Differential Privacy: Basics

CompSci 590.03
Instructor: Ashwin Machanavajjhala

Duke UNIVERSITY

Outline of lecture

- Differential Privacy
- Basic Algorithms
  - Laplace Mechanism
- Composition Theorems
- Exercise

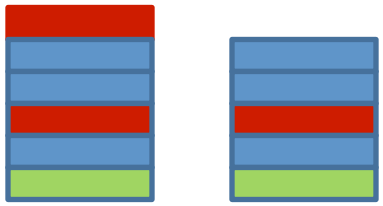
Duke UNIVERSITY

Differential Privacy

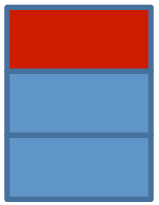
[Dwork ICALP 2006]

For every pair of inputs that differ in one row

For every output …



D_1 D_2



O

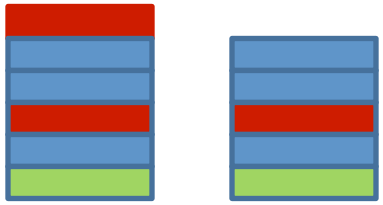
Adversary should not be able to distinguish between any D_1 and D_2 based on any O

$$\log\left(\frac{\Pr[A(D_1) = O]}{\Pr[A(D_2) = O]}\right) < \varepsilon \quad (\varepsilon>0)$$

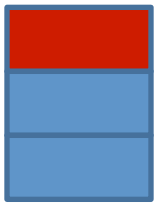
Why pairs of datasets *that differ in one row*?

For every pair of inputs that differ in one row

For every output …



D_1 D_2

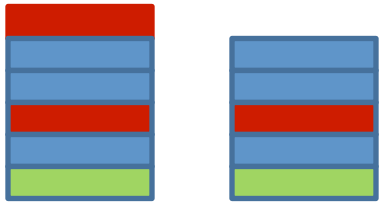


O

Simulate the presence or absence of a single record

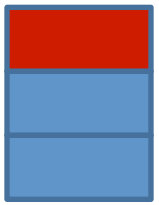
Why *all* pairs of datasets …?

