Privacy-Preserving Data Publishing

Interactive Sanitization : Introduction to Differential Privacy

Interactive querying without sanitization

Aggregate query

Example :

```
SELECT COUNT(*)
FROM PATIENT
WHERE zipcode=75002
     AND disease='Cold'
```

DBMS
(**trusted**)

Aggregate

User

Interactive querying ~~without~~ *with* sanitization

DBMS
(**trusted**)

Aggregate query

Example :

```
SELECT COUNT(*)
FROM PATIENT
WHERE zipcode=75002
     AND disease='Cold'
```

Sanitized *Aggregate*

Legitimate usage

Illegitimate usage

Illegitimate Usage ?

Example #1

Query	`SELECT COUNT(*) FROM PATIENT WHERE ssn='123' AND disease='Cold'`
Result	1

??

Illegitimate Usage ?

Example #1

Query	`SELECT COUNT(*) FROM PATIENT WHERE ssn='123' AND disease='Cold'`
Result	1

KO

Illegitimate Usage ?

Example #2

Query	`SELECT COUNT(*) FROM PATIENT WHERE (ssn='123' OR ssn='456') AND disease='Cold'`	`SELECT COUNT(*) FROM PATIENT WHERE (ssn='456' OR ssn='789') AND disease='Cold'`
Result	1	0

??

Illegitimate Usage ?

Example #2

Query	`SELECT COUNT(*) FROM PATIENT WHERE (ssn='123' OR ssn='456') AND disease='Cold'`	`SELECT COUNT(*) FROM PATIENT WHERE (ssn='456' OR ssn='789') AND disease='Cold'`
Result	1	0

KO

Illegitimate Usage ?

Example #3

Query	`SELECT COUNT(*)`
	`FROM PATIENT`
	`WHERE disease='Cold'`
		`SELECT COUNT(*)`
		`FROM PATIENT`
		`WHERE disease='Cold'`
		`AND ssn!='123'`

Result	400	399

??

Illegitimate Usage ?

Example #3

Query	`SELECT COUNT(*) FROM PATIENT WHERE disease='Cold'`	`SELECT COUNT(*) FROM PATIENT WHERE disease='Cold' AND ssn!='123'`
Result	400	399

KO

Illegitimate Usage ?

Example #4

***Connaissance auxiliaire** : l'individu cible a comme zipcode 75002.*

Query	`SELECT COUNT(*) FROM PATIENT WHERE zipcode=75002`	`SELECT COUNT(*) FROM PATIENT WHERE zipcode=75002 AND disease='Cold'`
Result	5	5

??

Illegitimate Usage ?

Example #4

Connaissance auxiliaire *: l'individu cible a comme zipcode 75002.*

Query	`SELECT COUNT(*) FROM PATIENT WHERE zipcode=75002`	`SELECT COUNT(*) FROM PATIENT WHERE zipcode=75002 AND disease='Cold'`
Result	5	5

KO

Various Methods for Sanitizing Aggregates

- **Idea 1 :** Analyze queries ? (Eg, refuse to answer to queries « *leading* » to a weak cardinality result.)
  - Costly ! (*e.g.,* **compute all the intersections** between the current query and the complete query history?) ;
  - Unsafe ! What means a refusal to answer to a given query?
- **Idea 2 : Perturb randomly the** aggregate : mechanism proposed for satisfying *differential privacy*

Differential Privacy : The Model

A Change in Paradigm

- With centralized publishing, we saw Paradigm #1, *aka* the uninformative principle : « Limit the knowledge gain of the attacker »

- Differential privacy says « Wrong way » :
  - Goes against the goal of data publishing: LEARN
    « Beer + Donuts = Diaper » (warning: this may be a myth)
    http://www.florent-masseglia.info/biere-et-couches-un-exemple-mythique-du-data-mining/
  - Is based on a hazardous before/after comparison:
    - Hard to know what the adversary knows
    - There **always** exists a possible auxiliary knowledge, possibly independent from the DB to protect, that leads to a *privacy breach*
      Ex: « John's **salary** *is twice the average salary in France. »*
      Ex: « Bob's **height** *is 5 cm less than three apples. »*

Design Rules

- Differential privacy :
  - Only takes into account what is completely known to the DBMS:
    - the DB + the aggregate query
    - Formulates as few assumption on the adversary as possible:
      - *Ad omnia* assumption: adversary knows all the records except one (worst-case)
      - *(must be nuanced…)*
  - Is a property of the sanitization algorithm, and not of a specific release
  - Does not limit explicitly the *knowledge gain*… *See Paradigm #2 next slide*

