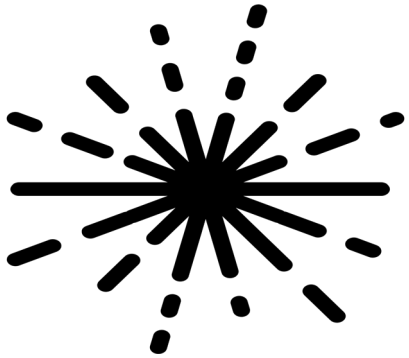


CS208: Applied Privacy for Data Science
Introduction to Differential Privacy

School of Engineering & Applied Sciences
Harvard University

February 8, 2022

Attacks on Aggregate Stats

For releasing d population proportions on a dataset of size n:

Questions:

- If we allow error greater than $\sqrt{d}/n$, can we prevent these attacks?
- Can we reason about unforeseen attacks?

Goals of Differential Privacy

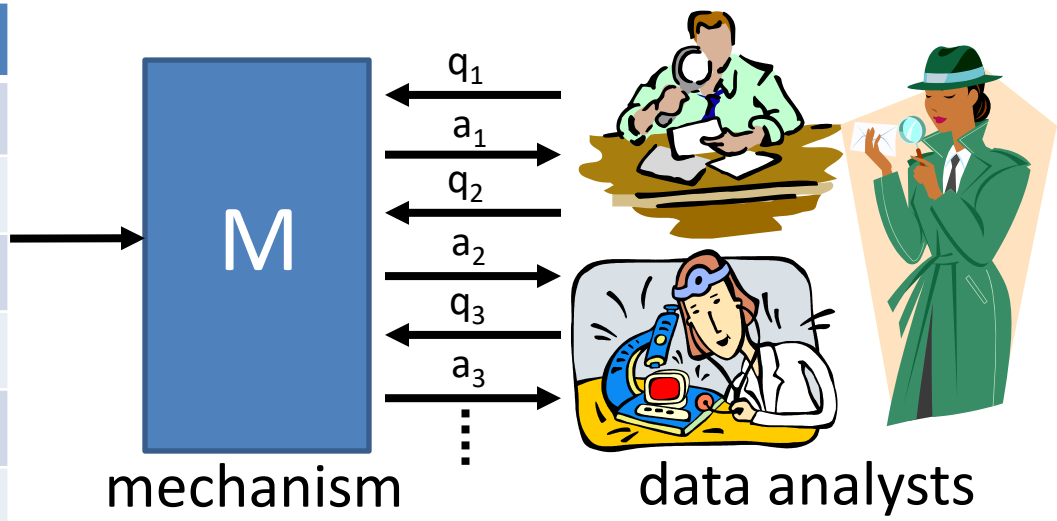
- Utility: enable “statistical analysis” of datasets
  - e.g. inference about population, ML training, useful descriptive statistics

- Privacy: protect individual-level data
  - against “all” attack strategies, auxiliary info.

Differential privacy

[Dinur-Nissim ’03+Dwork, Dwork-Nissim ’04, Blum-Dwork-McSherry-Nissim ’05, Dwork-McSherry-Nissim-Smith ’06]

Sex	Blood	…	HIV?
F	B	…	Y
M	A	…	N
M	O	…	N
M	O	…	Y
F	A	…	N
M	B	…	Y

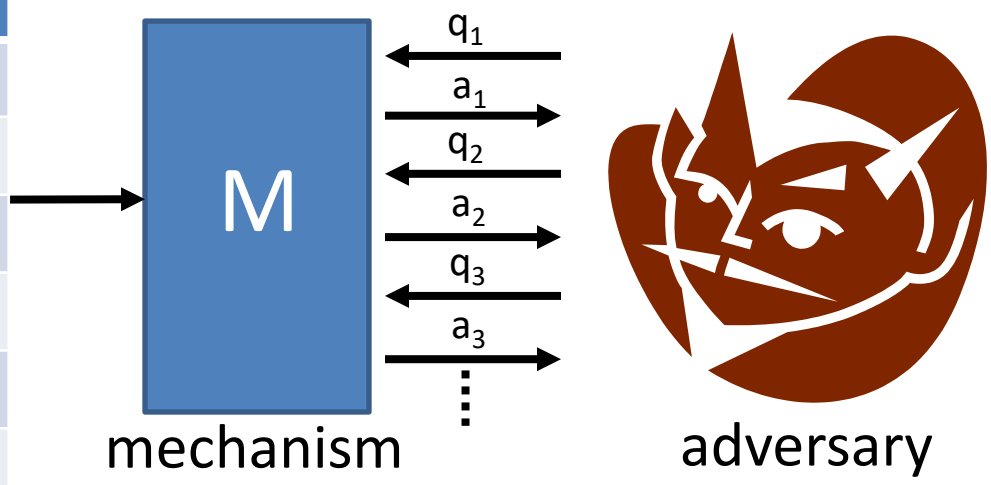


Requirement: effect of each individual should be “hidden”