For every pair of inputs that differ in one row



D_1 D_2

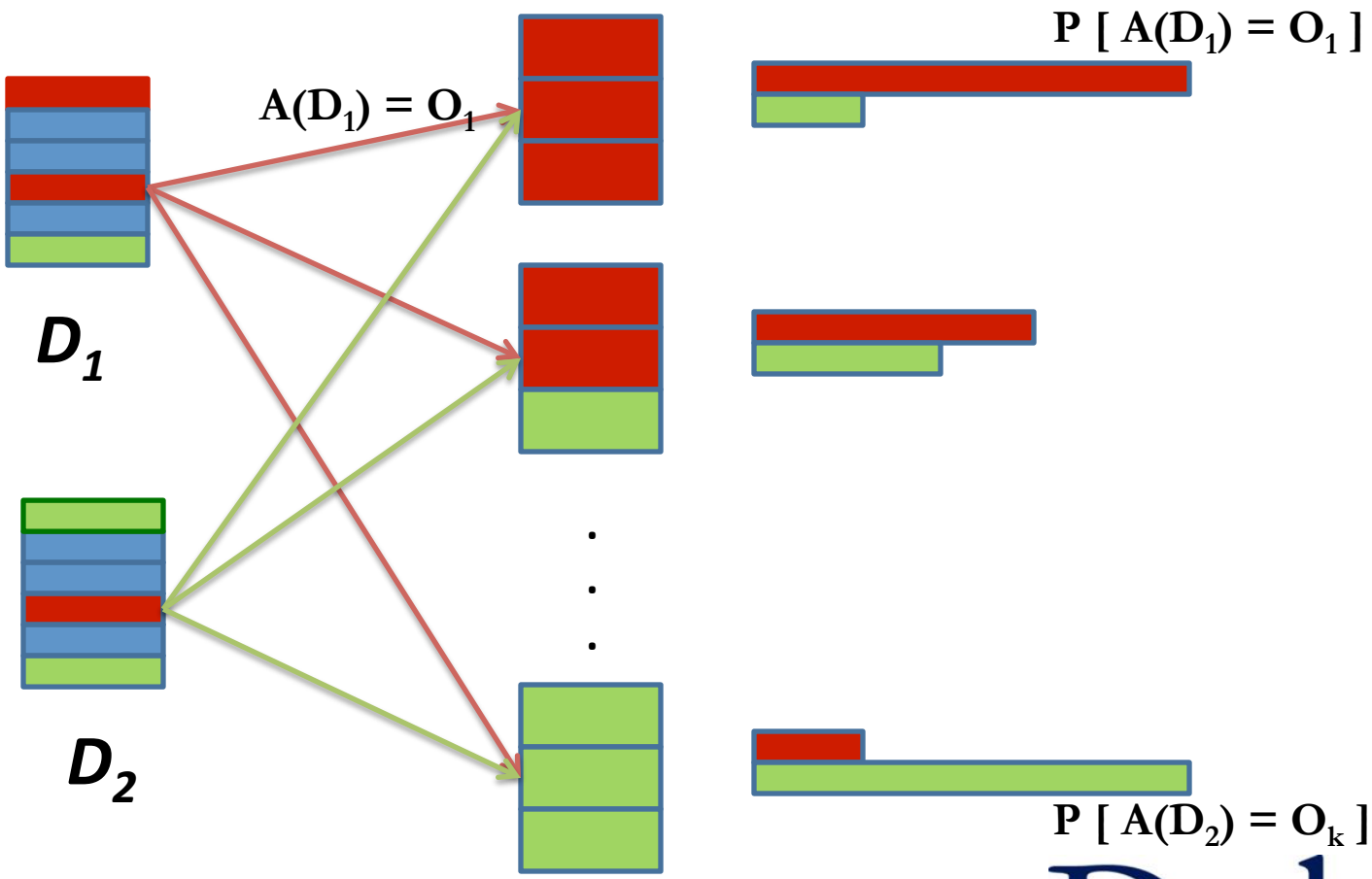
For every output …



O

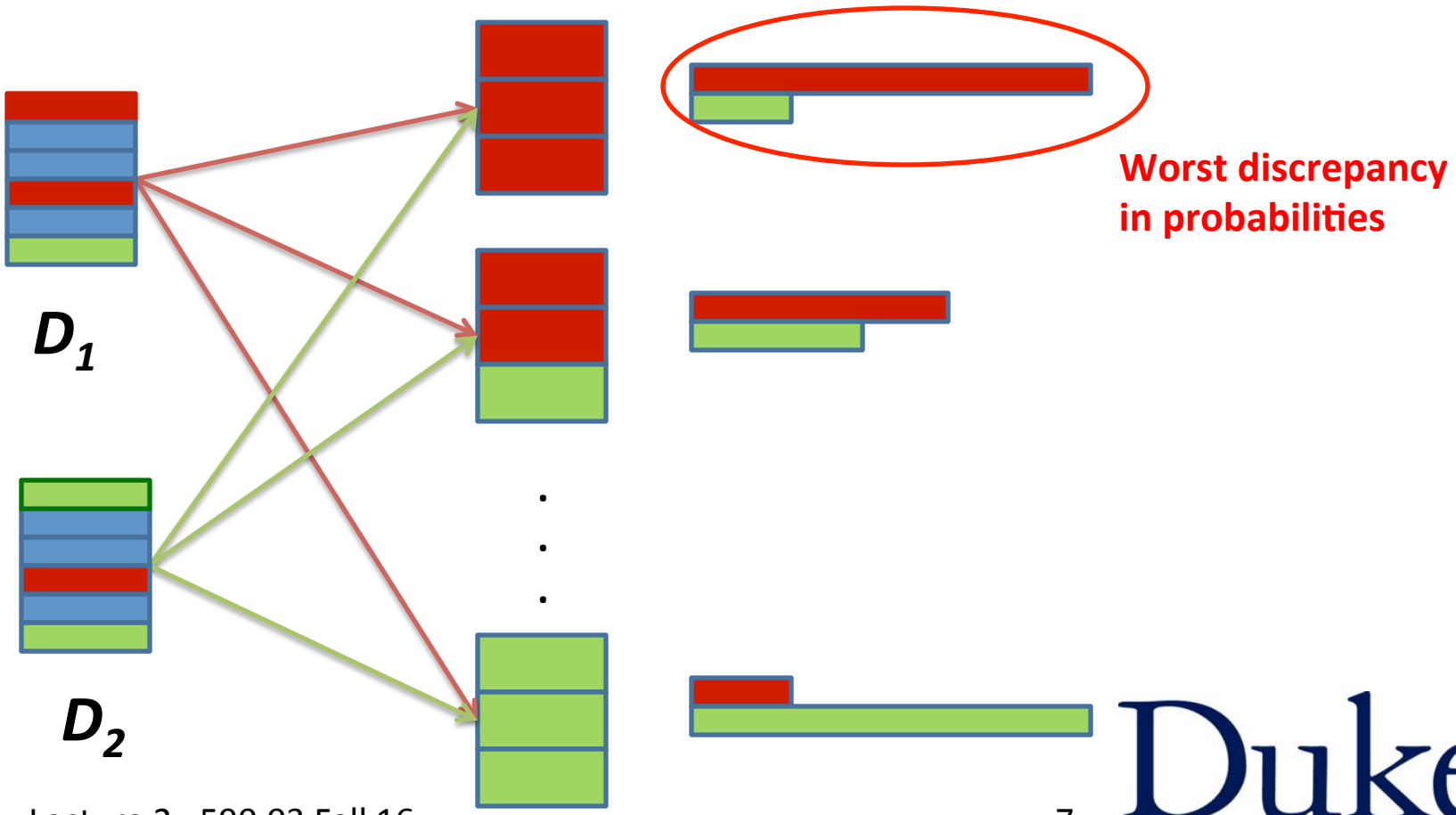
Guarantee holds no matter what the other records are.

