Administrative

- HW2 will be graded and statistics posted by the end of today
- HW3 due next Friday
- Recitation next Friday cancelled due to the mid-semester break
- Office hours next Friday cancelled because I'm out of town
  - Find me by next Monday for in-person questions on the homework

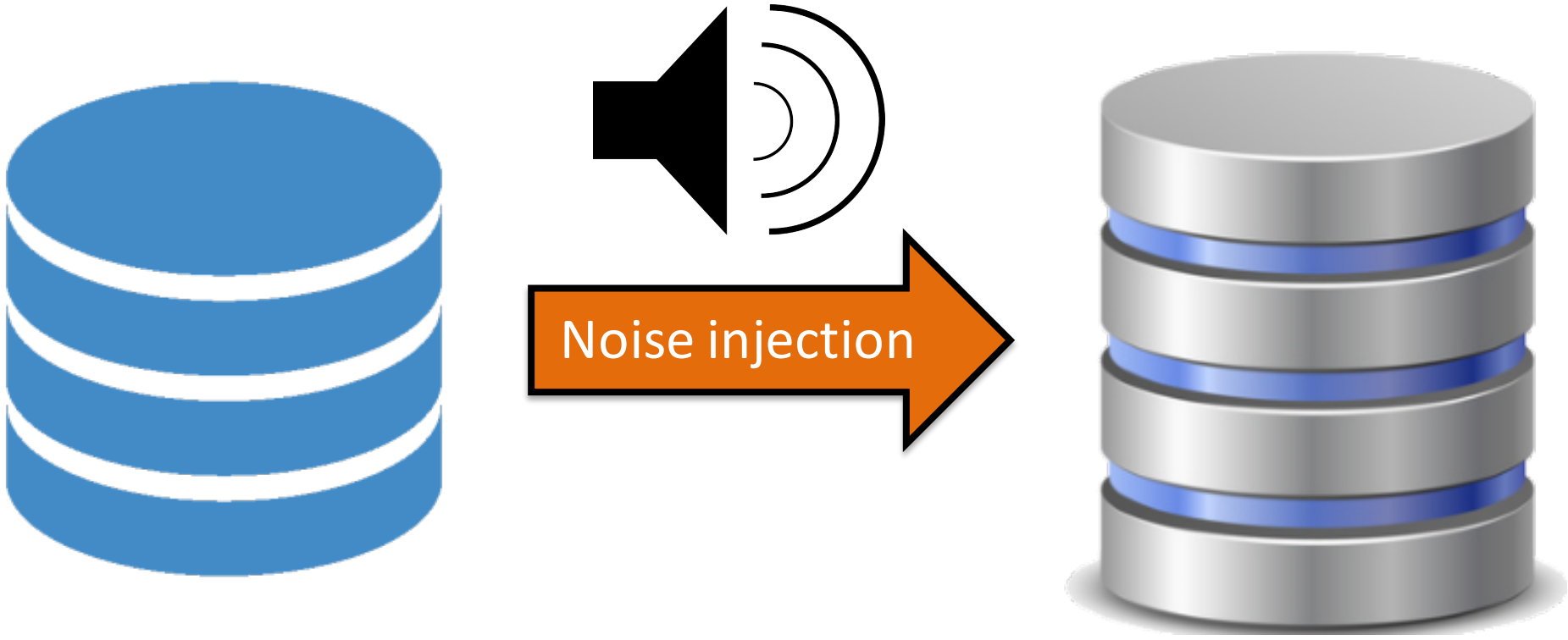
HW2 Statistics

Mean	86
Median	86
Standard Deviation	5.5
Max	94
Min	73

