Data Privacy
Hiding Data from the Database User II

Erman Ayday

Databases

- Many databases contain sensitive (personal) data
  - Hospital records, internet search information, the set of friends on different social sites, etc.
- It is a common scenario that the release of a function/ statistic on such data is socially beneficial
  - Used for apportioning resources, evaluating medical therapies, understanding the spread of disease, improving economic utility, and informing us about ourselves as a species
  - E.g., the usage of hospital records greatly helps medical research
- Hard to publish databases in a privacy-preserving way
- Crucial to ensure that the release of a function on a database does not leak too much information about the individuals
  - Differential privacy is a quite recent notion that tries to formalize this requirement

Privacy Mechanisms for Databases

- Non-interactive mechanisms
  - Database publishes a sanitized dataset
  - Researcher asks arbitrary queries on the sanitized dataset

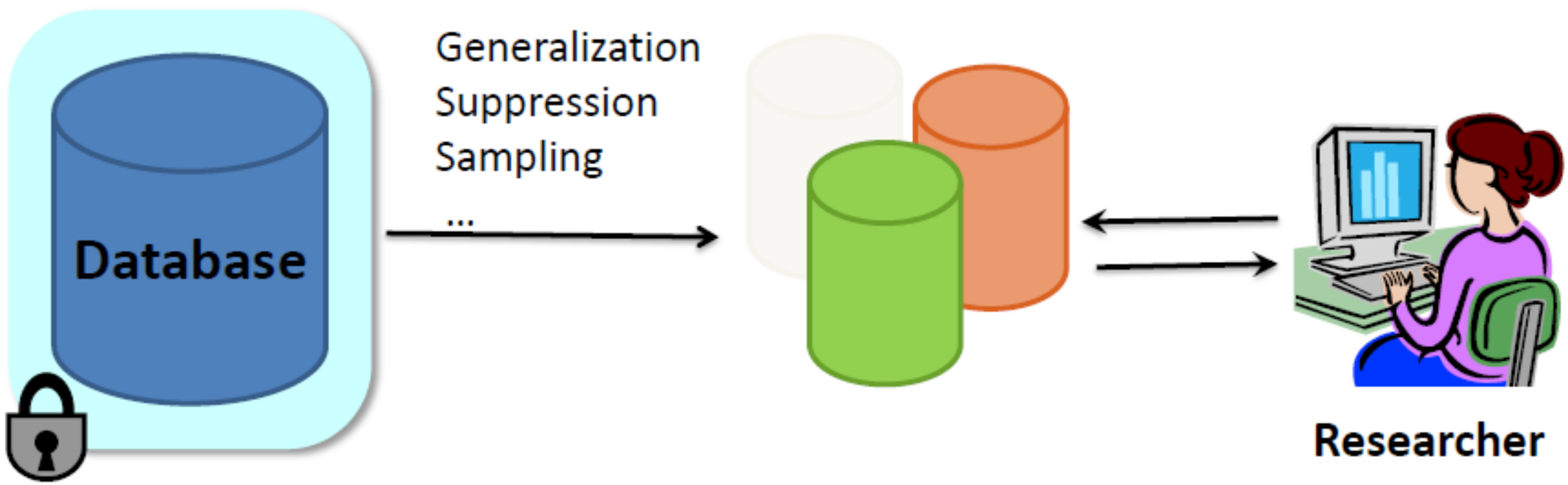


Figure: Ashwin Machanavajjhala

k-Anonymity [1]

- Each person contained in the database cannot be distinguished from at least k-1 other individuals whose information also appear in the released database

	Race	Birth	Gender	ZIP	Problem
t1	Black	1965	m	02141	short breath
t2	Black	1965	m	02141	chest pain
t3	Black	1964	f	02138	obesity
t4	Black	1964	f	02138	chest pain
t5	White	1964	m	02138	chest pain
t6	White	1964	m	02138	obesity
t7	White	1964	m	02138	short breath

[1] L. Sweeney. K-anonymity: A model for protecting privacy. Int. J. Uncertain. Fuzziness Knowl.-Based Syst., vol. 10, no. 5, pp. 557-570, Oct.

k-Anonymity - Limitation

- Does not provide privacy when sensitive values lack diversity

	Non-Sensitive			Sensitive
	Zip Code	Age	Nationality	Condition
1	13053	28	Russian	Heart Disease
2	13068	29	American	Heart Disease
3	13068	21	Japanese	Viral Infection
4	13053	23	American	Viral Infection
5	14853	50	Indian	Cancer
6	14853	55	Russian	Heart Disease
7	14850	47	American	Viral Infection
8	14850	49	American	Viral Infection
9	13053	31	American	Cancer
10	13053	37	Indian	Cancer
11	13068	36	Japanese	Cancer
12	13068	35	American	Cancer

(a)

	Non-Sensitive			Sensitive
	Zip Code	Age	Nationality	Condition
1	130**	< 30	*	Heart Disease
2	130**	< 30	*	Heart Disease
3	130**	< 30	*	Viral Infection
4	130**	< 30	*	Viral Infection
5	1485*	≥ 40	*	Cancer
6	1485*	≥ 40	*	Heart Disease
7	1485*	≥ 40	*	Viral Infection
8	1485*	≥ 40	*	Viral Infection
9	130**	3*	*	Cancer
10	130**	3*	*	Cancer
11	130**	3*	*	Cancer
12	130**	3*	*	Cancer

