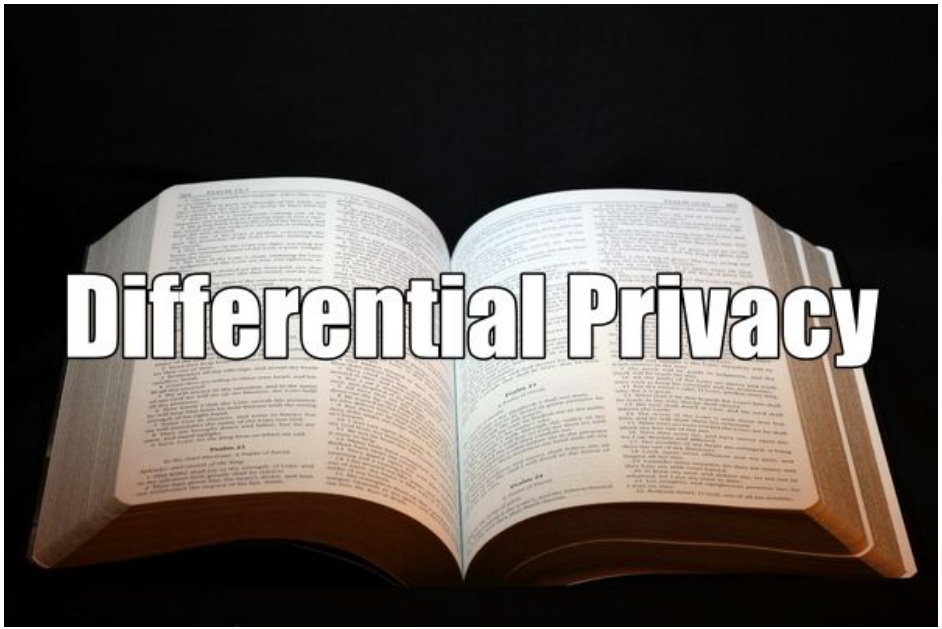
Introduction to Differential Privacy

Joseph “Joey” Knightbrook
knightbrook@google.com
15 April 2022

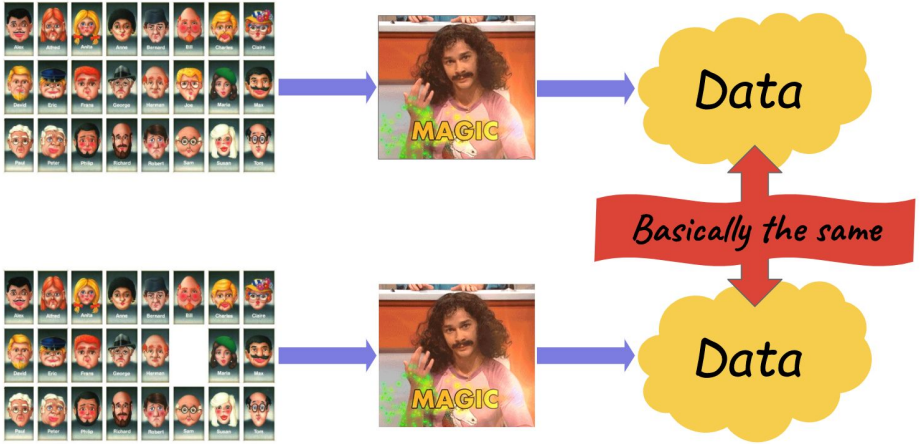
Why am I Here?



GOAL of Differential Privacy:

1) Protect Individuals' Privacy
2) Make Inferences about groups

How Does it Do This?



“Anonymized Data”?

Just get rid of the user-id?
...name, Social Security Number, etc

Not so fast...

ars TECHNICA BIZ & IT TECH SCIENCE POLICY CARS GAMING & CULTURE STORE

POLICY —

“Anonymized” data really isn’t—and here’s why not

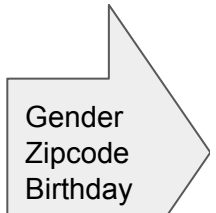
Companies continue to store and sometimes release vast databases of " ...

NATE ANDERSON - 9/8/2009, 4:25 AM

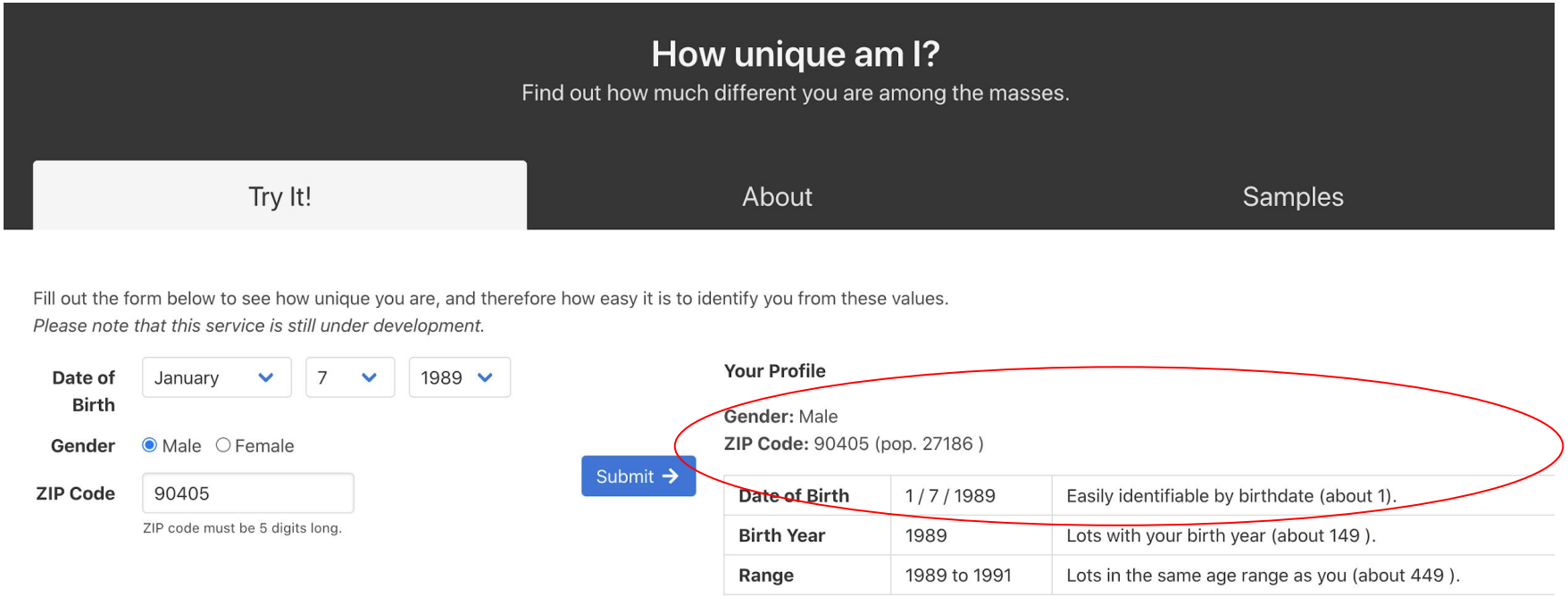
41

The Massachusetts Group Insurance Commission had a bright idea back in the mid-1990s—it decided to release "anonymized" data on state employees that showed every single hospital visit. The goal was to help researchers, and the state spent time removing all obvious identifiers such as name, address, and Social Security number. But a graduate student in computer science saw a chance to make a point about the limits of anonymization.

