Intro to Differential Privacy, Properties, Randomized Response

Katrina Ligett

What analyses on a database might violate privacy? What analyses are privacy-preserving?

what to promise?

only ask questions that pertain to large populations

maybe not

what to promise?

delete identifying information

maybe not

what to promise?

access to the output should not enable one to learn anything about an individual that could not be learned without access

is this possible?

hint: *either* privacy or utility separately is easy

what to promise?

access to the output should not enable one to learn anything about an individual that could not be learned without access

is this desirable?

what to promise?

access to the output should not enable one to learn anything about an individual that could not be learned without access

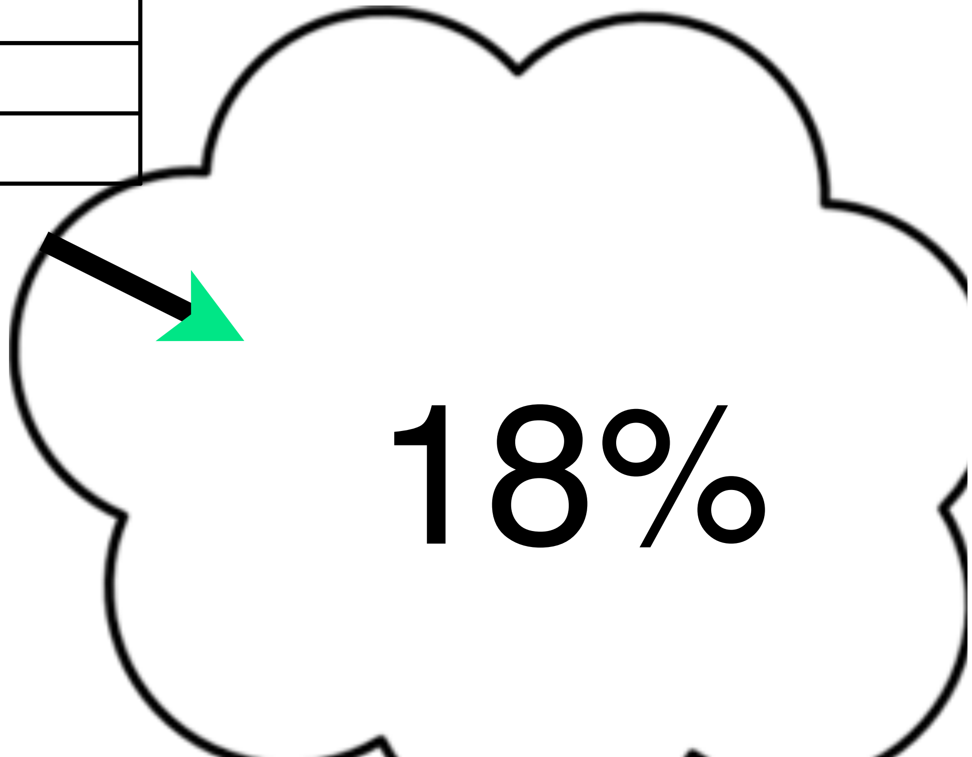
what to promise?

access to the output should not enable one to learn much more about an individual than could be learned via the same analysis omitting that individual from the database

what to promise?

