∞

$$D_\infty(P\|Q) < \varepsilon$$



Rényi Divergence

$$D_1(P\|Q) = \lim_{\alpha \to 1} D_\alpha(P\|Q) = E_P\left[\log \frac{P(x)}{Q(x)}\right]$$

$$D_\alpha(P\|Q) = \frac{1}{\alpha - 1} \log E_Q\left[\left(\frac{P(x)}{Q(x)}\right)^\alpha\right]$$

$$D_\infty(P\|Q) = \lim_{\alpha \to \infty} D_\alpha(P\|Q) = \log \max_x \frac{P(x)}{Q(x)}$$

Rényi Differential Privacy (RDP)

- (α, ϵ)-Rényi Differential Privacy (RDP):
$$\forall D, D': \mathrm{D}_\alpha(M(D) \| M(D')) \leq \epsilon$$

- (∞, ϵ)-RDP is ϵ-DP

- (α, ϵ)-RDP $\Rightarrow$ $(\epsilon + \frac{\log 1/\delta}{\alpha - 1}, \delta)$-DP for any δ

"Bad Outcomes" Interpretation

- ϵ-DP: $\forall S$

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

No Catastrophic Failure Mode!

- (α, ϵ)-Rényi DP: $\forall S$

$$\Pr[M(D) \in S] \leq \left(e^{\epsilon} \Pr[M(D') \in S]\right)^{1-1/\alpha}$$

- (ϵ, δ)-DP: $\forall S$

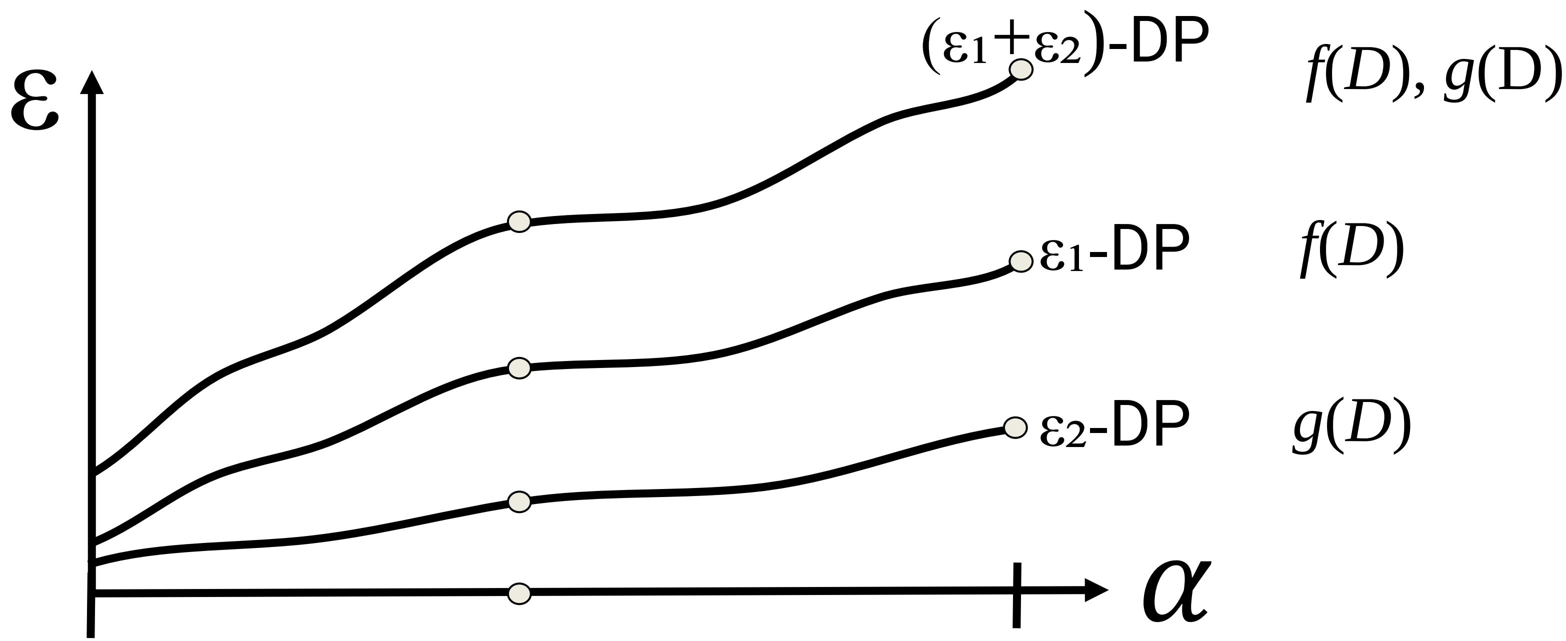
$$\Pr[M(D) \in S] \leq e^{\epsilon} \Pr[M(D') \in S] + \delta$$