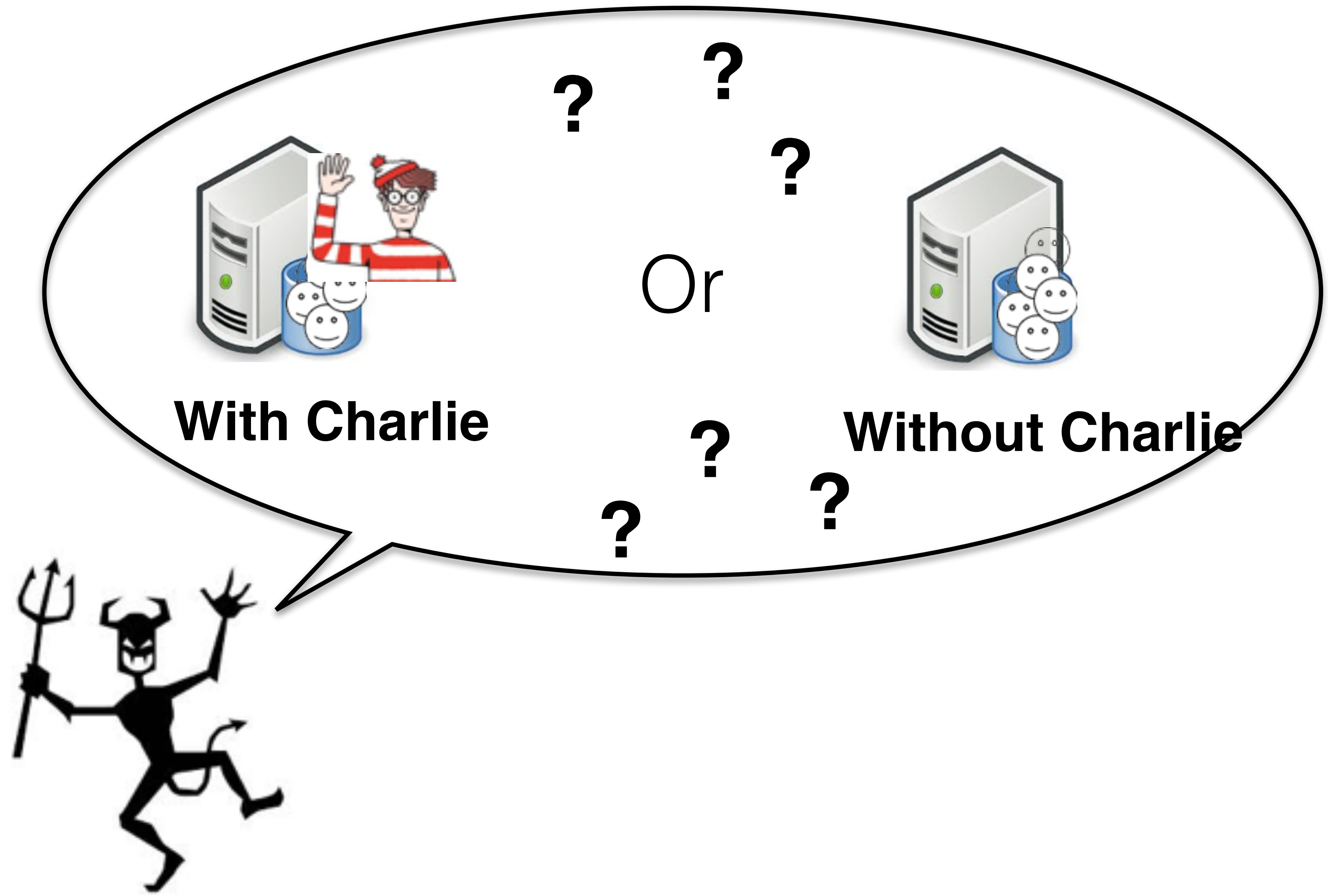
Privacy Paradigm #2

- Intuition:
  - Privacy does not concern global trends: a global trend *is not private* and *must* be learnt
  - Privacy concerns each individual value: in other words, each individual contribution to the global trend

Privacy Paradigm #2

- Intuition:
  - Privacy does not concern global trends: a global trend *is not private* and *must* be learnt
  - Privacy concerns each individual value: in other words, each individual contribution to the global trend
- **Paradigm #2 :** a function `f` satisfies differential privacy iif : **the possible impact of any individual on its result** (its possible outputs) **is limited**

Intuition



Sanitized result `r’`

Intuition

Pr [f ()] ≈ Pr [f ()]

(Limited impact of any possible Charlie*)*

Intuition

Pr [f ()] ≈ Pr [f ()]

Close to an e^{ε} factor
(ε is the privacy parameter, set by DBA)

(Limited impact of any possible Charlie*)*

Formal Model

- A *random function* `f` satisfies *ε-differential privacy* iif:

  For all D, D' differing in at most one record, and for all possible output S of `f` then it is true that :

  $\Pr[\, f(D) = S \,] \leq e^{\varepsilon} \times \Pr[\, f(D') = S \,]$

Formal Model

- A *random function* `f` satisfies *ε-differential privacy* iif:

  For all D, D' differing in at most one record, and for all possible output S of `f` then it is true that :

  $$\Pr[\, f(D) = S \,] \leq e^{\varepsilon} \times \Pr[\, f(D') = S \,]$$

Here, an **aggregate query** *with* **random** *perturbation*

Formal Model

- A *random function* `f` satisfies *ε-differential privacy* iif:

  For all D, D' differing in at most one record, and for all possible output S of `f` then it is true that :

  $$\Pr[\, \texttt{f}(D) = S \,] \leq e^{\varepsilon} \times \Pr[\, \texttt{f}(D') = S \,]$$

Every possible dataset

Formal Model

- A *random function* `f` satisfies *ε-differential privacy* iif:

  For all D, D' differing in at most one record, and for all possible output S of `f` then it is true that :

  $$\Pr[\, f(D) = S \,] \leq e^{\varepsilon} \times \Pr[\, f(D') = S \,]$$

Here : *D' is D with one more record (i.e., an individual) or one less record.*
Variant : *D' is D with one record that is different.*

Formal Model

- A *random function* `f` satisfies *ε-differential privacy* iif:

  For all D, D' differing in at most one record, and for all possible output S of `f` then it is true that :

  $$\Pr[\, f(D) = S \,] \leq e^{\varepsilon} \times \Pr[\, f(D') = S \,]$$

If one probability is 0, the other must be 0 too.

Formal Model

- A *random function* `f` satisfies *ε-differential privacy* iif:

  For all D, D' differing in at most one record, and for all possible output S of `f` then it is true that :

  $$\Pr[\, f(D) = S \,] \leq e^{\varepsilon} \times \Pr[\, f(D') = S \,]$$

Privacy parameter (e.g., 0.01, 0.1, ln(2), ln(3)), is public

Differential Privacy : The Algorithm (*aka* the Laplace Mechanism)

Perturbation of Results

Aggregate query

Example :

```
SELECT COUNT(*)
FROM PATIENT
WHERE zipcode=75002
     AND disease='Cold'
```