Differential privacy

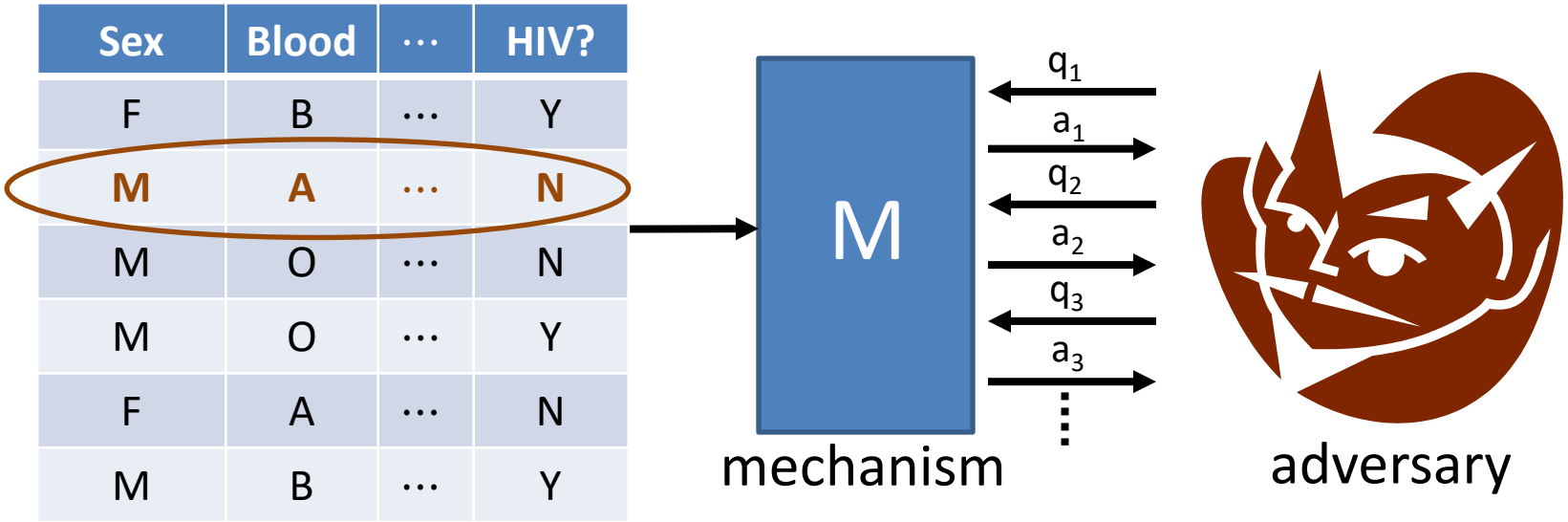
[Dinur-Nissim ’03+Dwork, Dwork-Nissim ’04, Blum-Dwork-McSherry-Nissim ’05, Dwork-McSherry-Nissim-Smith ’06]