An equivalence class

(b)

(a) A hospital records dataset
(b) The 4-anonymous version of the same hospital records dataset

l-diversity

- An equivalence class has ℓ-diversity if there are at least ℓ well-represented values for the sensitive attribute
- A database has ℓ-diversity if every equivalence class has ℓ-diversity

	ZIP Code	Age	Salary	Disease
1	476**	2*	3K	gastric ulcer
2	476**	2*	4K	gastritis
3	476**	2*	5K	stomach cancer
4	4790*	≥ 40	6K	gastritis
5	4790*	≥ 40	11K	flu
6	4790*	≥ 40	8K	bronchitis
7	476**	3*	7K	bronchitis
8	476**	3*	9K	pneumonia
9	476**	3*	10K	stomach cancer

A 3-diverse hospital records dataset

[1] A. Machanavajjhala, D. Kifer, J. Gehrke, and M. Venkitasubramaniam. L-diversity: Privacy beyond k-anonymity. ACM Trans. Knowl. Discov. Data, vol. 1, no. 1, Mar. 2007

l-diversity Limitations

- ℓ-diversity does not consider overall distribution of sensitive values
- ℓ-diversity does not consider semantics of sensitive values

t-Closeness

- An equivalence class has t-closeness if the distance between the distribution of a sensitive attribute in this class and the distribution of the attribute in the whole table is no more than a threshold t
- A table has t-closeness if all equivalence classes have t-closeness

N. Li and T. Li. t-closeness: Privacy beyond k-anonymity and l-diversity. IEEE 23rd Intl Conf. on Data Engineering (ICDE), 2007

Privacy Mechanisms for Databases

- Interactive mechanisms
  - Researcher directly asks queries to the database
  - Database can choose to answer truthfully or answer with noise
  - Auditor may keep track of all the queries pose to the database and deny queries

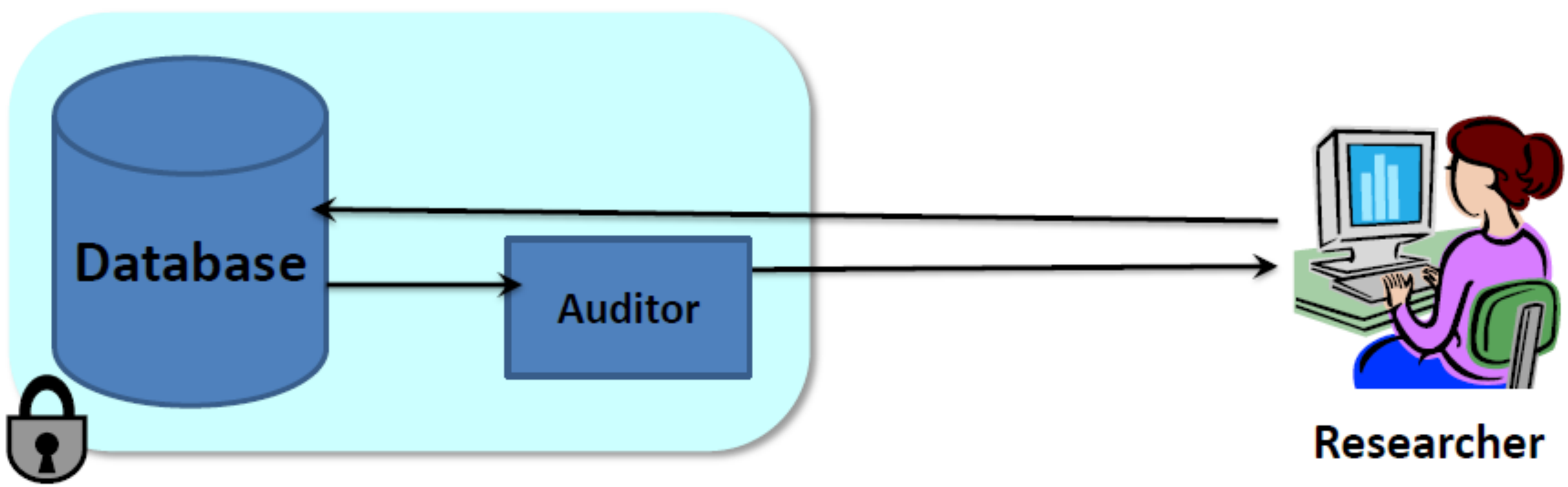


Figure: Ashwin Machanavajjhala

Defining Privacy for Interactive Mechanisms

- After learning the answer to a private query one should have no extra knowledge about any individual in comparison with the earlier situation
- Hard to achieve if we want the answer to have any utility
  - We must allow the leakage of some information
  - We can only demand a bound on the extent of leakage

Methods to Release Statistics

- Large query sets
  - Disallows queries about a specific individual or small set of individuals
  - But, how about the below queries?
    - "How many people in the database have the sickle cell trait?"
    - "How many people, not named X, in the database have the sickle cell trait?"

