

Differential Privacy

1

On note X le domaine (abstrait) des données
Fini

Une base de donnée $D \in X^n$ est un n -uplet d'éléments de X
ensemble

Il sera parfois pratique de représenter D par un histogramme

$$D \in \mathbb{N}^{|X|}$$

Si $X = \{v_1, \dots, v_k\}$, $D_k = |\{x \in D : x = v_k\}| \quad \forall k \in \{1, \dots, k\}$

En particulier, la taille de jauge de données est $n = \|D\|_1 = \sum_{k=1}^{|X|} D_k$

Si D, D' tels que $\|D - D'\|_1 \leq 1$, D et D' ne diffèrent que d'un plus un élément (On dit que D et D' sont voisins)

Def: (Algorithmme randomisé) ————— output

Un algorithme randomisé est une application $A: \mathbb{N}^{|X|} \rightarrow \mathcal{O}$,
où \mathcal{O} est un espace probabilisé

Pg: $\forall D \in \mathbb{N}^{|X|}$, $A(D)$ est une v.a. à valeurs dans \mathcal{O}

confidentialité différentielle d'un algorithme est sa sensibilité stochastique par rapport aux données:

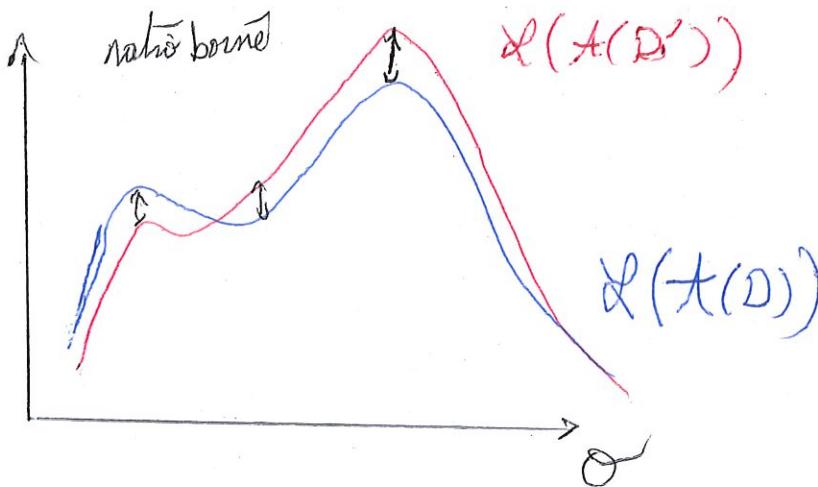
Def: (Differential privacy (DP))

Pour $\epsilon \geq 0$ et $0 < \delta < 1$, on dit que A est (ϵ, δ) -différentiellement confidentiel si $\forall D, D' \in \mathcal{N}^{|\mathcal{X}|}$ tels que $\|D - D'\|_1 \leq 1$

$$\forall S \subset \mathcal{O} \text{ mesurable}$$

$$P(A(D) \in S) \leq e^{\epsilon} P(A(D') \in S) + \delta,$$

où la probabilité est prise par rapport à l'algorithme A



• Ici, les données sont fixées

- La DP est une propriété de l'algorithme, pas des données
- Pour être non-triviale, un algo DP doit être randomisé (Sinon $A(D) = 0$)

(3)

⑩ $\delta = 0$)

• $(\epsilon, 0)$ -DP est appelé pure ϵ -DP

• $(\epsilon, 0)$ -DP garantit que, à chaque exécution de $A(D)$, la sortie presque aussi raisonnable d'être observée que pour n'importe quel jeu de données voisin.

• Si \mathcal{O} est fini, le rapport de raisonnabilités

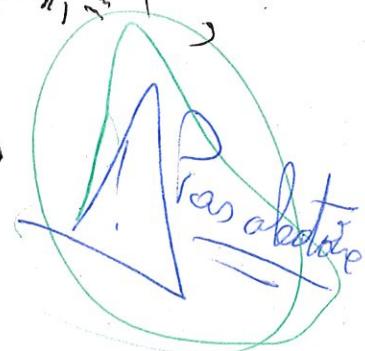
$$L_{A(D), A(D')}^{(\theta)} = \log \left(\frac{P(A(D) = \theta)}{P(A(D') = \theta)} \right)$$