Why *all* outputs?



Set of all outputs

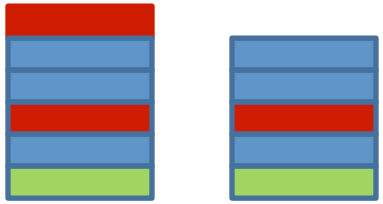
Should not be able to distinguish whether input was D_1 or D_2 no matter what the output



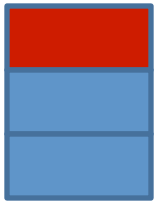
Privacy Parameter ε

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 better the utility)

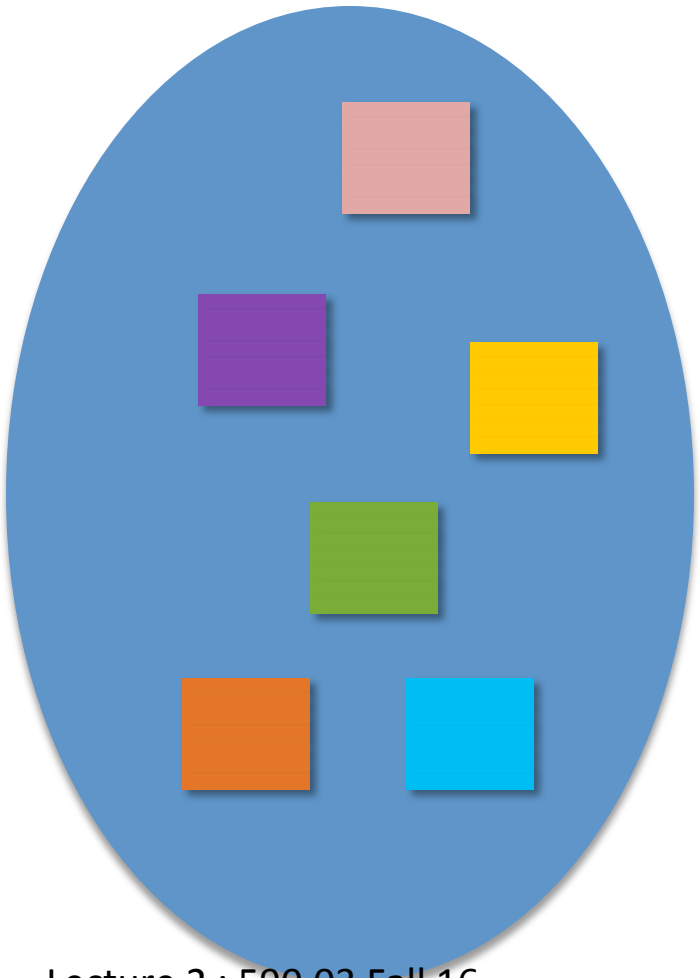
Outline of the Module 2

- Differential Privacy
- Basic Algorithms
  - Laplace Mechanism
- Composition Theorems

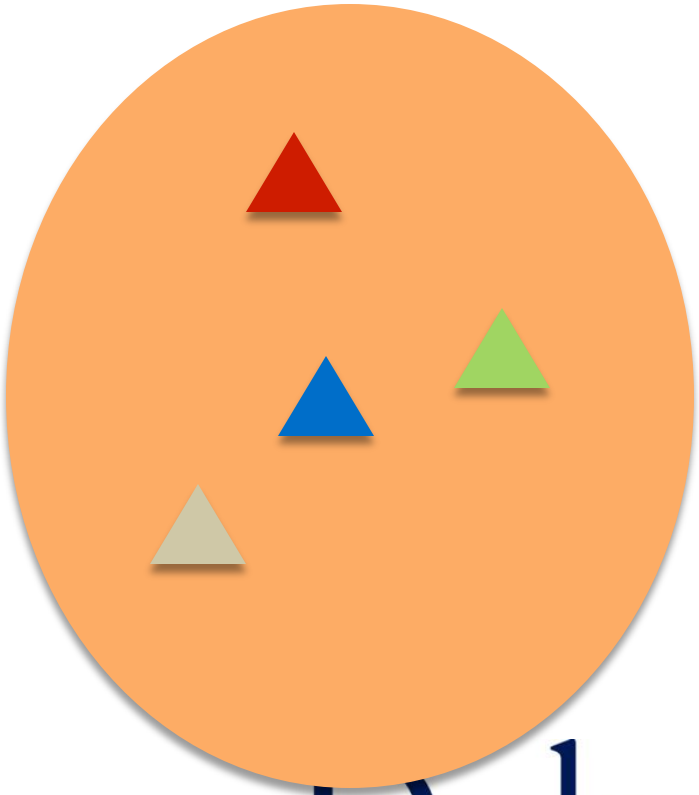
Can deterministic algorithms satisfy differential privacy?