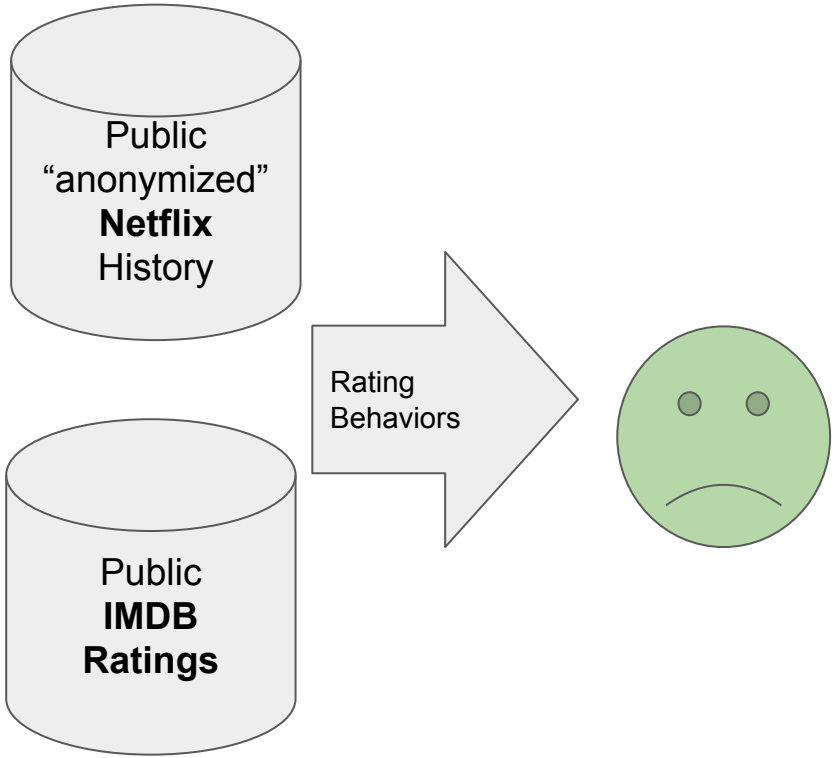
Latanya Sweeney requested a copy of the data and went to work on her "reidentification" quest. It didn't prove difficult. Law professor Paul Ohm describes Sweeney's work:



How Unique am I? Identity Website



Another Famous Example



“K-Anonymous” Data Release

K-Anonymity: Example

	ZIP Code	Age	Disease
1	47677	29	Heart Disease
2	47602	22	Heart Disease
3	47678	27	Heart Disease
4	47905	43	Flu
5	47909	52	Heart Disease
6	47906	47	Cancer
7	47605	30	Heart Disease
8	47673	36	Cancer
9	47607	32	Cancer

Table 1. Original Patients Table

3-Anonymous Table
(zip, age are identifiers)

	ZIP Code	Age	Disease
1	476**	2*	Heart Disease
2	476**	2*	Heart Disease
3	476**	2*	Heart Disease
4	4790*	≥ 40	Flu
5	4790*	≥ 40	Heart Disease
6	4790*	≥ 40	Cancer
7	476**	3*	Heart Disease
8	476**	3*	Cancer
9	476**	3*	Cancer

Table 2. A 3-Anonymous Version of Table 1

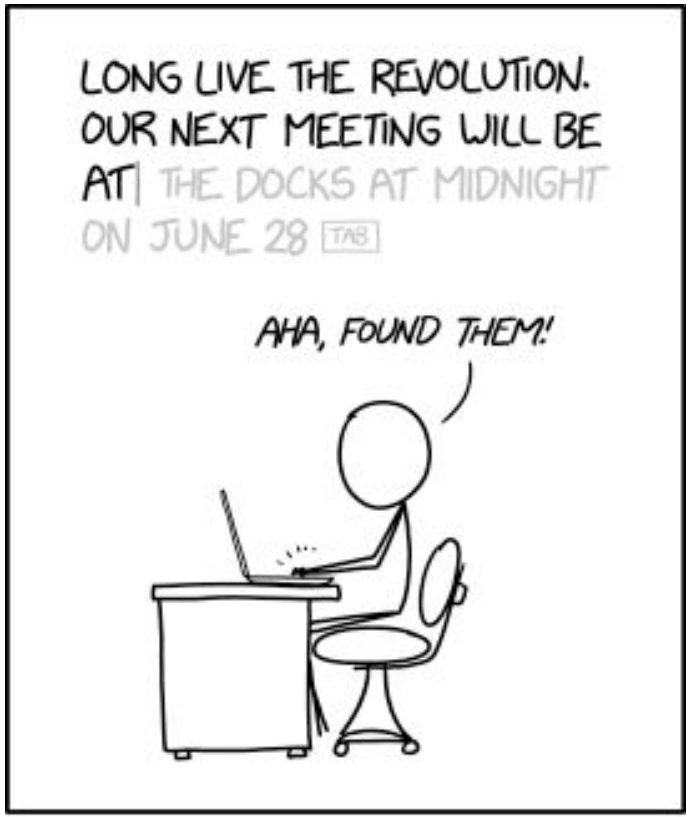
Problem: Prior Information!

	ZIP Code	Age	Disease
1	476**	2*	Heart Disease
2	476**	2*	Heart Disease
3	476**	2*	Heart Disease
4	4790*	≥ 40	Flu
5	4790*	≥ 40	Heart Disease
6	4790*	≥ 40	Cancer
7	476**	3*	Heart Disease
8	476**	3*	Cancer
9	476**	3*	Cancer

Table 2. A 3-Anonymous Version of Table 1

1. Basically full info
2. You had the flu, and your neighbor is Japanese
3. Your neighbor is Japanese

What if k is HUGE?



WHEN YOU TRAIN PREDICTIVE MODELS ON INPUT FROM YOUR USERS, IT CAN LEAK INFORMATION IN UNEXPECTED WAYS.

Maybe sample the data AND “anonymize”?


Home
The New York Times
Share
1086
"While The Times did not obtain the president's actual tax returns, it received the information contained in the returns from someone who had legal access to it. The Times was then able to **find matching results in the I.R.S. information on top earners — a publicly available database that each year comprises a one-third sampling of those taxpayers, with identifying details removed**."
TIMES INVESTIGATION
Decade in the Red: Trump Tax Figures Show Over $1 Billion in Business Losses

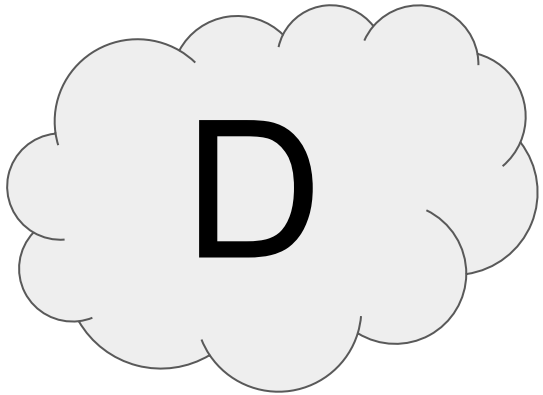
Questions?

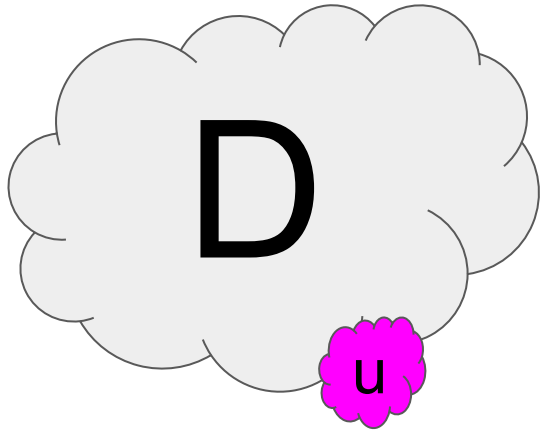
Fighting a Losing Battle...