Rapport de
raisonnalité

est appelé privacy loss

Ex: • (ϵ, δ) -DP assure que $\forall \theta \in \mathcal{O}, \|D - D'\| \leq 1$,

$$\boxed{P(L_{A(D), A(D')}^{(\theta)} \leq \epsilon) \geq 1 - \delta}$$



• En pratique:

$\hookrightarrow \epsilon = 1$ ($e^1 \approx 2.7$) et considéré comme raisonnable

$\hookrightarrow \epsilon = 0.1$ ($e^{0.1} \approx 1.1$) une garantie forte

- la DP est intrinsèquement robuste à des connaissances extérieures quelconques

↳ De tout D sauf une donnée !

- L'algorithme A peut être public : seul son algorithme doit rester secret

↳ Permet de débattre publiquement de A et de ses garanties

Thm: (Postprocessing)

Soit $A: \mathbb{N}^{|D|} \rightarrow \mathcal{O}$ un algo (ϵ, δ) -DP, et $f: \mathcal{O} \rightarrow \mathcal{O}'$ une fonction (randomisée) quelconque.

Alors, $f \circ A: \mathbb{N}^{|D|} \rightarrow \mathcal{O}'$ est (ϵ, δ) -DP

Dem: Soient $\|D - D'\|_1 \leq 1$, et $S' \subset \mathcal{O}'$

En notant $S = \{o \in \mathcal{O} \mid f(o) \in S'\}$, on a

$$\begin{aligned} P(f(A(D)) \in S' \mid f) &= P(A(D) \in S \mid f) \\ &\leq e^\epsilon P(A(D') \in S \mid f) + \delta \\ &= e^\epsilon P(f(A(D')) \in S' \mid f) + \delta \end{aligned}$$

f true comme déterministe

En intégrant par rapport à l'algorithme f , on obtient le résultat.

De façon similaire, on contrôle aisément le degré de confidentialité différentielle lorsque plusieurs analyses sont faites d'un même jeu de données

Prop. (Composition simple)

Si A_1, \dots, A_K sont des algorithmes (ϵ_k, δ_k) -DP, alors pour tout $D \in \mathbb{N}^{|\mathcal{X}|}$,

$$A(D) = (A_1(D), \dots, A_K(D))$$

est ϵ -DP avec $\begin{cases} \epsilon = \sum_{k=1}^K \epsilon_k & , \text{ dès que l'aléa des } A_1, \dots, A_K \\ \delta = \sum_{k=1}^K \delta_k & \text{ sont indépendants} \end{cases}$

Dem: Trivial pour $\delta = 0$. Voir Dwork pour le cas général

Rq: Si on compose avec des entrées différentes $(A_1(D_1), A_2(D_2))$ on obtient $(\max \epsilon_k, \max \delta_k)$ -DP

<u>E_D:</u>	Vit dans k93	Ne vit pas dans k93
♂	10	7
♀	13	8

Si chacune des entrées est (ϵ, δ) -DP, alors la publication du tableau tout entier est (ϵ, δ) -DP

(6)

cette définition de confidentialité, on peut aussi s'intéresser à ses conséquences pour des groupes de K individus. En effet, si K individus ont des données très corélées, ou bien qu'un individu unique contribue K fois aux données, ça a de l'intérêt

Prop: (Group-DP)

Tout algorithme (ϵ, δ) -DP A est $(K\epsilon, K e^{K\epsilon}\delta)$ -DP pour les groupes de taille K :

$\forall D, D' \in \mathbb{N}^{|X|}$ tels que $\|D - D'\|_1 \leq K$, $\forall S \subset \mathcal{O}$,

$$P(A(D) \in S) \leq e^{K\epsilon} P(A(D') \in S) + K e^{K\epsilon} \delta$$

Rq: Résultat différent/distinct de la stabilité pour la composition.

en: On note $D_0 = D, D_1, \dots, D_K = D'$ où pour tout $k \in \{0, \dots, K-1\}$

$$\|D_{k+1} - D_k\|_1 = 1$$

$$\begin{aligned}
 \forall S \subset \mathcal{O}, P(A(D_0) \in S) &\leq e^\epsilon P(A(D_1) \in S) + \delta \\
 &\leq e^\epsilon (e^\epsilon P(A(D_2) \in S) + \delta) + \delta \\
 &\quad \vdots \\
 &\leq e^{K\epsilon} P(A(D_K) \in S) + (1 + e^\epsilon + e^{2\epsilon} + \dots + e^{(K-1)\epsilon}) \delta \\
 &\leq e^{K\epsilon} P(A(D') \in S) + K e^{K\epsilon}
 \end{aligned}$$

