18734: Foundations of Privacy

Privacy-preserving Release of Statistics: Differential Privacy

Giulia Fanti

Slides by Anupam Datta

CMU

Fall 2019

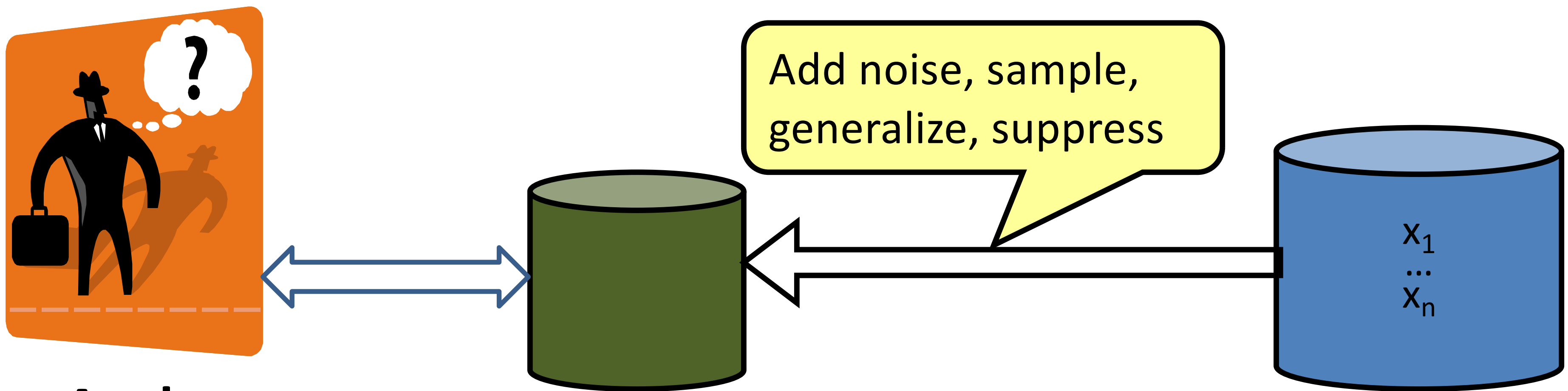
Administrative Stuff

- HW2 due tonight at midnight on Gradescope
  - Upload pdf with everything except AdFisher code and logs to Gradescope
  - Upload AdFisher code and logs to Canvas
- Note on Piazza use

Quiz

- On Canvas

Privacy-Preserving Statistics: Non-Interactive Setting



Analyst

Sanitized Database D'

Database D
maintained by trusted curator

- Census data
- Health data
- Network data
- …

Goals:

- Accurate statistics (low noise)
- Preserve individual privacy (what does that mean?)

Some possible approaches

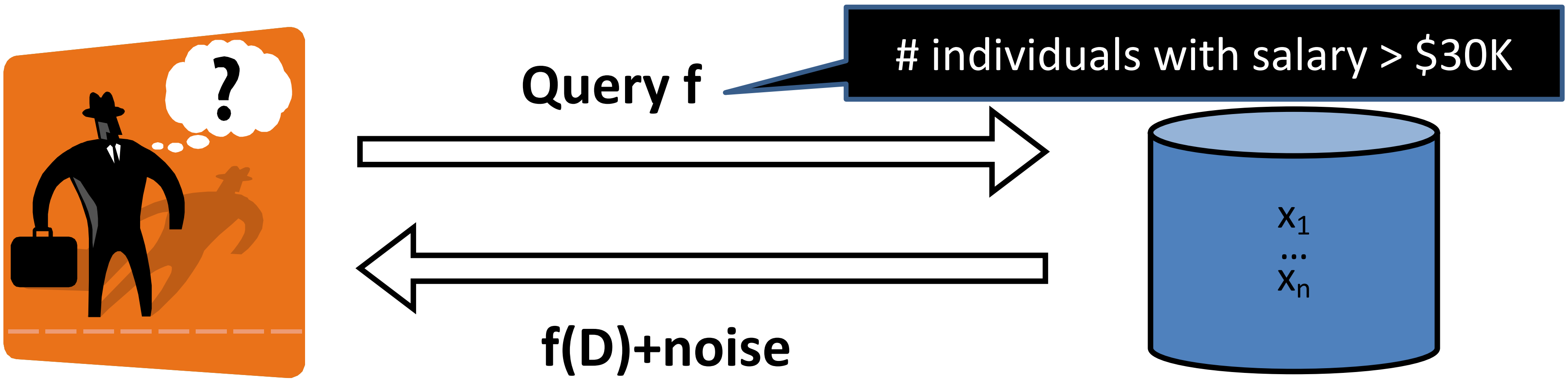
- Anonymize data
  - Re-identification, information amplification
- Summary statistics

