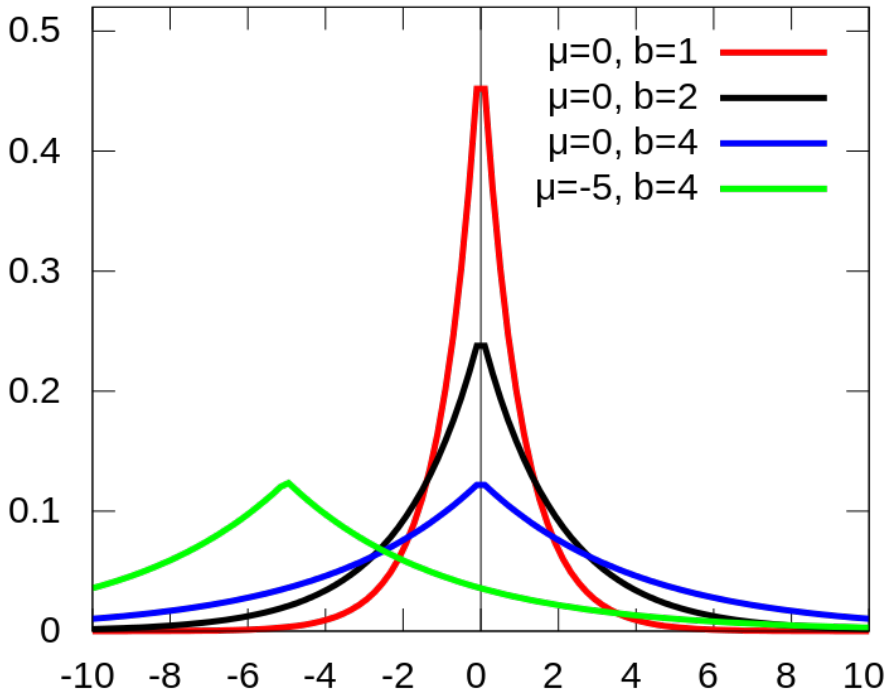
18734 Recitation

Laplace Mechanism

Laplace Distribution

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



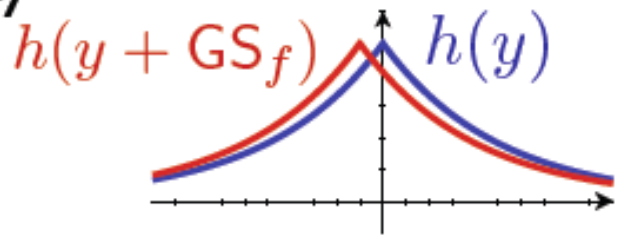
Source: Wikipedia

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

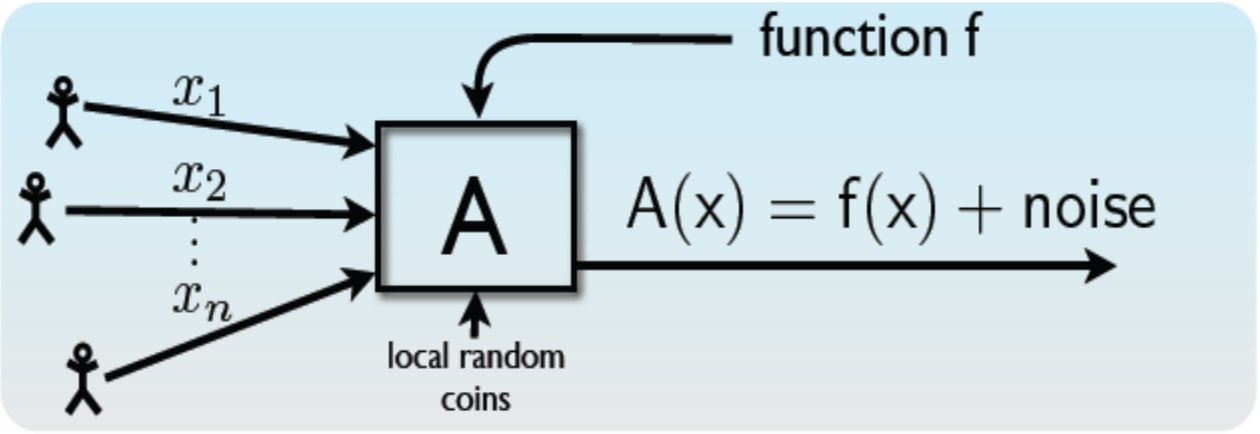
Differential Privacy: Definition

Randomized function A 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 A,

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

Slide: Adam Smith

Laplace Mechanism



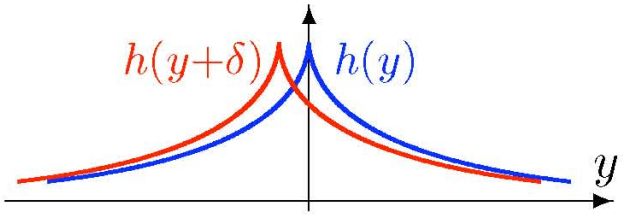
- Global Sensitivity: $\mathrm{GS}_f = \max_{\text{neighbors } x,x'} \|f(x) - f(x')\|_1$

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 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

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

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

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