***Perturbed** aggregate*

r' = r + noise

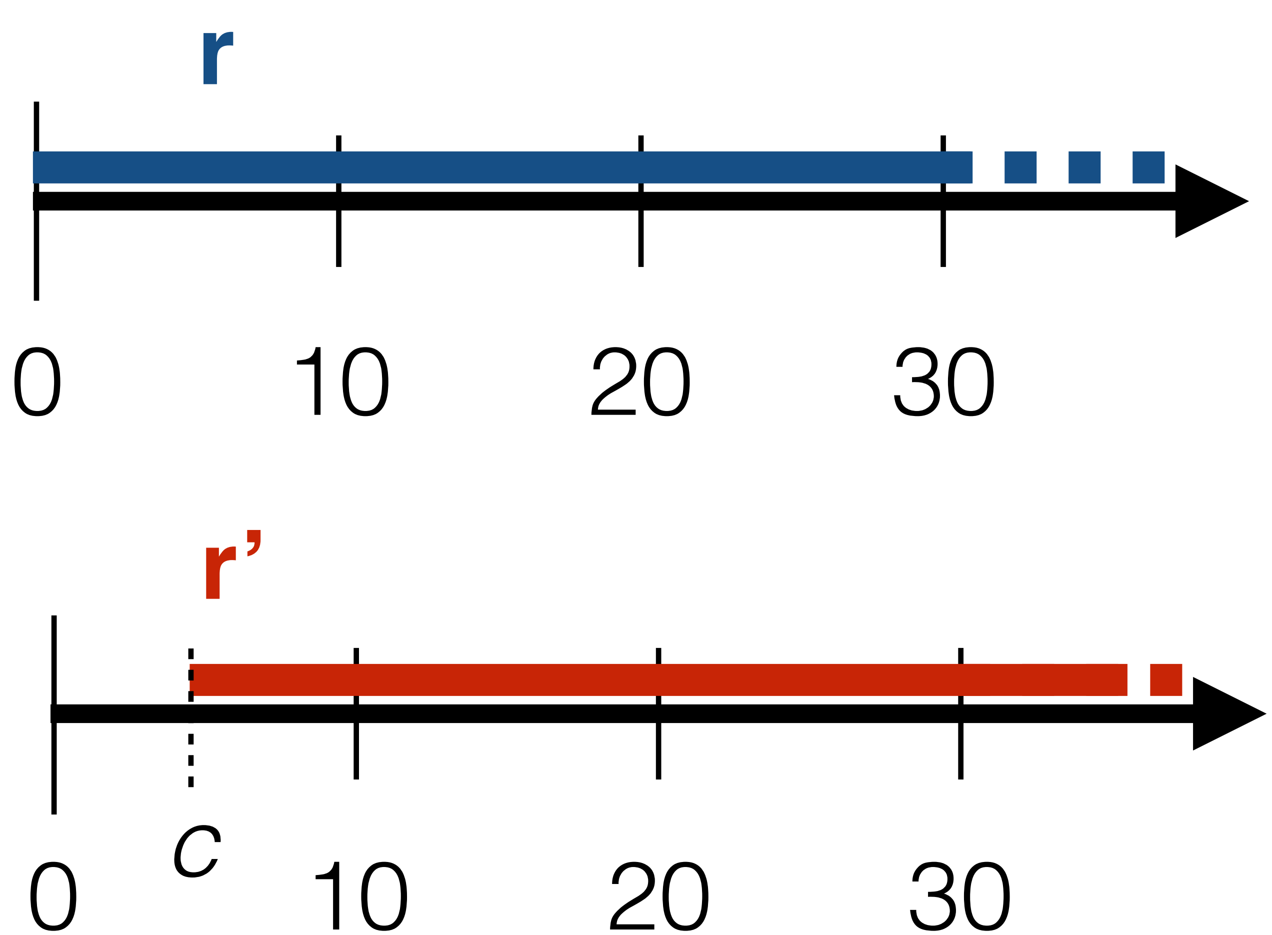
A random variable.

Which Distribution should it follow in order to hide the contribution of any possible individual?

Unsafe Mechanism: if *noise* were a constant value…

- r’ = r + c ;

Guarantees?
Good value for c?
Guessing c ?



All queries are not equally… *sensitive !*

- Differential privacy : « *the impact that any possible individual can have on the output is limited* » ;

- But… the possible impact of an individual is not the same whatever the query.
  -> Queries have a **sensitivity**:
  - A `COUNT` changes by +/- 1 *at most*
    depending on the presence/absence of an individual
  - A `SUM` of salaries changes by +/- max-value *at most*
    depending on the presence/absence of an individual

All queries are not equally… *sensitive !*

`COUNT` (‘Cold’) → 10

`COUNT` (‘Cold’) → 9 (at worst)

All queries are not equally… *sensitive !*

where salary ∈ [0, 1000]

`SUM` (salary) → 7 000

where salary ∈ [0, 1000]

`SUM` (salary) → 6 000 (at worst)

All queries are not equally… *sensitive !*

- The **sensitivity** of a function quantifies the impact that the presence/absence of an individual ***can have*** on its output ;

- Let `f` : domain → R, the sensitivity of `f` is :

$$S_f = \max_{D,D'} \| f(D) - f(D') \|_1$$
$$= \max_{D,D'} (|f(D) - f(D')|)$$

for any D,D' differing by exactly one record at most.

All queries are not equally… *sensitive !*

COUNT ('Cold') → 10

$$S_{\mathrm{COUNT}} = \max_{D,D'} \left| \mathrm{COUNT}(D) - \mathrm{COUNT}(D') \right|$$

$$S_{\mathrm{COUNT}} = 1$$

COUNT ('Cold') → 9 (au pire)

All queries are not equally… *sensitive !*

where salary ∈ [0, 1000]

SUM (salary) → 7 000

$$S_{\texttt{SUM}} = \texttt{max}_{D,D'} \mid \texttt{SUM}(D) - \texttt{SUM}(D') \mid$$
$$S_{\texttt{SUM}} = 1000$$

where salary ∈ [0, 1000]

SUM (salary) → 6 000 (at worst)

Computing the Distribution for Perturbing

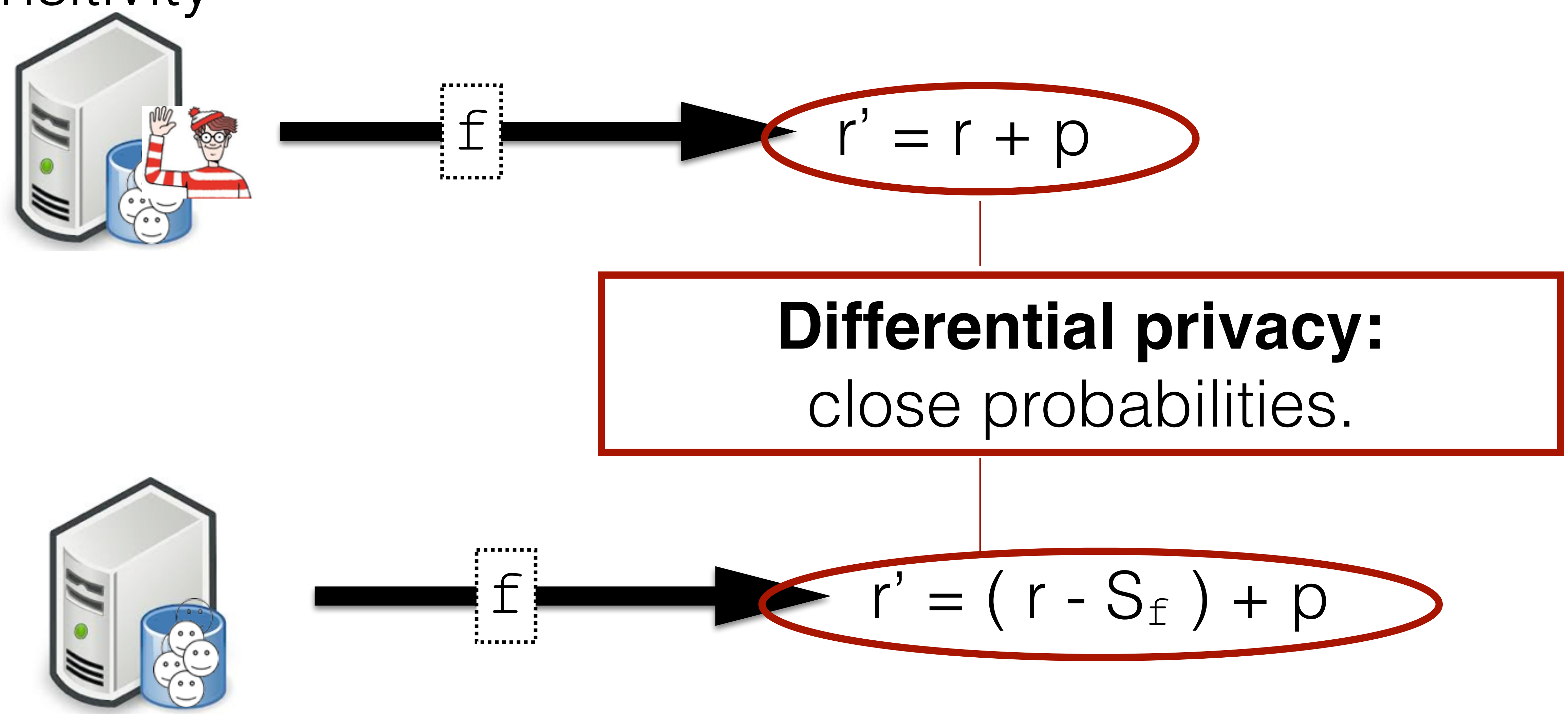
- **Objective :** hide any possible impact of the presence/absence of an individual, as quantified by the query sensitivity

