18734: Foundations of Privacy

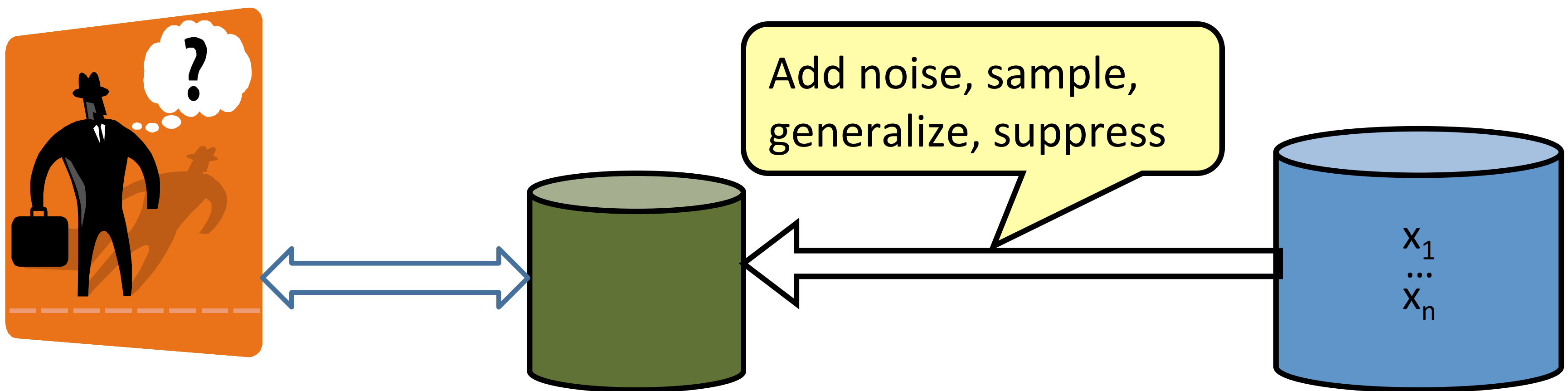
Privacy-preserving Release of Statistics: Differential Privacy

Anupam Datta

CMU

Fall 2015

Privacy-Preserving Statistics: Non-Interactive Setting



Analyst

Sanitized Database D'

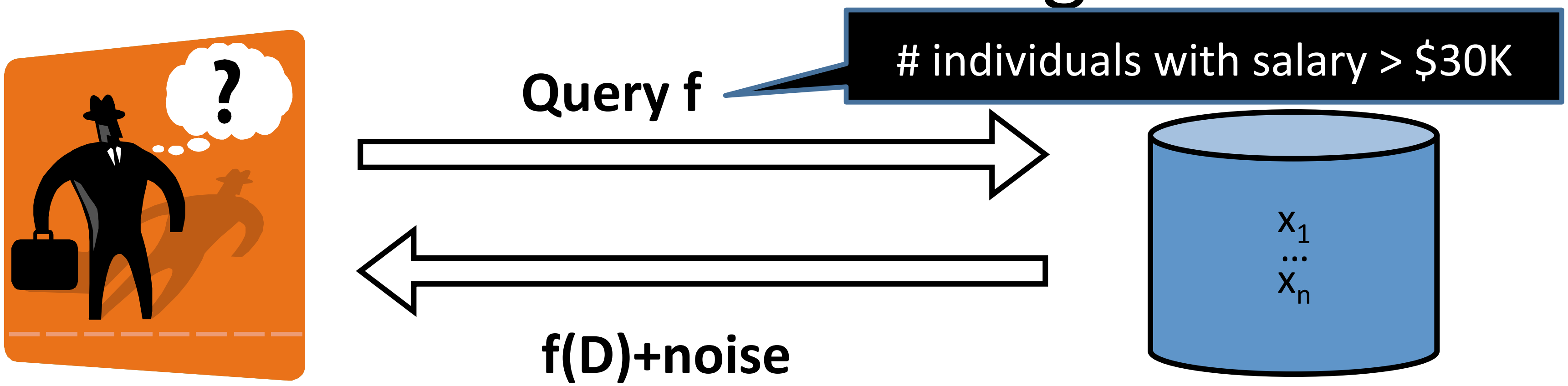
Database D maintained by trusted curator

- Census data
- Health data
- Network data
- …

Goals:

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

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?)

