

# Arbitrages statistiques dans l'apprentissage automatique confidentiel

Soutenance de stage

Alexi CANESSE sous la supervision d'Aurélien GARIVIER, Professeur,  
UMPA et École Normale Supérieure de Lyon

Stage de recherche effectué à l'UMPA dans le cadre de la  
L3 informatique fondamental de l'ÉNS de Lyon

2 août 2022

# Table des matières

- 1 Introduction
- 2 Méthode des histogrammes
- 3 Le mécanisme de sensibilité inverse
- 4 Comparaison entre le mécanisme de sensibilité inverse et la méthode des histogrammes
- 5 Conclusion
- 6 Références

# De l'importance de respecter la confidentialité

*“79% of adults assert they are very or somewhat concerned about how companies are using the data they collect about them, while 64% say they have the same level of concern about government data collection”*

*“a majority think the potential risks of data collection outweigh the benefits”*

Brooke AUXIER et al. *Americans and Privacy : Concerned, Confused and Feeling Lack of Control Over Their Personal Information*. 15 nov. 2019. URL : [https://www.pewresearch.org/internet/wp-content/uploads/sites/9/2019/11/Pew-Research-Center\\_PI\\_2019.11.15\\_Privacy\\_FINAL.pdf](https://www.pewresearch.org/internet/wp-content/uploads/sites/9/2019/11/Pew-Research-Center_PI_2019.11.15_Privacy_FINAL.pdf) (visit  le 20/07/2022)

# Anonymiser les données n'est pas suffisant



# Anonymiser les données n'est pas suffisant

European PARLIAMENT et Council of the EUROPEAN UNION. *Regulation on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (Data Protection Directive)*. 14 avr. 2016. URL : <https://gdpr-info.eu/recitals/no-26/> (visité le 20/07/2022)

# Anonymiser les données n'est pas suffisant

European PARLIAMENT et Council of the EUROPEAN UNION. *Regulation on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (Data Protection Directive)*. 14 avr. 2016. URL : <https://gdpr-info.eu/recitals/no-26/> (visité le 20/07/2022)

Yves-Alexandre de MONTJOYE et al. « Unique in the Crowd : The privacy bounds of human mobility ». In : *Nature* 3 (mars 2013). URL : <https://doi.org/10.1038/srep01376>

# Anonymiser les données n'est pas suffisant

European PARLIAMENT et Council of the EUROPEAN UNION. *Regulation on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (Data Protection Directive)*. 14 avr. 2016. URL : <https://gdpr-info.eu/recitals/no-26/> (visité le 20/07/2022)

Yves-Alexandre de MONTJOYE et al. « Unique in the Crowd : The privacy bounds of human mobility ». In : *Nature* 3 (mars 2013). URL : <https://doi.org/10.1038/srep01376>

Latanya SWEENEY. « Simple Demographics Often Identify People Uniquely ». In : (jan. 2000). DOI : 10.1184/R1/6625769.v1. URL : <https://dataprivacylab.org/projects/identifiability/paper1.pdf> (visité le 20/07/2022)

## *Background* essentiel sur la *differential privacy* (1/4)

La *differential privacy* [Dwo+06] quantifie la perte de confidentialité subit par un individu en participant à une étude.



## Background essentiel sur la *differential privacy* (1/4)

La *differential privacy* [Dwo+06] quantifie la perte de confidentialité subit par un individu en participant à une étude.

### Définition (Jeu de données voisins)

On dit que deux jeux de données  $x$  et  $y$  sont voisins et on note  $d_{\text{Ham}}(x, y) \leq 1$  s'ils diffèrent sur au plus une entrée *ie* la distance de HAMMING qui les sépare et majorée par 1.

## Background essentiel sur la *differential privacy* (1/4)

La *differential privacy* [Dwo+06] quantifie la perte de confidentialité subit par un individu en participant à une étude.

### Définition (Jeu de données voisins)

On dit que deux jeux de données  $x$  et  $y$  sont voisins et on note  $d_{\text{Ham}}(x, y) \leq 1$  s'ils diffèrent sur au plus une entrée ie la distance de HAMMING qui les sépare et majorée par 1.