$\xrightarrow{\quad f \quad}$ $r' = r + p$

$\xrightarrow{\quad f \quad}$ $r' = (r - S_{f}) + p$

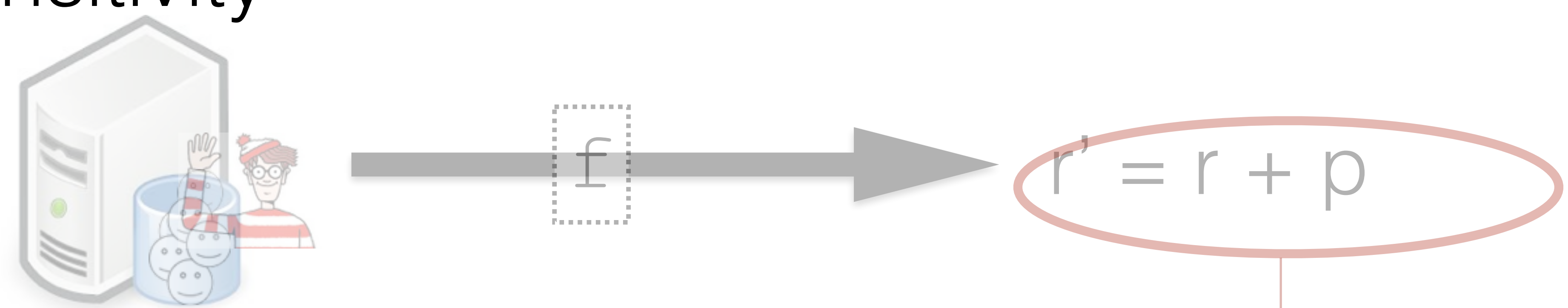
Computing the Distribution for Perturbing

- **Objective :** hide any possible impact of the presence/absence of an individual, as quantified by the query sensitivity



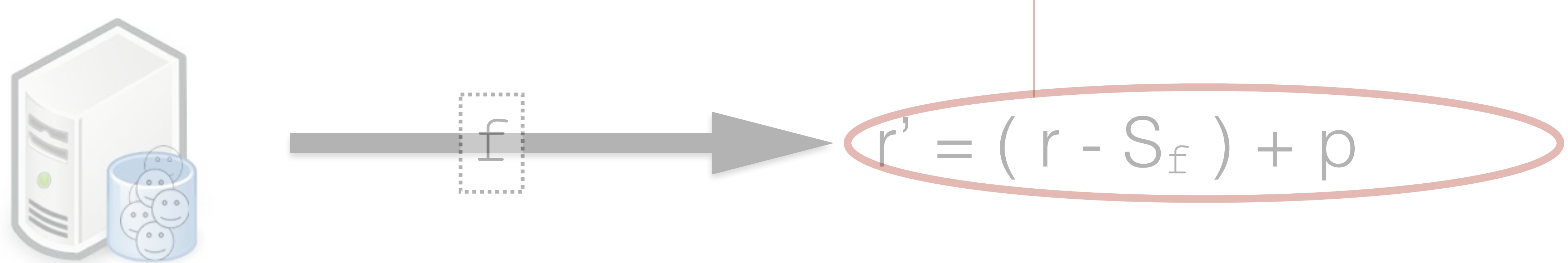
Computing the Distribution for Perturbing

- **Objective :** hide any possible impact of the presence/absence of an individual, as quantified by the query sensitivity



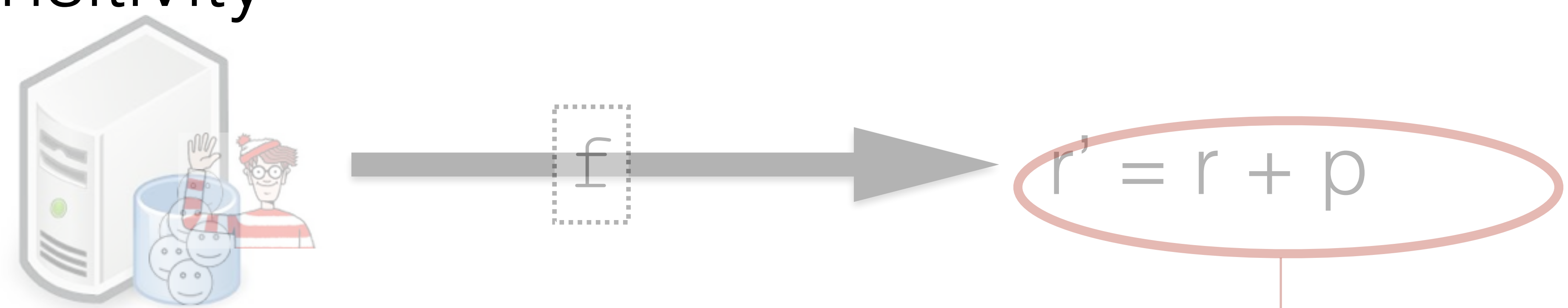
$$Pr[\, r' = r+p \,] \leq e^{\varepsilon} \times Pr[\, r' = r-S_f + p \,] \quad \textbf{(Eq. 1)}$$

$$Pr[\, r' = r-S_f + p \,] \leq e^{\varepsilon} \times Pr[\, r' = r+p \,] \quad \textbf{(Eq. 2)}$$