Name	Age
Alice	10
Bob	50
Carol	40

→ **Summary statistic**
Mean age = 33.33

  - Differencing attack

Privacy-Preserving Statistics: Interactive Setting



Analyst

Database D maintained by trusted curator

- Census data
- Health data
- Network data
- …

Goals:

- Accurate statistics (low noise)
- Preserve individual privacy (what does that mean?)

Classical Intuition for Privacy

- "If the release of statistics S makes it possible to determine the value [of private information] more accurately than is possible without access to S, a disclosure has taken place." [Dalenius 1977]
  - Privacy means that anything that can be learned about a respondent from the statistical database can be learned without access to the database

- Similar to semantic security of encryption

Impossibility Result

- "Theorem": For any reasonable definition of breach, if sanitized database contains information about database, then there exists an adversary and an auxiliary information generator that causes a breach with some nontrivial probability.

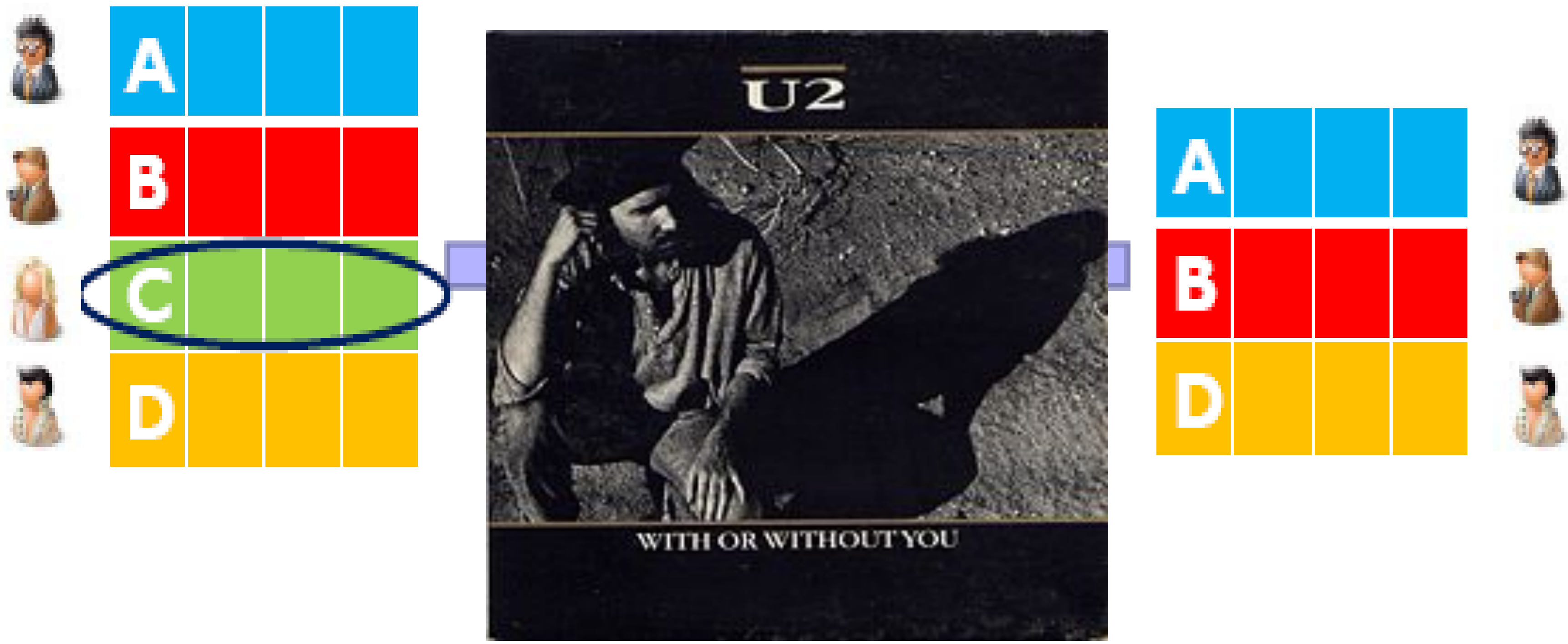
- Example
  - Terry Gross is two inches shorter than the average Lithuanian woman
  - DB allows computing average height of a Lithuanian woman
  - This DB breaks Terry Gross's privacy according to this definition… even if her record is not in the database!