### Définition (Differential privacy)

On dit qu'un mécanisme aléatoire  $\mathcal{M} : \mathcal{X}^{(\mathbb{N})} \rightarrow \mathcal{T}$  est  $(\varepsilon, \delta)$ -**differentially private** si pour tout  $\mathcal{S} \subset \mathcal{T}$  mesurable,

$$\forall x, y \in \mathcal{X}^{(\mathbb{N})} \quad d_{\text{Ham}}(x, y) \leq 1 \quad \Rightarrow \quad \mathbb{P}(\mathcal{M}(x) \in \mathcal{S}) \leq \exp(\varepsilon) \mathbb{P}(\mathcal{M}(y) \in \mathcal{S}) + \delta.$$

De plus, si  $\delta = 0$ , on dit que  $\mathcal{M}$  est  $\varepsilon$ -**differentially private**.

---

a. La notation  $\mathcal{X}^{(\mathbb{N})}$  désigne, de manière usuelle, l'ensemble des suites finies de  $\mathcal{X}$ .

## Background essentiel sur la *differential privacy* (2/4)

### Définition (Sensibilité d'une requête)

Soit  $\mathcal{X}$  un ensemble,  $(\mathcal{T}, \mathcal{N})$  un espace mesuré et  $f: \mathcal{X}^{(\mathbb{N})} \rightarrow \mathcal{T}$  une requête. On appelle **sensibilité de  $f$**  la grandeur  $\Delta f$  que l'on définit de la manière suivante :

$$\Delta f = \sup_{x, y \in \mathcal{X}^{(\mathbb{N})}} \{\mathcal{N}(f(x), f(y)) \mid d_{\text{Ham}}(x, y) = 1\}.$$

## Background essentiel sur la differential privacy (2/4)

### Définition (Sensibilité d'une requête)

Soit  $\mathcal{X}$  un ensemble,  $(\mathcal{T}, \mathcal{N})$  un espace mesuré et  $f: \mathcal{X}^{(\mathbb{N})} \rightarrow \mathcal{T}$  une requête. On appel **sensibilité de  $f$**  la grandeur  $\Delta f$  que l'on définit de la manière suivante :

$$\Delta f = \sup_{x, y \in \mathcal{X}^{(\mathbb{N})}} \{ \mathcal{N}(f(x), f(y)) \mid d_{\text{Ham}}(x, y) = 1 \}.$$

### Définition (Mécanisme de LAPLACE [Dwo+06])

Soit  $\mathcal{X}$  un ensemble de base,  $\varepsilon \in \mathbb{R}_+^*$ ,  $n \in \mathbb{N}$  et  $f: \mathcal{X}^{(\mathbb{N})} \rightarrow \mathbb{R}$  une requête. Notons Lap la fonction qui retourne une variable aléatoire suivant la loi de LAPLACE dont la paramètre est l'argument. On appel **mécanisme de LAPLACE** la fonction

$$\mathcal{M}_{f, \varepsilon} : \begin{cases} \mathcal{X}^{(\mathbb{N})} & \rightarrow & \mathbb{R} \\ x & \mapsto & f(x) + \text{Lap}(\Delta f / \varepsilon) \end{cases}$$

---

a. Il s'agit de la loi de densité  $x \mapsto 1/(2b) \exp(-|x|/b)$  où  $b$  est le paramètre.

## Background essentiel sur la *differential privacy* (3/4)

### Définition (Mécanisme de LAPLACE [Dwo+06])

Soit  $\mathcal{X}$  un ensemble de base,  $\varepsilon \in \mathbb{R}_+^*$ ,  $n \in \mathbb{N}$  et  $f: \mathcal{X}^{(n)} \rightarrow \mathbb{R}$  une requête. Notons  $\text{Lap}$  la fonction qui retourne une variable aléatoire suivant la loi de LAPLACE dont la paramètre est l'argument. On appelle **mécanisme de LAPLACE** la fonction

$$\mathcal{M}_{f,\varepsilon} : \begin{cases} \mathcal{X}^{(n)} & \rightarrow \mathbb{R} \\ x & \mapsto f(x) + \text{Lap}(\Delta f/\varepsilon) \end{cases}$$

### Théorème

*Le mécanisme de LAPLACE est differentially private.*

## Background essentiel sur la *differential privacy* (4/4)

### Théorème (Théorème de composition (simple) [Dwo+06])

Soit  $\mathcal{X}$  un ensemble de base,  $n \in \mathbb{N}$  un nombre de mécanismes,  $(\mathcal{T}_i)_{i \leq n}$  des ensembles d'arrivé et  $(\mathcal{M}_i : \mathcal{X}^{(\mathbb{N})} \rightarrow \mathcal{T}_i)_{i \leq n}$  des mécanismes mutuellement indépendants respectivement  $\varepsilon_i$ -differentially private. L'exécution des  $n$  mécanismes est  $\left(\sum_{i=1}^n \varepsilon_i\right)$ -differentially private.