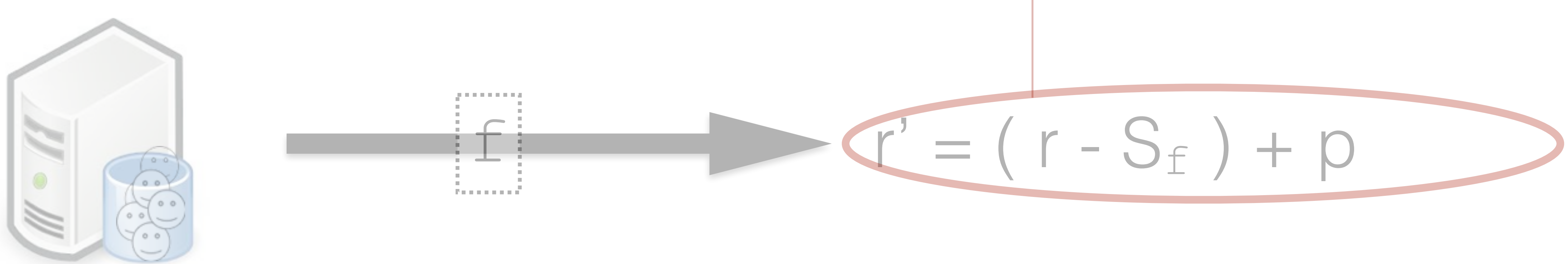
Computing the Distribution for Perturbing

- **Objective :** hide any possible impact of the presence/absence of an individual, as quantified by the query sensitivity



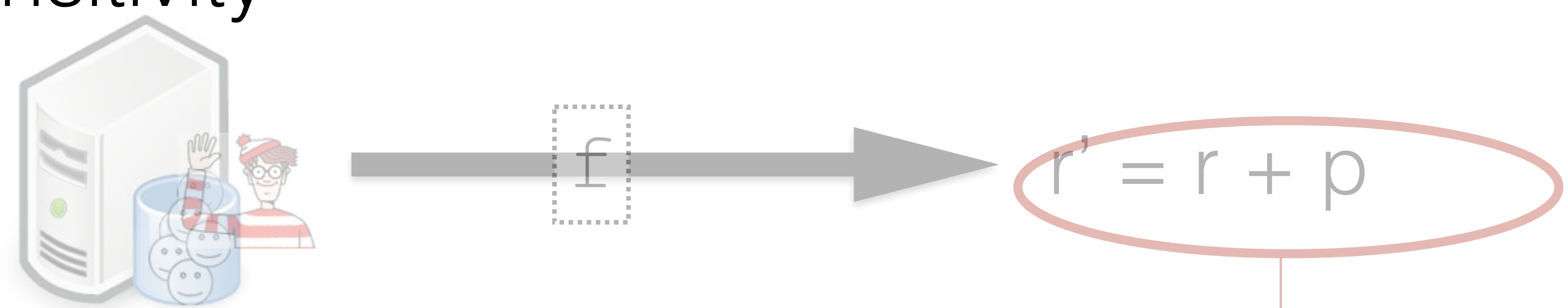
$Pr[\, p = (r' - r) \,] \leq e^{\varepsilon} \times Pr[\, p - S_f = (r' - r) \,]$ **(Eq. 1)**

... **(Eq. 2)**



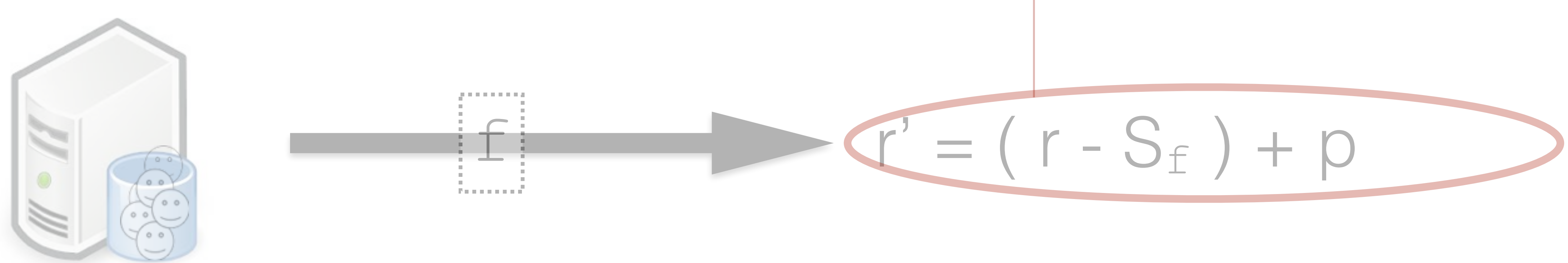
Computing the Distribution for Perturbing

- **Objective :** hide any possible impact of the presence/ absence of an individual, as quantified by the query sensitivity



$\Pr[\,p\,] \le e^{\varepsilon} \times \Pr[\,p - S_f\,]$ **(Eq. 1)**

... **(Eq. 2)**



Computing the Distribution for Perturbing

- **Objective :** hide any possible impact of the presence/absence of an individual, as quantified by the query sensitivity

$$\Pr[\, p \,] \leq e^{\varepsilon} \times \Pr[\, p - S_{f} \,] \qquad \textbf{(Eq. 1)}$$

... **(Eq. 2)**

In which Distribution can *p* be sampled so that
(Eq. 1) and (Eq. 2) are satisfied ?

Laplace Distribution

$\Pr[\,p\,] \leq e^{\varepsilon} \times \Pr[\,p - S_f\,]$ **(Eq. 1)**

... **(Eq. 2)**

- Probability of any x value (PDF of Laplace):
  $\text{Laplace}(x \mid 0, b) = 1/2b * e^{-|x|/b}$

- Hence :
  - $\Pr[\,p\,] = 1/2b * e^{-|p|/b}$ **and** $\Pr[\,p - S_f\,] = 1/2b * e^{-|p-S_f|/b}$
  - What should be b if we want (Eq. 1) and (Eq. 2) ?
    (Eq. 1) => $1/2b * e^{-|p|/b} \leq e^{\varepsilon} * 1/2b * e^{-|p-S_f|/b}$
    $\Rightarrow e^{-|p|/b} \leq e^{\varepsilon} * e^{-|p-S_f|/b} \Rightarrow e^{(|p-S_f|-|p|)/b} \leq e^{\varepsilon}$
    $\Rightarrow (|p-S_f|-|p|)/\varepsilon \leq b$ **=>** $\mathbf{|p-S_f| / \varepsilon \leq |p| / \varepsilon + b}$

