Data Privacy Course Exercise 2

PB18061352 郑龙韬

1. Concept of DP (15')

1.1 Prove that the Laplace mechanism preserves $(\varepsilon, 0)$-DP.

Proof. Given any function $f : \mathbb{N}^{|\mathcal{X}|} \to \mathbb{R}^k$, the Laplace mechanism is defined as:

$$\mathcal{M}_L(x, f(\cdot), \varepsilon) = f(x) + (Y_1, \ldots, Y_k)$$

where Y_i are i.i.d. random variables drawn from $Lap(\frac{\Delta f}{\varepsilon})$.

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, p_y denote the probability density function of $\mathcal{M}_L(x, f, \varepsilon)$ and $\mathcal{M}_L(y, f, \varepsilon)$, respectively. Compare them at some arbitrary point $z \in \mathbb{R}^k$:

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

Similarly, $\frac{p_x(z)}{p_y(z)} \geq \exp(-\varepsilon)$ by symmetry.

Therefore, for every run of the Laplace mechanism, the output observed is (almost) equally likely to be observed on every neighboring database, simultaneously. That is, the Laplace mechanism preserves $(\varepsilon, 0)$-DP. □

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

Answer: There are theoretical distinctions between $(\varepsilon, 0)$- and (ε, δ)-DP. $(\varepsilon, 0)$-DP ensures that for all adjacent x, y, the absolute value of the privacy loss is bounded by ε, while (ε, δ)-DP ensures this condition with probability at least $1 - \delta$.

We are interested in values of δ that are less than the inverse of any polynomial in the size of the database. The reason is that values of δ on the order of $\frac{1}{\|x\|_1}$ are very dangerous. (ε, δ)-DP provides privacy protection for typical members of a data set and compromise the privacy of some other

participants. It can be achieved by randomly selecting a subset of rows and releasing them in their entirety.

1.3 Please explain the difference between DP and Local DP.

Answer: The difference is about **what** and **when** the data are disturbed.

- For DP, every users send their raw data to a centralized server, the DP algorithm generates the set of disturbed data based on the whole set of raw data (globally).
- For Local DP, the noise is added locally. Each user runs local DP algrithm and obtains noisy data from their own raw data, then sends them to the server and become part of the set of disturbed data.

2. Basics of DP (30')

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

表 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 $\{\text{Male, Female}\} \times \{0, 1, 2, \ldots, 100\}^6$. **Two inputs X and Y are neighbouring inputs if X only has one element different from that in Y.**

2.1 What is the sensitivity of the following queries:

(1) $q_1 = \frac{1}{2000}\Sigma_{\text{ID}=1}^{2000}\text{Mathematics}_{\text{ID}}$

Answer: $f = q_1$, Δf is the sensitivity of function f, $\Delta f = \frac{100}{2000} = 0.05$ (since the range of the math score is $[0, 100]$, the difference between the sum of two neighbouring inputs i s 100 at most, then the average is 0.05).

(2) $q_2 = \max_{\text{ID}\in[1,2000]} \text{English}_{\text{ID}}$

Answer: $f = q_2$, for max operation, the sensitivity $\Delta f = 100$ (consider two neighbouring inputs: $X = \{0, \ldots, 0, 0\}$ and $Y = \{0, \ldots, 0, 100\}$).

2.2 Design ε-differential privacy mechanisms corresponding to the two queries in question 2.1 where $\varepsilon = 0.1$ (using laplace mechanism for q_1, exponential mechanism for q_2).

Answer:

- Laplace mechanism for q_1: $\Delta f = 0.05$, $\frac{\Delta f}{\varepsilon} = 0.5$, $\mathcal{M}_L(x, f(\cdot), \varepsilon) = f(x) + Y$, where Y are i.i.d random variables drawn from $Lap(0.5)$.
- Exponential mechanism for q_2: define the score function $u(x, r) = I(r = \max(x))$, where $I(\cdot)$ is an indicator function, $\Delta u = 1$. By exponential mechanism, the output is $r \in \{0, 1, 2, \ldots, 100\}$ with probability $\propto \exp(\frac{\varepsilon u(x,r)}{2\Delta u}) = \exp(0.05u(x, r))$.