Dwork and Naor. On the Difficulties of Disclosure Prevention in Statistical Databases or The Case for Differential Privacy. 2016

Takeaway message

- Our privacy definitions must account for auxiliary information.
- Recall: Netflix paper

Differential Privacy: Idea
[Dwork, McSherry, Nissim, Smith 2006]



Released statistic is about the same
if any individual's record is
removed from the database

An Information Flow Idea

Changing input databases in a specific way
changes output statistic by a small amount

Not Absolute Confidentiality

Does not guarantee that Terry Gross's height won't be learned by the adversary

Differential Privacy: Definition

Randomized sanitization function κ has ε-differential privacy if for all data sets D_1 and D_2 differing by at most one element and all subsets S of the range of κ,

$$\Pr[\kappa(D_1) \in S] \leq e^{\varepsilon} \Pr[\kappa(D_2) \in S]$$

Answer to query # individuals with salary > $30K is in range [100, 110] with approximately the same probability in D_1 and D_2

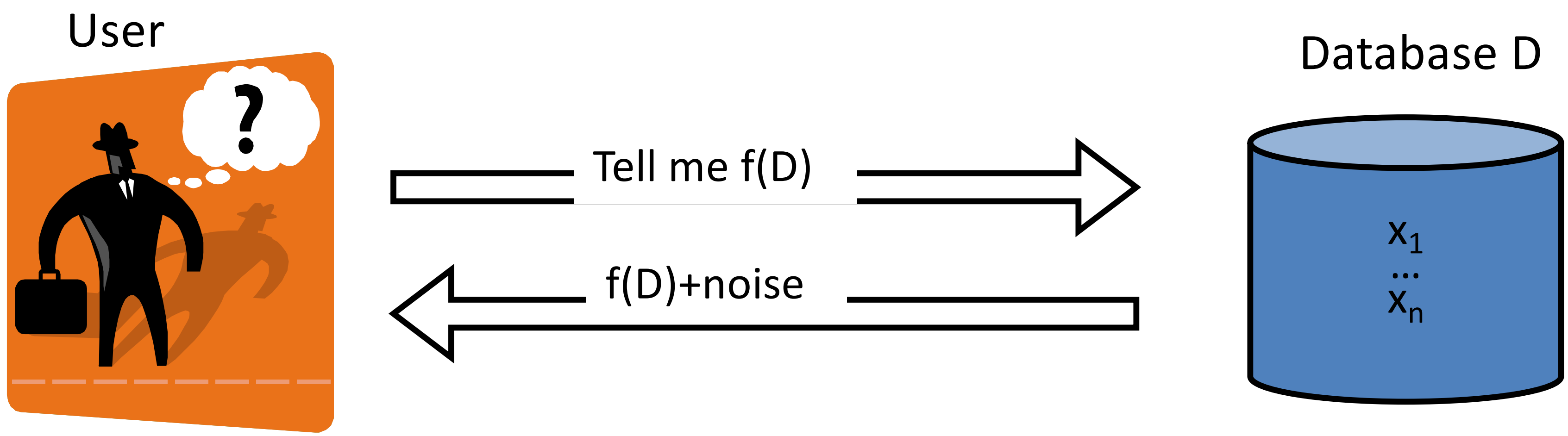
Check your understanding

Randomized sanitization function κ has ε-differential privacy if for all data sets D_1 and D_2 differing by at most one element and all subsets S of the range of κ,

$$\Pr[\kappa(D_1) \in S] \leq e^{\varepsilon} \Pr[\kappa(D_2) \in S]$$

- What does differential privacy mean when $\epsilon = 0$?
- What range of values can ϵ take?