18734 Recitation

Differential Privacy and the Laplace Mechanism

Database Sanitization



Differential Privacy

Control

Experimental

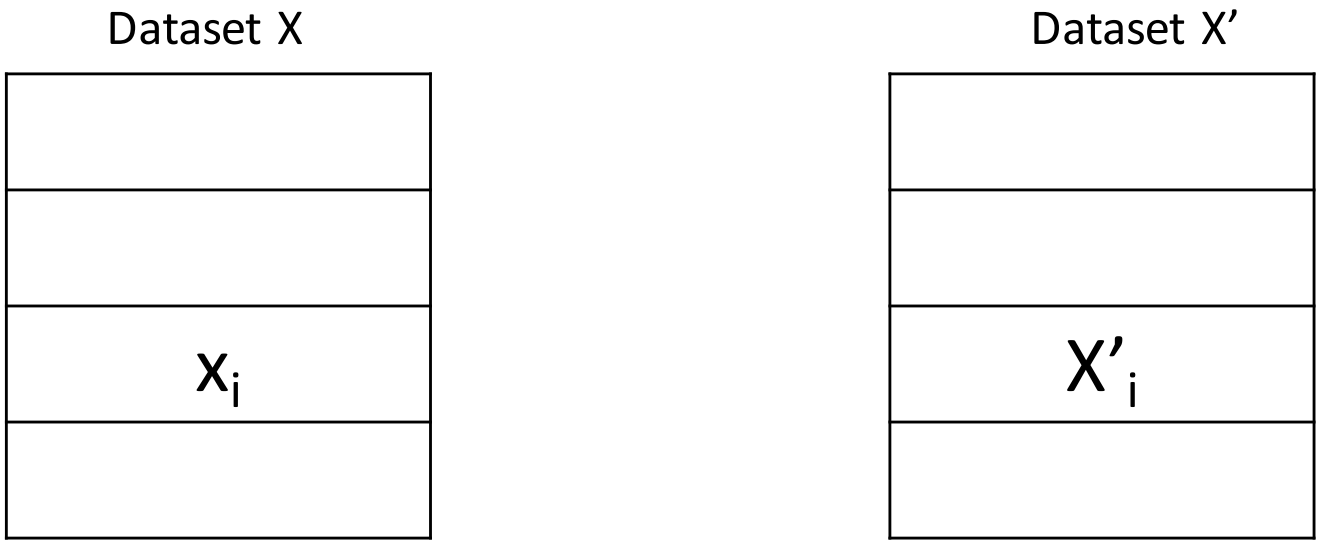
Does smoking increase your chances of getting cancer?

Control

Experimental

Does smoking increase your chances of getting cancer?

Differential Privacy: Definition



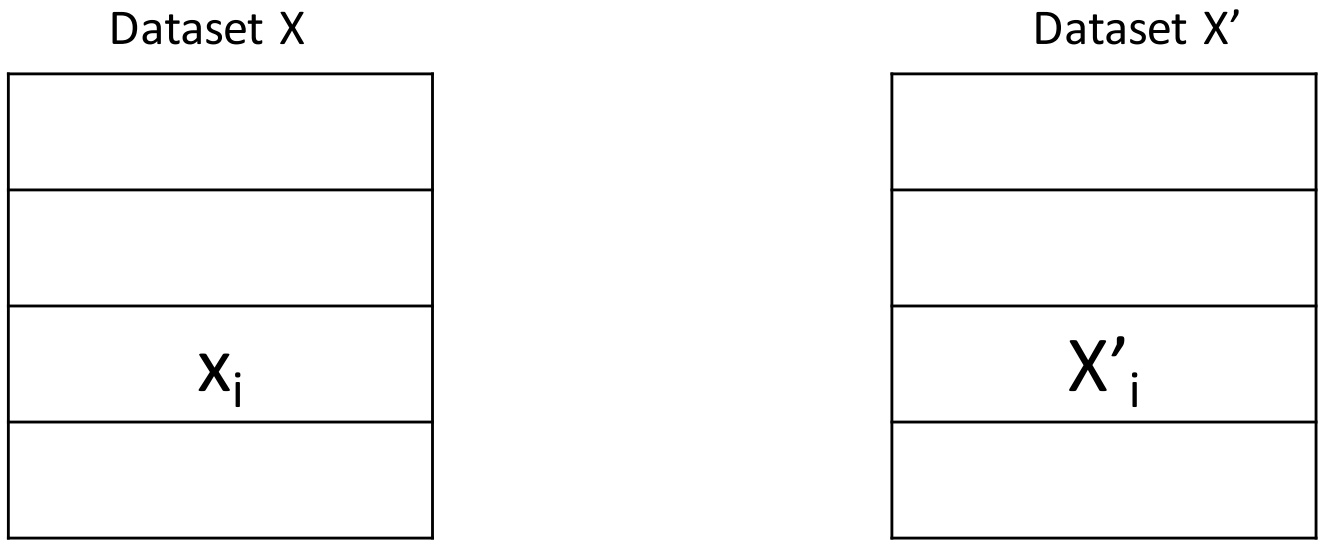
Function K runs on X to produce K(X)