Enter
Differential Privacy

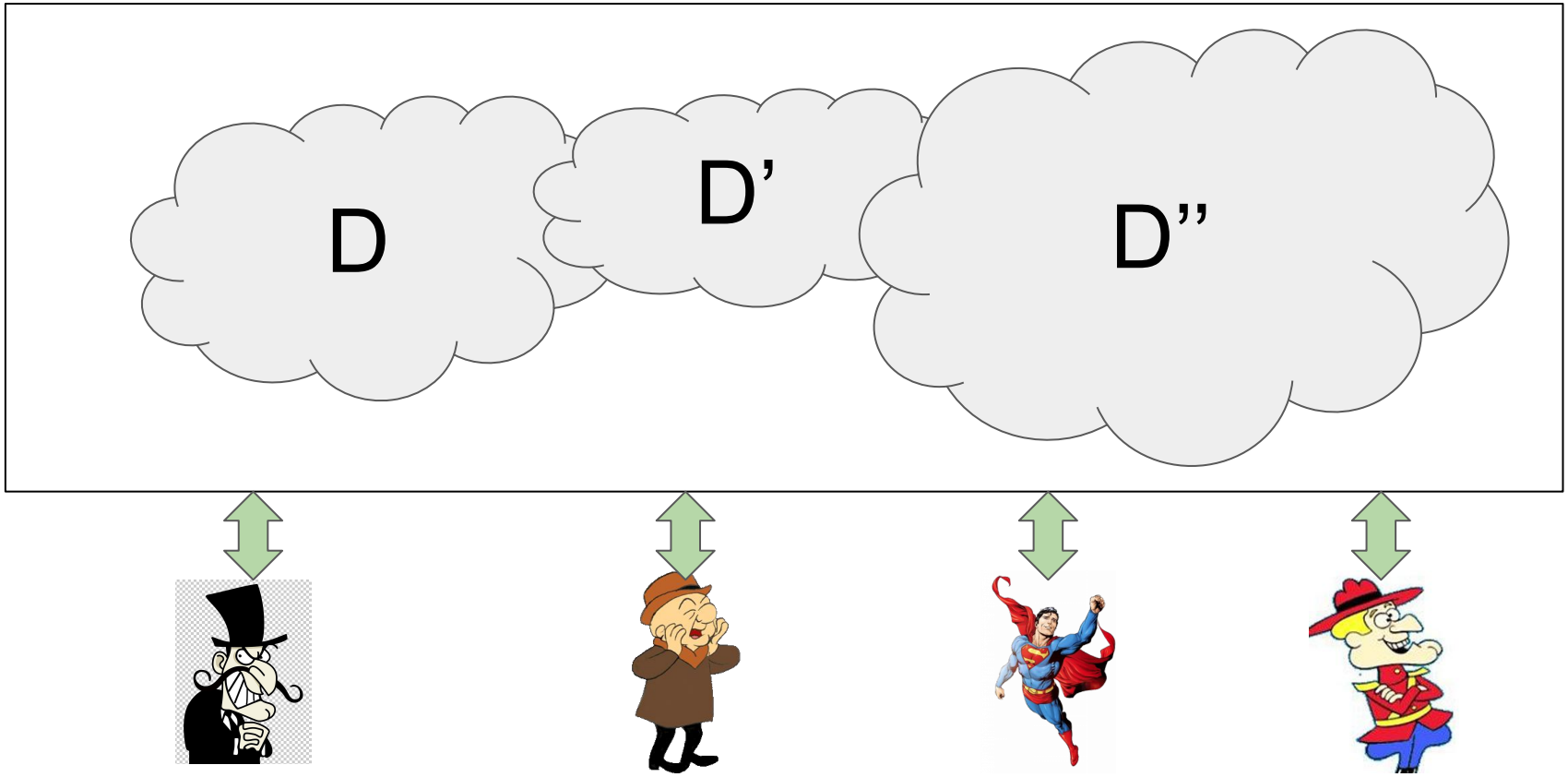
Goal: Minimize Effect on ‘u’

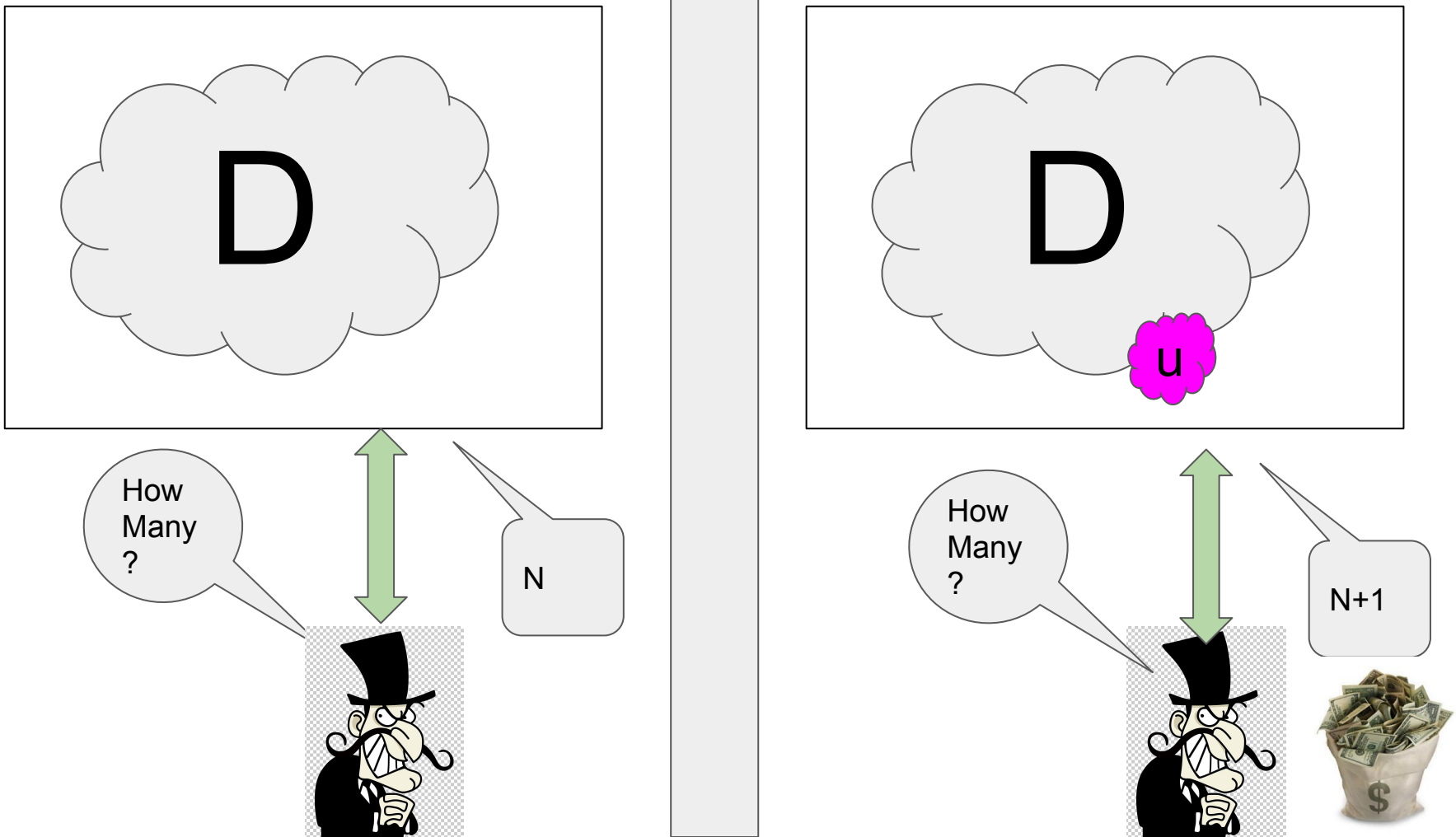
(pun intended)



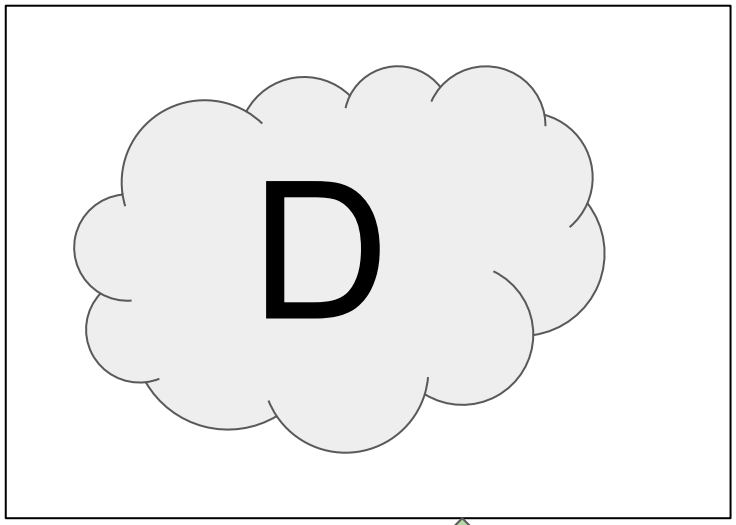


Trusted Entity (Company, Government, NGO etc.