Sex	Blood	…	HIV?
F	B	…	Y
M	A	…	N
M	O	…	N
M	O	…	Y
F	A	…	N
M	B	…	Y



Differential privacy

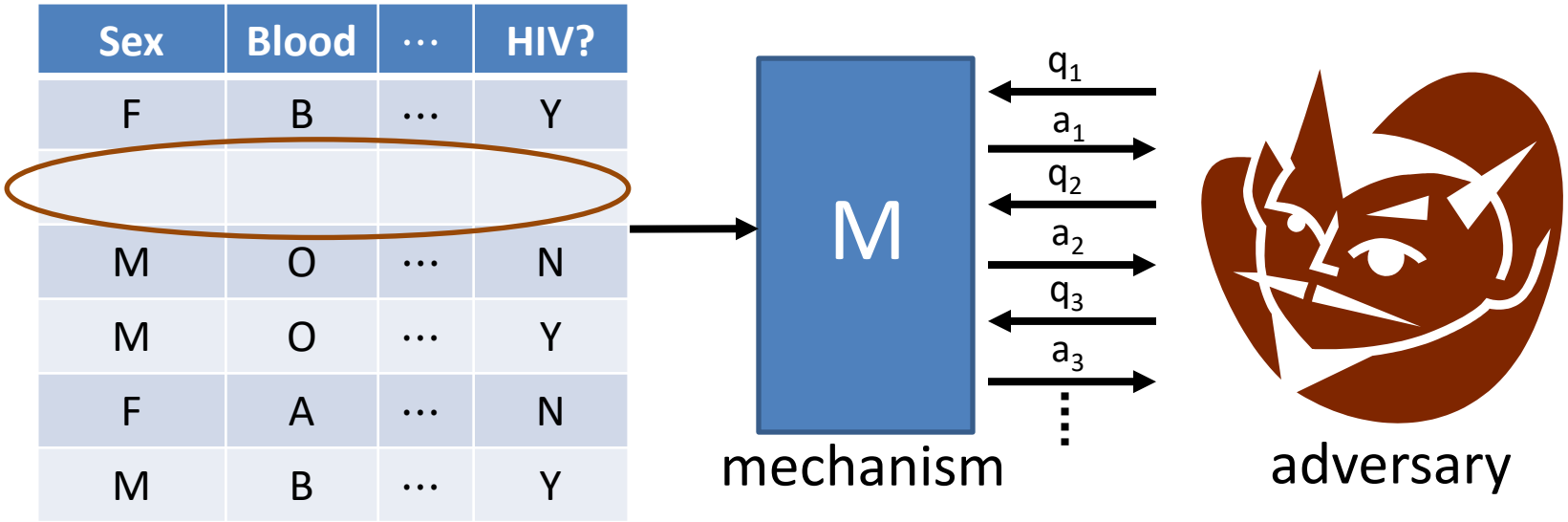
[Dinur-Nissim ’03+Dwork, Dwork-Nissim ’04, Blum-Dwork-McSherry-Nissim ’05, Dwork-McSherry-Nissim-Smith ’06]



Requirement: an adversary shouldn’t be able to tell if any one person’s data were changed arbitrarily

Differential privacy

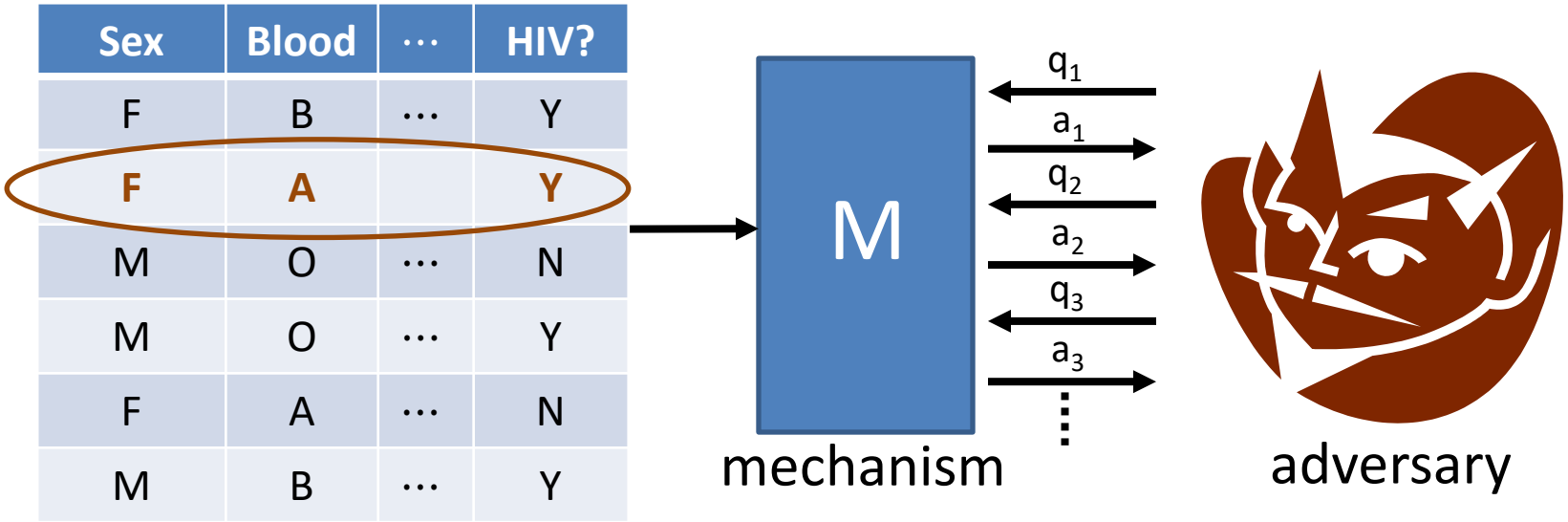
[Dinur-Nissim ’03+Dwork, Dwork-Nissim ’04, Blum-Dwork-McSherry-Nissim ’05, Dwork-McSherry-Nissim-Smith ’06]