Non-trivial deterministic Algorithms do not satisfy differential privacy

Space of all inputs

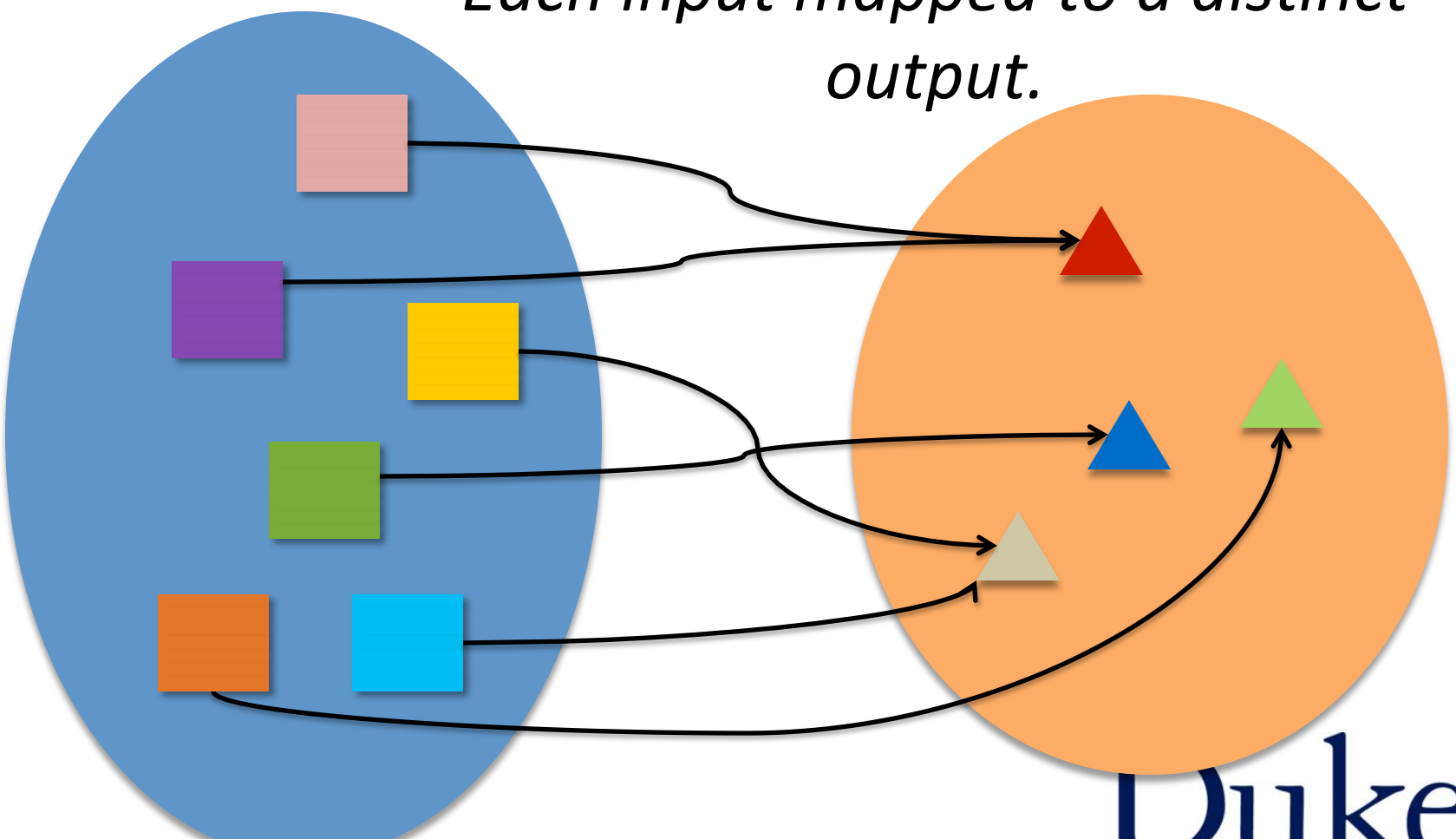


Space of all outputs (at least 2 distinct ouputs)