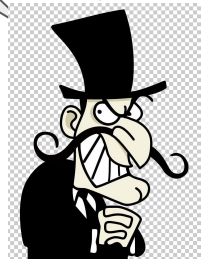



D
How Many ?
N
D
u
How Many ?
N+1

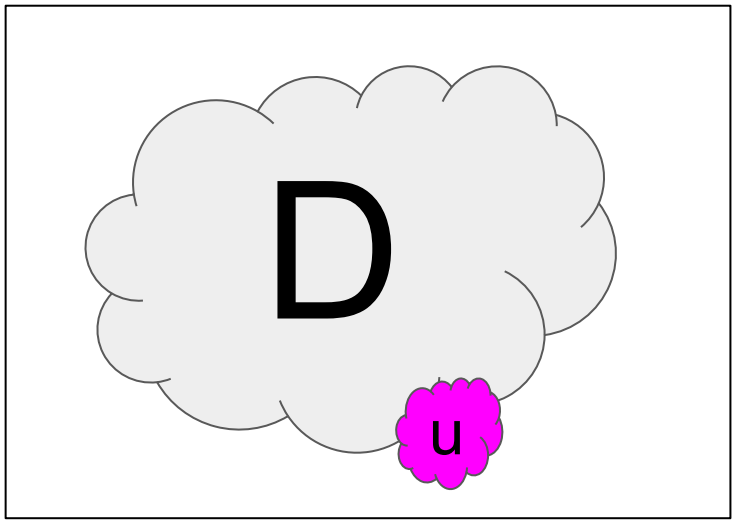
Key Insight: Adding Noise



How Many ?



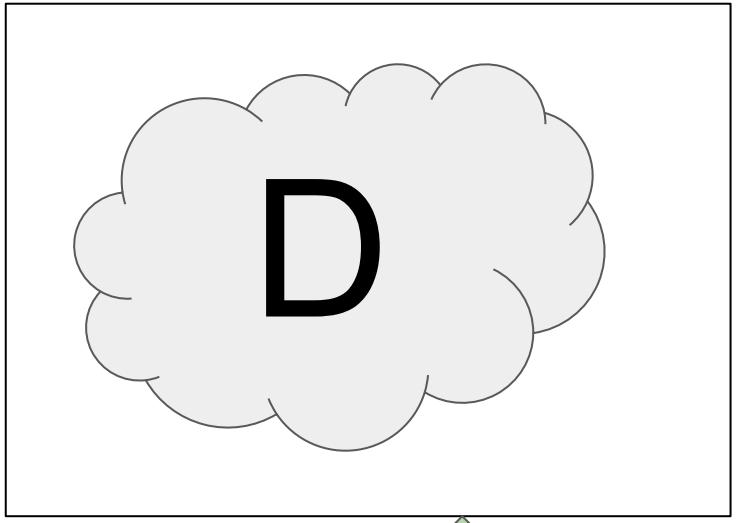
$N+x_0$



How Many ?

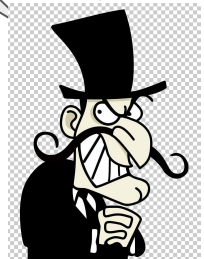
$N+x_1+1$

??

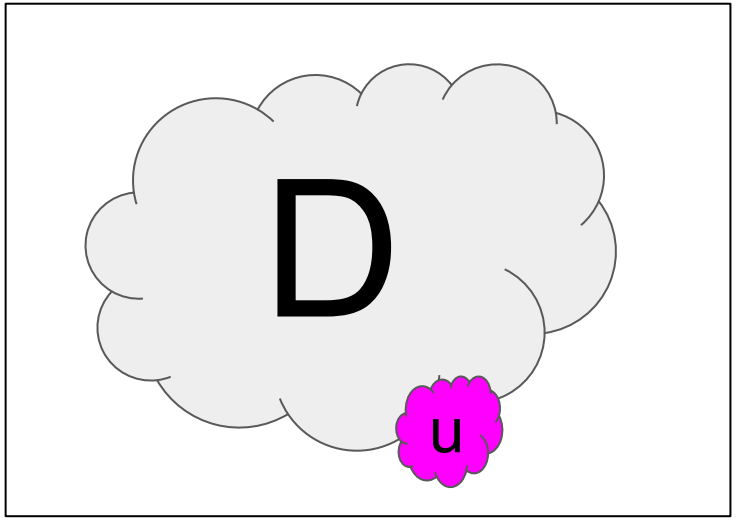


How Many ?

$N+x_0$



$$P\left[M\left(D\right)=l\right]$$



How Many ?

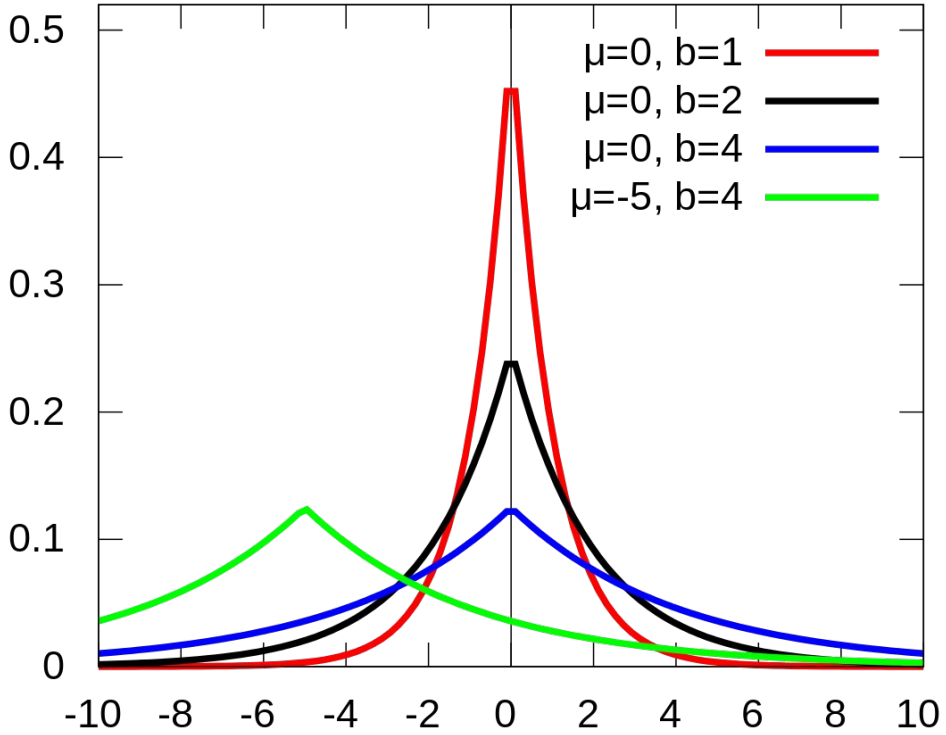
$N+x_1+1$

$$P\left[M\left(D+u\right)=l\right]$$

??

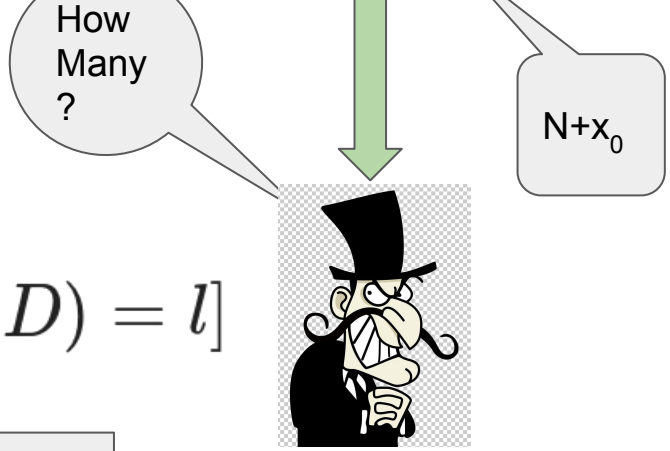
Add Noise!

$x_i \sim$



Often Laplace:

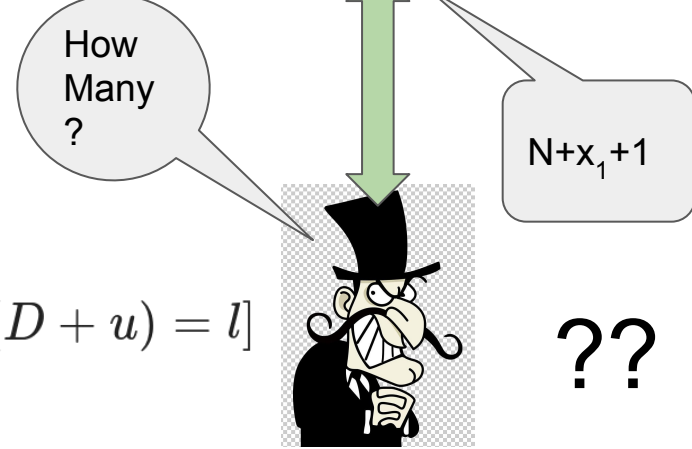
$$p = \frac{1}{2b} e^{-\frac{|x-\mu|}{b}}$$



$$P\left[M\left(D\right)=l\right]$$

Laplace_PDF = N - L

$$p=\frac{1}{2b}e^{-\frac{|N-l|}{b}}$$



$$P\left[M\left(D+u\right)=l\right]$$

??