## Background essentiel sur la differential privacy (4/4)

### Théorème (Théorème de composition (simple) [Dwo+06])

Soit  $\mathcal{X}$  un ensemble de base,  $n \in \mathbb{N}$  un nombre de mécanismes,  $(\mathcal{T}_i)_{i \leq n}$  des ensembles d'arrivé et  $(\mathcal{M}_i : \mathcal{X}^{(\mathbb{N})} \rightarrow \mathcal{T}_i)_{i \leq n}$  des mécanismes mutuellement indépendants respectivement  $\varepsilon_i$ -differentially private. L'exécution des  $n$  mécanismes est  $\left(\sum_{i=1}^n \varepsilon_i\right)$ -differentially private.

### Démonstration.

Considérons  $(\mathcal{S}_i)_{i \leq n} \subset \prod_{i=1}^n \mathcal{T}_i$  et  $x, x' \in \mathcal{X}^{(\mathbb{N})}$  tel que  $d_{\text{Ham}}(x, x') \leq 1$ .

$$\begin{aligned} \mathbb{P}(\mathcal{M}_1(x) \in \mathcal{S}_1 \wedge \mathcal{M}_2(x) \in \mathcal{S}_2 \wedge \dots) &\stackrel{\text{indé.}}{=} \prod_{i=1}^n \mathbb{P}(\mathcal{M}_i(x) \in \mathcal{S}_i) \\ &\stackrel{\text{DP}}{\leq} \prod_{i=1}^n e^{\varepsilon_i} \mathbb{P}(\mathcal{M}_i(x') \in \mathcal{S}_i) \\ &\stackrel{\text{indé.}}{=} e^{\sum_{i=1}^n \varepsilon_i} \mathbb{P}(\mathcal{M}_1(x') \in \mathcal{S}_1 \wedge \dots) \end{aligned}$$



# L'échec de la méthode naïve

Soit  $x \in [0, 1]^n$ . On a  $\Delta_{\text{méd}} = 1$ .

0	0
0	0
1	0
1	1
1	1



# L'échec de la méthode naïve

Soit  $x \in [0, 1]^n$ . On a  $\Delta_{\text{méd}} = 1$ .

Or,

$$|\text{méd}(x) - \mathcal{M}_{\text{méd}, \varepsilon}| = |\text{Lap}(1/\varepsilon)|.$$

# L'échec de la méthode naïve

Soit  $x \in [0, 1]^n$ . On a  $\Delta_{\text{méd}} = 1$ .

Or,

$$|\text{méd}(x) - \mathcal{M}_{\text{méd}, \varepsilon}| = |\text{Lap}(1/\varepsilon)|.$$

D'où,

$$\mathbb{E} (|\text{méd}(x) - \mathcal{M}_{f, \varepsilon}|) = \int_{\mathbb{R}} \mathbb{P} (|\text{Lap}(1/\varepsilon)| > t) \, dt = \int_{\mathbb{R}_+} e^{-\varepsilon t} dt = \frac{1}{\varepsilon}.$$

# L'échec de la méthode naïve

Soit  $x \in [0, 1]^n$ . On a  $\Delta_{\text{méd}} = 1$ .

Or,

$$|\text{méd}(x) - \mathcal{M}_{\text{méd}, \varepsilon}| = |\text{Lap}(1/\varepsilon)|.$$

D'où,

$$\mathbb{E} (|\text{méd}(x) - \mathcal{M}_{f, \varepsilon}|) = \int_{\mathbb{R}} \mathbb{P} (|\text{Lap}(1/\varepsilon)| > t) dt = \int_{\mathbb{R}_+} e^{-\varepsilon t} dt = \frac{1}{\varepsilon}.$$

Néanmoins nous avons forcément  $\text{méd}(x) \in [0, 1]$ . **Cette méthode ne peut donc pas convenir.**

# Table des matières

- 1 Introduction
- 2 Méthode des histogrammes
- 3 Le mécanisme de sensibilité inverse
- 4 Comparaison entre le mécanisme de sensibilité inverse et la méthode des histogrammes
- 5 Conclusion
- 6 Références