Requirement: an adversary shouldn’t be able to tell if any one person’s data were changed arbitrarily

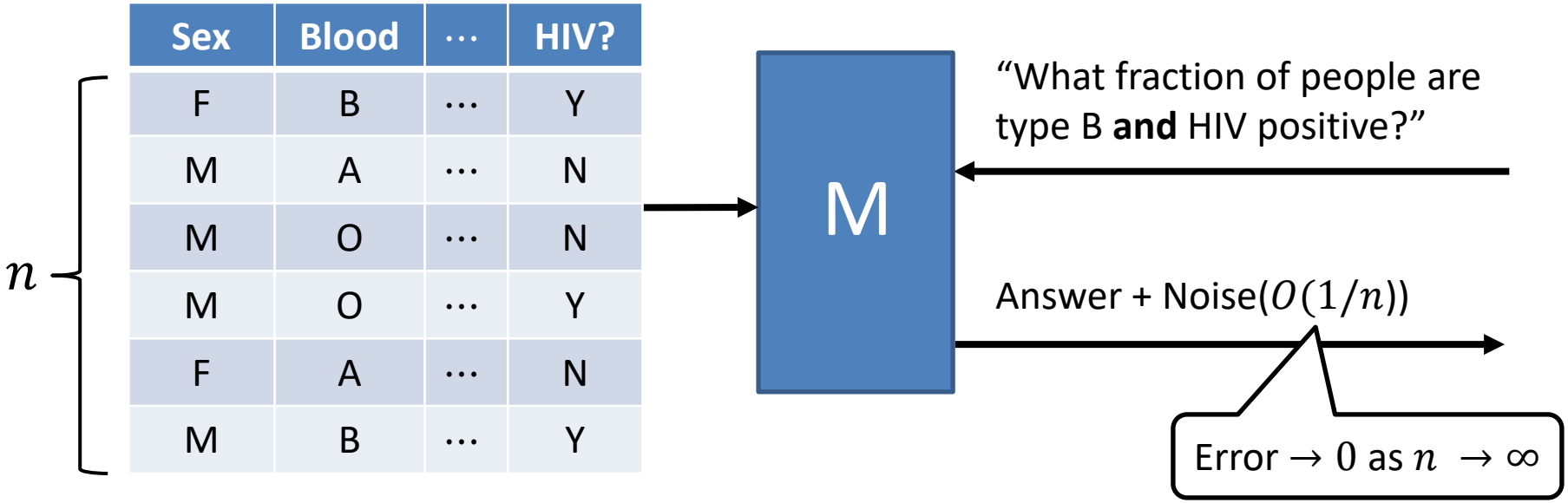
Differential privacy

[Dinur-Nissim ’03+Dwork, Dwork-Nissim ’04, Blum-Dwork-McSherry-Nissim ’05, Dwork-McSherry-Nissim-Smith ’06]



Requirement: an adversary shouldn’t be able to tell if any one person’s data were changed arbitrarily