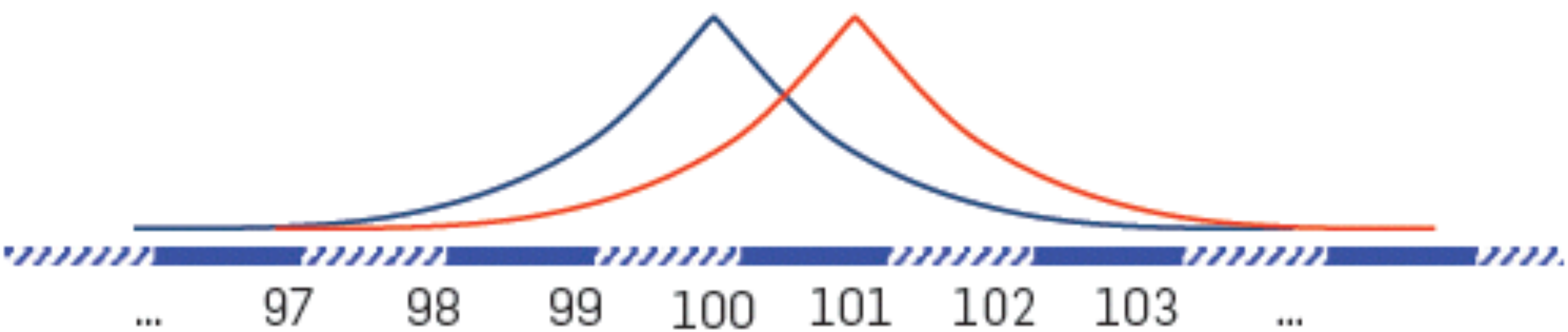
    And (Eq. 2) => … **=>** $\mathbf{|p| / \varepsilon - b \leq |p-S_f| / \varepsilon}$

    Since we always have $|p| - |S_f| \leq |p-S_f| \leq |p| + |S_f|$

    Then (Eq. 1) and (Eq. 2) are both satisfied with: $\mathbf{b = |S_f| / \varepsilon}$

Illustration of the Perturbed Output Probabilities

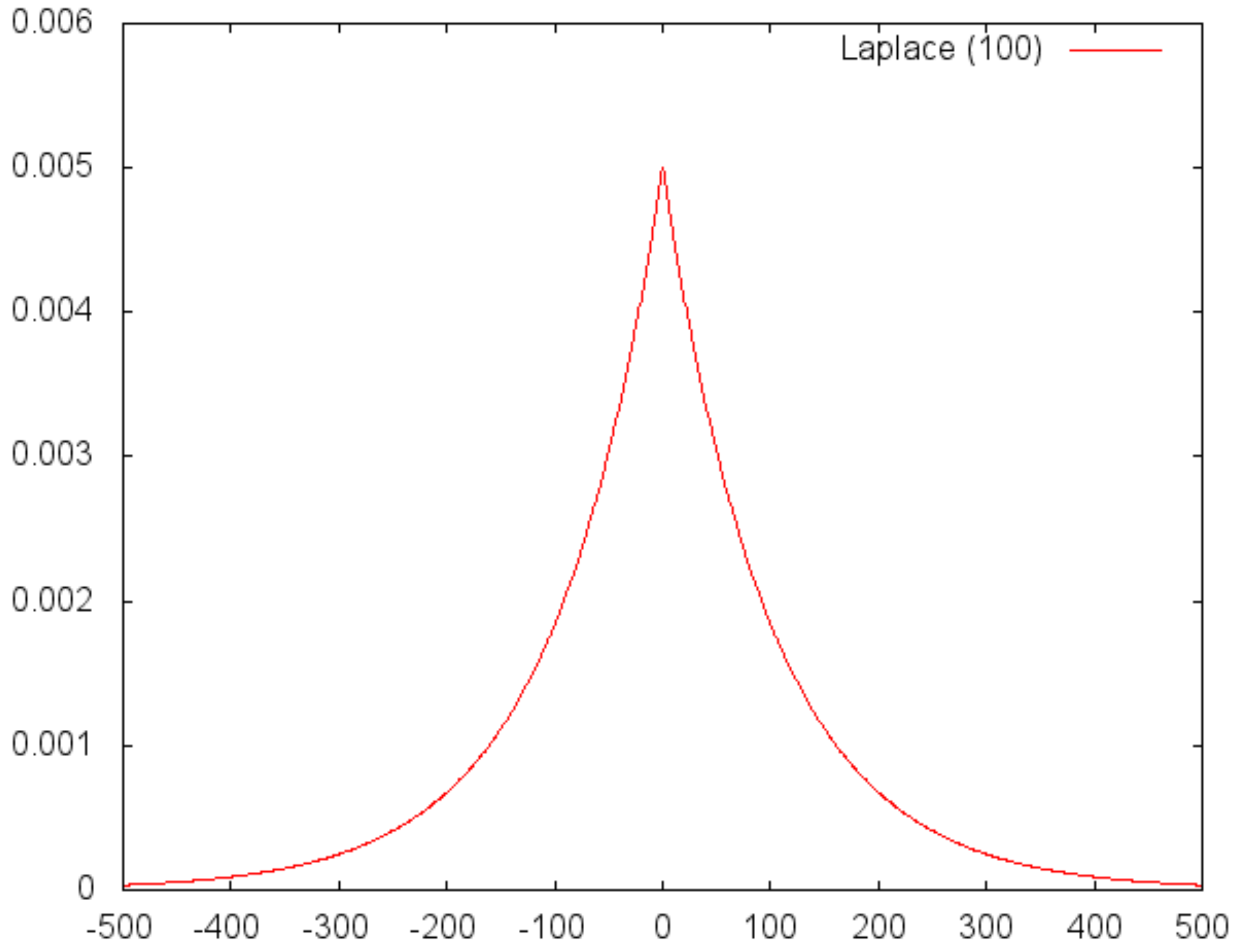
- Assume that the true Count is 100
  - In red, proba. of perturbed outputs (r+p) when Bob is in
  - In blue, *idem* when Bob is out