Laplace_PDF = (N + 1) - L

$$p=\frac{1}{2b}e^{-\frac{|N+1-l|}{b}}$$

$$p = \frac{1}{2b} e^{-\frac{|N-l|}{b}}$$

???

$$p = \frac{1}{2b} e^{-\frac{|N+1-l|}{b}}$$

$$p(D) - P(D+u) = \frac{1}{2b} e^{-\frac{|N-l|}{b}} - \frac{1}{2b} e^{-\frac{|N+1-l|}{b}}$$

$$p = \frac{1}{2b} e^{-\frac{|N-l|}{b}}$$??? $$p = \frac{1}{2b} e^{-\frac{|N+1-l|}{b}}$$

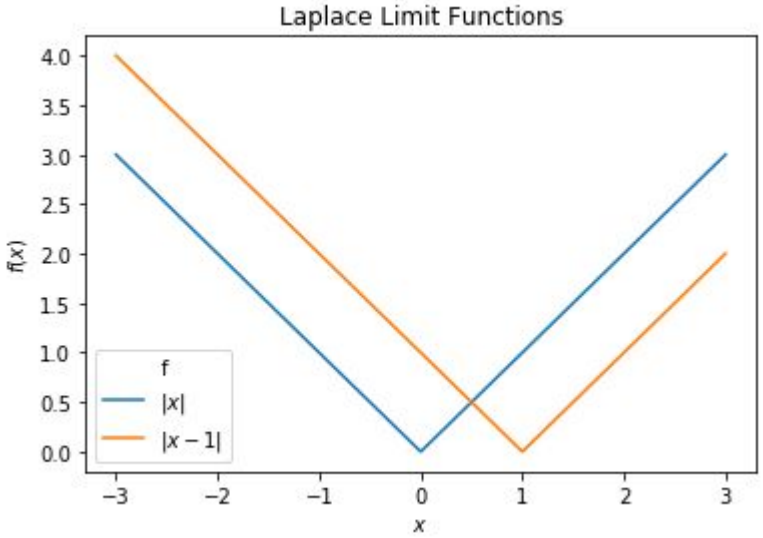
$$p(D)/P(D+u) = \frac{e^{-\frac{|N-l|}{b}}}{e^{-\frac{|N+1-l|}{b}}}$$

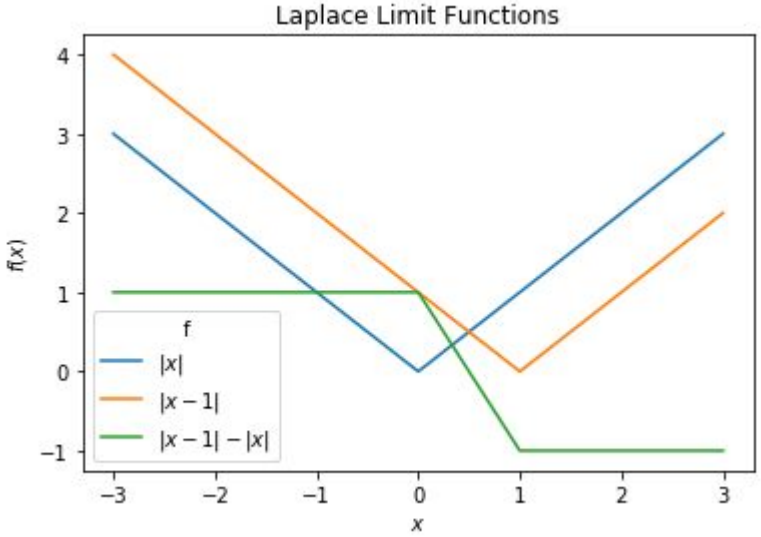
$$p(D)/P(D+u) = e^{-\frac{1}{b}[|N-l|-|N+1-l|]}$$

$$p(D)/P(D+u) = e^{[|x|-|x+1|]}$$

Laplace Limit

$$p(D)/P(D+u) = e^{[|x|-|x+1|]}$$

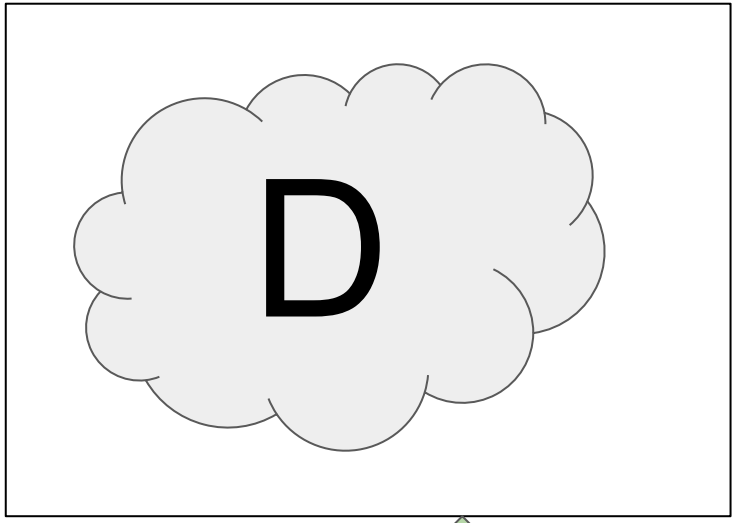




$$p(D)/P(D+u) = e^{-\frac{1}{b}[|N-l|-|N+1-l|]}$$

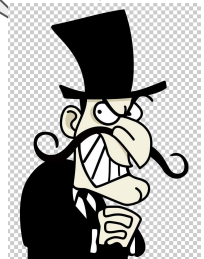
$$p(D)/P(D+u) \leq e^{-\frac{1}{b}}$$

NO MATTER THE OUTPUT
(L is gone!)
Attacker only gets a maximum “hint” of $e^{1/b}$
That a TARGET user was in the database

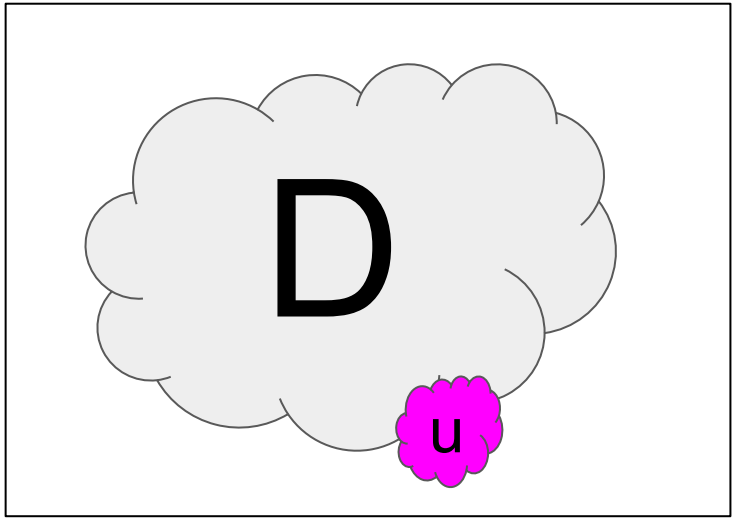


How Many ?

$N+x_0$



$$P\left[M\left(D\right)=l\right]$$



How Many ?

$N+x_1+1$

$$P\left[M\left(D+u\right)=l\right]$$

$$e^{1/b}$$

Questions?