Some possible defenses

- Anonymize data
  - Re-identification, information amplification
- Queries over large data sets
  - Differencing attack
- Query auditing
  - Refusal leaks, computational tractability
- Summary statistics
  - Frequency lists

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 [Dwork, Naor 2006]

- Result: For reasonable "breach", if sanitized database contains information about database, then some adversary breaks this definition

- 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!

Very Informal Proof Sketch

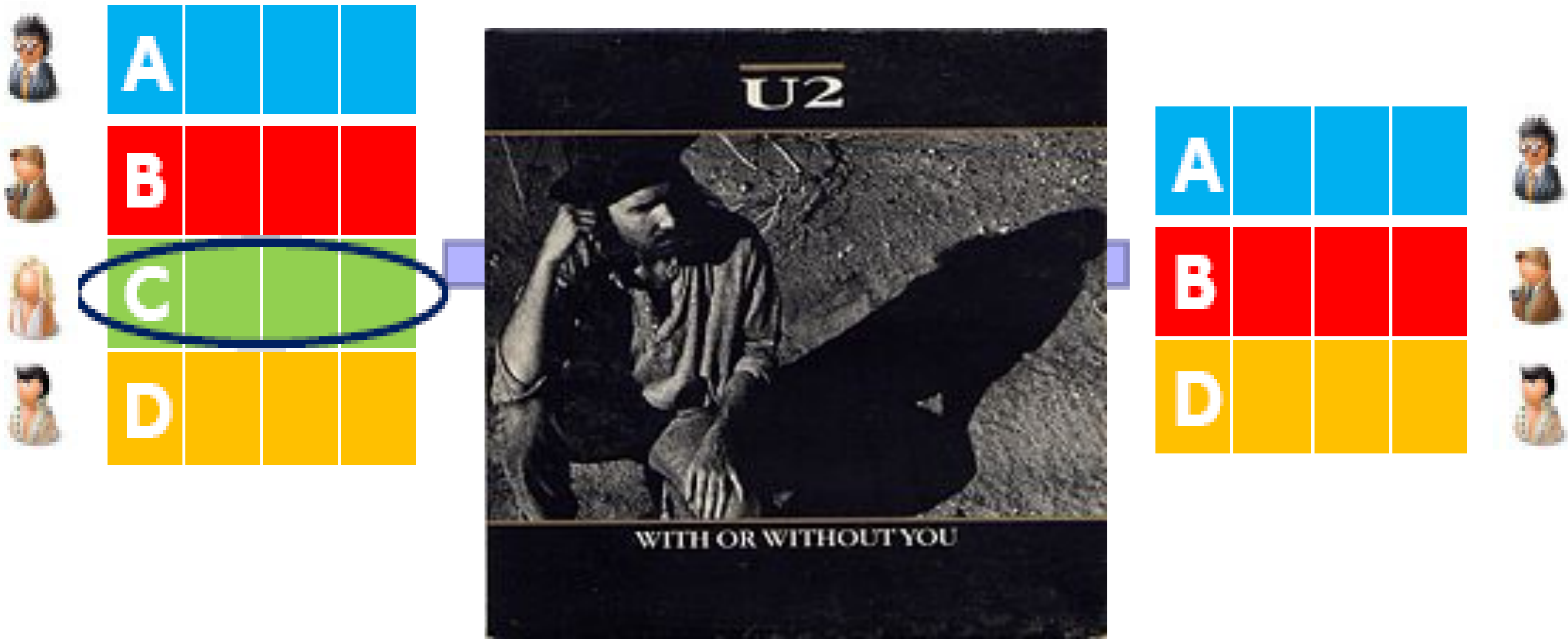
- Suppose DB is uniformly random
- "Breach" is predicting a predicate g(DB)
- Adversary's background knowledge:

  $r, H(r\ ;\ San(DB)) \oplus g(DB)$

  where H is a suitable hash function, r=H(DB)
- By itself, does not leak anything about DB
- Together with San(DB), reveals g(DB)