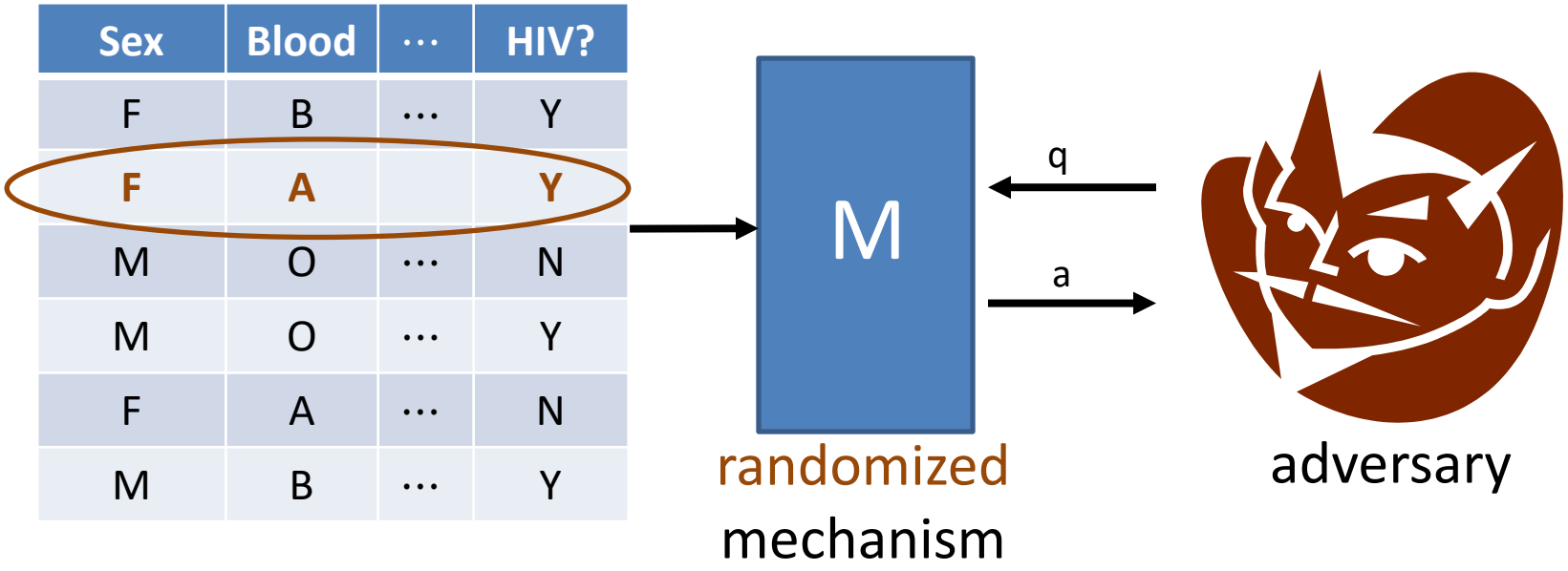
Simple approach: random noise



- Very little noise needed to hide each person as $n \to \infty$.
- **Note:** this is just for one query

DP for one query/release

[Dinur-Nissim ’03+Dwork, Dwork-Nissim ’04, Blum-Dwork-McSherry-Nissim ’05, Dwork-McSherry-Nissim-Smith ’06]