think of output as randomized

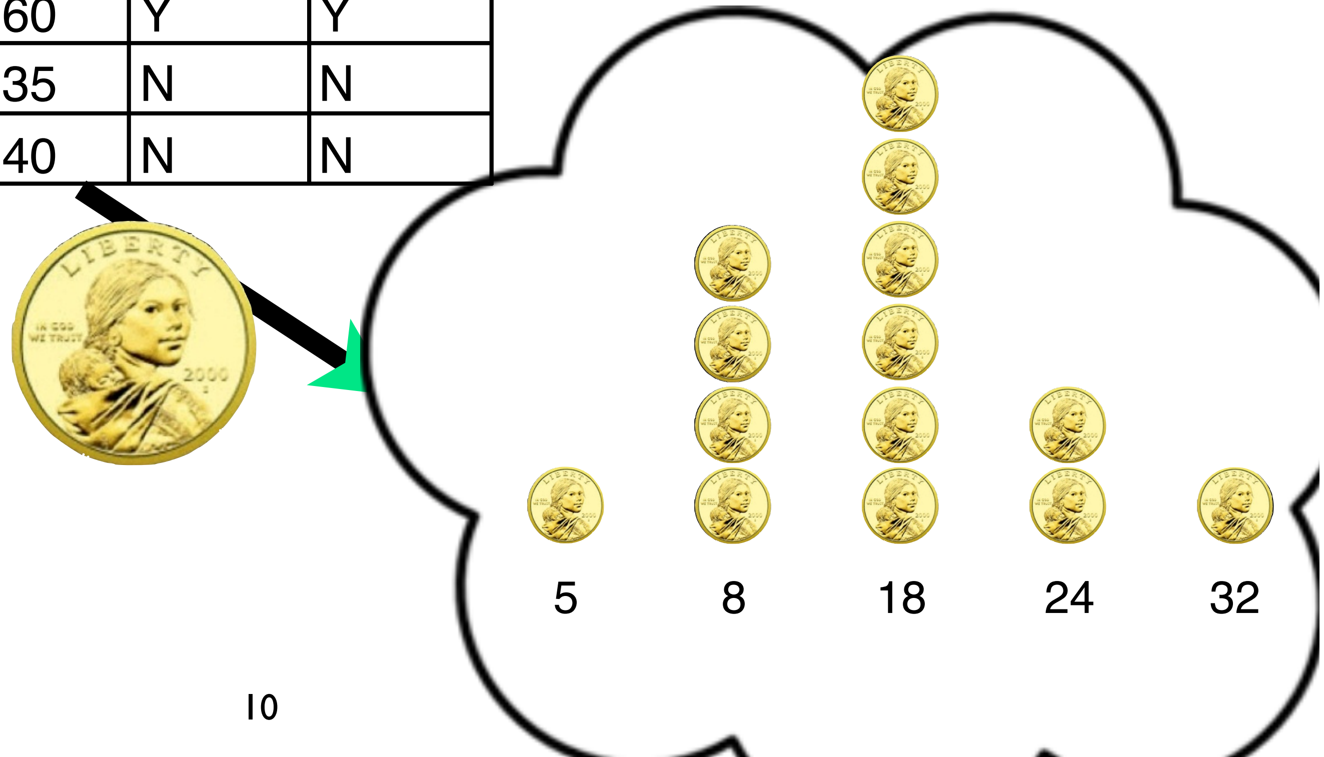
name	DOB	sex	weight	smoker	lung cancer
John Doe	12/1/51	M	185	Y	N
Jane Smith	3/3/46	F	140	N	N
Ellen Jones	4/24/59	F	160	Y	Y
Jennifer Kim	3/1/70	F	135	N	N
Rachel Waters	9/5/43	F	140	N	N



what to promise?