2.3 Let $M_1, M_2, \ldots, M_{100}$ be 100 Gaussian mechanisms that satisfy $(\varepsilon_0, \delta_0)$-DP, respectively, with respect to the database. Given $(\varepsilon, \delta) = (1.25, 10^{-5})$, calculate σ (parameter of Gaussian distribution) for every query with the composition theorem and the advanced composition theorem (choose $\delta' = \delta_0$) such that the total query satisfies (ε, δ)-DP.

Answer: We first compute δ_0 and ε_0 for the two theorems.

- By composition theorem, for $\mathcal{M}_i$ $(1 \le i \le 100)$ with $(\varepsilon_0, \delta_0)$-differential privacy, and $\mathcal{M}_{[100]} = (\mathcal{M}_1(x), \ldots, \mathcal{M}_{100}(x))$, $\mathcal{M}_{[100]}$ is $(100\varepsilon_0, 100\delta_0)$-differentially private.

$$\delta_0 = \frac{\delta}{100} = 10^{-7}$$
$$\varepsilon_0 = \frac{\varepsilon}{100} = 0.0125$$

- By advanced composition, $\forall \varepsilon, \delta, \delta \ge 0$, the class of $(\varepsilon_0, \delta_0)$-differentially private mechanisms satisfies $(\varepsilon', k\delta_0 + \delta')$-differential privacy under k-fold adaptive composition for

$$\varepsilon = \varepsilon' = \sqrt{2k \ln(1/\delta')} \cdot \varepsilon_0 + k\varepsilon_0(e^{\varepsilon_0} - 1)$$
$$\delta = k\delta_0 + \delta' = (k + 1)\delta_0$$

  Given target privacy parameters $k = 100$, $\varepsilon = 1.25$ and $\delta = 10^{-5}$, to ensure $(\varepsilon', k\delta + \delta')$ cumulative privacy loss over k mechanisms, it suffices that each mechanism is $(\varepsilon_0, \delta_0)$-differentially private, where

$$\delta_0 = \frac{\delta}{101} \approx 9.9010 \times 10^{-8}$$
$$\varepsilon_0 \approx 0.0212$$

For $c > 2\ln(1.25/\delta_0)$, the Gaussian mechanism with parameter $\sigma \ge c\Delta_2(f)/\varepsilon$ is (ε, δ)-DP. Without loss of generality, we assume $f = q_1$. Therefore, $\Delta_2(f) = \Delta_1(f) = 0.05$ since $f(x)$ is a scalar.

- By composition theorem,

$$c > 2\ln(1.25/\delta_0) = 2\ln(1.25/10^{-7}) \approx 32.6825$$
$$\sigma \geq \frac{c \times 0.05}{0.0125} \approx 163.4124$$

- By advanced composition,

$$c > 2\ln(1.25/\delta_0) = 2\ln(1.25/(9.9010 \times 10^{-8})) \approx 32.7024$$
$$\sigma \geq \frac{c \times 0.05}{0.0212} \approx 77.1283$$

3. Local DP (30')

Algorithm 1

Input: tuple $t_i \in [-1, 1]$ and privacy parameter ε.

Output: tuple $t_i^* \in [-C, C]$.

1: Sample x uniformly at random from $[0, 1]$;
2: $C = \frac{\exp(\varepsilon/2)+1}{\exp(\varepsilon/2)-1}$;
3: $l(t_i) = \frac{C+1}{2} \cdot t_i - \frac{C-1}{2}$;
4: $r(t_i) = l(t_i) + C - 1$;
5: **if** $x < \frac{\exp(\varepsilon/2)}{\exp(\varepsilon/2)+1}$ **then**
6: Sample t_i^* uniformly at random from $[l(t_i), r(t_i)]$;
7: **else**
8: Sample t_i^* uniformly at random from $[-C, l(t_i)] \cup [r(t_i), C]$;
9: **end if**
10: **return** t_i^*

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. The output t_i^* is sampled based on x uniformly sampled from $[0, 1]$, then the probability density function of t_i^* is a piecewise constant function as follows:

$$Pr[t_i^* = x|t_i] = \begin{cases} \frac{\exp(\varepsilon)-\exp(\varepsilon/2)}{2\exp(\varepsilon/2)+2}, & \text{if } x \in [l(t_i), r(t_i)] \\ \frac{\exp(\varepsilon)-\exp(\varepsilon/2)}{\exp(\varepsilon)(2\exp(\varepsilon/2)+2)}, & \text{if } x \in [-C, l(t_i)] \cup [r(t_i), C] \end{cases}$$