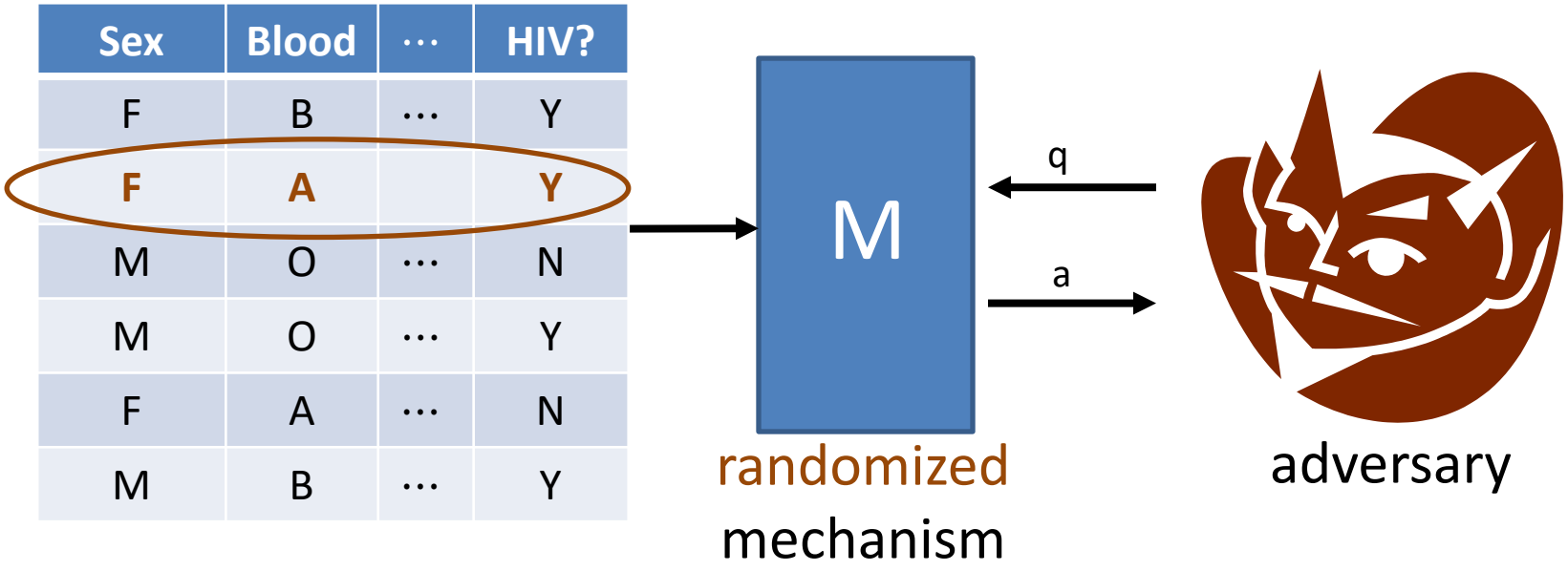


Requirement: for all x, x' differing on one row, and all q

Distribution of $M(x, q) \approx_\varepsilon$ Distribution of $M(x', q)$

DP for one query/release

[Dinur-Nissim ’03+Dwork, Dwork-Nissim ’04, Blum-Dwork-McSherry-Nissim ’05, Dwork-McSherry-Nissim-Smith ’06]

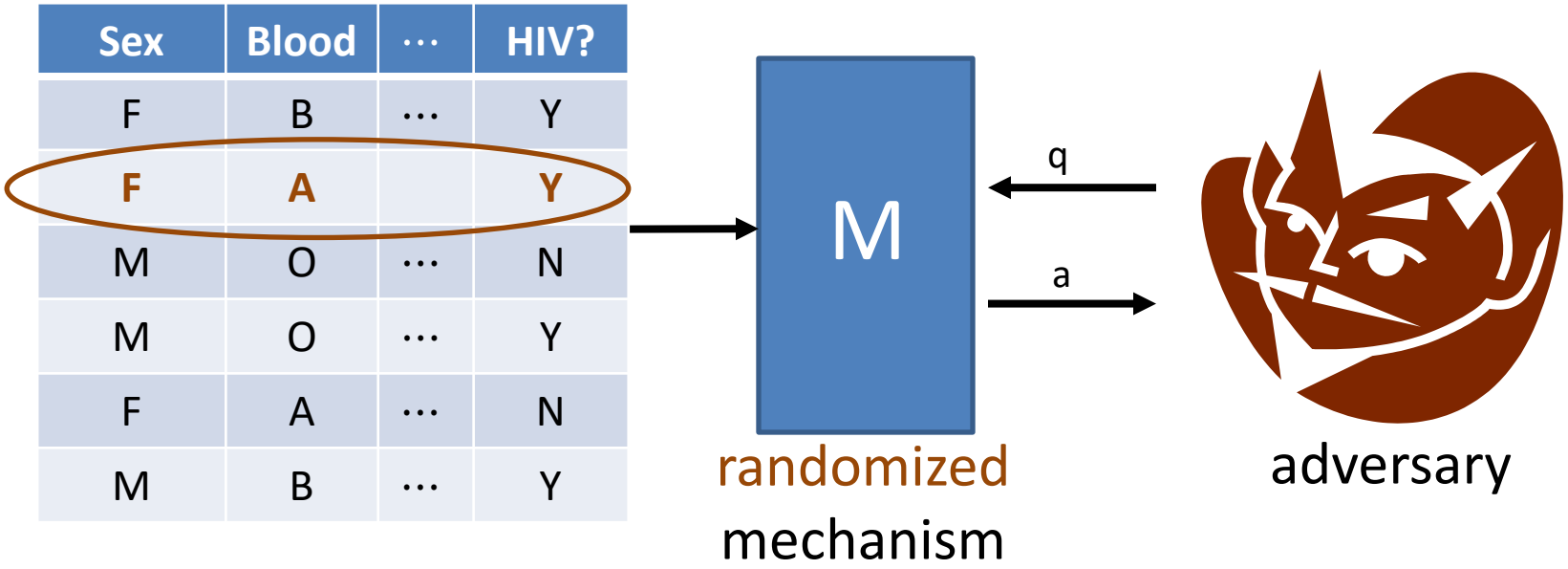


Requirement: for all x, x' differing on one row, and all q

$\forall$ sets T, $\Pr[M(x,q) \in T] \lesssim (1+\varepsilon) \cdot \Pr[M(x',q) \in T]$