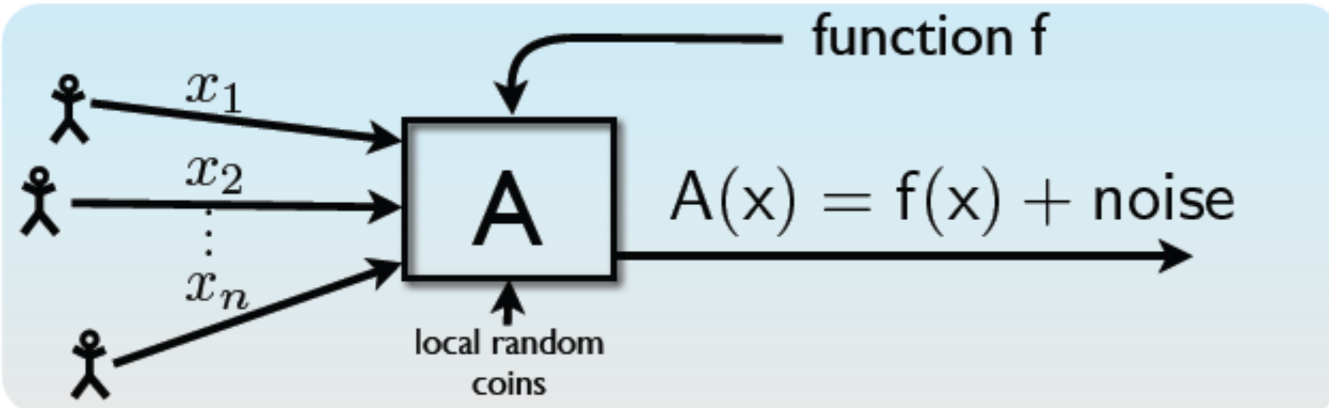
Achieving Differential Privacy: Interactive Setting



How much and what type of noise should be added?