think of output as randomized

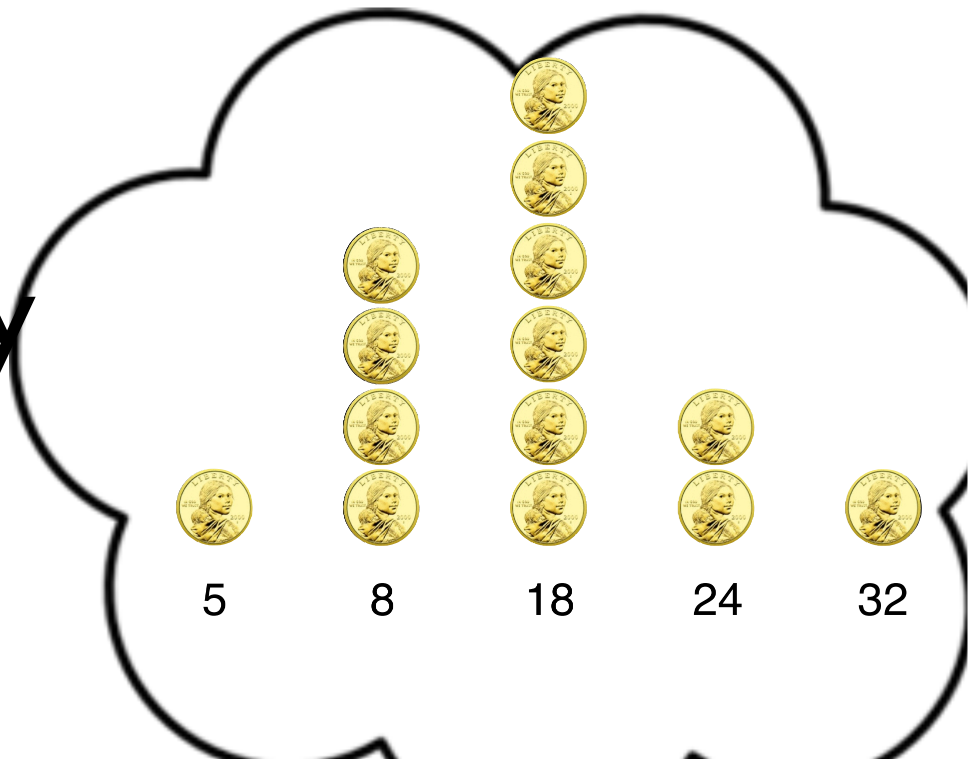
name	DOB	sex	weight	smoker	lung cancer
John Doe	12/1/51	M	185	Y	N
Jane Smith	3/3/46	F	140	N	N
Ellen Jones	4/24/59	F	160	Y	Y
Jennifer Kim	3/1/70	F	135	N	N
Rachel Waters	9/5/43	F	140	N	N



what to promise?

think of output as randomized

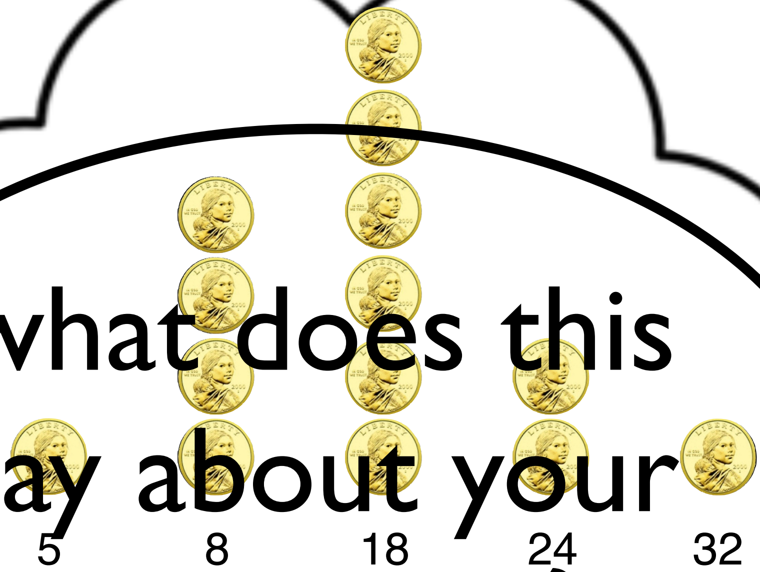
promise: if you leave the database, no outcome will change probability by very much



what to promise?

think of output as randomized

promise: if you leave the database, no outcome will change probability by very much



what does this say about your incentives?

statistical database model

X set of possible entries/rows

one row per person

database x a set of rows; $x \in \mathbb{N}^{|X|}$

(histogram)

name	DOB	sex	weight	smoker	lung cancer
John Doe	12/1/51	M	185	Y	N
Jane Smith	3/3/46	F	140	N	N
Ellen Jones	4/24/59	F	160	Y	Y
Jennifer Kim	3/1/70	F	135	N	N
Rachel Waters	9/5/43	F	140	N	N