Inégalité arithmético-géométrique

$$\frac{x_0 + \dots + x_{K-1}}{K} \leq (x_0 - x_{K-1})^{\frac{1}{K}}$$

$$x_0 + \dots + x_{K-1} \leq (x_0 - x_{K-1})^{\frac{1}{K}}$$

Algorithmes DP via perturbation de la sortie

Supposons que l'on souhaite construire/calculer une fonction

$$f: \mathbb{N}^{|X|} \rightarrow \mathbb{R}^K$$

qui dépend de données confidentielles ①

- ↳ Comment le faire en satisfaisant la condition de DP ?
- ↳ Combien ajouter d'incertitude ?

Def: (Sensibilité Δ_1 , globale)

La sensibilité Δ_1 , globale de $f: \mathbb{N}^{|X|} \rightarrow \mathbb{R}^K$ est

$$\Delta_1(f) := \max_{D, D': \|D - D'\| \leq 1} \|f(D) - f(D')\|,$$

dont

Dans \mathbb{R}^k

- $\Delta_1(f)$ traduit combien f est affectée par le changement d'une valeur.
- Elle donne une échelle pour l'incertitude qu'il faut ajouter afin de cacher la contribution de l'individu.

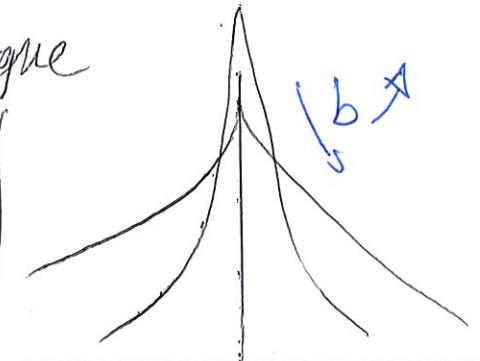
$$\text{Ex: } f(D) = \#\{\text{♂ dans } D\} \quad \Delta_1(f) = 1$$

$$\text{Ex: } f(D) = \text{Salaire moyen} \quad \Delta_1(f) \leq \frac{\text{moye salaire}}{m}$$

(8)

Rappel: la loi de Laplace $\text{Lap}(b)$, $b > 0$ est la loi sur \mathbb{R} ayant pour densité par rapport à Lebesgue

$$\forall y \in \mathbb{R}, P(y, b) = \frac{1}{2b} e^{-\frac{|y|}{b}}$$



Si $Y \sim \text{Lap}(b)$

$$\therefore \mathbb{E}[Y] = 0, \mathbb{E}[|Y|] = b, \mathbb{E}[Y^2] = 2b^2$$

$$\therefore \forall t \geq 0, P(|Y| > tb) \leq e^{-t}$$

Def: (Mechanisme de Laplace)

Pour $f: \mathbb{N}^{|\mathcal{X}|} \rightarrow \mathbb{R}^K$, on note $\mathcal{A}_{\text{lap}}(\Delta, f, \epsilon)$ l'algorithme suivant.

1) Calculer $\Delta = \Delta_1(f)$

2) Pour $k \in \{1, -, K\}$, tirer $Y_k \sim \text{Lap}\left(\frac{\Delta}{\epsilon}\right)$ indépendamment

3) Renvoyer $f(D) + Y$, où $Y = (Y_1, -, Y_K)$

L'idée est simplement de perturber la sortie aléatoirement par un ajout dans chaque composante, à une échelle donnée par la sensibilité Δ ,

Thm: $\forall \epsilon > 0$, $\mathcal{A}_{\text{lap}}(\Delta, f, \epsilon)$ est $(\epsilon, 0)$ -DP

Dom: Soient $\|D - D'\| \leq 1$ et $S \subset \mathbb{R}^K$ mesurable.