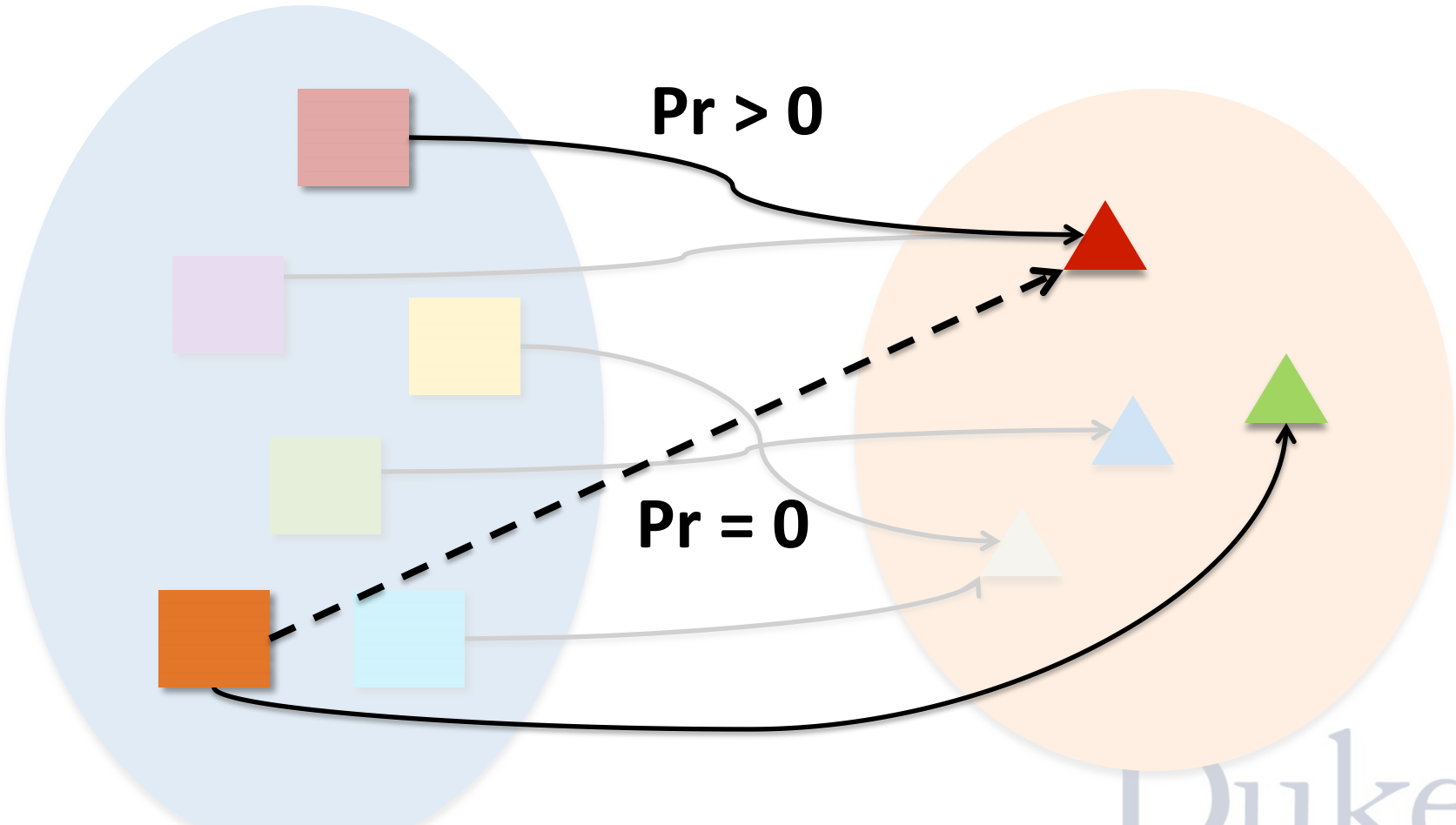


Non-trivial deterministic Algorithms do not satisfy differential privacy

Each input mapped to a distinct output.



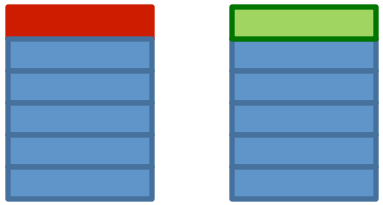
There exist two inputs that differ in one entry mapped to different outputs.