Name	Sickle cell trait
A	Yes
B	Yes
C	No
D	No
X	No
Y	Yes
Z	No

Methods to Release Statistics

- Query auditing
  - Keeps the query history to determine if a response would be disclosive
  - Computationally infeasible
  - Refusal to respond to a query may itself be disclosive

- Example:
  - Max sensitive value of males?
    => 2
  - Max sensitive value of 1st year PhD students?
    =>3
  - Xi: only female 1st year PhD student
    - Sensitive value of Xi: 3

Name	1st year PhD	Gender	Sensitive value
Ben	Y	M	1
Bha	N	M	1
Ios	Y	M	1
Jan	N	M	2
Jian	Y	M	2
Jie	N	M	1
Joe	N	M	2
Moh	N	M	1
Son	N	F	1
Xi	Y	F	3
Yao	N	M	2

Figure: Ashwin Machanavajjhala

Methods to Release Statistics

- Subsampling
  - A subset of the rows is chosen at random and released and statistics are computed on the subsample
  - Appearing in a subsample may have terrible consequences
    - Every time subsampling occurs some individual suffers
- Input perturbation
  - Data or the queries are modified before a response is generated
  - Repeating the same query yields the same answer
  - Generalization of subsampling (has the same disadvantage)

Methods to Release Statistics

- Randomized response
  - Respondents to a query flip a coin and, based on the outcome they decide between honestly reporting a value and responding randomly
  - Privacy comes from the uncertainty of how to interpret a reported value
- Adding random noise to the output
  - If done naively this approach will fail
    - E.g., if the same query is asked repeatedly, then the responses can be averaged, and the true answer will eventually emerge
  - Cannot be fixed by recording each query and providing the same response each time a query is re-issued
    - Syntactically different queries may be semantically equivalent, and, if the query language is sufficiently rich, then the equivalence problem itself is undecidable

Problems About Naïve Noise Addition

- **Theorem:** Let M be a mechanism that adds noise bounded by E. Then there exists an adversary that can re-construct the database to within $4E$ positions (Dinur and Nissim 2003)
- **Example:** Consider a database of n entries
  - Adding noise with magnitude always bounded by $n/401$ is blatantly non-private against an adversary that can ask all 2^n possible queries
    - Query all the possible subsets of the database
  - Adversary can construct a candidate database that agrees with the real database in 99% of the entries
- **Another result:** Noise of magnitude $o(\sqrt{n})$ is blatantly non-private against a series of $n \log^2 n$ randomly generated queries (Dinur and Nissim 2003)
- **(Hard to Achieve) Goal:** Generate a noisy table that will permit highly accurate answers to be derived for computations that are not specified at the outset

Dalenius' s Desideratum (1977)

- Tore Dalenius, statistician
- Articulated a privacy goal for statistical databases:
- *"anything that can be learned about a respondent from the statistical database should be learnable without access to the database"*
- Many papers in the literature attempt to formalize Dalenius' goal by requiring that
  - the adversary' s prior and posterior views about an individual (i.e., before and after having access to the statistical database) shouldn't be too different or
  - that access to the statistical database shouldn't change the adversary's views about any individual too much
- But, if the statistical database teaches us anything at all, then it should change our beliefs about individuals

Differential Privacy [1]

- **A new privacy goal:** minimize the increased risk to an individual incurred by joining (or leaving) the database
  - Move from comparing an adversary's prior and posterior views of an individual to comparing the risk to an individual when included in, versus when not included in, the database
  - There are attempts to weaken this definition to increase utility (e.g., membership privacy)
- **Motivation:** A privacy guarantee that limits risk incurred by joining therefore encourages participation in the dataset, increasing social utility
- ***Differential privacy:*** privacy-preserving statistical analysis of data

[1] C. Dwork. Differential Privacy. ICALP, 2006

Differential Privacy

- **Basic philosophy:** instead of the real answer to a query, output a random answer, such that by a small change in the database (someone joins or leaves), the distribution of the answer does not change much

Example

Query #1
avg blood sugar level
of the group?

Alice	4.2
Bob	5.9
Cathy	5.2
Diana	6.9
Ellen	5.7
Avg:	**5.58**

Query #2
avg blood sugar level
of female members?

Alice	4.2
-	
Cathy	5.2
Diana	6.9
Ellen	5.7
Avg:	**5.50**

Differentially private approach:
let's add some noise of `unif(-2,2)`

Alice	4.5
Bob	5.1
Cathy	4.41
Diana	6.2
Ellen	5.7
Avg:	**5.23**

Err. ~7%

Alice	3.0
-	
Cathy	3.7
Diana	7.5
Ellen	7.5
Avg:	**5.46**

Err. <1%

Blood sugar level of Bob?
`5*5.58-4*5.5 = 5.9`

Blood sugar level of Bob?
`5*5,23-4*5,46 = 4,3`

Err. ~27%

Figure: Gabor Gorgy Gulyas

Differential Privacy - Definitions

- $\mathcal{D}$: The set of input databases
- R: Output space of the query
- F: Query function