"Nuclear Option":

- With probability δ publish everything
- With probability 1 publish δ fraction of inputs

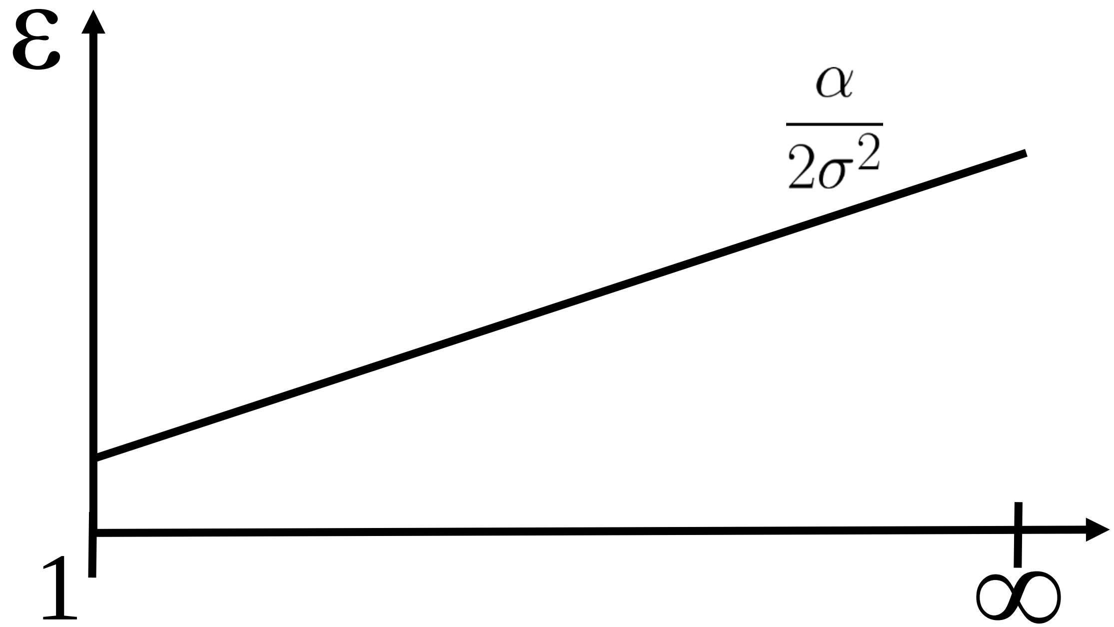
Composition

- Simultaneous release of (α, ϵ_1)-RDP and (α, ϵ_2)-RDP is $(\alpha, \epsilon_1 + \epsilon_2)$-RDP