On note g et g' les densités de $A_{\text{lap}}(D, f, \varepsilon)$ et $A_{\text{lap}}(D', f, \varepsilon)$ par rapport à $\lambda_{\mathbb{R}^K}$. Ainsi,

$$\frac{\mathbb{P}(A_{\text{lap}}(D, f, \varepsilon) \in S)}{\mathbb{P}(A_{\text{lap}}(D', f, \varepsilon) \in S)} = \frac{\int_S g(s) ds}{\int_S g'(s) ds} \leq \max_{s \in S} \frac{g(s)}{g'(s)}$$

Or, par construction de A_{lap} , si $p : \mathbb{R}^K \rightarrow \mathbb{R}$ désigne la densité de $\text{Lap}\left(\frac{\Delta}{\varepsilon}\right)$, on a $\forall s = (s_1, \dots, s_K) \in S$,

$$g(s) = \prod_{k=1}^K p(s_k - f_k(D)) \text{ et } g'(s) = \prod_{k=1}^K p(s_k - f_k(D'))$$

Ainsi, $\forall s \in S$,

$$\frac{g(s)}{g'(s)} = \frac{\prod_{k=1}^K \exp\left(-\frac{\varepsilon}{\Delta} |s_k - f_k(D)|\right)}{\exp\left(-\frac{\varepsilon}{\Delta} |s_k - f_k(D')|\right)}$$

$$= \exp\left(\frac{\varepsilon}{\Delta} \sum_{k=1}^K \{|s_k - f_k(D')| - |s_k - f_k(D)|\}\right)$$

$$\leq \exp\left(\frac{\varepsilon}{\Delta} \sum_{k=1}^K |f_k(D) - f_k(D')|\right)$$

$= \|f(D) - f(D')\|, \Delta = \overline{\|f(D)\|}$

$$\leq \exp\left(\frac{\varepsilon}{\Delta} \Delta\right) = \exp(\varepsilon) \square$$

De plus, le mécanisme de Laplace reste relativement proche de f

$$\boxed{\text{Prop: } \mathbb{E} \left[\|t_{\text{lap}}(D, f, \epsilon) - f(D)\|_1 \right] \leq K \frac{\Delta_1(f)}{\epsilon}}$$

$$\boxed{\text{Prop: } \forall \beta \in (0, 1], P \left(\|t_{\text{lap}}(D, f, \epsilon) - f(D)\|_\infty \leq \log \left(\frac{K}{\beta} \right) \frac{\Delta_1(f)}{\epsilon} \right) \geq 1 - \beta}$$

Dem: Borne d'union \oplus Borne de queue de la loi de Laplace

Mécanisme Gaussien

Rq: Parfois, il est plus pratique de manipuler des perturbations gaussiennes \rightarrow Somme de gaussiennes est gaussienne
 \rightarrow Même type de biais que les données elles-mêmes

Déf: (Sensibilité Δ_2)

$$\boxed{\text{Pour } f: \mathbb{N}^{|D|} \rightarrow \mathbb{R}^k, \Delta_2(f) = \max_{\|D - D'\|_1 \leq 1} \|f(D) - f(D')\|_2}$$

Déf: (Mécanisme Gaussien) $t_{\text{gauss}}(D, f, \epsilon, \delta)$ est la sortie de

1) Calculer $\Delta = \Delta_2(f)$

2) $\forall i \in \{1, \dots, k\}, Y_i \stackrel{\text{iid}}{\sim} N(0, \sigma^2)$ où $\sigma^2 = \sqrt{2 \log \left(\frac{1.25}{\delta} \right)} \cdot \frac{\Delta}{\epsilon}$

3) Sortir $f(D) + Y$, où $Y = (Y_1, \dots, Y_k)$

(11)

Rq: Similaire à Laplace, mais avec Δ_ϵ au lieu de Δ ,
 { dépendance en ϵ ET δ

Thm: $A_{\text{guar}}(\cdot, f, \epsilon, \delta)$ est (ϵ, δ) -DP

Rq: Pour $\delta \ll 1$, la différence entre ϵ -DP et (ϵ, δ) -DP ne se constate pas en pratique.