K is ε-differentially privacy if for any two neighbor datasets X, X' that differ in just one element at position i:

For all outcomes v,

$$e^{-\varepsilon} \leq \frac{\Pr(K(X)=v)}{\Pr(K(X')=v)} \leq e^{\varepsilon}$$

Approximate Differential Privacy: Definition



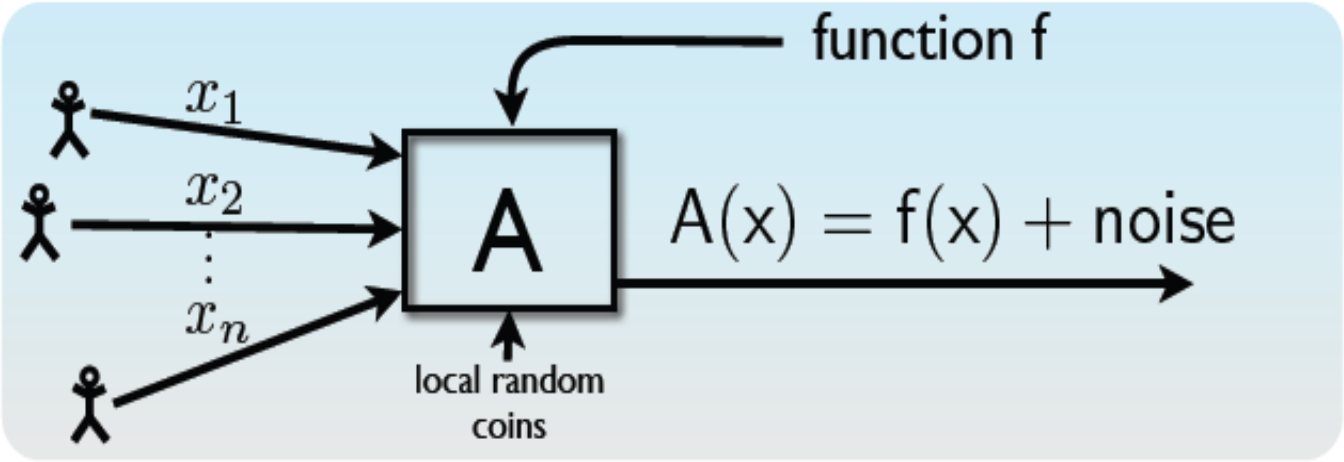
Function K runs on X to produce K(X)

K has ε-differentially privacy if for any two neighbor datasets X, X' that differ in just one element at position i:

For all outcomes v,

$$1 - \varepsilon \approx e^{-\varepsilon} \leq \frac{\Pr(K(X)=v)}{\Pr(K(X')=v)} \leq e^{\varepsilon} \approx 1 + \varepsilon$$