$$F\colon \mathcal{D} \rightarrow R$$

- d: Distance function on the set of databases
- Neighboring databases: Pairs of databases (D, D') differing only in one row (e.g., individual)

$$d(D - D') = 1$$

ε-Differential Privacy – Formal Definition

- *Let* $\mathcal{D}$ be a set *of databases with distance function* d *and an image set* R. *We call a randomized function* M ε-differentially private *if for all* $D_1, D_2 \in \mathcal{D}$ *with* $d(D_1, D_2) \leq 1$ *and for all* $C \subseteq R$ *we have*

$$\Pr(M(D_1) \in C) \leq \exp(\varepsilon) \cdot \Pr(\mathrm{M}(\mathrm{D}_2) \in C)$$

ε-Differential Privacy

- Ensures, that even if the adversary knows each record in the database except for the record of a person x, he cannot learn much about the record of x
- Guarantees a strong protection against the adversary learning information based on others' data and the output

Differential Privacy – Weaker Notion

- Approximate differential privacy:
- *Let* $\mathcal{D}$ be a set *of databases with distance function* d *and an image set* R. *We call a randomized function* M (ε, δ)-differentially private *if for all* $D_1, D_2 \in \mathcal{D}$ *with* $d(D_1, D_2) \leq 1$ *and for all* $C \subseteq R$ *we have*

$$\Pr(M(D_1) \in \boldsymbol{C}) \leq \exp(\varepsilon) \cdot \Pr(\mathrm{M}(\mathrm{D}_2) \in \boldsymbol{C}) + \delta$$

Achieving Differential Privacy

- Output Randomization
- Add noise to the answer of a query such that
  - Answer does not leak too much information about the database
  - Noisy answers are close to the original answers

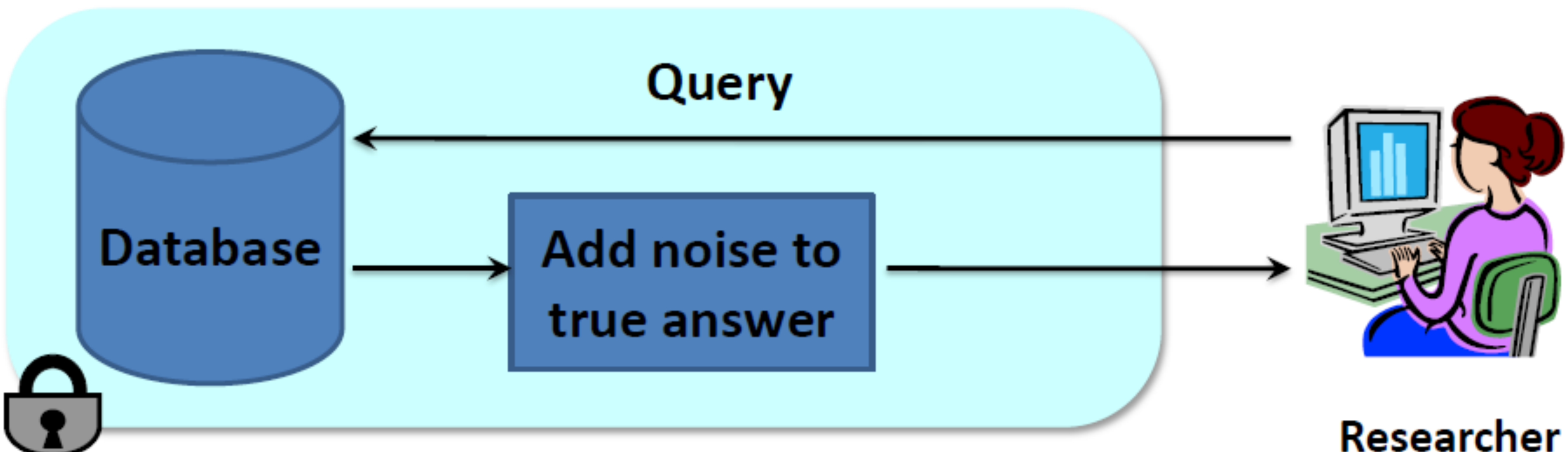
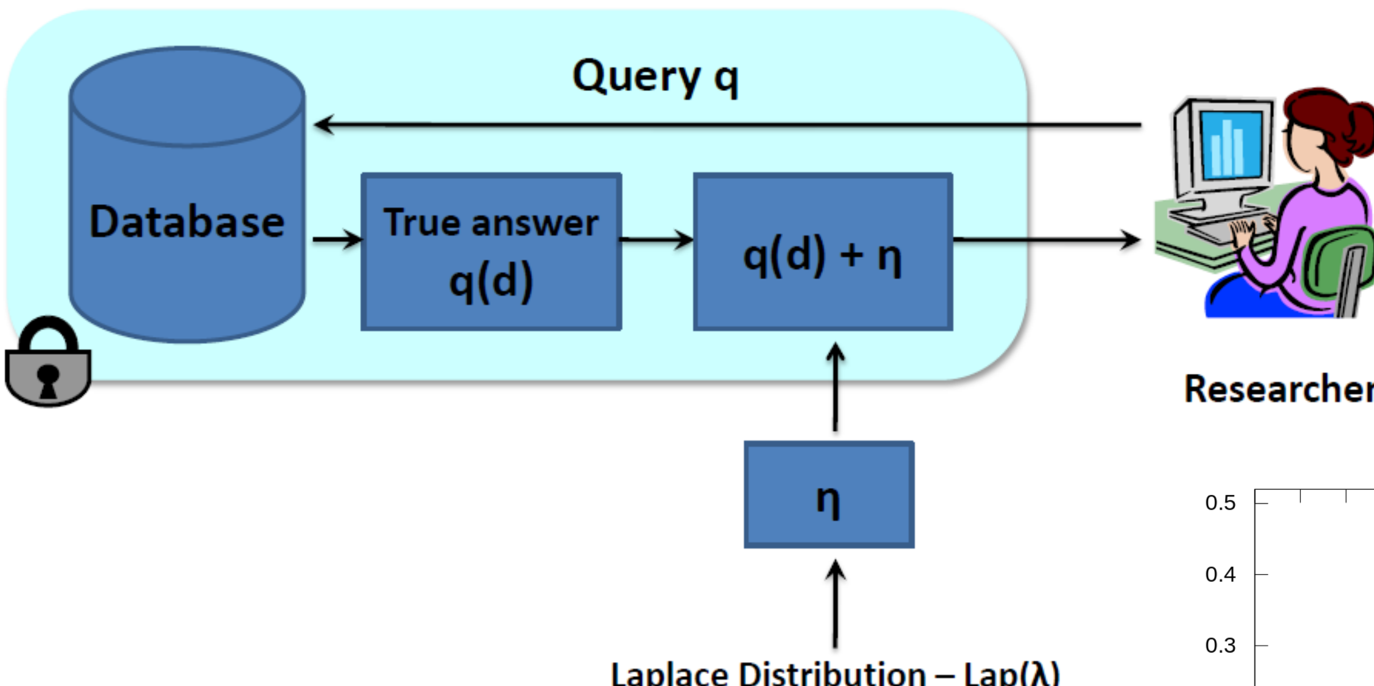