Prop: $\bullet P\left(\|A_{\text{guar}}(D, f, \epsilon, \delta) - f(D)\|_\infty < \sqrt{2 \log(\frac{125}{\delta})} \log\left(\frac{K}{\epsilon}\right) \frac{\Delta_\epsilon(f)}{\epsilon}\right) \geq 1 - \beta$

↳ Dépendance $\log(\frac{1}{\delta})$

Rq: Pour $f(D) \in \mathbb{N}$ ces mécanismes produisent des réels, donc sont peu satisfaisants

↳ Possibilité de construire un mécanisme à valeur entière avec des lois géométriques tronquées

Nisme Exponentiel

(12)

↳ les mécanismes précédents ne fonctionnent que si f est numerique!

↳ Ils ne sont précis que pour f assez régulière / stable en D

$$\Delta_1(f), \Delta_2(f) \\ \text{petit}$$

On considère $f: N^{|X|} \rightarrow \mathcal{O}$, où \mathcal{O} est abstrait fini

• Une fonction de score $s: N^{|X|} \times \mathcal{O} \rightarrow \mathbb{R}$,
 $s(D, o)$ représentant à quel point il est satisfaisant de
 retourner o lorsque $f(D)$ est requête.

↳ $o = f(D)$ maximise $o \mapsto s(D, o)$ Dépend du problème

f : (Sensibilité d'un score)

$$\Delta(s) = \max_{\theta \in \mathcal{O}} \max_{\|D - D'\| \leq 1} |s(D, \theta) - s(D', \theta)|$$

Si $f: N^{|X|} \rightarrow \mathbb{R}$ et $s(D, \theta) = |\theta - f(D)|$, $\Delta(s) = \Delta_1(f)$
 $\{\theta = \mathbb{R}$

L'idée du mécanisme exponentiel est de renvoyer $\theta \in \Theta$ avec une probabilité croissante en son score $\Delta(D, \theta)$

↳ Des scores hauts associés exponentiellement à leur vraisemblance

Def: (Mécanisme exponentiel) $A_{\exp}(D; N \xrightarrow{f} \Theta, \Delta; N \xrightarrow{h} \Theta \rightarrow \mathbb{R}, \epsilon)$

1) Calculer $\Delta = \Delta(\Delta)$

2) Renvoyer $\theta \in \Theta$ avec probabilité

$$\frac{\exp\left(\frac{\Delta(D, \theta) \epsilon}{2\Delta}\right)}{\sum_{\theta' \in \Theta} \exp\left(\frac{\Delta(D, \theta') \epsilon}{2\Delta}\right)}$$

$$\sum_{\theta' \in \Theta} \exp\left(\frac{\Delta(D, \theta') \epsilon}{2\Delta}\right)$$

Thm: $A_{\exp}(\cdot, f, \Delta, \epsilon)$ est ϵ -DP

En notant $\Delta^*(D) = \max_{\theta \in \Theta} \Delta(D, \theta)$ et $D^* = \{\theta \in \Theta, \Delta(D, \theta) = \Delta^*(D)\}$

$$P\left(\Delta(A_{\exp}(D, \Delta, \epsilon)) \leq \Delta^*(D) - \frac{2\Delta(\Delta)}{\epsilon} \left(\log\left(\frac{|\Theta|}{|\Theta^*|}\right) + 1 \right)\right) \leq e^{-t}$$

↳ A_{\exp} renvoie app le meilleur score à $-\frac{\Delta(\Delta)}{\epsilon}$ près

↳ les garanties sont d'autant meilleures que $|D^*|$ est grand

Autres formalismes de confidentialité

1) Local Differential Privacy

② Définition

Lorsqu'un tiers de confiance n'est pas disponible, il faut être en mesure de garantir la confidentialité

Def: Un randomiseur local est une fonction randomisée $R: \mathcal{X} \rightarrow \mathcal{Z}$

Def: (Local Differential Privacy)

$R: \mathcal{X} \rightarrow \mathcal{Z}$ est dit (ϵ, δ) - localement différentiellement confidentiel lorsque $\forall x, x' \in \mathcal{X}, \forall y \in \mathcal{Z}$,

$$P(R(x)=y) \leq e^\epsilon P(R(x')=y) + \delta$$

Rq: Équivalent à (ϵ, δ) -DP pour les données de taille 1

, LDP est bien plus contraignant que DP

⑤ Un mécanisme LDP vs DP

Supposons que $X = \{v_1, \dots, v_K\}$

Def: (K-Réponse Randomisée) $R_{RR,K}(x, \epsilon)$

1) Tirer $B \sim \text{Ber}\left(\frac{K}{e^\epsilon + K - 1}\right)$

2) $\begin{cases} \text{Si } B = 0, \text{ renvoyer } x \\ \text{Sinon, renvoyer } y \in \text{Unif}(X) \end{cases}$

↳ Généralisation du tirage pile/face en intro à des variables plus que binaires

Thm: $R_{RR,K}(\cdot, \epsilon)$ est $(\epsilon, 0)$ -LDP

Dém: Si $\begin{cases} x \neq y \\ x' \neq y \end{cases}$ ou $x = x' = y$, $P(R_{RR,K}(x, \epsilon) = y) = P(R_{RR,K}(x', \epsilon) = y)$

$$= \frac{1-K}{e^\epsilon + K - 1}$$

$$\text{Si } \begin{cases} x = y \\ x' \neq y \end{cases}, \quad \left\{ \begin{array}{l} P(R_{RR,K}(x) = y) = \frac{e^\epsilon - 1}{e^\epsilon + K - 1} \\ P(R_{RR,K}(x') = y) = \frac{1}{e^\epsilon + K - 1} \end{array} \right.$$

$$+ \frac{1}{e^\epsilon + K - 1} \times \frac{1}{K} = \frac{e^\epsilon}{e^\epsilon + K - 1}$$

$$P(R_{RR,K}(x) = y) = \frac{1}{e^\epsilon + K - 1}$$

d'où la ratio $\leq e^\epsilon$

Ex: Soit $h = (h_1, \dots, h_k)$ l'histogramme de données confidentielles :

$$h_k = \frac{1}{m} \sum_{i=1}^m \mathbb{1}_{g_i = v_k}$$

LDP

En notant $p = \frac{e^\epsilon - 1}{e^\epsilon + k - 1}$, K-RR permet d'obtenir un estimateur non-biaisé \hat{h}_k de h_k , donné par

$$\hat{h}_k^{\text{LDP}} = \frac{\left(\frac{1}{m} \sum_{i=1}^m \mathbb{1}_{g_i = v_k} \right) - \frac{1-p}{K}}{p - \frac{1-p}{K}} = \frac{(e^\epsilon + k - 1)}{(e^\epsilon - 1)} \left(\frac{1}{m} \sum_{i=1}^m \mathbb{1}_{g_i = v_k} - 1 \right)$$

Prop: $E\left[\left(\hat{h}_k^{\text{LDP}} - h_k\right)^2\right] = \frac{K-2+e^\epsilon}{m(e^\epsilon - 1)^2}$

DP Si $f: X \rightarrow [0, 1]$, la sensibilité de $\bar{f}(x) = \frac{1}{m} \sum_{i=1}^m f(x_i)$

$$\Delta_f(f) = \max_{x \neq x'} |\bar{f}(x) - \bar{f}(x')| = \frac{1}{m}$$

Ainsi, le mécanisme de Laplace ($b = \frac{\epsilon}{m \cdot \Delta_f}$) donne un ϵ -DP avec

$$E\left[\left(\hat{h}_k^{\text{LDP}} - h_k\right)^2\right] \approx \frac{1}{m^2 \epsilon^2}$$

(17)

On constate ici une différence notable de précision / utilité entre ϵ -DP et ϵ -LDP : On perd un facteur $\frac{1}{n}$.

↳ Inévitables ici.

Rq: Cela montre que la LDP n'est utilisable que pour n grand

2) Statistical Queries

E2S

$$\mathcal{F} = \{f: X \rightarrow [0, 1]\}$$

$$\phi(P) = b \text{ où } P = \text{Unif}[0,1]$$

Standard: $\max_i X_i$

$$DP = \max_i |X_i| + \gamma$$

$$LDP = \max_i |X_i| + \gamma \sqrt{3}$$

$$SQ = D \times C + \log(n)$$

Ce formalisme s'intéresse non pas aux données mais aux distributions

→ Restriction aux requêtes linéaires / moyennes

τ = tolérance

Thm: Si A_{SQ} fait au plus t requêtes, alors il existe un mécanisme ϵ -LDP simulant A_{SQ} basé sur $n = \frac{t}{\epsilon^2 \tau^2}$ échantillons

Rq: LDP $\xrightleftharpoons[\text{poly}]{} SQ$