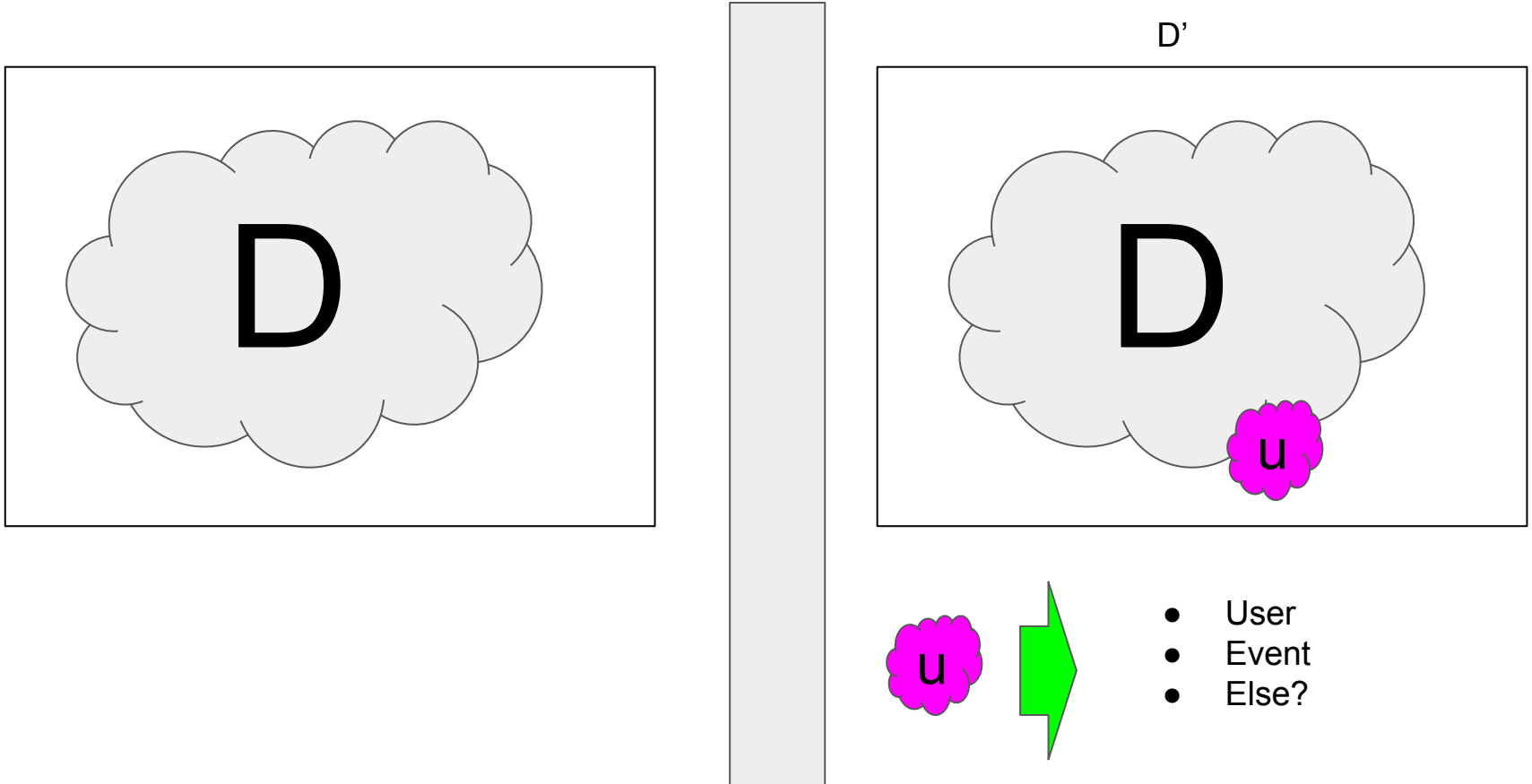
$$r = \frac{P\left[M\left(D+u\right)=l\right]}{P\left[M\left(D\right)=l\right]}$$

$$-\epsilon \leq \ln\left[\frac{P[M(D)=z]}{P[M(D')=z]}\right] \leq \epsilon$$

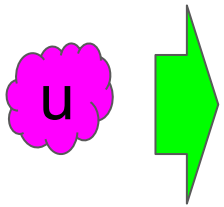
All possible z
All “neighbors” D, D’

$$P[M(D)=l] \leq e^{\epsilon} P[M(D')=l]$$

What is a Neighbor?



Effect of a Neighbor...



- User
- Event
- Else?

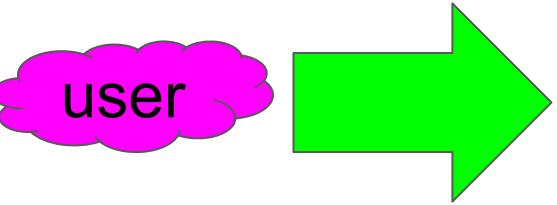
“Sensitivity”:

The maximum amount a neighbor can change the output under consideration

What is the Output?

- User Count
- Event Count
- Average Spent

Example:



IMPRESSION COUNT

- MAX_IMPRESSIONS_PER_USER
- (filtering like this comes up all the time)

$$\epsilon_{\text{added}} = \frac{\epsilon_{\text{goal}}}{S}$$

$$N_{\text{added}} = S \cdot N_{\text{goal}}$$

Questions?

Epsilon Intuition

$$-\epsilon \leq \ln\left[\frac{P[M(D) = z]}{P[M(D') = z]}\right] \leq \epsilon$$

- Epsilon as “Information Release”
- Higher:
  - More Information
  - Less private
  - More Accurate
- Epsilon ~ 1/Noise_Added

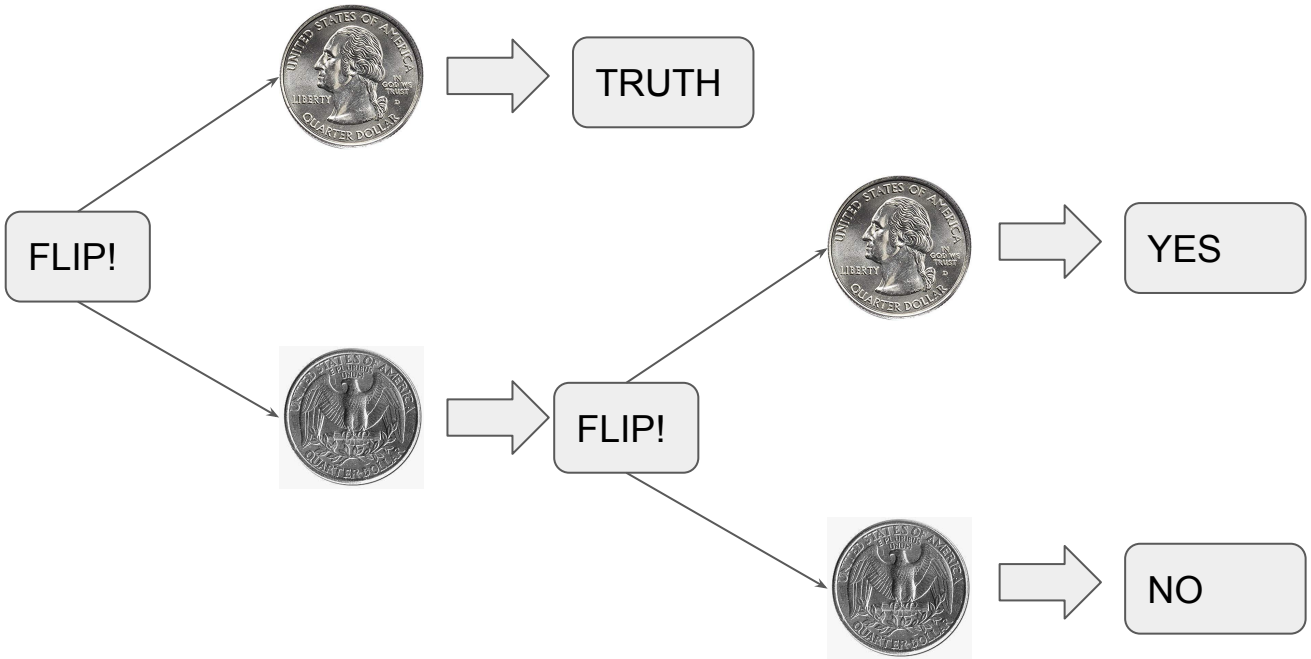
WHAT EPSILON IS OK!?

ln(3)
1.0986122887
….
1.1?

Ask: Are you a criminal?

TRUTH:	Y	N
ANSWER:	?	N

Randomized Response: Are you a Criminal?



Ask: Are you a criminal?

TRUTH:	Y	N
ANSWER:	?	N

Randomized Response: Are you a Criminal?

YES, p=.5

FLIP!

YES, p=.25

FLIP!

NO, p=.25

Randomized Response: Are you a Criminal?

NO, p=.5

FLIP!