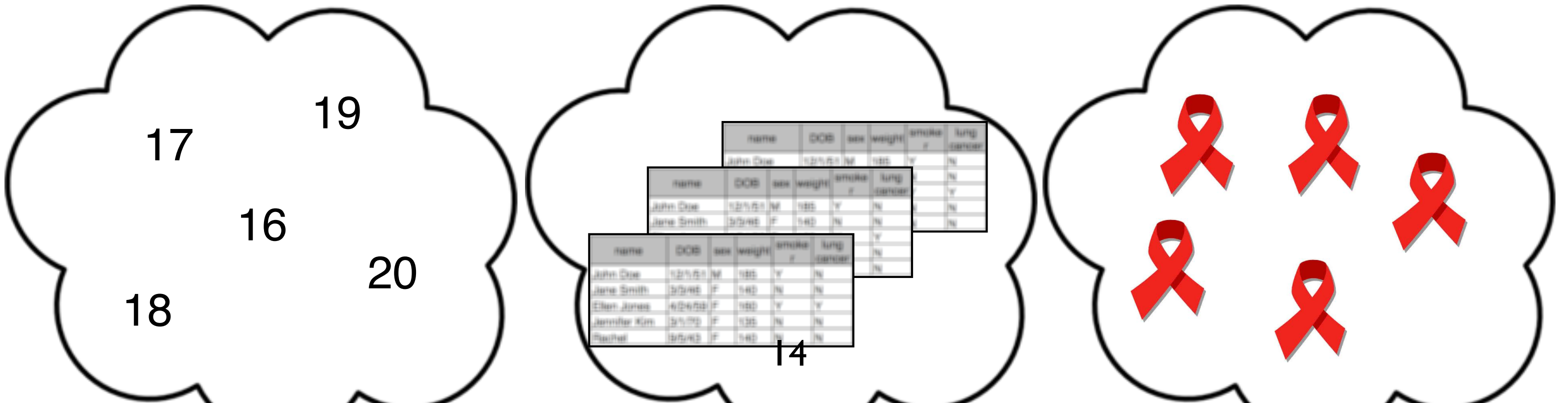
analyst objective

wishes to compute on $x \in \mathbb{N}^{|X|}$

fit a model, compute a statistic, share "sanitized" data

preserve privacy of individuals

design randomized algorithm M mapping x to into outcome space, that masks small changes in x



neighboring databases

what's a small change?

require nearly identical behavior on neighboring databases differing by the addition or removal of a single row:

$$||x - y||_1 \leq 1$$

for $x, y \in \mathbb{N}^{|X|}$

differential privacy

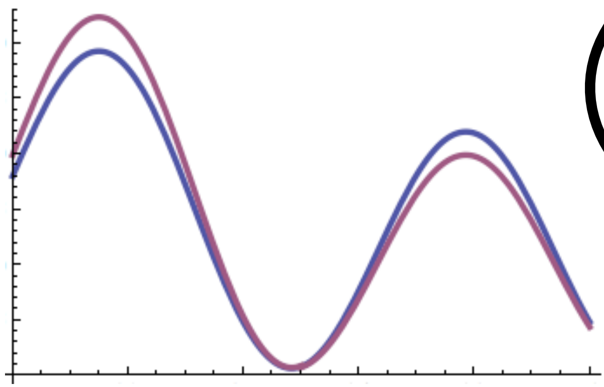
[DinurNissim03, DworkNissimMcSherrySmith06, Dwork06]

ε-Differential Privacy for algorithm M:

for any two neighboring data sets x_1, x_2, differing by the addition or removal of a single row