Figure: Ashwin Machanavajjhala

Laplacian Noise

- Output randomization can be implemented by adding noise drawn from some distribution
- Add noise from a Laplacian distribution



$$P(\eta|\lambda) = \frac{1}{2\lambda}\exp(-|\eta|/\lambda)$$

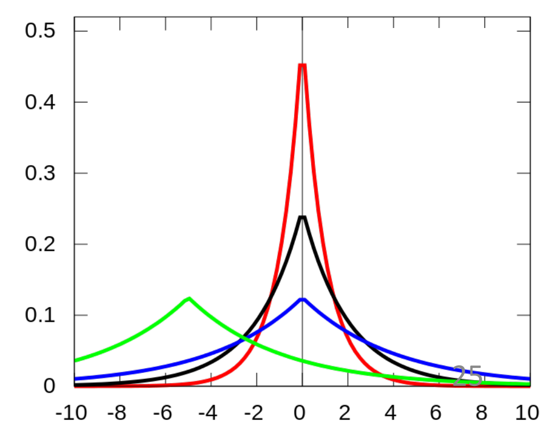
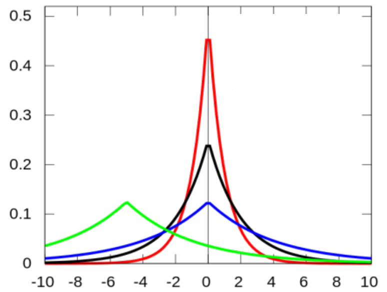


Figure: Ashwin Machanavajjhala

Why Laplace Noise?



- The Laplace distribution with parameter λ, denoted $Lap(\lambda)$, has density function $P(\eta|\lambda) = \frac{1}{2\lambda}\exp(-|\eta|/\lambda)$ with variance $2\lambda^2$
  - Taking $\lambda = 1/\varepsilon$ the density at η is proportional to $e^{-\varepsilon|\eta|}$
- This distribution has highest density at 0 (good for accuracy)
- For any η, η' such that $|\eta - \eta'| \leq 1$ the density at η is at most e^{ε} times the density at η', satisfying the differential privacy requirement
- It is symmetric about 0 and has a heavy tail

How Much Noise for Privacy?

- Selecting ε
  - The parameter ε is public, and its selection is a social question
  - Selection of ε by Cynthia Dwork:
    - *"We tend to think of ε as, 0.01, 0.1, or in some cases, $ln2$ or $ln3$"*
  - Smaller ε means better privacy
  - But, what about the utility?