YES, p=.25

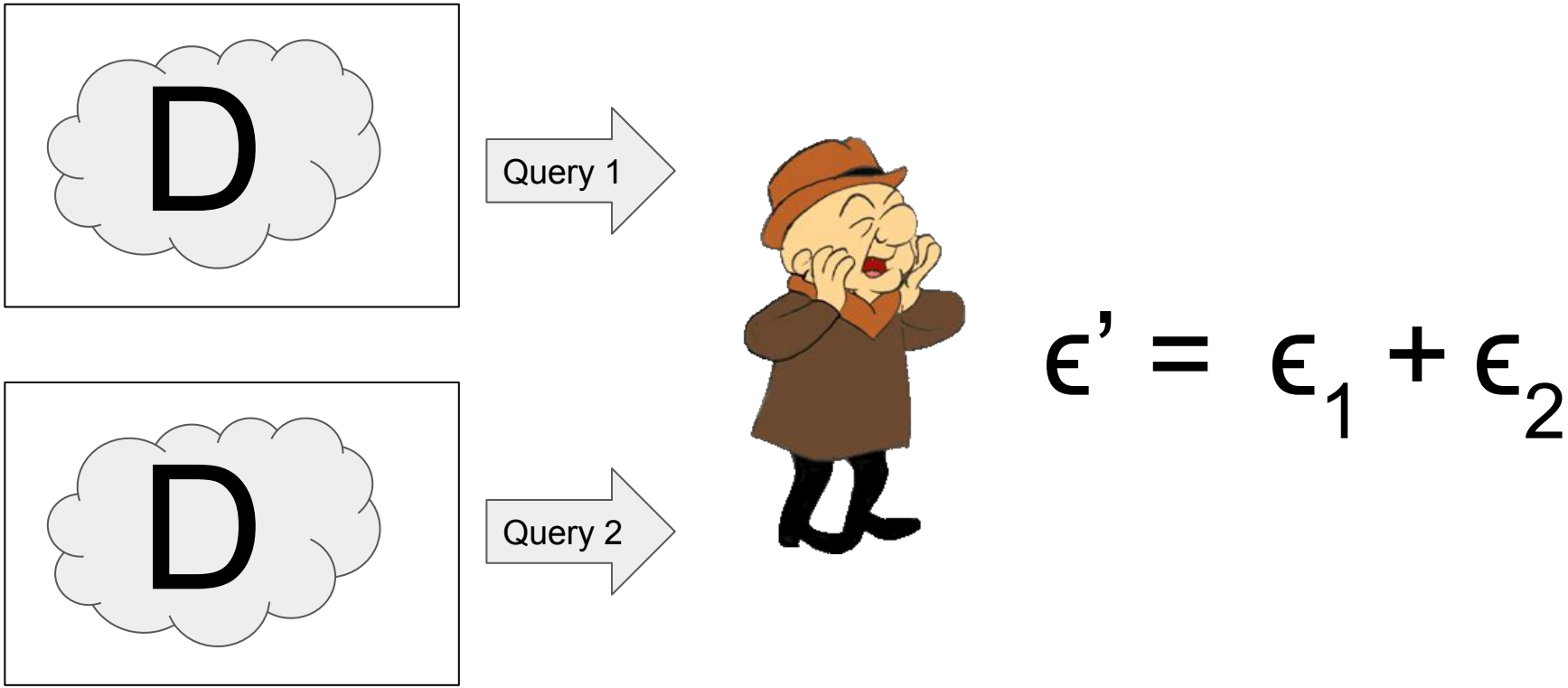
FLIP!

NO, p=.25

$$
\begin{aligned}
r_1 &= \frac{P[M(\text{criminal}) = \text{YES}]}{P[M(\text{not-criminal}) = \text{YES}]} = \frac{.75}{.25} = 3 \\
r_2 &= \frac{P[M(\text{criminal}) = \text{NO}]}{P[M(\text{not-criminal}) = \text{NO}]} = \frac{.25}{.75} = \frac{1}{3} \\
\epsilon &= \ln(3)
\end{aligned}
$$

More Practical Considerations

Epsilons Add Together

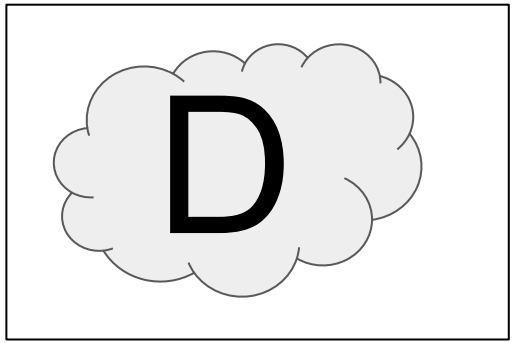


$$\epsilon' = \epsilon_1 + \epsilon_2$$

...Because Probabilities Multiply

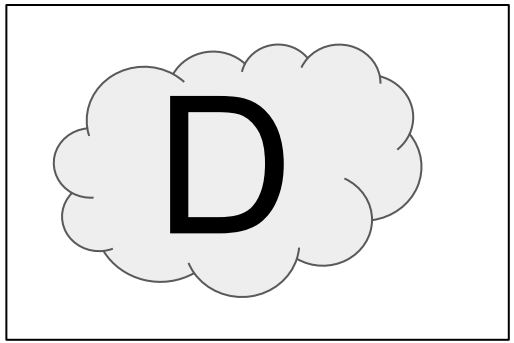
$$
\begin{aligned}
r &= \frac{P\left[M_1\left(D+u\right)=l_1\right]P\left[M_2\left(D+u\right)=l_2\right]}{P\left[M_1\left(D\right)=l_1\right]P\left[M_2\left(D\right)=l_2\right]} \\
&\leq e^{\epsilon_1}e^{\epsilon_2} \\
&\leq e^{\epsilon_1+\epsilon_2}
\end{aligned}
$$

“Privacy Budget”



Query 1

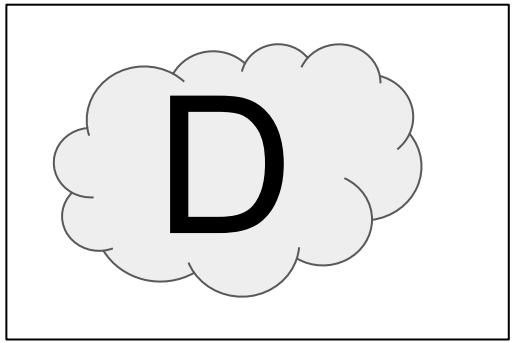
. . . .



Query n

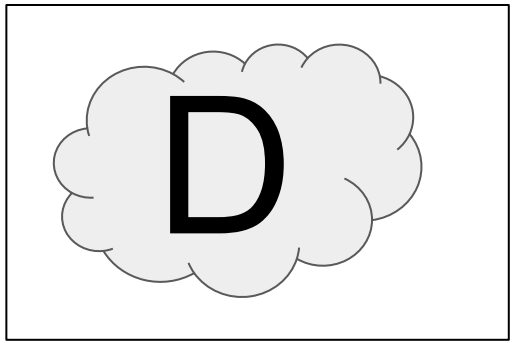
$$\epsilon_{LIMIT} < \epsilon_1 + \ldots + \epsilon_n$$

“Budget” is an overloaded term!



Query 1

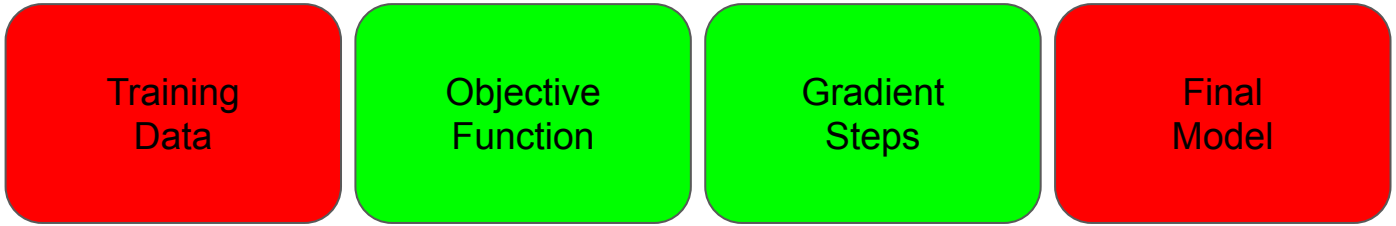
. . . .