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

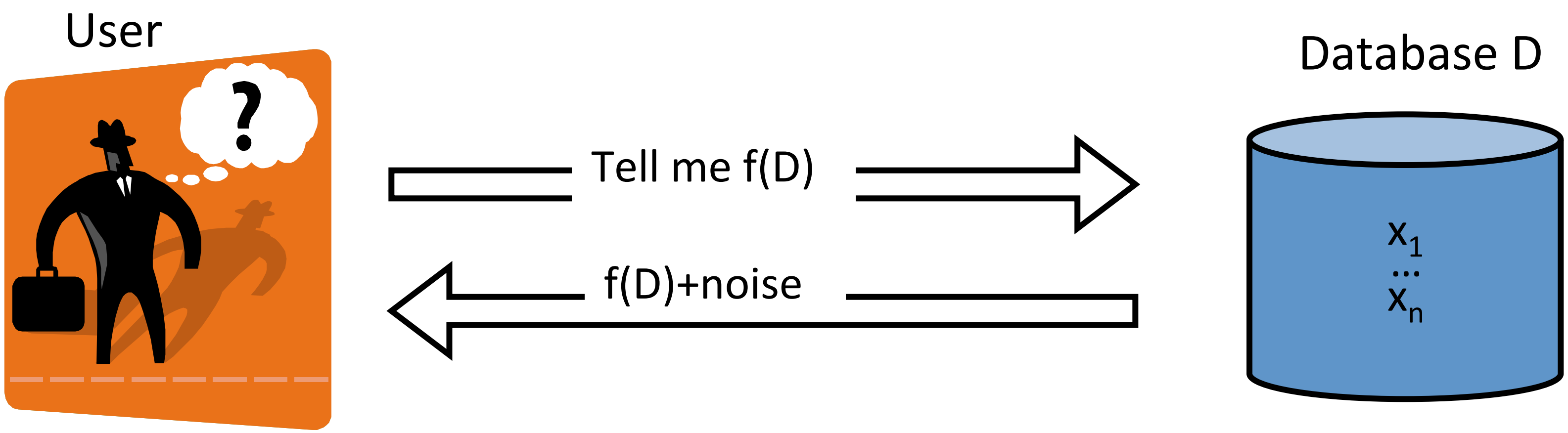
Differential Privacy: Definition

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

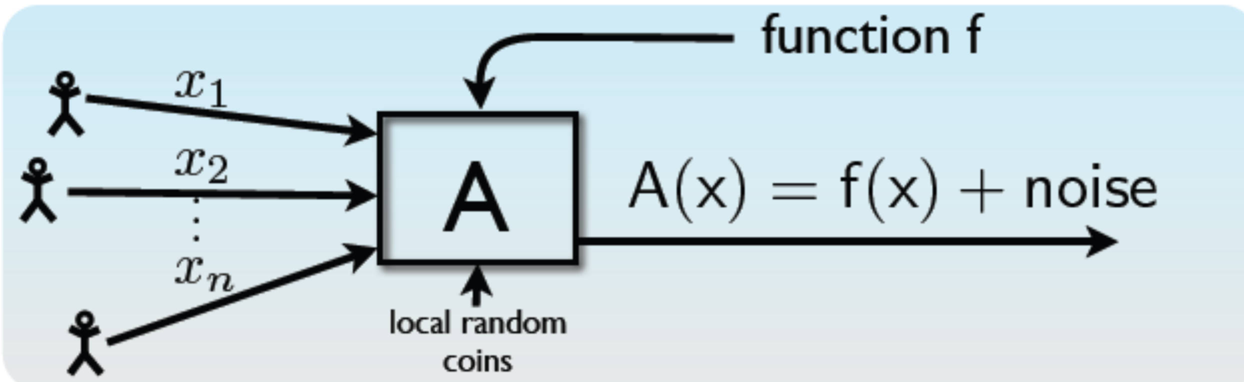
Achieving Differential Privacy: Interactive Setting



How much and what type of noise should be added?