# AboveThreshold [Dwo+06]

```
1 AboveThreshold(database, queries, threshold, epsilon){
2   Assert("les requêtes sont toutes de sensibilité 1");
3   result = 0;
4   noisyThreshold = threshold + Lap(2/epsilon);
5   for(querie in queries){
6     nu = Lap(4/epsilon);
7     if(querie(D) + nu > noisyThreshold)
8       return result;
9     else
10      ++result;
11  }
12  return -1; /* Aucune requête n'a dépassé le seuil */
13 }
```

## Théorème

*AboveThreshold est  $\epsilon$ -differentially private.*

# Présentation de la méthode des histogrammes

```
1 HistogramMethod(database, epsilon, a, b){
2   steps = 1.5*n/log(n);
3   epsilon /= 9; /* composition theorem */
4   result = {};
5   for(d in {1 ... 9}){ /* which decile */
6     T = d*card(database)/10;
7     for(i in {1 ... steps}){
8       fi = x -> card({element in x | element < i*(b-a)
9                                     /steps});
10      queries.push_back(fi);
11    }
12    T = d*card(database)/10;
13    result.push_back(AboveThreshold(database, queries, T,
14                                     epsilon)*(b-a)/steps));
15  }
16  return result;
17 }
```

# Analyse de précision - le cas de la loi uniforme standard (1/5)

## Définition (Fonction Beta incomplète (régularisée))

On appelle respectivement fonction beta incomplète et fonction beta incomplète régularisée les fonctions

$$\begin{aligned} B : \begin{cases} [0, 1] \times (\mathbb{R}_+^*)^2 & \rightarrow \mathbb{R}_+ \\ (x, \alpha, \beta) & \mapsto \int_0^x t^{\alpha-1} (1-t)^{\beta-1} dt \end{cases} \\ I_{\bullet} : \begin{cases} (\mathbb{R}_+^*)^2 & \rightarrow \mathbb{R}_+ \\ (\alpha, \beta) & \mapsto \frac{B(\bullet, \alpha, \beta)}{B(1, \alpha, \beta)} \end{cases} . \end{aligned}$$

# Analyse de précision - le cas de la loi uniforme standard (1/5)

## Définition (Fonction Beta incomplète (régularisée))

On appelle respectivement fonction beta incomplète et fonction beta incomplète régularisée les fonctions

$$B : (x, \alpha, \beta) \mapsto \int_0^x t^{\alpha-1} (1-t)^{\beta-1} dt \quad \text{et} \quad I_{\bullet} : (\alpha, \beta) \mapsto \frac{B(\bullet, \alpha, \beta)}{B(1, \alpha, \beta)}.$$

## Définition (Loi beta)

On appelle loi beta de paramètre  $(\alpha, \beta) \in \mathbb{R}_+^*$  la loi de densité

$$f_{\alpha, \beta} : [0, 1] \ni x \mapsto \frac{x^{\alpha-1} (1-x)^{\beta-1}}{B(1, \alpha, \beta)}$$



# Analyse de précision - le cas de la loi uniforme standard (2/5)

## Définition (Statistique d'ordre)

Soit  $X$  un échantillon statistique de cardinal  $n \in \mathbb{N}$ . Pour tout  $k \in \llbracket 1, n \rrbracket$  on note  $X_{(i)}$  et on appelle **statistique d'ordre** de rang  $k$  la  $k$ -ème plus petite valeur de l'échantillon.

## Théorème (Loi des statistiques d'ordre d'un échantillon issue de $\mathcal{U}(0, 1)$ )

*Soit  $X$  un ensemble de  $n$  variables aléatoires  $(X_i)_i$  indépendantes et suivant toutes la loi uniforme sur  $[0, 1]$  et  $k \in \llbracket 1, n \rrbracket$ . La  $k$ -ème statistique d'ordre de  $X$ ,  $X_{(k)}$ , est distribuée suivant la loi beta de paramètre  $(k, n - k + 1)$ .*