Slide: Adam Smith

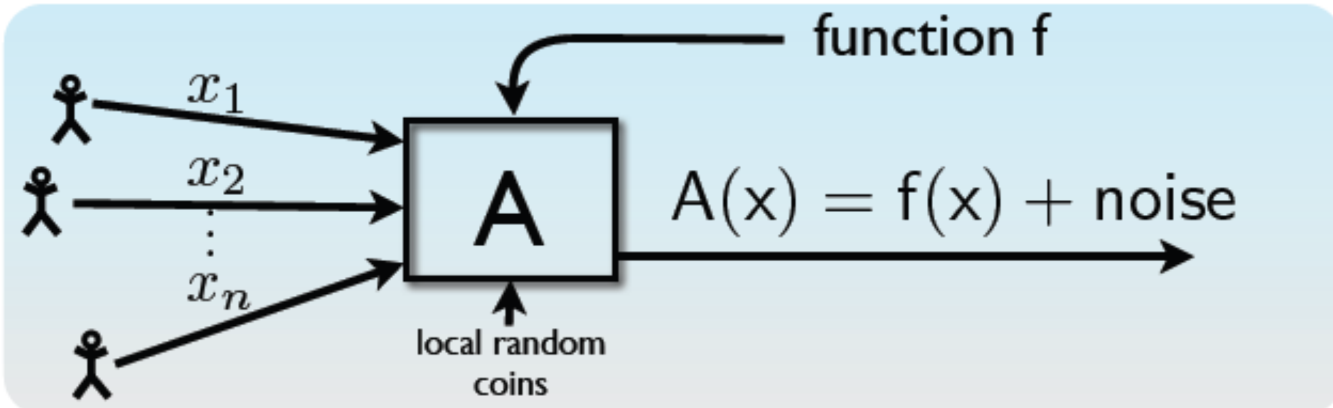
Example: Noise Addition



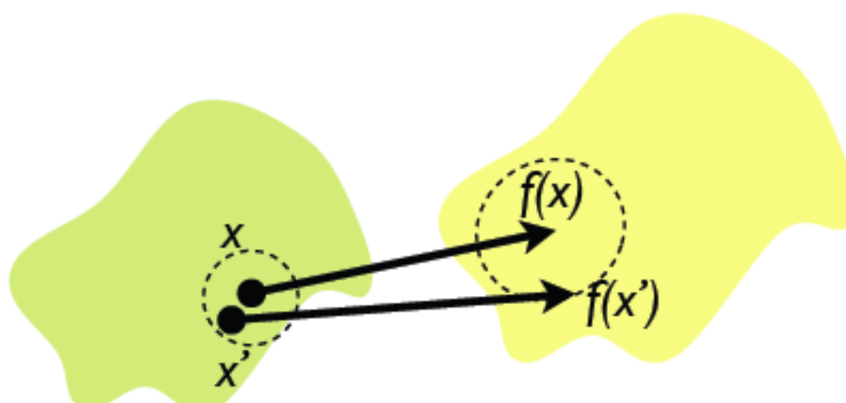
- Say we want to release a summary $f(x) \in \mathbb{R}^p$
  - e.g., proportion of diabetics: $x_i \in \{0,1\}, \; f(x) = \frac{1}{n}\sum x_i$
- Simple approach: add noise to f(x)
  - How much noise is needed?
- Intuition: f(x) can be released accurately when f is insensitive to individual entries $x_1, x_2, \ldots, x_n$

Slide: Adam Smith

Global Sensitivity



- **Global Sensitivity:** $\mathrm{GS}_f = \max_{\text{neighbors } x, x'} \| f(x) - f(x') \|_1$



Exercise

- Function f: # individuals with salary > $30K
- Global Sensitivity of f = ?

- Global Sensitivity: $\mathsf{GS}_f = \max_{\text{neighbors } x, x'} \|f(x) - f(x')\|_1$

- Answer: 1

Exercise 2

- Function $f(x) = \frac{1}{n}\sum_{i=1}^{n} x_i$, where $x_i \in S$
- Global Sensitivity of f = ?

- Global Sensitivity: $\mathrm{GS}_f = \max_{\text{neighbors } x,x'} \|f(x) - f(x')\|_1$