Slide: Adam Smith

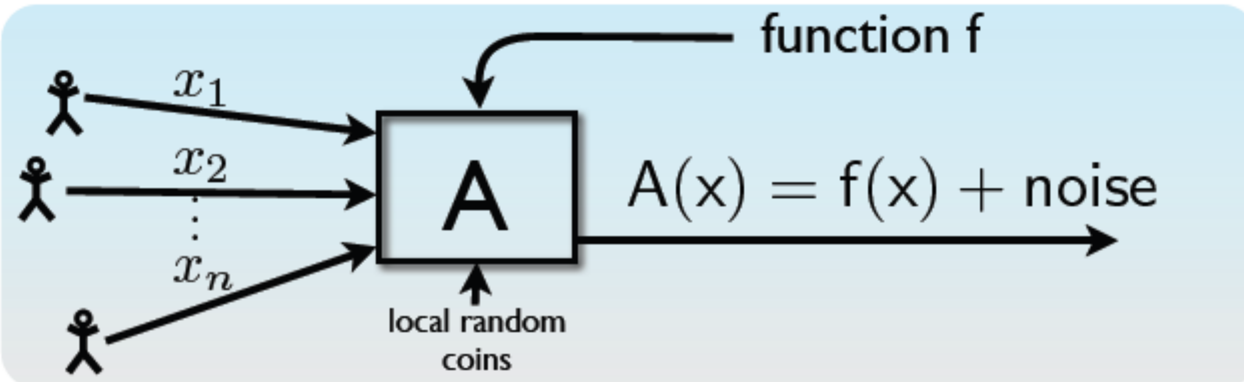
Example: Noise Addition



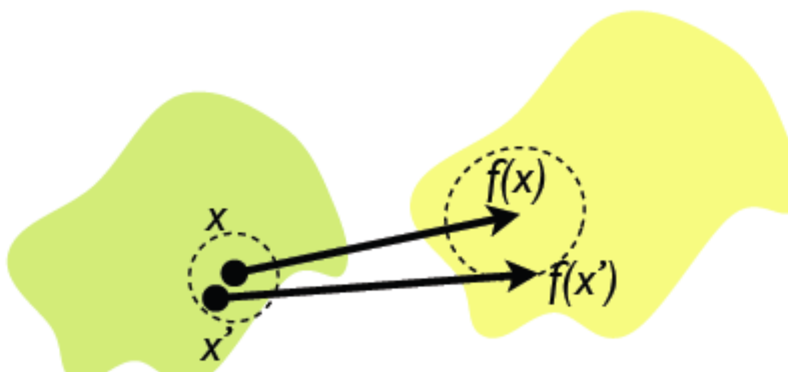
- Say we want to release a summary $f(x) \in \mathbb{R}^p$
  - e.g., proportion of diabetics: $x_i \in \{0,1\}$, $f(x) = \frac{1}{n}\sum x_i$
- Simple approach: add noise to f(x)
  - How much noise is needed?
- Intuition: f(x) can be released accurately when f is insensitive to individual entries $x_1,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

Slide: Adam Smith

Global Sensitivity



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



Exercise

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

- Answer: 1

Background on Probability Theory
(see Oct 11, 2013 recitation)

Continuous Probability Distributions

- Probability density function (PDF), f_X

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

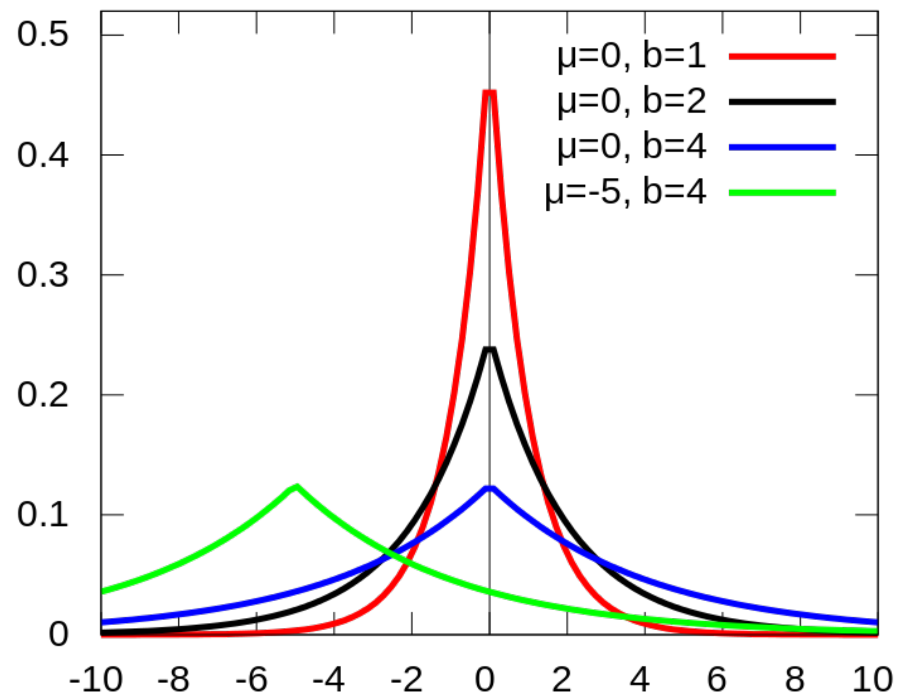
- Example distributions
  - Normal, exponential, Gaussian, Laplace