Random Sampling …

… also does not satisfy differential privacy

Input



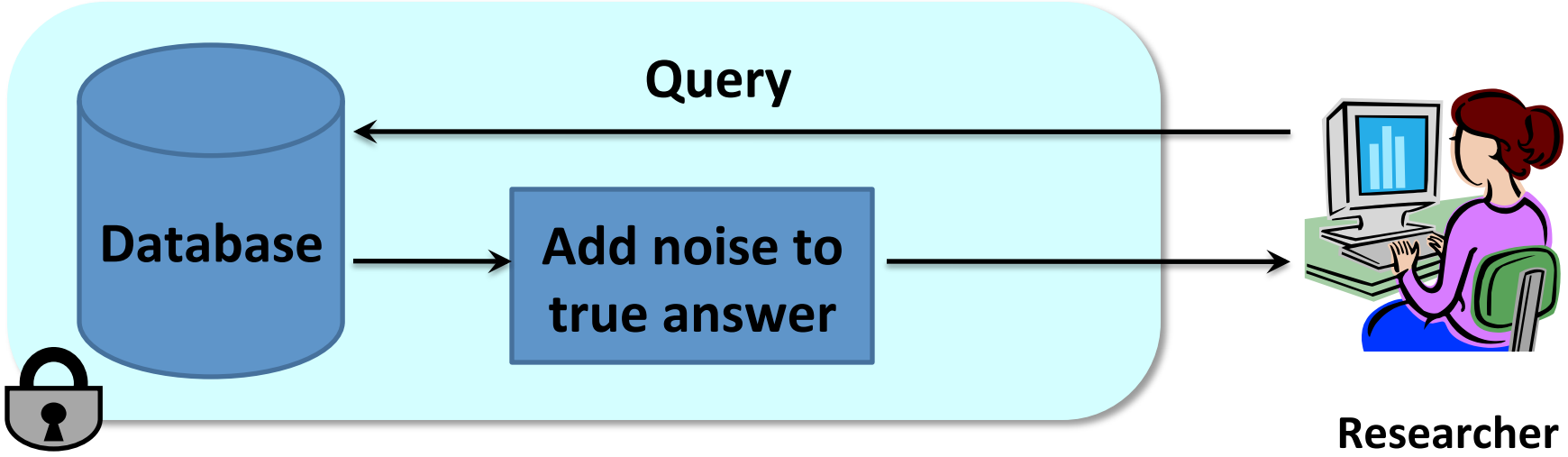
D_1 ***D_2***

Output

O

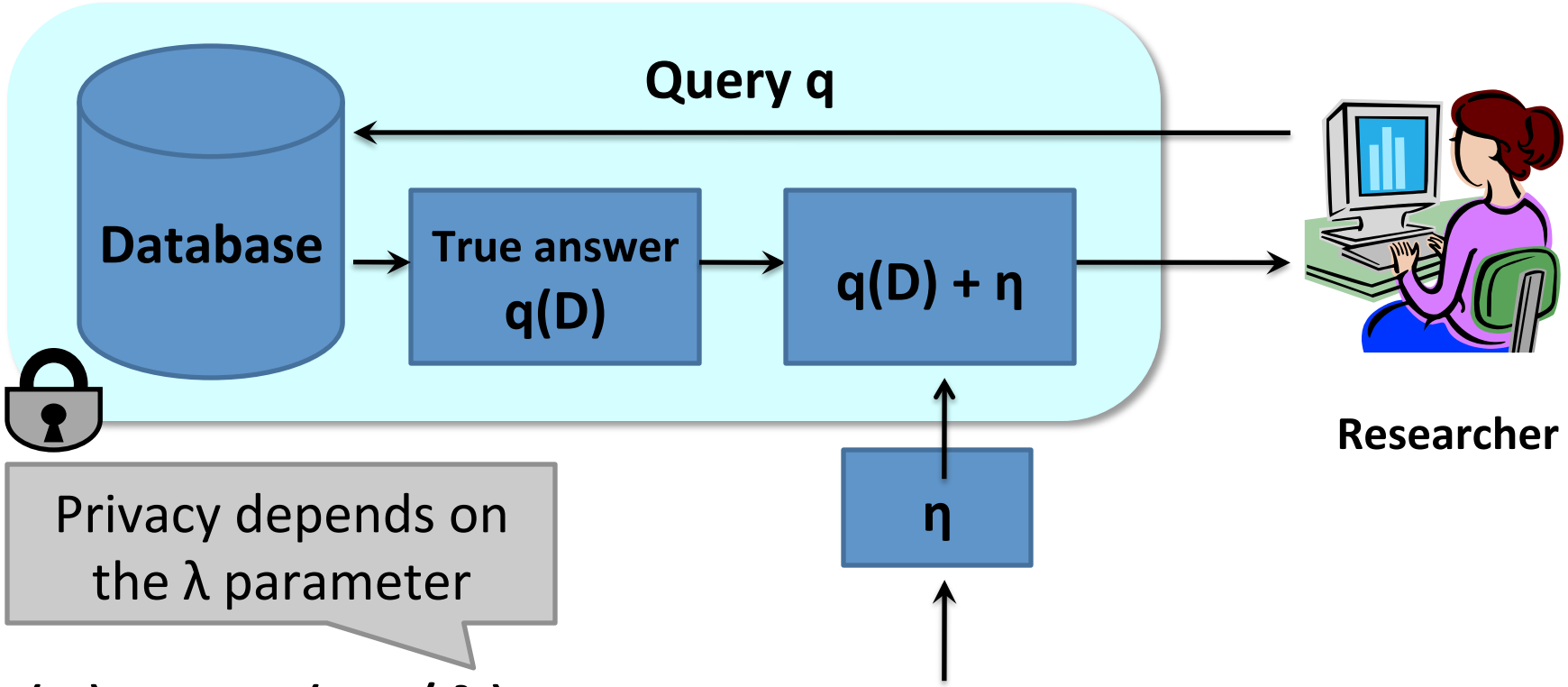
$\Pr[D_2 \rightarrow O] = 0$ implies $\log\left(\frac{\Pr[D_1 \rightarrow O]}{\Pr[D_2 \rightarrow O]}\right) = \infty$

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.