Query n

$$\epsilon_{LIMIT} < \epsilon_1 + \ldots + \epsilon_n$$

Differential Privacy in Machine Learning



Success isn’t Certain

Appendices

Epsilons Adding is Worst-Case

$$
\begin{aligned}
r &= \frac{P\left[M_1\left(D+u\right)=l_1\right]P\left[M_2\left(D+u\right)=l_2\right]}{P\left[M_1\left(D\right)=l_1\right]P\left[M_2\left(D\right)=l_2\right]} \\
&\leq e^{\epsilon_1}e^{\epsilon_2} \\
&\leq e^{\epsilon_1+\epsilon_2}
\end{aligned}
$$

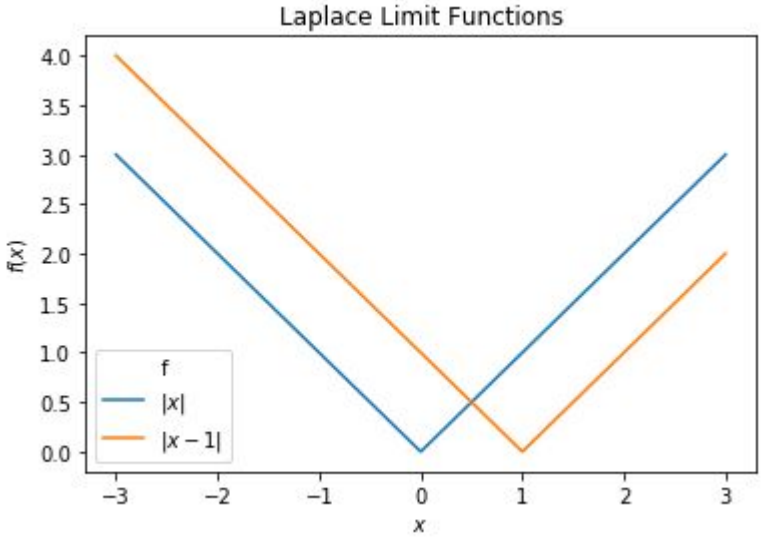
There's a Universe of ways to be clever here

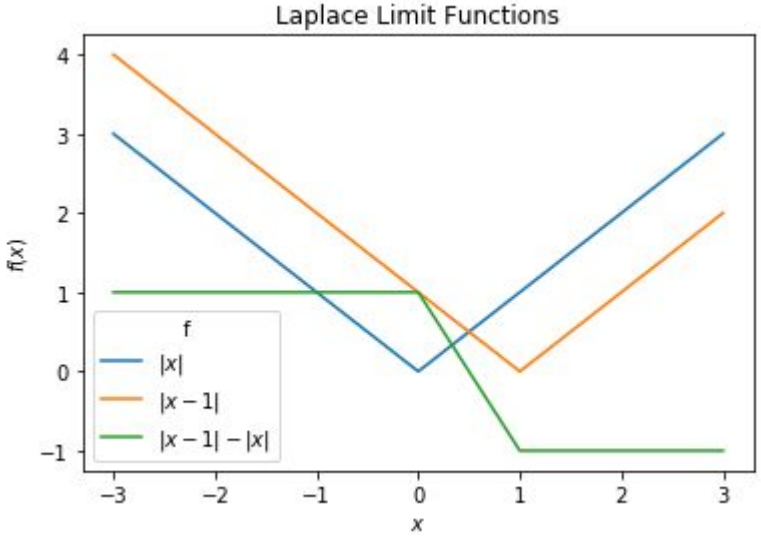
"Advanced Composition Theorems"

Is Failure an Option?

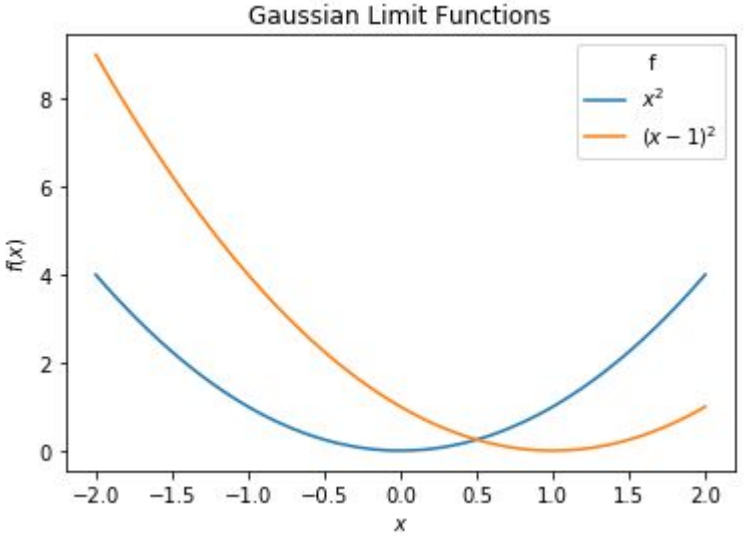
Laplace Limit

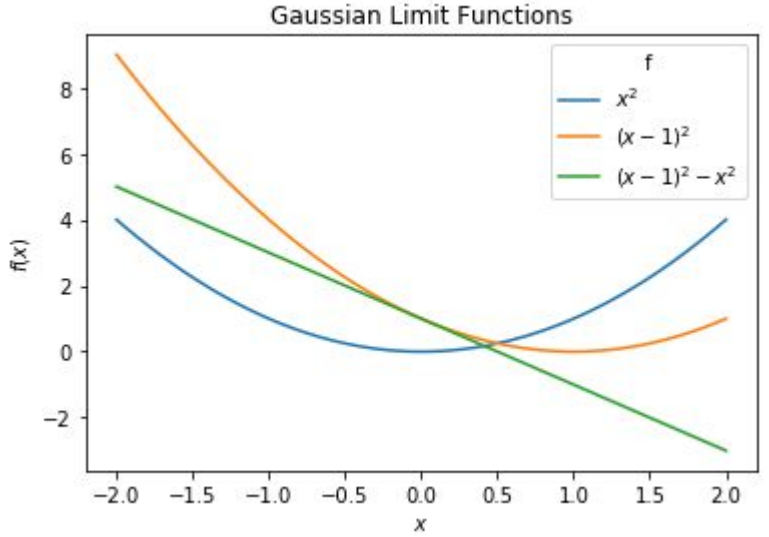
$$p(D)/P(D+u) = e^{[|x|-|x+1|]}$$





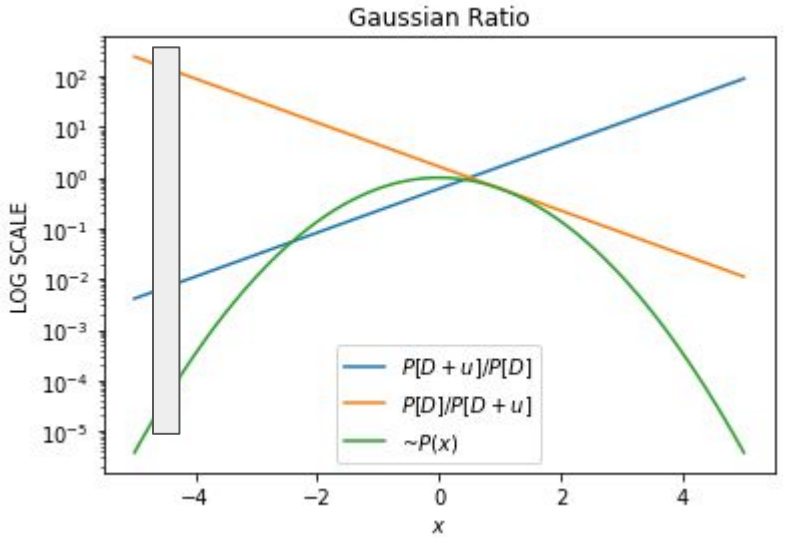
Gaussian Limit $p(D)/P(D+u) = e^{-\left[x^2-(x+1)^2\right]}$





Gaussian Limit $p(D)/P(D+u) = e^{-\left[x^2-(x+1)^2\right]}$

10^{-5} failure probability
$r < \sim 10^{2.5}$



$$r = \frac{P\left[M\left(D+u\right)=l\right]}{P\left[M\left(D\right)=l\right]}$$

$$-\epsilon \leq \ln\left[\frac{P[M(D)=z]}{P[M(D')=z]}\right] \leq \epsilon$$

All possible z
All “neighbors” D, D’
δ of the time!

$$P[M(D)=z] \leq P[M(D')=z]e^{\epsilon}+\delta$$

Lightning Round of Relevant Facts

- Deltas Add too!
- Sampling Can Reduce Epsilon Spent
- No Information from Post-Processing (epsilon stays epsilon)
- Smarter Epsilon, Delta Accounting