Laplace Distribution

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

Mean = μ

Variance = $2b^2$



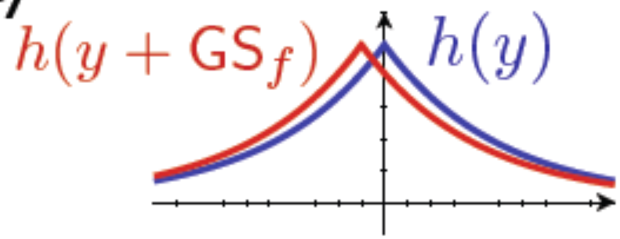
Source: Wikipedia

Laplace Distribution

- Laplace distribution $\mathsf{Lap}(\lambda)$ has density

$$h(y) \propto e^{-|y|/\lambda}$$

- Changing one point translates curve



Change of notation from previous slide:

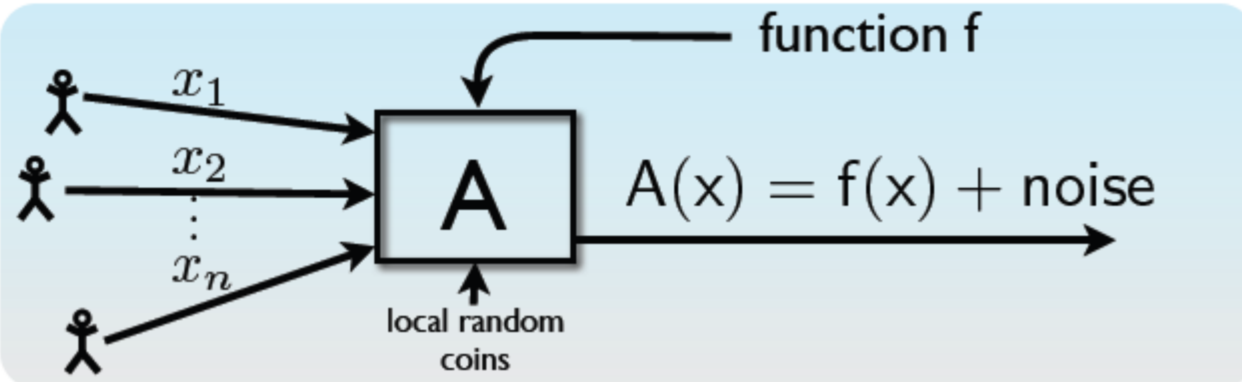
x → y μ → 0

b → λ

Achieving Differential Privacy

Slide: Adam Smith

Laplace Mechanism



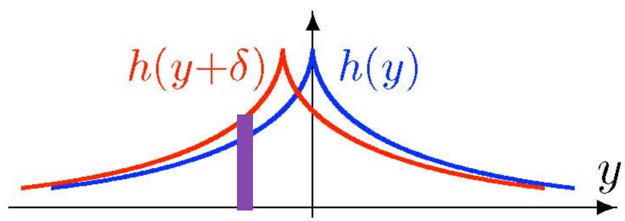
- Global Sensitivity: $GS_f = \max_{\text{neighbors } x, x'} \|f(x) - f(x')\|_1$
  - Example: $GS_{\text{proportion}} = \frac{1}{n}$