# Analyse de précision - le cas de la loi uniforme standard (3/5)

## Théorème (Précision moyenne de HistogramMethod)

Soit  $X$  un ensemble de  $n$  (tel que  $8 \log(3n/(\beta \log n))/\varepsilon) \leq n/10$ ) variables aléatoires  $(X_i)_i$  indépendantes et suivant toutes la loi uniforme standard. Soit  $i \in \llbracket 1, 9 \rrbracket$ . Notons  $(d_i^l)_i$  les déciles de la loi. Posons  $A$  la variable aléatoire  $\text{HistogramMethod}(X, \text{epsilon}, 0, 1)$  et  $\alpha = 8 \log(3n\sqrt{n})/\varepsilon$ . On a

$$\begin{aligned} \mathbb{E}(|A_i - d_i^l|) &\leq \int_0^{d_i^l} \left(1 - I_{d_i^l+t}(\lceil in/10 + \alpha \rceil, n - \lceil in/10 + \alpha \rceil + 1)\right) dt \\ &\quad + \int_0^{d_i^l} I_{d_i^l-t}(\lfloor in/10 - \alpha \rfloor, n - \lfloor in/10 - \alpha \rfloor + 1) dt + \beta \\ &\quad + I_{d_i^l-0.1}(\lfloor in/10 - \alpha \rfloor, n - \lfloor in/10 - \alpha \rfloor + 1) \\ &\quad + I_{1-d_i^l-0.1}(n - \lceil in/10 + \alpha \rceil + 1, \lceil in/10 + \alpha \rceil) \\ &\quad + \frac{2 \log n}{3n} + d_i^l \beta. \end{aligned}$$

# Analyse de précision - le cas de la loi uniforme standard (4/5)

## Corollaire ((im)Précision moyenne de HistogramMethod)

Soit  $X$  un ensemble de  $n$  (tel que  $8 \log(3n\sqrt{n})/\varepsilon) \leq n/10$ ) variables aléatoires  $(X_i)_i$  indépendantes et suivant toutes la loi uniforme standard. Soit  $i \in \llbracket 1, 9 \rrbracket$ . Notons  $(d_i^l)_i$  les déciles de la loi. Posons  $A$  la variable aléatoire  $\text{HistogramMethod}(X, \text{epsilon}, 0, 1)$ ,  $\alpha = 8 \log(3n\sqrt{n})/\varepsilon$ . On a

$$\begin{aligned} \mathbb{E}(|A_i - d_i^l|) &\leq \int_0^{d_i^l} \left(1 - I_{d_i^l+t}(\lceil in/10 + \alpha \rceil, n - \lceil in/10 + \alpha \rceil + 1)\right) dt \\ &\quad + \int_0^{d_i^l} I_{d_i^l-t}(\lfloor in/10 - \alpha \rfloor, n - \lfloor in/10 - \alpha \rfloor + 1) dt + \frac{1}{\sqrt{n} \log n} \\ &\quad + I_{d_i^l-0.1}(\lfloor in/10 - \alpha \rfloor, n - \lfloor in/10 - \alpha \rfloor + 1) \\ &\quad + I_{1-d_i^l-0.1}(n - \lceil in/10 + \alpha \rceil + 1, \lceil in/10 + \alpha \rceil) \\ &\quad + \frac{2 \log n}{3n} + \frac{d_i^l}{\sqrt{n} \log n}. \end{aligned}$$

# Analyse de précision - le cas de la loi uniforme standard (5/5)

## Théorème ((im)Précision moyenne de HistogramMethod)

Soit  $X$  un ensemble de  $n$  (tel que  $0 \leq 8 \log(3n\sqrt{n})/\varepsilon) \leq n/10$ ) variables aléatoires  $(X_i)_i$  indépendantes et suivant toutes la loi uniforme sur  $[0,1]$ . Soit  $i \in \llbracket 1, 9 \rrbracket$  et  $k \in \mathbb{N}$ . Notons  $(d_i)_i$  les décile de la loi. Posons  $A$  la variable aléatoire  $\text{HistogramMethod}(X, \text{epsilon}, 0, 1)$  et  $\alpha = 8 \log(3n\sqrt{n})/\varepsilon$ . On a,

$$\mathbb{E}(|A_i - d_i^l|) \leq 2\sqrt{\frac{\pi}{2n}} + \frac{d_i^l + 1}{\sqrt{n} \log n} + \frac{\log n}{n} \left( \frac{2}{3} + \frac{16}{\varepsilon} \log(3) \right) \\ + 2 \exp \left( -2n \left( 0.1 - \frac{\alpha}{n} \right)^2 \right).$$

# Table des matières

- 1 Introduction
- 2 Méthode des histogrammes
- 3 **Le mécanisme de sensibilité inverse**
- 4 Comparaison entre le mécanisme de sensibilité inverse et la méthode des histogrammes
- 5 Conclusion
- 6 Références