any $S \subseteq \text{range}(M)$,

$\Pr[M(x_1) \in S] \leq e^{\varepsilon} \Pr[M(x_2) \in S]$

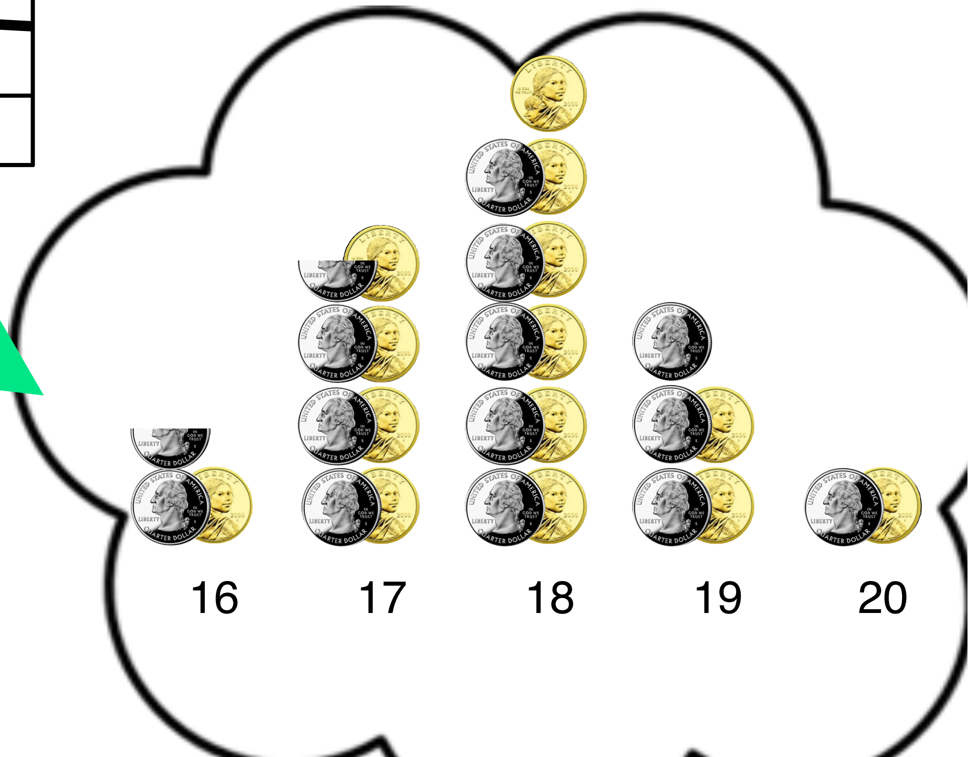


$\mathbf{e^{\varepsilon} \sim (1 + \varepsilon)}$

differential privacy

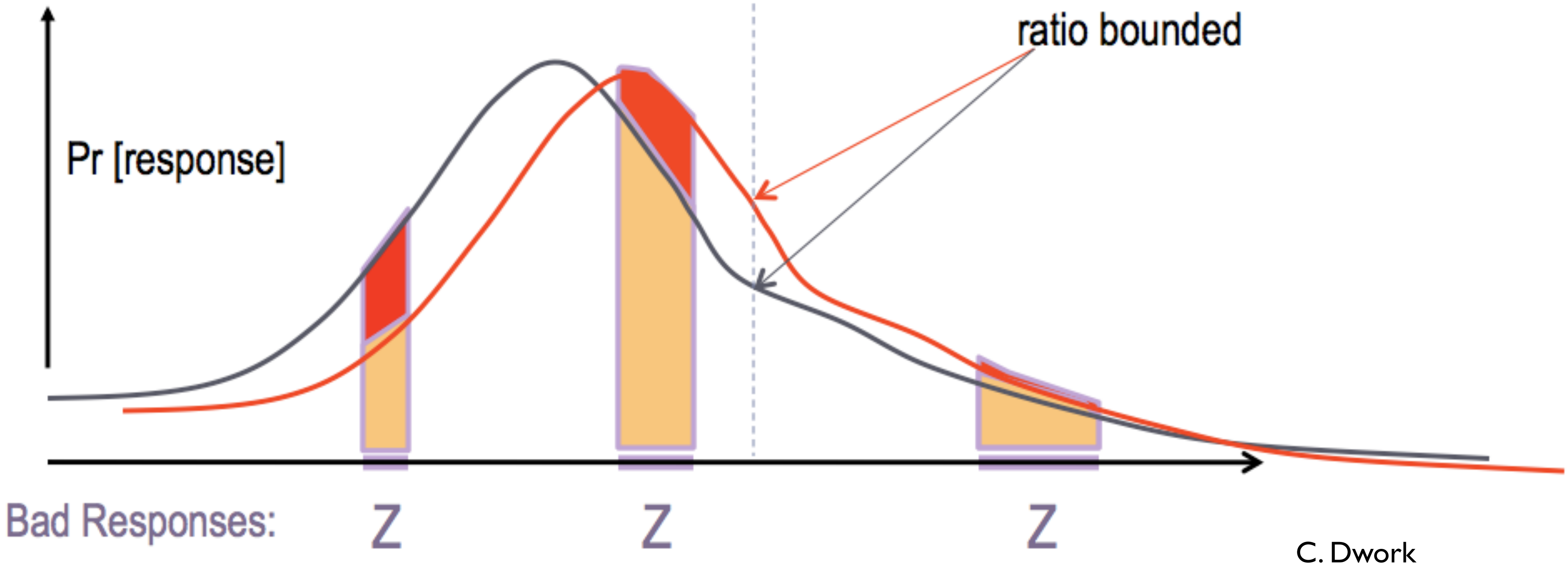
$$\Pr[M(x_1) \in S] \leq e^{\varepsilon} \Pr[M(x_2) \in S]$$