Theorem: If $A(x) = f(x) + \text{Lap}\left(\frac{GS_f}{\epsilon}\right)$, 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 $h(y) \propto e^{-\frac{\|y\|_1}{\lambda}}$



$$\frac{\Pr[A(x) = t]}{\Pr[A(x') = t]}$$

Sliding property of $\mathsf{Lap}\left(\frac{\mathsf{GS}_f}{\varepsilon}\right)$: $\frac{h(y)}{h(y+\delta)} \leq e^{\varepsilon \cdot \frac{\|\delta\|}{\mathsf{GS}_f}}$ for all y, δ

Proof idea:

$A(x)$: blue curve

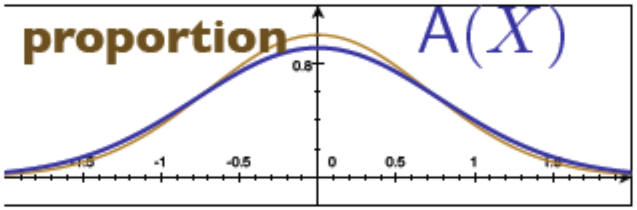
$A(x')$: red curve

$\delta = f(x) - f(x') \leq \mathsf{GS}_f$

Slide: Adam Smith

Example: Noise Addition

- Example: proportion of diabetics
  - $\mathrm{GS}_{\text{proportion}} = \frac{1}{\mathsf{n}}$
  - Release $\mathrm{A(x)} = \text{proportion} \pm \frac{1}{\epsilon n}$
- Is this **a lot**?
  - If x is a random sample from a large underlying population, then **sampling noise** $\approx \frac{1}{\sqrt{n}}$
  - A(x) "as good as" real proportion



Using Global Sensitivity

- Many natural functions have low global sensitivity
  - Histogram, covariance matrix, strongly convex optimization problems

Composition Theorem

- If A_1 is ε_1-differentially private and A_2 is ε_2-differentially private and they use independent random coins then $< A_1, A_2 >$ is $(\varepsilon_1+\varepsilon_2)$-differentially private

- Repeated querying degrades privacy; degradation is quantifiable

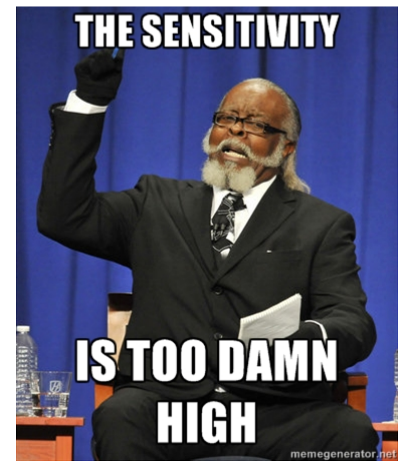
Applications

- Netflix data set [McSherry, Mironov 2009; MSR]
  - Accuracy of differentially private recommendations (wrt one movie rating) comparable to baseline set by Netflix
- Network trace data sets [McSherry, Mahajan 2010; MSR]

Packet-level analyses		High accuracy
Packet size and port dist.	(§5.1.1)	strong privacy
Worm fingerprinting [27]	(§5.1.2)	weak privacy
Flow-level analyses		
Common flow properties [30]	(§5.2.1)	strong privacy
Stepping stone detection [33]	(§5.2.2)	medium privacy
Graph-level analyses		
Anomaly detection [13]	(§5.3.1)	strong privacy
Passive topology mapping [9]	(§5.3.2)	weak privacy

Challenge: High Sensitivity

- Approach: Add noise proportional to sensitivity to preserve ε-differential privacy



- Improvements:
  - Smooth sensitivity [Nissim, Raskhodnikova, Smith 2007; BGU-PSU]
  - Restricted sensitivity [Blocki, Blum, Datta, Sheffet 2013; CMU]

Challenge: Identifying an Individual's Information

- Information about an individual may not be just in their own record
  - Example: In a social network, information about node A also in node B *influenced* by A, for example, because A may have caused a link between B and C

Differential Privacy: Summary

- An approach to releasing privacy-preserving statistics
- A rigorous privacy guarantee
  - Significant activity in theoretical CS community
- Several applications to real data sets
  - Recommendation systems, network trace data,..
- Some challenges
  - High sensitivity, identifying individual's information, repeated querying