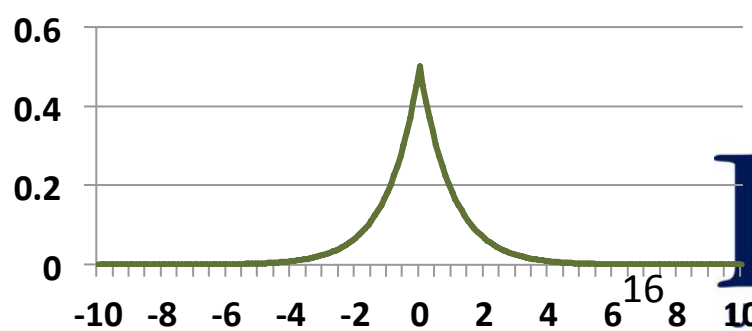
Laplace Mechanism



Privacy depends on the λ parameter

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

Laplace Distribution – Lap(λ)



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

How much noise for privacy?

[Dwork et al., TCC 2006]

Sensitivity: Consider a query q: *I* → 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

Duke UNIVERSITY

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

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

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

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

Utility of Laplace Mechanism

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

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

$$= \mathrm{Var}(\ \mathrm{Lap}(S(q)/\varepsilon)\ )$$

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

Outline of the Module 2

- Differential Privacy
- Basic Algorithms
  - Laplace & Exponential Mechanism
  - Randomized Response
- Composition Theorems

Why Composition?

- Reasoning about privacy of a complex algorithm is hard.

- Helps software design
  - If building blocks are proven to be private, it about privacy of a complex algorithm built e blocks.

A bound on the number of queries

- In order to ensure utility, a statistical database must leak some information about each individual
- We can only hope to bound the amount of disclosure

- Hence, there is a limit on number of queries that can be answered

Dinur Nissim Result

- A vast majority of records in a database of size n can be reconstructed when $n \log(n)^2$ queries are answered by a statistical database … **[Dinur-Nissim PODS 2003]**

  … even if each answer has been arbitrarily altered to have up to $o(\sqrt{n})$ error

.

Sequential Composition

- If M_1, M_2, ..., 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 $\varepsilon=\varepsilon_1+...+\varepsilon_k$

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?

Duke UNIVERSITY

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,

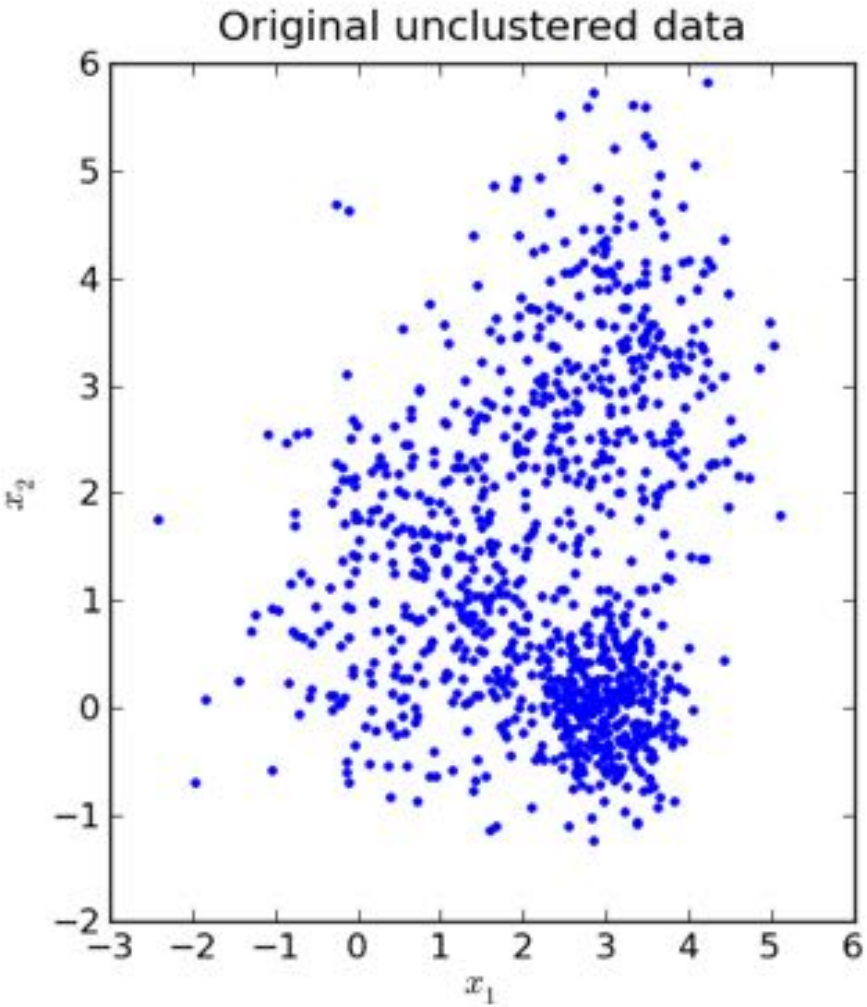
  then the combination of their outputs satisfies ε-differential privacy with $\varepsilon = \max\{\varepsilon_1,\ldots,\varepsilon_k\}$