- Sensitivity of a Query (Dwork et al., TCC 2006)
  - If the sensitivity of a query is S, then the following guarantees ε-differential privacy:

$$\lambda = S/\varepsilon$$

Sensitivity of a Query – S(F)

- For any two neighboring databases (D, D')

$$S(F) = \max_{D,D'} ||F(D) - F(D')||$$

- Sensitivity of counting queries:
  - The number of elements in the database that have a given property P
  - By adding or deleting one element of the database, F can change by at most 1
  - $S(counting) = 1$
- Sensitivity of histogram queries:
  - Suppose each entry in d takes values in $\{c_1, c_2, \dots, c_n\}$
  - $Histogram(d) = \{m_1, \dots, m_n\}$
    - $m_i = (\#\ entries\ in\ d\ with\ value\ c_i)$
  - $S(histogram) = 2$

Sensitivity - Exercise

- Consider a database of n numbers in which each entry is an integer from the set [0,100]

- Sensitivity of mean?
  - 100/n
- Sensitivity of median?
  - 100

Differential Privacy – Proof (1)

- **Theorem:** Adding noise drawn from a Laplacian distribution guarantees ε-differential privacy if $\lambda \geq S(F)/\varepsilon$

- **Proof:**

Let $D = \{\boldsymbol{x}_1, x_2, \ldots, x_n\}$ & $D' = \{\boldsymbol{y}_1, x_2, \ldots, x_n\}$ be 2 inputs databases

Let F be a query with sensitivity $S(F)$

– $F(D) = F(x_1, x_2, \ldots, x_n) = a$ $\quad F(D') = F(y_1, x_2, \ldots, x_n) = b$

– $|a - b| \leq S(F)$

Let be $o = a + \eta$ the perturbed output for $F(D)$

– η is sampled i.i.d from $Lap(S(F)/\varepsilon)$

Differential Privacy – Proof (2)

$$\log\left(\frac{\Pr(F(D)=o)}{\Pr(F(D')=o)}\right)=\log\left(\frac{\Pr(\eta=a-o)}{\Pr(\eta=b-o)}\right)$$

$$=\log\left(\frac{\Pr(\eta=a-o)}{\Pr(\eta=b-o)}\right)=\log\left(\frac{\exp(-|a-o|/\lambda)}{\exp(-|b-o|/\lambda)}\right)$$

$$=\frac{|a-o|}{\lambda}-\frac{|b-o|}{\lambda}$$

$$\leq\frac{|a-b|}{\lambda}\leq\frac{S(F)}{\lambda}\leq\varepsilon$$

Composability

- $F_1(D)$ – guarantees some privacy definition with parameter ε_1
- $F_2(D)$ – guarantees some privacy definition with parameter ε_2
- Then releasing both $F_1(D)$ and $F_2(D)$ satisfies the same privacy definition with parameter $f(\varepsilon_1, \varepsilon_2)$

Composability of Differential Privacy

- **Theorem:** If algorithms $F_1, F_2, \ldots, F_k$ use independent randomness and each F_i satisfies $\boldsymbol{\varepsilon_i}$-differential privacy, respectively. Then, outputting all the answers together satisfies differential privacy with

$$\boldsymbol{\varepsilon = \varepsilon_1 + \varepsilon_2 + \ldots + \varepsilon_k}$$

When Output Perturbation Doesn't Make Sense

- What if we have a non-numeric valued query?
  - "What is the most common eye color in this room?"
- What if the perturbed answer isn't almost as good as the exact answer?
  - "Which price would bring the most money from a set of buyers?"

Example: Apples for Sale

$1.00

$1.00

$1.00

$4.01

Set the price of apples at $1.00 for profit: $4.00

Set the price of apples at $4.01 for profit $4.01

Best price: $4.01

2nd best price: $1.00

Profit if you set the price at $4.02: $0

Profit if you set the price at $1.01: $1.01

Exponential Mechanism [1] - Overview

- Generalization of ε-differential privacy
- For a query F on a dataset D:

The exponential mechanism $\mathcal{E}$ takes a score function q_F, a parameter ε and does the following:

- $\mathcal{E}(D, q_F, \varepsilon)$= output r with probability proportional to $\exp\left(\frac{\varepsilon}{2\Delta_q} q_F(D, r)\right)$
- $q_F(D, r)$ is the score function for query F
- Δ_q is the sensitivity of the score function q

[1] F. McSherry and K. Talwar. Mechanism design via differential privacy. 48th Annual Symposium on Foundations of Computer Science, 2007

Score Function

- The score function $q_F : \mathcal{D} \times R \rightarrow \mathbb{R}$ corresponding to function F determines how good a given output is for a given input
- $q_F(D, r) \in \mathbb{R}$ means the value of output r on input D
  - Intuitively it means how close is $F(D)$ to r
- Higher values mean better result

$$OPT_F(D) := \max\{q_F(D, r) : r \in R\}$$

Score Function – Examples

- If a function takes its values from $\mathbb{R}^k$, then $q_F(D,r) = -||F(D) - r||$ is a natural score function
  - $||.||$ is a norm on $\mathbb{R}^k$
- F = counting query
  - $q_F(D,r) = -|F(D) - r|$
- F = average
  - $q_F(D,r) = -|(\sum_{x \in D} x)/|D| - r|$