For any output $t^* \in [-C, C]$ and anttwo input tuples $t_1, t_2 \in [-1, 1]$ of algorithm 1, we have

$$Pr[t^*|t_1] \leq (\frac{\exp(\varepsilon) - \exp(\varepsilon/2)}{2\exp(\varepsilon/2) + 2})/(\frac{\exp(\varepsilon) - \exp(\varepsilon/2)}{\exp(\varepsilon)(2\exp(\varepsilon/2) + 2)})Pr[t^*|t_2] \leq \exp(\varepsilon)Pr[t^*|t_2]$$

Upon analysis, $Pr[t^*|t_1] \leq \exp(\varepsilon)Pr[t^*|t_2]$. Thus, algorithm 1 satisfies ε-LDP. □

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

Proof.

$$\begin{aligned}
\mathbb{E}[t_i^*] &= \int_{-C}^{C} x \cdot Pr(t_i^* = x)dx \\
&= \int_{-C}^{l(t_i)} x\frac{\exp(\varepsilon)-\exp(\varepsilon/2)}{\exp(\varepsilon)(2\exp(\varepsilon/2)+2)}dx + \int_{l(t_i)}^{r(t_i)} x\frac{\exp(\varepsilon)-\exp(\varepsilon/2)}{2\exp(\varepsilon/2)+2}dx + \int_{r(t_i)}^{C} x\frac{\exp(\varepsilon)-\exp(\varepsilon/2)}{\exp(\varepsilon)(2\exp(\varepsilon/2)+2)}dx \\
&= \frac{\exp(\varepsilon)-\exp(\varepsilon/2)}{\exp(\varepsilon)(4\exp(\varepsilon/2)+4)}(l^2(t_i)-r^2(t_i)) + \frac{\exp(\varepsilon)-\exp(\varepsilon/2)}{4\exp(\varepsilon/2)+4}(r^2(t_i)-l^2(t_i)) \\
&= \frac{\exp(\varepsilon)-\exp(\varepsilon/2)}{2\exp(\varepsilon/2)+2} \cdot (-\frac{1}{2\exp(\varepsilon)} + \frac{1}{2}) \cdot (C+1)t_i \cdot (C-1) \\
&= t_i
\end{aligned}$$

$$\begin{aligned}
Var[t_i^*] &= \mathbb{E}[(t_i^*)^2] - (\mathbb{E}[t_i^*])^2 \\
&= \int_{-C}^{l(t_i)} x^2\frac{\exp(\varepsilon)-\exp(\varepsilon/2)}{\exp(\varepsilon)(2\exp(\varepsilon/2)+2)}dx + \int_{l(t_i)}^{r(t_i)} x^2\frac{\exp(\varepsilon)-\exp(\varepsilon/2)}{2\exp(\varepsilon/2)+2}dx \\
&\quad + \int_{r(t_i)}^{C} x^2\frac{\exp(\varepsilon)-\exp(\varepsilon/2)}{\exp(\varepsilon)(2\exp(\varepsilon/2)+2)}dx - t_i^2 \\
&= \frac{\exp(\varepsilon)-\exp(\varepsilon/2)}{\exp(\varepsilon)(6\exp(\varepsilon/2)+6)}(l^3(t_i)-r^3(t_i)+2C^3) + \frac{\exp(\varepsilon)-\exp(\varepsilon/2)}{6\exp(\varepsilon/2)+6)}(r^3(t_i)-l^3(t_i)) - t_i^2 \\
&= \frac{\exp(\varepsilon)-\exp(\varepsilon/2)}{2\exp(\varepsilon/2)+2)}((\frac{1}{3}-\frac{1}{3\exp(\varepsilon)})\frac{6\exp(\varepsilon)t_i^2+2}{(\exp(\varepsilon/2)-1)^3} + \frac{2}{3\exp(\varepsilon)}(\frac{\exp(\varepsilon/2)+1}{\exp(\varepsilon/2)-1})^3) - t_i^2 \\
&= \frac{t_i^2}{e^{\varepsilon/2}-1} + \frac{e^{\varepsilon/2}+3}{3(e^{\varepsilon/2}-1)^2}
\end{aligned}$$

□

4. Random Subsampling (25')

Given a dataset $x \in \mathcal{X}^n$, and $m \in \{0, 1, \ldots, n\}$, a random m-subsample 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 $n \in N$, $|\mathcal{X}| \geq 2$, $m \in \{1, \ldots, n\}$, $\varepsilon > 0$ and $\delta < \frac{m}{n}$, the mechanism $M(x)$ that outputs a random m-subsample of $x \in \mathcal{X}^n$ is not (ε, δ)-DP.