- Answer: $\dfrac{|\max(S)|}{n}$

Background on Probability

Continuous Probability Distributions

- Probability density function (PDF), f_X

$$\Pr[a \leq X \leq b] = \int_a^b f_X(x)\,dx.$$

- Example distributions
  - Normal, exponential, Gaussian, Laplace

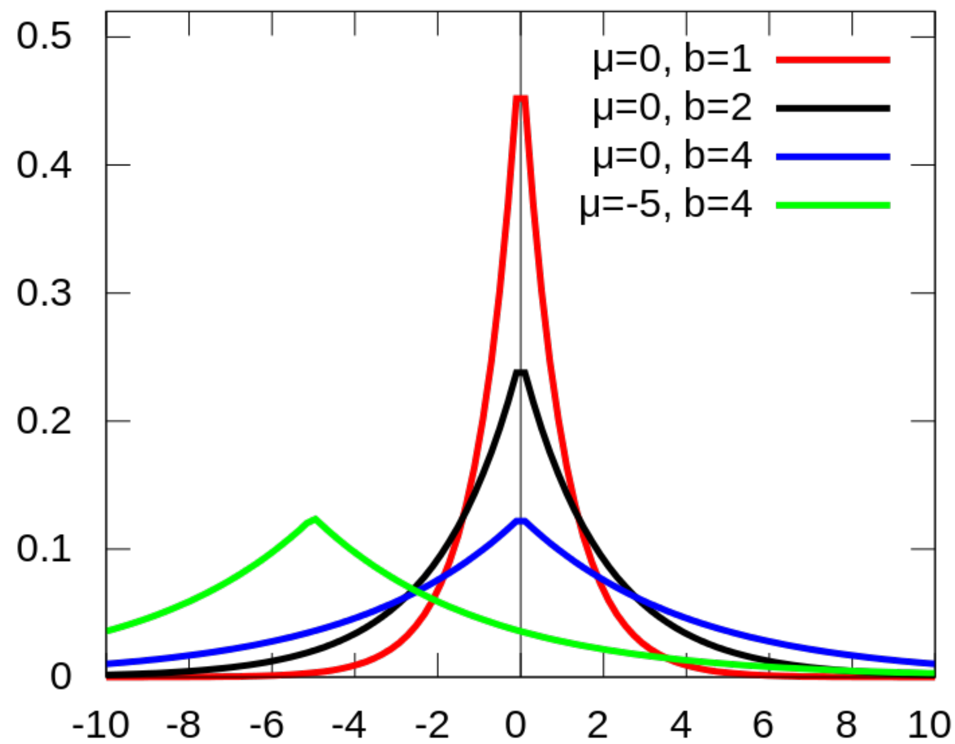
Laplace Distribution

Laplace random variable has PDF

$$\frac{1}{2\,b}\exp\left(-\frac{|x-\mu|}{b}\right)$$

Mean = μ

Variance = $2b^2$



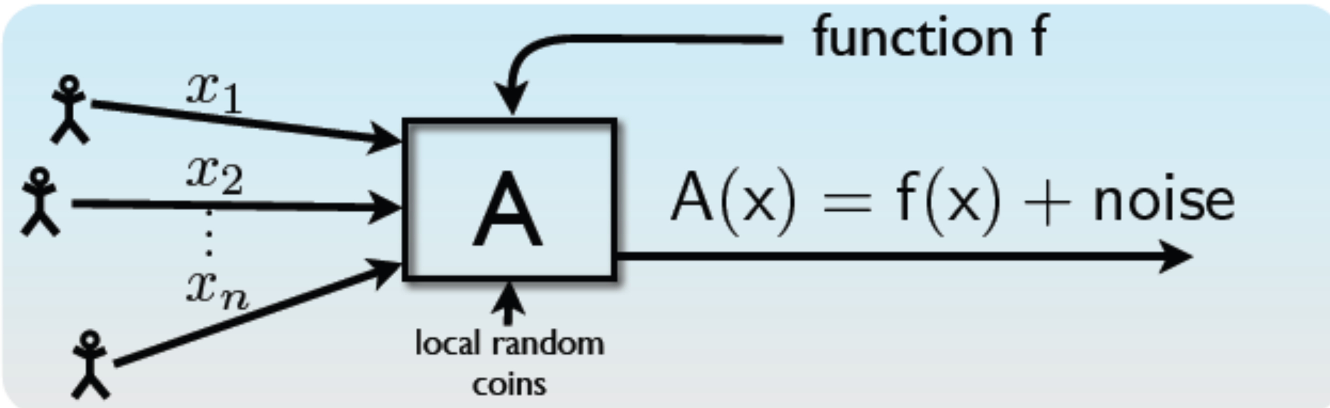
Source: Wikipedia

We use $\mathrm{Lap}(b)$ to denote the 0-mean version of this

Achieving Differential Privacy

Slide: Adam Smith

Laplace Mechanism



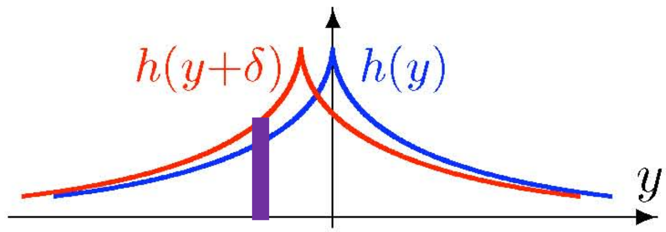
- **Global Sensitivity:** $GS_f = \max_{\text{neighbors } x, x'} \|f(x) - f(x')\|_1$
  - Example: $\mathrm{GS}_{\text{proportion}} = \frac{1}{n}$

Theorem: If $A(x) = f(x) + \mathrm{Lap}\left(\frac{GS_f}{\epsilon}\right)$, then A is ϵ-differentially private.

Laplace Mechanism: Proof Idea

Theorem: If $\mathsf{A}(\mathsf{x}) = f(\mathsf{x}) + \mathsf{Lap}\left(\frac{\mathsf{GS}_f}{\epsilon}\right)$, then A is ϵ-differentially private.

Laplace distribution $\mathsf{Lap}(\lambda)$ has density $h(y) \propto e^{-\frac{\|y\|_1}{\lambda}}$



Work with your neighbors to prove the Theorem.

Hint: Compute $\frac{f_{A(x)}(t)}{f_{A(x')}(t)}$