name	DOB	sex	weight	smoker	lung cancer
John Doe	12/1/51	M	185	Y	N
Jane Smith	3/3/46	F	140	N	N
~~Ellen Jones~~	~~4/24/59~~	~~F~~	~~160~~	~~Y~~	~~Y~~
Jennifer Kim	3/1/70	F	135	N	N
Rachel Waters	9/5/43	F	140	N	N



differential privacy

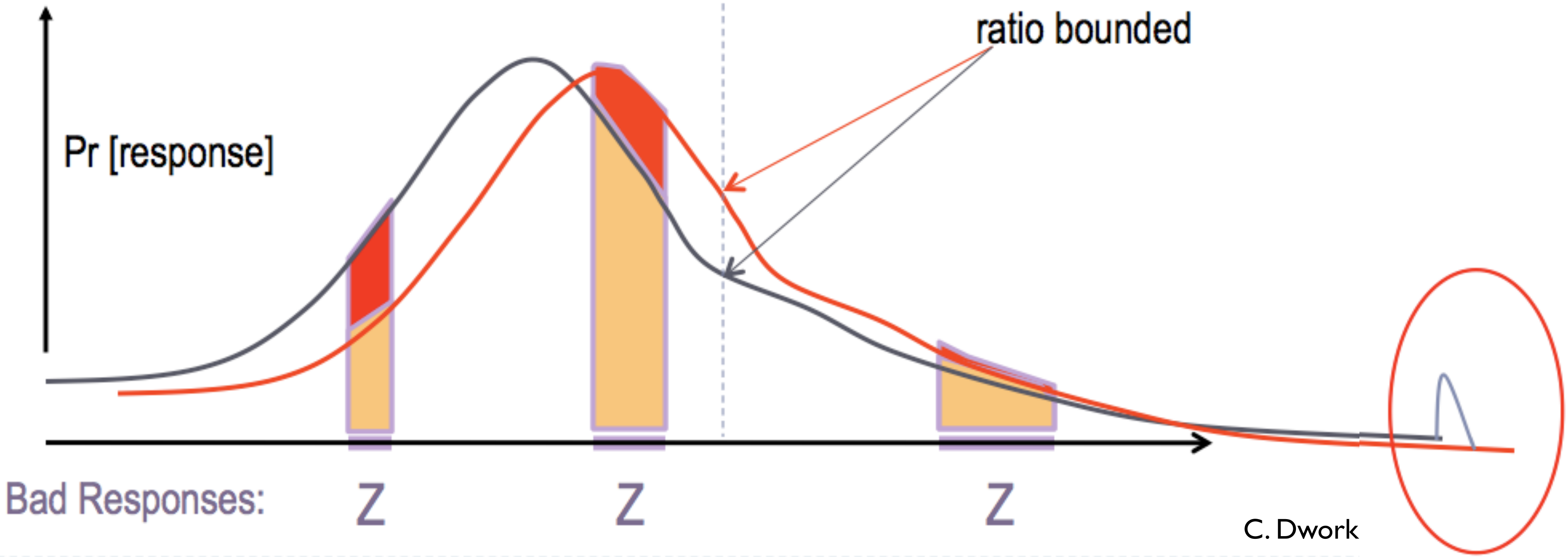
$$\Pr[M(x_1) \in S] \leq e^{\varepsilon} \Pr[M(x_2) \in S]$$