Laplace Mechanism



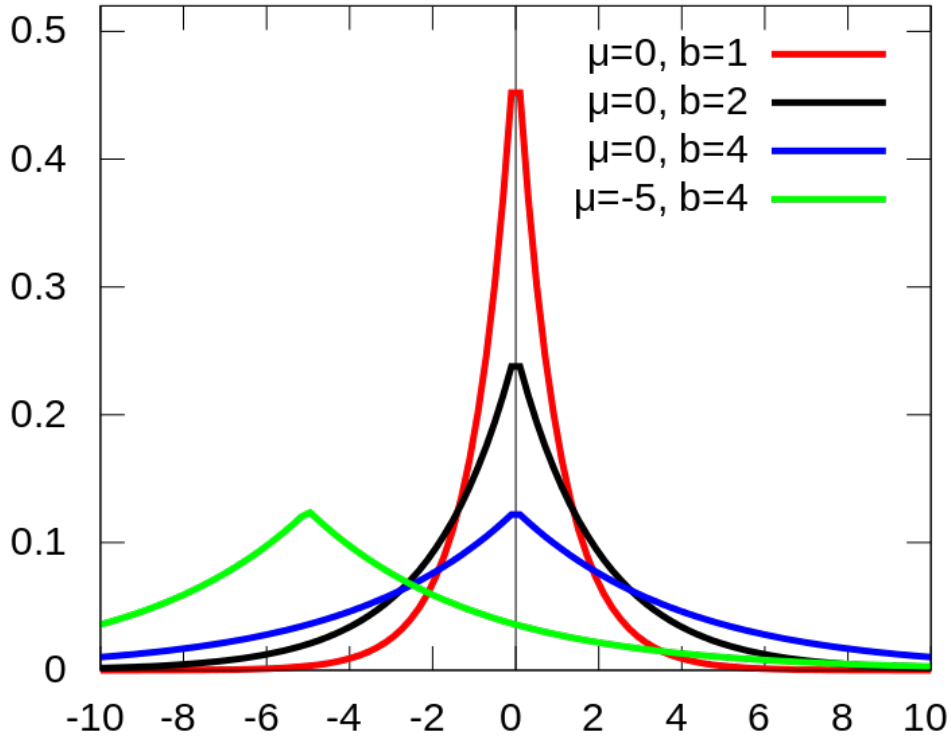
- Global Sensitivity: $GS_f = \max_{\text{neighbors } x,x'} \|f(x) - f(x')\|_1$

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

Slide: Adam Smith

Laplace Distribution

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



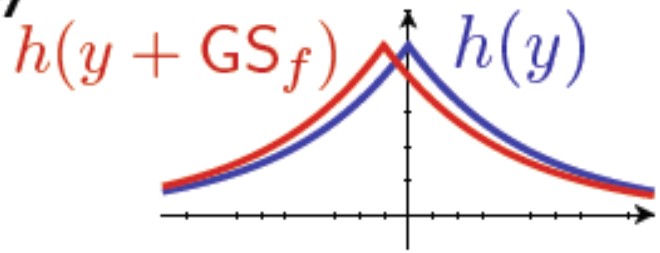
Source: Wikipedia

Laplace Distribution

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

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

- Changing one point translates curve



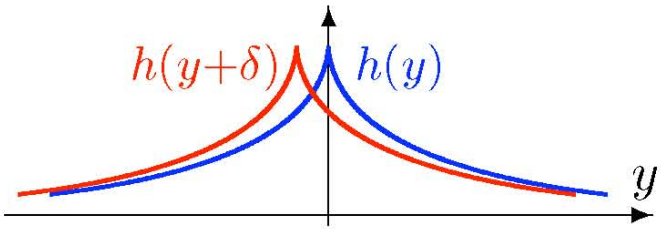
$$\frac{1}{2b} exp(-\frac{|y-\mu|}{b})$$

Change of notation from previous slide:
μ → 0
b → λ

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 $h(y) \propto e^{-\frac{\|y\|_1}{\lambda}}$



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

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

Laplace Mechanism: Proof

- To Prove:

$$\Pr[A(x) \in S] \leq e^{\varepsilon} \Pr[A(x') \in S]$$

  - $A(x) = f(x) + \text{Laplace}(GS_f/\varepsilon)$
  - $A(x') = f(x') + \text{Laplace}(GS_f/\varepsilon)$

- Steps:
  - Distribution of $A(x)$: $\text{Laplace}(f(x), GS_f/\varepsilon)$
  - Distribution of $A(x')$: $\text{Laplace}(f(x'), GS_f/\varepsilon)$

Laplace Mechanism: Proof

$$\frac{Pr[A(x) \in S]}{Pr[A(x') \in S]} = \frac{e^{-|y-f(x)|/\lambda}}{e^{-|y-f(x')|/\lambda}}$$

$$= e^{\frac{|y-f(x')|-|y-f(x)|}{\lambda}}$$

$$\leq e^{\frac{|f(x)-f(x')|}{\lambda}} \leq e^{\frac{GS_f \epsilon}{GS_f}} = e^{\epsilon}$$