# Présentation du mécanisme (1/2)

## Définition (Longueur)

Soit  $x \in \mathcal{X}^{(\mathbb{N})}$ ,  $f: \mathcal{X}^{(\mathbb{N})} \rightarrow \mathcal{T}$  et  $t \in \mathcal{T}$ . La longueur est le nombre minimum de valeurs à modifier dans  $x$  pour obtenir  $x'$  tel que  $f(x') = t$ . On a

$$\text{len}_f(x, t) := \inf_{x' \in \mathcal{X}^{(\mathbb{N})}} \{ \|x - x'\|_1 \mid f(x') = t \}.$$

## Définition (Mécanisme de sensibilité inverse [AD20])

Soit  $f: \mathcal{X}^{(\mathbb{N})} \rightarrow \mathcal{T}$  et  $\varepsilon \in \mathbb{R}_+$ . Pour une mesure  $\mu$  sur  $\mathcal{T}$ , on définit le mécanisme aléatoire  $\mathcal{M}(x)$  par sa fonction de densité

$$t \mapsto \frac{\exp(-\text{len}_f(x, t)\varepsilon/2)}{\int_{\mathcal{T}} \exp(-\text{len}_f(x, s)\varepsilon/2) d\mu(s)}.$$

# Présentation du mécanisme (2/2)

## Définition (Longueur lisse)

Soit  $x \in \mathcal{X}^{(\mathbb{N})}$ ,  $f: \mathcal{X}^{(\mathbb{N})} \rightarrow \mathcal{T}$  et  $\rho \in \mathbb{R}_+$ . Si  $\mathcal{N}$  est une norme sur  $\mathcal{T}$ ,

$$\text{len}_f^\rho: \begin{cases} \mathcal{T} & \rightarrow \mathbb{N} \\ t & \mapsto \inf_{s \in \mathcal{T}, \mathcal{N}(s,t) \leq \rho} \{\text{len}_f(x, s)\} \end{cases}.$$

## Définition (Mécanisme de sensibilité inverse $\rho$ -lisse [AD20])

Soit  $f: \mathcal{X}^{(\mathbb{N})} \rightarrow \mathcal{T}$  et  $\rho, \varepsilon \in \mathbb{R}_+$ . Pour une mesure  $\mu$  sur  $\mathcal{T}$ , on définit le mécanisme aléatoire  $\mathcal{M}_{\text{cont}}(x)$  par sa fonction de densité

$$t \mapsto \frac{\exp(-\text{len}_f^\rho(x, t)\varepsilon/2)}{\int_{\mathcal{T}} \exp(-\text{len}_f^\rho(x, s)\varepsilon/2) d\mu(s)}.$$

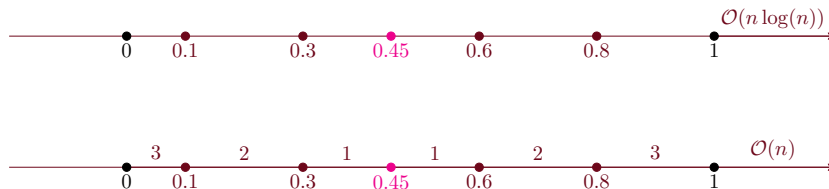
# Analyse de complexité



La complexité de l'algorithme est donc au pire en  $\mathcal{O}(n \log(n))$ .