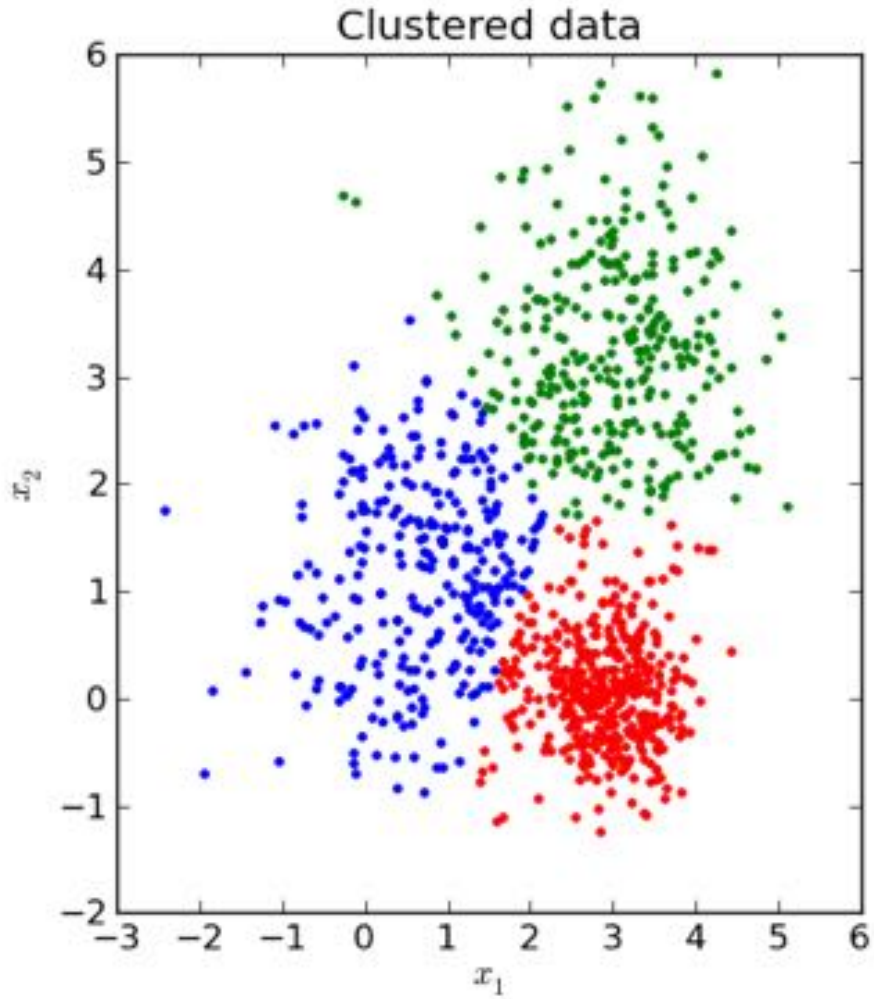
Postprocessing

- 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.

Case Study: K-means Clustering





Kmeans

- Partition a set of points $x_1, x_2, \ldots, x_n$ into k clusters $S_1, S_2, \ldots, S_k$ such that the following is minimized:

$$\sum_{i=1}^{k} \sum_{x_j \in S_i} \|x_j - \mu_i\|_2^2$$

Mean of the cluster S_i

Kmeans

Algorithm:

- Initialize a set of k centers
- Repeat

  Assign each point to its nearest center

  Recompute the set of centers

  Until convergence …

- Output the final k centers

Exercise

- What is a differentially private algorithm for releasing a k-means clustering (i.e., outputting the final set of k centers)?

Summary

- Differentially private algorithms ensure an attacker can't infer the presence or absence of a single record in the input based on any output.

- Building blocks
  - Laplace mechanism

- Composition rules help build complex algorithms using building blocks

Duke UNIVERSITY