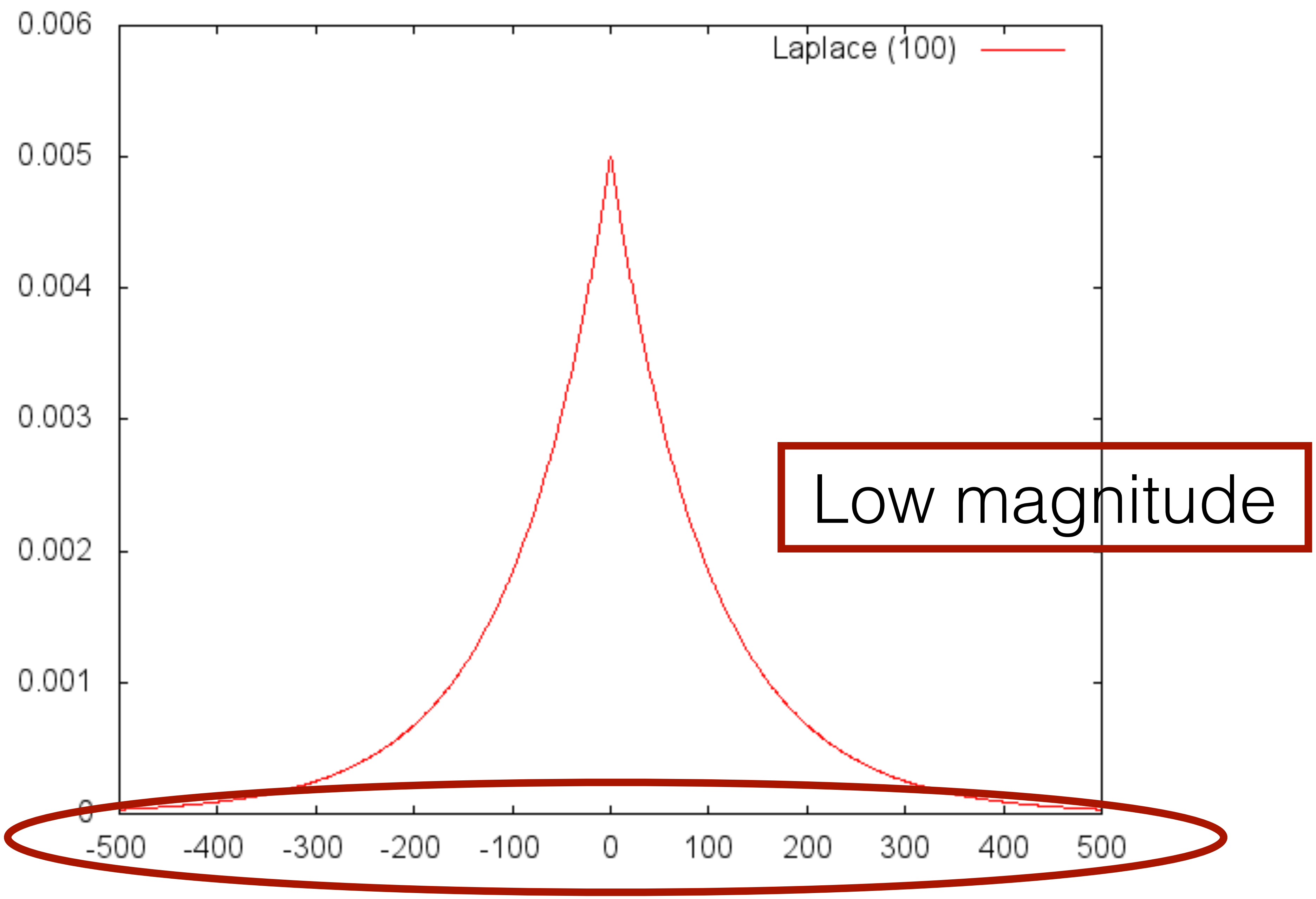
Example: Perturbing

- `COUNT` over DB of salaries ;
- Let ε = 0.01 the privacy parameter (e^{ε}=1.01) ;
- The sensitivity of a `COUNT` is: $S_{\mathtt{COUNT}} = 1$;
- Hence in order to perturb a `COUNT` , *p* is sampled in : `Laplace(`$S_{\mathtt{COUNT}}$ / ε) = `Laplace(`100`)`

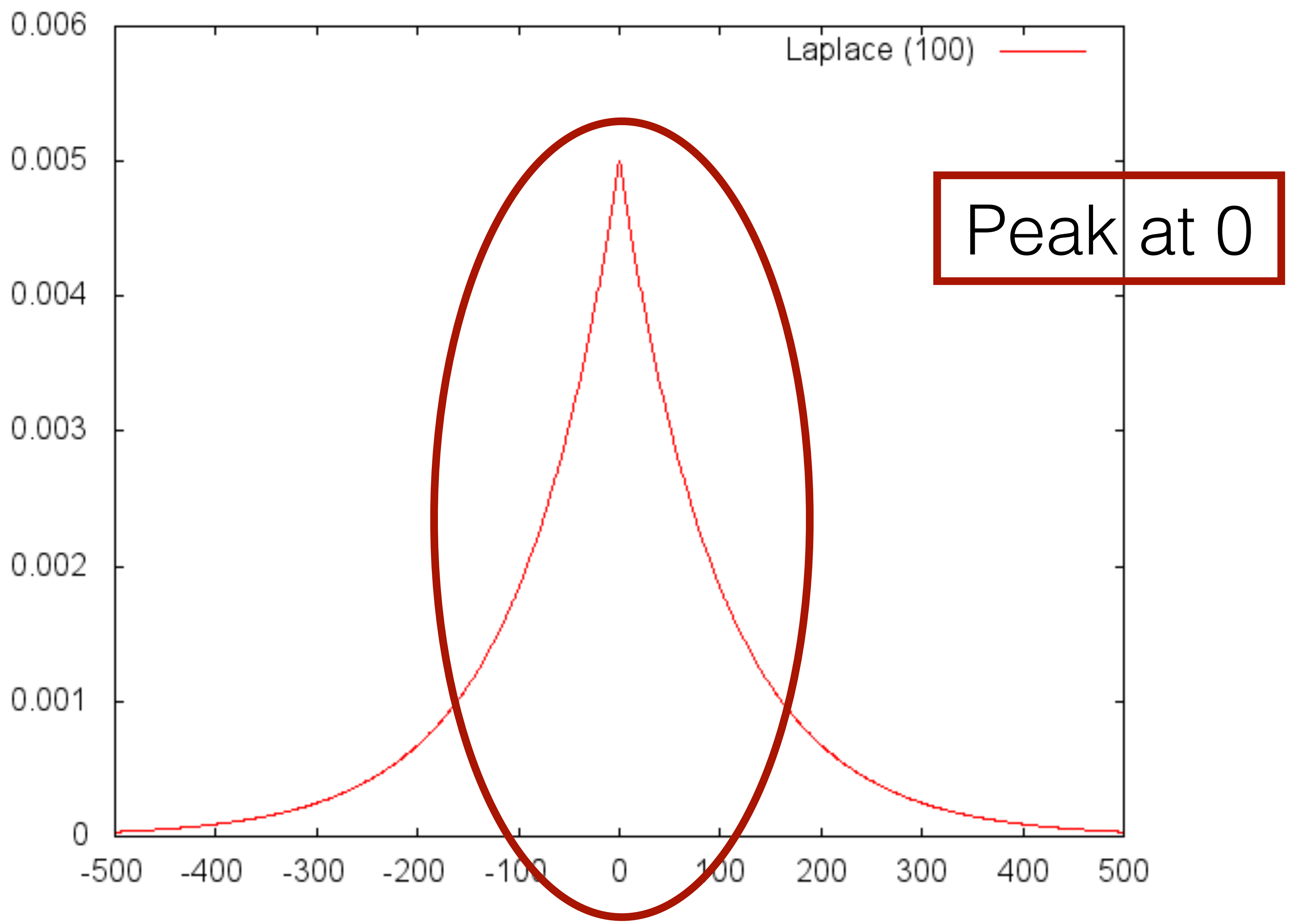
Example: Perturbing



Example


Laplace (100)
0.006
0.005
0.004
0.003
0.002
0.001
0
-500
-400
-300
-200
-100
0
100
200
300
400
500
Low magnitude