DP for one query/release

[Dwork-McSherry-Nissim-Smith '06]



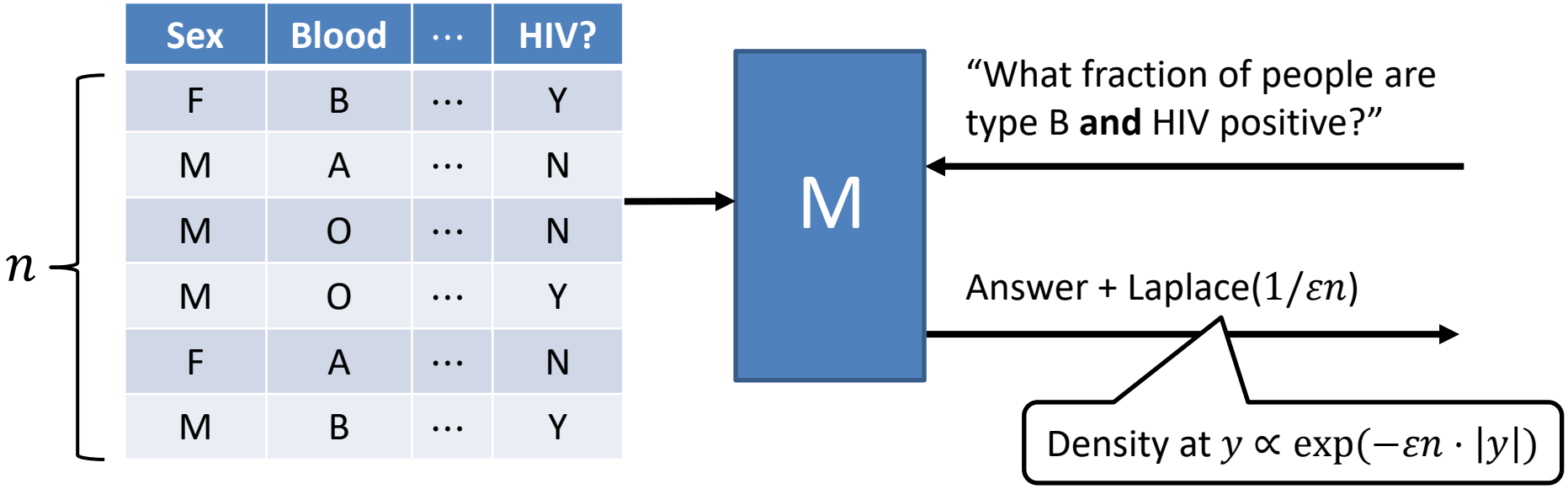
Def: M is ε-DP if for all x, x' differing on one row, and all q

$\forall$ sets T, $\quad \Pr[M(x,q) \in T] \leq e^{\varepsilon} \cdot \Pr[M(x',q) \in T]$

(Probabilities are (only) over the randomness of M.)