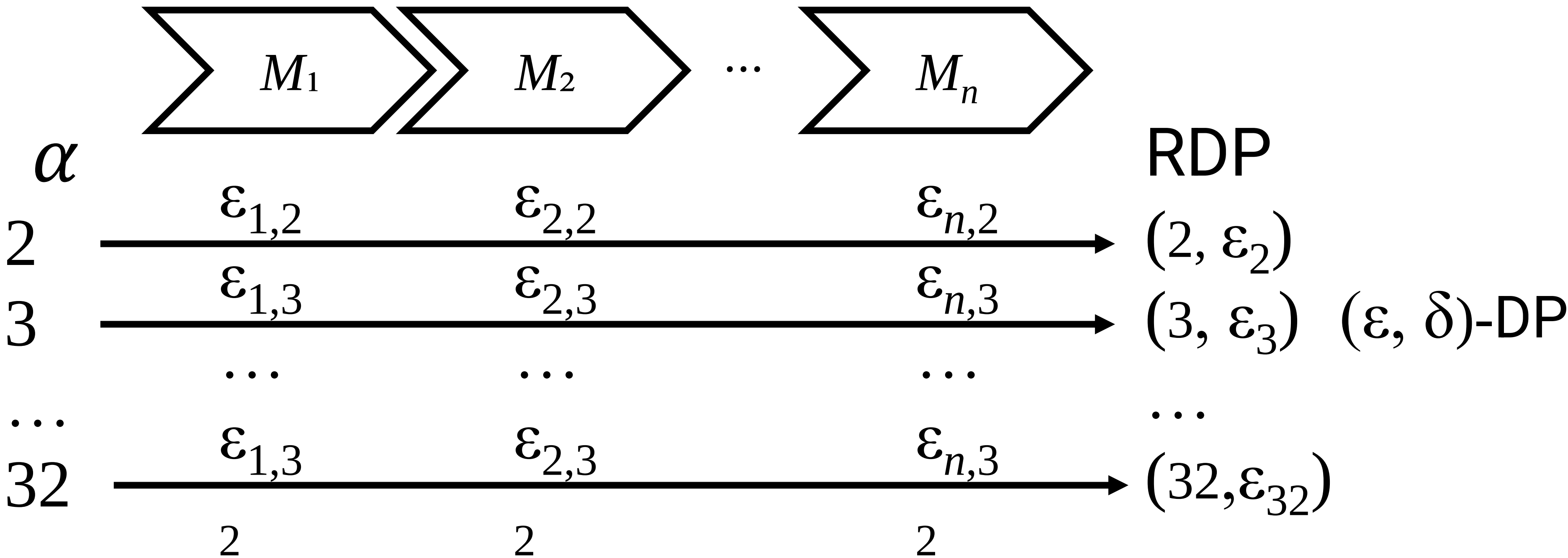


Rényi Budget Curve: Gaussian Mechanism

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



RDP as Privacy Accountant (e.g., DPGD)



(α, ϵ)-RDP $\Rightarrow (\epsilon + \frac{\log 1/\delta}{\alpha - 1}, \delta)$-DP for any δ

RDP as Privacy Accountant (e.g., DPSGD)

- Tight analysis of Gaussian noise
- Privacy amplification via sub-sampling

Reference	Conditions	Privacy bound
Abadi et al.[ACG$^+$16]	$q < \frac{1}{16\sigma}$ $\alpha \leq 1 + \sigma^2 \ln \frac{1}{q\sigma}$	$(\alpha, q^2 \frac{\alpha}{(1-q)\sigma^2} + O(q^3\alpha^3/\sigma^3))$-RDP for $q \to 0$
Abadi et al.[ACG$^+$16]	integer α	Numerical procedure
Bun et al. [BDRS18]	$q \leq \frac{1}{10}, \sigma \geq \sqrt{5}$ $\alpha \leq \frac{1}{2}\sigma^2 \ln \frac{1}{q}$	$(\alpha, q^2 \cdot \frac{6\alpha}{\sigma^2})$-RDP for fixed-size sample
Mironov et al. [MTZ19]	$q < \frac{1}{5}, \sigma \geq 4$ $\alpha \leq \frac{1}{2}\sigma^2 L - 2\ln\sigma$ $\alpha \leq \frac{\frac{1}{2}\sigma^2 L^2 - \ln 5 - 2\ln\sigma}{L + \ln(q\alpha) + \frac{1}{2\sigma^2}}$	$(\alpha, q^2 \cdot \frac{2\alpha}{\sigma^2})$-RDP for i.i.d. (Poisson) sample
Mironov et al. [MTZ19]	arbitrary $\alpha \geq 1$	Numerical procedure

Summary

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

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

- Designing complex DP algorithms
  - Composition
  - Answer multiple queries
  - DPSGD