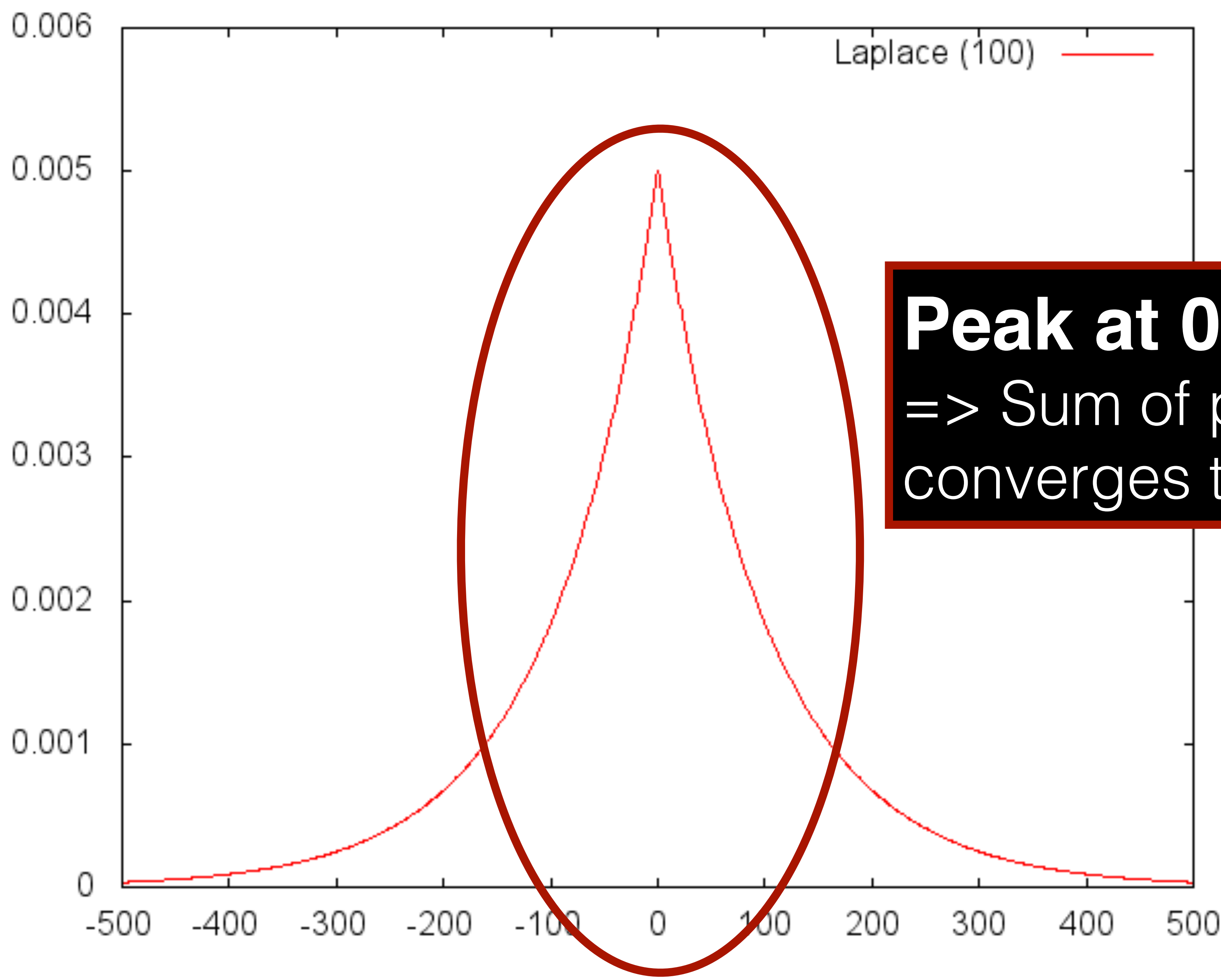
Example


Laplace (100)
0.006
0.005
0.004
0.003
0.002
0.001
0
-500
-400
-300
-200
-100
0
100
200
300
400
500
Peak at 0

Unlimited Querying? NO

- We have seen how to guarantee differential privacy with respect to a single query on the DB
- What do the privacy guaranties become against an *infinite number* of queries?

Progressive Reduction of Privacy Guarantees



Peak at 0

=> Sum of perturbations converges towards 0 !

Unlimited Querying? NO

- We have seen how to guarantee differential privacy with respect to a single query on the DB
- What do the privacy guaranties become against an *infinite number* of queries?
  => NULL
- **Requirement:** limit the number of queries on each DB

Several Queries (limited): Composability

- Composability:
  - **Sequential:** what about two queries over two ***non disjoint*** sets of records, each satisfying independently ε_i - differential privacy?
  - **Parallel :** what about two queries over **disjoint** sets of records, each satisfying independently ε_i - differential privacy?

Several Queries (limited): Composability

- Composability:
  - **Sequential:** what about two queries over two ***non disjoint*** sets of records, each satisfying independently ε_i - differential privacy?
    => ($2*\varepsilon_i$) - differential privacy is satisfied (and in general for n queries, $\sum\varepsilon_i$ - differential privacy is satisfied)
  - **Parallel :** what about two queries over **disjoint** sets of records, each satisfying independently ε_i - differential privacy?
    ε_i - differential privacy is satisfied (also for n queries)

Budget

- As a result, ε can be considered as a **privacy budget** :
  - To be distributed between queries (e.g., uniformly with *n* queries will yield $\varepsilon_i = \varepsilon / n$) ;
  - Stop answering to queries when the privacy budget is completely consumed: $\sum \varepsilon_i \leq \varepsilon$ must always be true !

Differential Privacy : Conclusion

A Gold Standard

- The current *de facto* standard, thanks to its sound guarantees and its composability properties:
  - Robust against an attacker knowing all records except one
  - Self-composable
- Intense research activities : studying its mathematical properties, proposing new mechanisms, proposing variants of the model
- Some real-life uses (eg http://onthemap.ces.census.gov)