Sensitivity of Score Function

- Sensitivity of the scoring function q_F:

$$\Delta_q = \max_{r \in R, D, D'} |q_F(D,r) - q_F(D',r)|$$

- The sensitivity tells the maximum change in the scoring function for any pair of datasets D, D' such that $d(D, D') \leq 1$
- Intuitively, it tells how large can be a change in the "goodness" of an output after an elementary change in the input database

Exponential Mechanism

- An exponential mechanism $\mathcal{E}$ belonging to a query F with score function q_F gives an output r with the following probability on an input database D

$$\Pr(r) = \frac{\exp\left(\frac{\varepsilon.\, q_F(D,r)}{2\Delta_q}\right)}{\sum_{s\in R} \exp\left(\frac{\varepsilon.\, q_F(D,s)}{2\Delta_q}\right)}$$

- **Idea:** Make high quality outputs exponentially more likely at a rate that depends on the sensitivity of the quality score (and the privacy parameter)

Privacy of Exponential Mechanism

- **Theorem:** The exponential mechanism $\mathcal{E}(D, q_F, \varepsilon)$ corresponding to a function $F: \mathcal{D} \rightarrow R$, with score function $q_F : \mathcal{D} \times R \rightarrow \mathbb{R}$ gives $\varepsilon-$ differential privacy
- **Proof:**

Fix any $D, D' \in \mathcal{D}$ with $d(D, D') \leq 1$ and any $r \in R$

Let Δ_q be the sensitivity of score function q_F

$$\Delta_q = \max_{r \in R, D, D'} |q_F(D, r) - q_F(D', r)|$$

Privacy of Exponential Mechanism - Proof

$$\frac{\Pr[\mathcal{E}(D, q_F, \varepsilon) = r]}{\Pr[\mathcal{E}(D', q_F, \varepsilon) = r]} = \frac{\left(\frac{\exp\left(\frac{\varepsilon.\, q_F(D,r)}{2\Delta_q}\right)}{\sum_{s\in R}\exp\left(\frac{\varepsilon.\, q_F(D,s)}{2\Delta_q}\right)}\right)}{\frac{\exp\left(\frac{\varepsilon.\, q_F(D',r)}{2\Delta_q}\right)}{\sum_{s\in R}\exp\left(\frac{\varepsilon.\, q_F(D',s)}{2\Delta_q}\right)}}$$

$$= \left(\frac{\exp\left(\frac{\varepsilon.\, q_F(D,r)}{2\Delta_q}\right)}{\exp\left(\frac{\varepsilon.\, q_F(D',r)}{2\Delta_q}\right)}\right) \times \left(\frac{\sum_{s\in R}\exp\left(\frac{\varepsilon.\, q_F(D',s)}{2\Delta_q}\right)}{\sum_{s\in R}\exp\left(\frac{\varepsilon.\, q_F(D,s)}{2\Delta_q}\right)}\right)$$

$*$ $\qquad$ $**$

Privacy of Exponential Mechanism - Proof

$$\left(\frac{\exp\left(\frac{\varepsilon.\, q_F(D,r)}{2\Delta_q}\right)}{\exp\left(\frac{\varepsilon.\, q_F(D',r)}{2\Delta_q}\right)}\right) = \exp\left(\frac{\varepsilon(q_F(D,r) - q_F(D',r))}{2\Delta_q}\right)$$

$$\leq \exp\left(\frac{\varepsilon\Delta_q}{2\Delta_q}\right) = \exp\left(\frac{\varepsilon}{2}\right)$$

*

$$\left(\frac{\sum_{s\in R}\exp\left(\frac{\varepsilon.\, q_F(D',s)}{2\Delta_q}\right)}{\sum_{s\in R}\exp\left(\frac{\varepsilon.\, q_F(D,s)}{2\Delta_q}\right)}\right) = \left(\frac{\sum_{s\in R}\exp\left(\frac{\varepsilon(q_F(D',s) + q_F(D,s) - q_F(D,s))}{2\Delta_q}\right)}{\sum_{s\in R}\exp\left(\frac{\varepsilon.\, q_F(D,s)}{2\Delta_q}\right)}\right)$$

$$\leq \left(\frac{\sum_{s\in R}\exp\left(\frac{\varepsilon\left(q_F(D,s) + \Delta_q\right)}{2\Delta_q}\right)}{\sum_{s\in R}\exp\left(\frac{\varepsilon.\, q_F(D,s)}{2\Delta_q}\right)}\right) = \left(\frac{\exp\left(\frac{\varepsilon}{2}\right)\sum_{s\in R}\exp\left(\frac{\varepsilon\left(q_F(D,s)\right)}{2\Delta_q}\right)}{\sum_{s\in R}\exp\left(\frac{\varepsilon.\, q_F(D,s)}{2\Delta_q}\right)}\right) = \exp\left(\frac{\varepsilon}{2}\right)$$

**

Privacy of Exponential Mechanism - Proof

- Using $*$ and $**$:

$$\frac{\Pr[\mathcal{E}(D, q_F, \varepsilon) = r]}{\Pr[\mathcal{E}(D', q_F, \varepsilon) = r]} \leq \exp\left(\frac{\varepsilon}{2}\right) \exp\left(\frac{\varepsilon}{2}\right)$$

$$= \exp \varepsilon$$

Utility of Exponential Mechanism

- Probability of obtaining a highly suboptimal output is exponentially small
- **Theorem** (Gupta et al., 2010): Let R be finite, and $r^* = \mathcal{E}(D, q_F, \varepsilon)$. Let also $R_{OPT}(D)$ be the set of optimal outputs for input D such that