The Laplace Mechanism

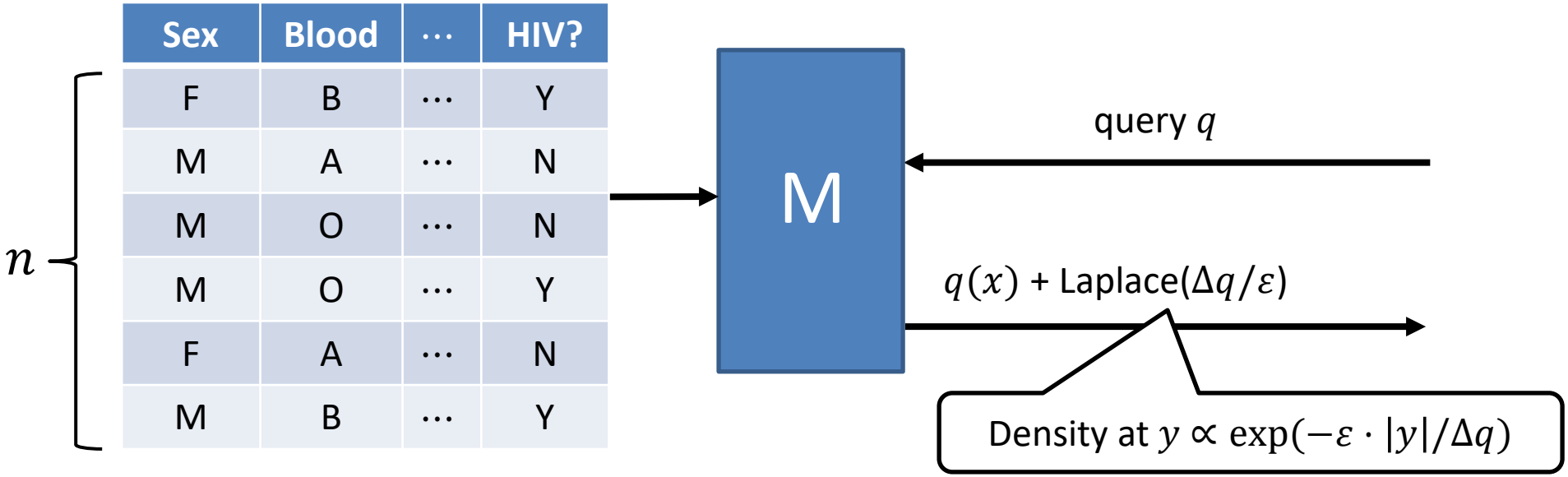
[Dwork-McSherry-Nissim-Smith '06]



- Very little noise needed to hide each person as $n \to \infty$.

The Laplace Mechanism

[Dwork-McSherry-Nissim-Smith '06]



- Very little noise needed to hide each person as $n \to \infty$.

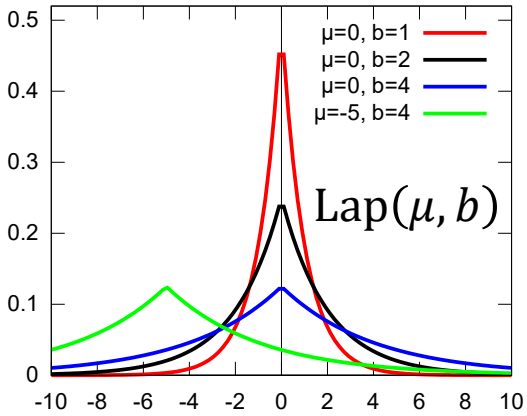
The Laplace Mechanism

[Dwork-McSherry-Nissim-Smith '06]

- Let $\mathcal{X}$ be a data universe, and $\mathcal{X}^n$ a space of datasets.
  - This is the Bounded DP setting: n known and public.
- For $x, x' \in \mathcal{X}^n$, write $x \sim x'$ if x and x' differ on ≤ 1 row.
- For a query $q : \mathcal{X}^n \to \mathbb{R}$, the global sensitivity is

$$\Delta q = \mathrm{GS}_q = \max_{x \sim x'} |q(x) - q(x')|.$$

- The Laplace distribution with scale s, $\mathrm{Lap}(s)$:
  - Has density function $f(y) = e^{-|y|/s}/2s$.
  - Mean 0, standard deviation $\sqrt{2} \cdot s$.



Theorem: $M(x, q) = q(x) + \mathrm{Lap}(\Delta q/\varepsilon)$ is ε-DP.

Calculating Global Sensitivity

1. $\mathcal{X} = \{0,1\}, q(x) = \sum_{i=1}^{n} x_i, \Delta q =$
2. $\mathcal{X} = \mathbb{R}, \ q(x) = \sum_{i=1}^{n} x_i \,, \Delta q =$
3. $\mathcal{X} = [0,1], \ q(x) = \text{mean}(x_1, x_2, \dots, x_n), \Delta q =$
4. $\mathcal{X} = [0,1], \ q(x) = \text{median}(x_1, x_2, \dots, x_n), \Delta q =$
5. $\mathcal{X} = [0,1], \ q(x) = \text{variance}(x_1, x_2, \dots, x_n), \Delta q =$

Q: for which of these queries is the Laplace Mechanism “useful”?

A: