中国科学技术大学计算机学院
《数据隐私的方法伦理和实践》作业

2021.06.06

学生姓名：胡毅翔

学生学号：PB18000290

计算机实验教学中心制

2019 年 9 月

1 Concept of DP

1.1

Prove that the Laplace mechanism preserves (ϵ, 0)-DP.

Proof. Let $x \in \mathbb{N}^{|\mathcal{X}|}$ and $y \in \mathbb{N}^{|\mathcal{X}|}$ be such that $\|x-y\|_1 \leq 1$, and let $f(\cdot)$ be some function $f : \mathbb{N}^{|\mathcal{X}|} \to \mathbb{R}^k$. Let p_x denote the probability density function of $\mathcal{M}_L(x, f, \varepsilon)$, and let p_y denote the probability density function of $\mathcal{M}_L(y, f, \varepsilon)$. We compare the two at some arbitrary point $z \in \mathbb{R}^k$

$$\begin{aligned}
\frac{p_x(z)}{p_y(z)} &= \prod_{i=1}^{k}\left(\frac{\exp\left(-\frac{\varepsilon|f(x)_i - z_i|}{\Delta f}\right)}{\exp\left(-\frac{\varepsilon|f(y)_i - z_i|}{\Delta f}\right)}\right) \\
&= \prod_{i=1}^{k}\exp\left(\frac{\varepsilon\left(|f(y)_i - z_i| - |f(x)_i - z_i|\right)}{\Delta f}\right) \\
&\leq \prod_{i=1}^{k}\exp\left(\frac{\varepsilon\left|f(x)_i - f(y)_i\right|}{\Delta f}\right) \\
&= \exp\left(\frac{\varepsilon \cdot \|f(x) - f(y)\|_1}{\Delta f}\right) \\
&\leq \exp(\varepsilon)
\end{aligned}$$

where the first inequality follows from the triangle inequality, and the last follows from the definition of sensitivity and the fact that $\|x-y\|_1 \leq 1$. That $\frac{p_x(z)}{p_y(z)} \geq \exp(-\varepsilon)$ follows by symmetry. □

1.2

Please explain the difference between $(\epsilon, 0) -$ DP and (ϵ, δ) -DP. Typically, what range of δ we're interested in? Explain the reason.

Solution. Even δ is negligible, there are theoretical distinctions between $(\varepsilon, 0)$ - and (ε, δ) - differential privacy.

$(\varepsilon, 0)$ -differential privacy: for every run of the mechanism $M(x)$, the output observed is (almost) equally likely to be observed on every neighboring database, simultaneously.

(ε, δ) - differential privacy: given an output $\xi \sim M(x)$ it may be possible to find a database y such that ξ is much more likely to be produced on y than it is when the database is x. The privacy loss (divergence) incurred by observation ξ :

$$\mathcal{L}^{(\xi)}_{\mathcal{M}(x)\|\mathcal{M}(y)} = \ln\left(\frac{\Pr[\mathcal{M}(x) = \xi]}{\Pr[\mathcal{M}(y) = \xi]}\right)$$

(ε, δ) - differential privacy ensures that for all adjacent x, y, the absolute value of the privacy loss will be bounded by ε with probability at least $1 - \delta$.

Typically, we are interested in values of δ that are less than the inverse of any polynomial in the size of the database.

Because, for each piece of data in data set, there is a probability that it will be released. Each piece of different data in this ralease is independent, so this mechanism can release $n\delta$ sample. So in order to prevent such leakage, it must be less than $1/n$.

□

1.3

Please explain the difference between DP and Local DP.

Solution. Definition of ϵ -local differential privacy is that a randomized function f satisfies ϵ local differential privacy if and only if for any two input tuples t and t' in the domain of f, and for any output t^* of f, we have:

$$\Pr\left[f(t)=t^*\right] \leq \exp(\epsilon) \cdot \Pr\left[f\left(t'\right)=t^*\right]$$

1. The notation $\Pr[\cdot]$ means probability. If f 's output is continuous, $\Pr[\cdot]$ is replaced by the probability density function.

2. Basically, local differential privacy is a special case of differential privacy where the random perturbation is performed by the users, not by the aggregator.

3. According to the above definition, the aggregator, who receives the perturbed tuple t, cannot distinguish whether the true tuple is t or another tuple t' with high confidence (controlled by parameter ϵ), regardless of the background information of the aggregator.

4. This provides plausible deniability to the user.

While the definition of differential privacy is that A randomized algorithm M with domain $\mathbb{N}^{|X|}$ is (ϵ, δ) -differentially private if for all $S \subset$ Range (M) and for all $x, y \in \mathbb{N}|X|$ such that $\|x-y\|_1 \leq 1$:

$$\Pr[M(x) \in S] \leq \exp(\epsilon) \Pr[M(y) \in S]+\delta$$

where the probability space is over the coin flips of the mechanism M. If $\delta=0$, we say that M is δ -differentially private.

We can find out the difference between LDP and DP is that DP restrictions on tuple $x, y \in \mathbb{N}|X|$ such that $\|x-y\|_1 \leq 1$, while LDP restrictions on any two input tuples t and t'.

□

2 Basics of DP

ID	Sex	Chinese	Mathematics	English	Physics	Chemistry	Biology
1	Male	96	58	80	53	56	100
2	Male	60	63	77	50	59	75
3	Female	83	86	98	69	80	100
...							
2000	Female	86	83	98	87	82	92

Table 1: Scores of students in School A

Table 1 is the database records scores of students in School A in the final exam. We need to help teacher query the database while protecting the privacy of students' scores. The domain of this database is $\{$ Male, Female $\} \times \{0,1,2,\ldots,100\}^6$. In this question, assume that two inputs X and Y are neighbouring inputs if X can be obtained from Y by removing or adding one element. Answer the following questions.

2.1

What is the sensitivity of the following queries:

1. $q_1 = \frac{1}{2000}\sum_{ID=1}^{2000}$ Mathematics $_{ID}$
2. $q_2 = \max_{ID\in[1,2000]}$ English $_{ID}$

Solution. 1. $q_1 = \frac{1}{2000}\sum_{ID=1}^{2000}$ Mathematics $_{ID} = \frac{100}{2000} = 0.05$

2. $q_2 = \max_{ID\in[1,2000]}$ English $_{ID} = 100$

□

2.2

Design ϵ -differential privacy mechanisms corresponding to the two queries in 2.1 where $\epsilon = 0.1$. (Using Laplace mechanism for q_1, Exponential mechanism for q_2.)

Solution. 1.

$$q_1 = \frac{1}{2000}\sum_{ID=1}^{2000} Mathematics_{ID} + Y$$

where Y is random variable drawn from $Lap\,(0.5)$

2. Output y with probability $\propto exp\left(\frac{0.1*u(x,y)}{2*100}\right)$, $y = q_2$

□

2.3

Let $M_1, M_2, \ldots, M_{100}$ be 100 Gaussian mechanisms that satisfy $(\epsilon_0, \delta_0) - \mathrm{DP}$, respectively, with respect to the database. Given $(\epsilon, \delta) = (1.25, 10^{-5})$, calculate σ for every query with the composition theorem (Theorem 3.16 in the textbook) and the advanced composition theorem (Theorem 3.20 in the textbook, choose $\delta' = \delta$) such that the total query satisfies (ϵ, δ) - DP.

Solution. 1.

$$\sum k = 1100\epsilon_0 = 1.25, \sum k = 1100\delta_0 = 10^{-5}$$

$$\epsilon_0 = 0.0125, \delta_0 = 10^{-7}$$

2.

$$k\delta_0 + \delta_0 = 10^{-5}$$

$$\epsilon_0 = \frac{1.25}{2\sqrt{2kln(\frac{1}{\delta_0})}}$$

$$\delta_0 = 9.9 \times 10^{-8}$$

$$\epsilon_0 = 0.011$$

□

3 Local DP

This question focuses on the problem of estimating the mean value of a numeric attributes by collecting data from individuals under ϵ -LDP. Assume that each user u_i 's data record t_i contains a single numeric attribute whose value lies in range $[-1, 1]$. Answer the following questions.

3.1

Prove that Algorithm 1 satisfies ϵ -LDP.

Proof.

$$l(t_i) = \frac{e^{\epsilon/2}t_i - 1}{e^{\epsilon/2} - 1}$$

$$r(t_i) = \frac{e^{\epsilon/2}t_i + 1}{e^{\epsilon/2} - 1}$$

$\forall t_i, t_j \in [-1, 1]$

$$Pr[f(t_i) = t^*] = \frac{(e^{\epsilon/2} - 1)e^{\epsilon/2}}{2(e^{\epsilon/2} + 1)}, t^* \in [l(t_i), r(t_i)]$$

$$Pr[f(t_i)=t^*]=\frac{(e^{\epsilon/2}-1)}{2(e^{\epsilon/2}+1)e^{\epsilon/2}}, t^*\in[-C,l(t_i)]\cup[r(t_i),C]$$

Thus,

$$\Pr\left[f(t_i)=t^*\right]\leq\exp(\epsilon)\cdot\Pr\left[f\left(t_j\right)=t^*\right]$$

□

3.2

Prove that given an input value t_i, Algorithm 1 returns a noisy value t_i^* with $\mathbb{E}\left[t_i^*\right]=t_i$ and $\operatorname{Var}\left[t_i^*\right]=\frac{t_i^2}{e^{\epsilon/2}-1}+\frac{e^{\epsilon/2}+3}{3\left(e^{\epsilon/2}-1\right)^2}$

Proof.

$$E[t_i^*]=\int_{-C}^{l(t_i)}x\frac{(e^{\epsilon/2}-1)}{2(e^{\epsilon/2}+1)e^{\epsilon/2}}dx+\int_{r(t_i)}^{C}x\frac{(e^{\epsilon/2}-1)}{2(e^{\epsilon/2}+1)e^{\epsilon/2}}dx+\int_{l(t_i)}^{r(t_i)}x\frac{(e^{\epsilon/2}-1)e^{\epsilon/2}}{2(e^{\epsilon/2}+1)}dx=t^*$$

$$\begin{aligned}Var[t_i^*]=&E[(t_i^*)^2]-(E[t_i^*])^2\\=&\int_{-C}^{l(t_i)}x^2\frac{(e^{\epsilon/2}-1)}{2(e^{\epsilon/2}+1)e^{\epsilon/2}}dx+\int_{r(t_i)}^{C}x^2\frac{(e^{\epsilon/2}-1)}{2(e^{\epsilon/2}+1)e^{\epsilon/2}}dx+\int_{l(t_i)}^{r(t_i)}x^2\frac{(e^{\epsilon/2}-1)e^{\epsilon/2}}{2(e^{\epsilon/2}+1)}dx-(t^*)^2\\=&\frac{t_i^2}{e^{\epsilon/2}-1}+\frac{e^{\epsilon/2}+3}{3\left(e^{\epsilon/2}-1\right)^2}\end{aligned}$$

□

4 Random Subsampling

Given a dataset $x\in\mathcal{X}^n$, and $m\in\{0,1,\ldots,n\}$, a random m -sumsample of x is a new (random) dataset $x'\in\mathcal{X}^m$ formed by keeping a random subset of m rows from x and throwing out the remaining $n-m$ rows.

4.1

Show that for every $n\in\mathbb{N},\mathcal{X}\geq 2,m\in\{1,\ldots,n\},\epsilon>0$ and $\delta<m/n$ the mechanism $M(x)$ that outputs a random m -subsample of $x\in\mathcal{X}^n$ is not $(\epsilon,\delta)-\mathrm{DP}$

Proof. Let $\mathcal{X}=\{0,1\}$ and consider the two datasets $x=0^n$ and $x'=10^{n-1}$. Now define $S=\{z\in\{0,1\}^m\mid z\neq 0^m\}$. Then for every ϵ and every $\delta<m/n$

$$e^{\varepsilon}\Pr[A(x)\in S]+\delta=\delta<\frac{m}{n}=\Pr\left[A\left(x'\right)\in S\right]$$

contradicting $(\varepsilon,\delta)-\mathrm{dp}$ of M. □

4.2

Although random subsamples do not ensure differential privacy on their own, a random subsample dose have the effect of "amplifying" differential privacy. Let $M : \mathcal{X}^m \rightarrow \mathcal{R}$ be any algorithm. We define the algorithm $M' : \mathcal{X}^n \rightarrow \mathcal{R}$ as follows: choose x' to be a random m -subsample of x, then output $M(x')$. Prove that if M is (ϵ, δ) -DP, then M' is $((e^\epsilon - 1) \cdot m/n, \delta m/n)$ -DP. Thus, if we have an algorithm with the relatively weak guarantee of 1 -DP, we can get an algorithm with ϵ -DP by using a random subsample of a database that is larger by a factor of $1/(e^\epsilon - 1) = O(1/\epsilon)$.

Proof. We'll use $T \subseteq \{1, \ldots, n\}$ to denote the identities of the m -subsampled rows (i.e. their row number, not their actual contents). Note that T is a random variable, and that the randomness of M' includes both the randomness of the sample T and the random coins of M. Let $x \sim x'$ be adjacent databases and assume that x and x' differ only on some row t. Let x_T (or x'_T) be a subsample from x (or x') containing the rows in T. Let S be an arbitrary subset of the range of M'. For convenience, define $p = m/n$ To show $(p(e^\varepsilon - 1), p\delta) - \text{dp}$, we have to bound the ratio

$$\frac{\Pr[M'(x) \in S] - p\delta}{\Pr[M'(x') \in S]} = \frac{p\Pr[M(x_T) \in S \mid i \in T] + (1-p)\Pr[M(x_T) \in S \mid i \notin T] - p\delta}{p\Pr[M(x'_T) \in S \mid i \in T] + (1-p)\Pr[M(x'_T) \in S \mid i \notin T]}$$

by $e^{p(e^\varepsilon - 1)}$. For convenience, define the quantities

$$C = \Pr[M(x_T) \in S \mid i \in T]$$
$$C' = \Pr[M(x'_T) \in S \mid i \in T]$$
$$D = \Pr[M(x_T) \in S \mid i \notin T] = \Pr[M(x'_T) \in S \mid i \notin T]$$

We can rewrite the ratio as

$$\frac{\Pr[M'(x) \in S]}{\Pr[M'(x') \in S]} = \frac{pC + (1-p)D - p\delta}{pC' + (1-p)D}$$

Now we use the fact that, by (ε, δ) -dp, $A \leq e^{\varepsilon} \min\{C', D\} + \delta$. The rest is a calculation:

$$\begin{aligned}
& pC + (1-p)D - p\delta \\
\leq & p\left(e^{\varepsilon} \min\{C', D\} + \delta\right) + (1-p)D - p\delta \\
\leq & \; p\left(\min\{C', D\} + \left(e^{\varepsilon} - 1\right) \min\{C', D\}\right) + \delta) + (1-p)D - p\delta \\
\leq & p\left(\min\{C', D\} + \left(e^{\varepsilon} - 1\right)\left(pC' + (1-p)D\right) + \delta\right) + (1-p)D - p\delta \\
& (Because \min\{x, y\} \leq \alpha x + (1-\alpha)y \; for \; every \; 0 \leq \alpha \leq 1) \\
\leq & p\left(C' + \left(e^{\varepsilon} - 1\right)\left(pC' + (1-p)D\right) + \delta\right) + (1-p)D - p\delta \quad (Because \; \min\{x, y\} \leq x) \\
\leq & p\left(C' + \left(e^{\varepsilon} - 1\right)\left(pC' + (1-p)D\right)\right) + (1-p)D \\
\leq & \left(pC' + (1-p)D\right) + \left(p\left(e^{\varepsilon} - 1\right)\right)\left(pC' + (1-p)D\right) \\
\leq & \left(1 + p\left(e^{\varepsilon} - 1\right)\right)\left(pC' + (1-p)D\right) \\
\leq & e^{p\left(e^{\varepsilon} - 1\right)}\left(pC' + (1-p)D\right)
\end{aligned}$$

So we've succeeded in bounding the necessary ratio of probabilities. Note, if you are willing to settle for $(\mathrm{O}(\mathrm{cm/n}), \mathrm{O}(\delta m/\mathrm{n})) - \mathrm{dp}$ the calculation is much simpler. All this algebra is mostly just to get the tight bound. $\square$