D: $R_{OPT}(D) = \{r \in R : q_F(D, r) = OPT_F(D)\} \implies$

$$\Pr\left[q_F(D, r^*) \leq OPT_F(D) - \frac{2\Delta}{\varepsilon}\left(\log\left(\frac{|R|}{|R_{OPT}|}\right) + t\right)\right] \leq e^{-t}$$

Remember: $OPT_F(D) := \max\{q_F(D, r) : r \in R\}$

Utility of Exponential Mechanism

- **Proof:**

$$x = OPT_F(D) - \frac{2\Delta}{\varepsilon}\left(\log\left(\frac{|R|}{|R_{OPT}|}\right) + t\right)$$

$$\Pr[q_F(D, r^*) \leq x] \leq \frac{\Pr[q_F(D, r^*) \leq x]}{\Pr[q_F(D, r^*) = OPT_F(D)]}$$

Replace x

$$\leq \frac{|R| \exp\left(\frac{\varepsilon x}{2\Delta_q}\right)}{|R_{OPT}| \exp\left(\frac{\varepsilon OPT_F(D)}{2\Delta_q}\right)} = \left(\frac{|R|}{|R_{OPT}|}\right) \exp\left(-\log\left(\frac{|R|}{|R_{OPT}|}\right) - t\right)$$

$$= \left(\frac{|R|}{|R_{OPT}|}\right)\left(\frac{|R_{OPT}|}{|R|}\right) e^{-t} = e^{-t}$$

Utility of Exponential Mechanism

- **Theorem:**

$$\Pr\left[q_F(D,r^*) \le OPT_F(D) - \frac{2\Delta}{\varepsilon}\left(\log\left(\frac{|R|}{|R_{OPT}|}\right) + t\right)\right] \le e^{-t}$$

$R_{OPT} \ge 1$ by definition

- **Corollary:**

$$\Pr\left[q_F(D,r^*) \le OPT_F(D) - \frac{2\Delta}{\varepsilon}(\log(|R|) + t)\right] \le e^{-t}$$

Exponential Mechanism - Examples

- "What is the most common nationality?"
  - Suppose there are 4 nationalities
  - R = {Chinese, Indian, American, Greek}
  - $|R| = 4$
- $q_F(D, nationality)$ = # people in D having that nationality
  - Sensitivity of q_F is 1.
- $OPT_F(D)$ = nationality with the max score

$$\Pr\left[q_F(D, r^*) \leq OPT_F(D) - \frac{2\Delta}{\varepsilon}(\log(|R|) + t)\right] \leq e^{-t}$$

- Exponential mechanism will output some nationality that is shared by at least K people with probability $1 - e^{-3}(= 0.95)$
- $K \geq OPT_F(D) - 2(\log(4) + 3)/\varepsilon = OPT_F(D) - 6.8/\varepsilon$

Exponential Mechanism - Examples

- "What is the most common eye color in this room?"
  - R={Red, Blue, Green, Brown, Purple}

- $K \geq OPT_F(D) - \frac{2(\log 5+3)}{\varepsilon} < OPT_F(D) - 7.4\epsilon$
  - With probability $1 - e^{-3} (= 0.95)$
- *Independent* of the number of people in the room
- Very small error if n is large

Summary

- Differential privacy:
  - Strong adversary (who may know exact information about all but one individual in the data)
  - Adversary can't distinguish between two worlds with different values for an individual (or if an individual is in the table or not)
  - Satisfies composability
- Adding noise from a Laplace distribution guarantees differential privacy
- Exponential mechanism can be used to ensure differential privacy when range of algorithm is not a real number
- Every differentially private algorithm is captured by exponential mechanism
  - By choosing the appropriate score function

References

- L. Sweeney. K-anonymity: A model for protecting privacy. Int. J. Uncertain. Fuzziness Knowl.-Based Syst., vol. 10, no. 5, pp. 557-570, Oct. 2002
- A. Machanavajjhala, D. Kifer, J. Gehrke, and M. Venkitasubramaniam. L-diversity: Privacy beyond k-anonymity. ACM Trans. Knowl. Discov. Data, vol. 1, no. 1, Mar. 2007
- N. Li and T. Li. t-closeness: Privacy beyond k-anonymity and l-diversity. IEEE 23rd Intl Conf. on Data Engineering (ICDE), 2007
- C. Dwork. Differential Privacy. ICALP, 2006
- C. Dwork. A firm foundation for private data analysis. CACM, 2010
- F. McSherry and K. Talwar. Mechanism design via differential privacy. 48th Annual Symposium on Foundations of Computer Science, 2007
- A. Machanavajjhala. Privacy in a Mobile-Social World, Duke, Fall 2012
- Aaron Roth. The Algorithmic Foundations of Data Privacy, University of Pennsylvania, Fall 2011