C. Dwork

(ε, δ)-differential privacy

$$\Pr[M(x_1) \in S] \leq e^{\varepsilon} \Pr[M(x_2) \in S] + \delta$$



differential privacy

$$\Pr[M(x_1) \in S] \leq e^{\varepsilon} \Pr[M(x_2) \in S]$$

Is a statistical property of mechanism behavior

unaffected by auxiliary information

independent of adversary's computational power

differential privacy

$$\Pr[M(x_1) \in S] \leq e^{\varepsilon} \Pr[M(x_2) \in S]$$

promise: if you leave the database, no outcome will change probability by very much

is this achievable?

yes!

chrome

source: http://www.recode.net/2016/6/15/11940908/mossberg-apple-is-still-a-world-of-its-own

source: http://edge.alluremedia.com.au/m/l/2015/03/ApplePrivacy.jpg

today

Formalizing privacy

Privacy properties and basic tools

Randomized Response

differential privacy

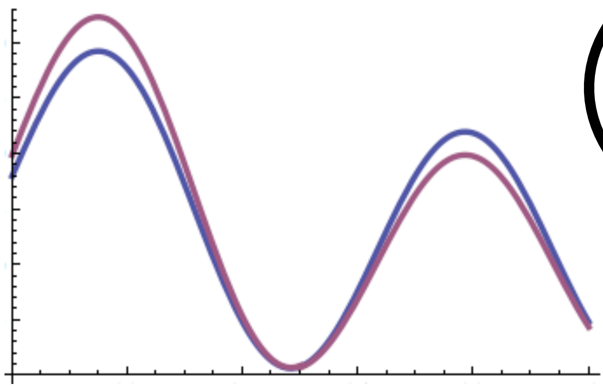
[DinurNissim03, DworkNissimMcSherrySmith06, Dwork06]

ε-Differential Privacy for algorithm M:

for any two neighboring data sets x_1, x_2, differing by the addition or removal of a single row

any $S \subseteq$ range(M),

$$\Pr[M(x_1) \in S] \leq e^{\varepsilon} \Pr[M(x_2) \in S]$$



$e^{\varepsilon} \sim (1 + \varepsilon)$

group privacy

Thm. Any $(\varepsilon, 0)$-DP mechanism M is $(k\,\varepsilon, 0)$-DP for groups of size k. i.e., for all

$$||x - y||_1 \leq k$$

and any $S \subseteq$ range(M),

$$\Pr[M(x) \in S] \leq e^{\varepsilon k} \Pr[M(y) \in S]$$

post-processing

Thm. Let $M : \mathbb{N}^{|X|} \to R$ be (ε, δ)-DP.

Let f: $R \to R'$ be an arbitrary randomized mapping.

Then $f \circ M : \mathbb{N}^{|X|} \to R'$ is (ε, δ)-DP.

composition

[DworkKenthapadiMcSherryMironovNaor06,DworkLei09]

Thm. For $i \in [k]$, let $M_i : \mathbb{N}^{|X|} \to R_i$ be $(\varepsilon_\mathrm{i}, \delta_\mathrm{i})$-DP. Then the mechanism $(M_1(x), \ldots, M_\mathrm{k}(x))$ is $(\sum_i \varepsilon_i, \sum_i \delta_i)$-DP.

actually, holds even if subsequent computations chosen as function of previous results

"advanced" version

Is bigger delta better for privacy, or worse?

What about epsilon?

today

Formalizing privacy

Privacy properties and basic tools

Randomized Response

Randomized Response
[Warner65]

flip a coin

if tails, respond truthfully

if heads, flip a second coin and respond "yes" if heads; respond "no" if tails

Claim. Randomized Response is (ln 3, 0)-DP.

Proof.
$$\frac{\Pr[\text{Response} = \text{Yes} | \text{Truth} = \text{Yes}]}{\Pr[\text{Response} = \text{Yes} | \text{Truth} = \text{No}]}$$
$$= \frac{3/4}{1/4} = \frac{\Pr[\text{Response} = \text{No} | \text{Truth} = \text{No}]}{\Pr[\text{Response} = \text{No} | \text{Truth} = \text{Yes}]} = 3.$$