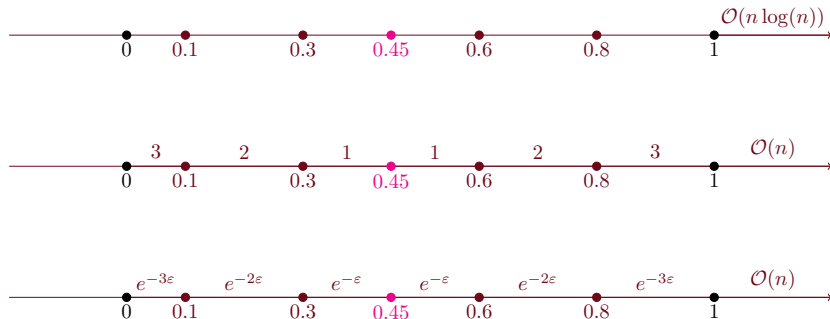


# Analyse de complexité



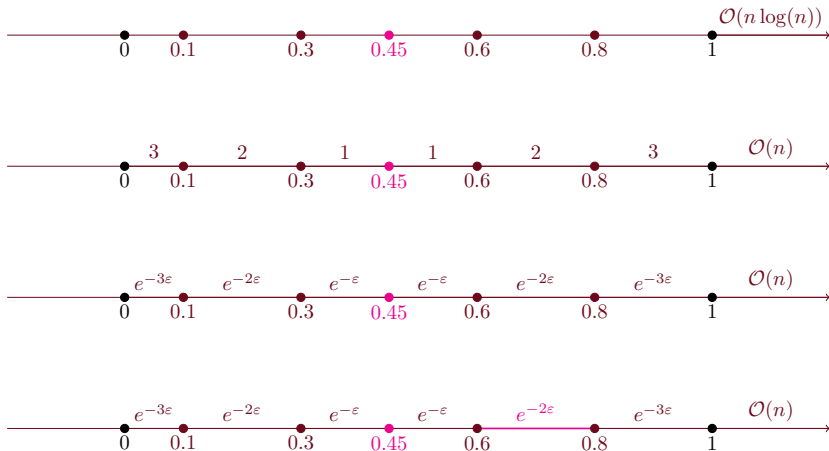
La complexité de l'algorithme est donc au pire en  $\mathcal{O}(n \log(n))$ .

# Analyse de complexité



La complexité de l'algorithme est donc au pire en  $\mathcal{O}(n \log(n))$ .

# Analyse de complexité



La complexité de l'algorithme est donc au pire en  $\mathcal{O}(n \log(n))$ .

# Précision pour l'estimation de déciles

# Table des matières

- 1 Introduction
- 2 Méthode des histogrammes
- 3 Le mécanisme de sensibilité inverse
- 4 Comparaison entre le mécanisme de sensibilité inverse et la méthode des histogrammes
- 5 Conclusion
- 6 Références

# Comparaison des bornes obtenues

# Résultats expérimentaux

# Table des matières

- 1 Introduction
- 2 Méthode des histogrammes
- 3 Le mécanisme de sensibilité inverse
- 4 Comparaison entre le mécanisme de sensibilité inverse et la méthode des histogrammes
- 5 Conclusion**
- 6 Références



# Conclusion

# Table des matières

- 1 Introduction
- 2 Méthode des histogrammes
- 3 Le mécanisme de sensibilité inverse
- 4 Comparaison entre le mécanisme de sensibilité inverse et la méthode des histogrammes
- 5 Conclusion
- 6 **Références**

# Référence

- [AD20] Hilal ASI et John C. DUCHI. « Near Instance-Optimality in Differential Privacy ». In : *ArXiv abs/2005.10630* (mai 2020). URL : <https://arxiv.org/pdf/2005.10630.pdf>.
- [Aux+19] Brooke AUXIER et al. *Americans and Privacy : Concerned, Confused and Feeling Lack of Control Over Their Personal Information*. 15 nov. 2019. URL : [https://www.pewresearch.org/internet/wp-content/uploads/sites/9/2019/11/Pew-Research-Center\\_PI\\_2019.11.15\\_Privacy\\_FINAL.pdf](https://www.pewresearch.org/internet/wp-content/uploads/sites/9/2019/11/Pew-Research-Center_PI_2019.11.15_Privacy_FINAL.pdf) (visité le 20/07/2022).
- [Dwo+06] Cynthia DWORK et al. « Calibrating Noise to Sensitivity in Private Data Analysis ». In : *Theory of Cryptography*. Sous la dir. de Shai HALEVI et Tal RABIN. Berlin, Heidelberg : Springer Berlin Heidelberg, 2006, p. 265-284. ISBN : 978-3-540-32732-5.
- [Mon+13] Yves-Alexandre de MONTJOYE et al. « Unique in the Crowd : The privacy bounds of human mobility ». In : *Nature* 3 (mars 2013). URL : <https://doi.org/10.1038/srep01376>.
- [PE16] European PARLIAMENT et Council of the EUROPEAN UNION. *Regulation on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (Data Protection Directive)*. 14 avr. 2016. URL : <https://gdpr-info.eu/recitals/no-26/> (visité le 20/07/2022).
- [Swe00] Latanya SWEENEY. « Simple Demographics Often Identify People Uniquely ». In : (jan. 2000). DOI : [10.1184/R1/6625769.v1](https://doi.org/10.1184/R1/6625769.v1). URL :