Answer: Let $\mathcal{X} = \{0, 1\}$, two adjacent databases $x = 0^n$ and $x' = 1\|0^{n-1}$, and $S = \{0, 1\}^m$ ($0^m \notin S$). $\forall \varepsilon > 0$ and $\delta < \frac{m}{n}$,

$$e^{\varepsilon} Pr[M(x) \in S] + \delta = 0 + \delta = \delta < \frac{m}{n} = Pr[M(x') \in S]$$

Namely,

$$\frac{Pr[M(x') \in S] - \delta}{Pr[M(x) \in S]} > e^{\varepsilon}$$

which contradicts (ε, δ)-DP.

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 $M : \mathcal{X}^m \to \mathcal{R}$ be any algorithm. We define the algorithm $M' : \mathcal{X}^n \to \mathcal{R}$ 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 $((e^{\varepsilon} - 1) \cdot m/n, \delta m/n)$-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 $1/(e^{\varepsilon}-1) = O(1/\varepsilon)$.

Proof. Let S, T denote two arbitrary subsamples, x, x' denote two adjacent databases, and t, t' for their m-subsamples. For convenience, we define probabilities as follows:

$$\begin{aligned}
p_1 &= Pr[A(t) \in S | i \in T] \\
p_2 &= Pr[A(t') \in S | i \in T] \\
p_3 &= Pr[A(t) \in S | i \notin T] = Pr[A(t') \in S | i \notin T]
\end{aligned}$$

If M satisfies (ε, δ)-DP, we have $p_1 \leq e^{\varepsilon} \min(p_2, p_3) + \delta$.

$$\begin{aligned}
\frac{Pr[M'(x) \in S] - \frac{m}{n}\delta}{Pr[M'(x') \in S]} &= \frac{\frac{m}{n}p_1 + (1 - \frac{m}{n})p_3 - \frac{m}{n}\delta}{\frac{m}{n}p_2 + (1 - \frac{m}{n})p_3} \\
&\leq \frac{\frac{m}{n}(e^{\varepsilon}\min(p_2, p_3) + \delta) + (1 - \frac{m}{n})p_3 - \frac{m}{n}\delta}{\frac{m}{n}p_2 + (1 - \frac{m}{n})p_3} \qquad \text{(DP)} \\
&= \frac{\frac{m}{n}(\min(p_2, p_3) + (e^{\varepsilon} - 1)\min(p_2, p_3) + \delta) + (1 - \frac{m}{n})p_3 - \frac{m}{n}\delta}{\frac{m}{n}p_2 + (1 - \frac{m}{n})p_3} \\
&\leq \frac{\frac{m}{n}(\min(p_2, p_3) + (e^{\varepsilon} - 1)(\frac{m}{n}p_2 + (1 - \frac{m}{n})p_3) + \delta) + (1 - \frac{m}{n})p_3 - \frac{m}{n}\delta}{\frac{m}{n}p_2 + (1 - \frac{m}{n})p_3} \\
&\leq \frac{\frac{m}{n}(p_2 + (e^{\varepsilon} - 1)(\frac{m}{n}p_2 + (1 - \frac{m}{n})p_3) + \delta) + (1 - \frac{m}{n})p_3 - \frac{m}{n}\delta}{\frac{m}{n}p_2 + (1 - \frac{m}{n})p_3} \\
&= \frac{\frac{m}{n}(p_2 + (e^{\varepsilon} - 1)(\frac{m}{n}p_2 + (1 - \frac{m}{n})p_3)) + (1 - \frac{m}{n})p_3}{\frac{m}{n}p_2 + (1 - \frac{m}{n})p_3} \\
&= \frac{(\frac{m}{n}p_2 + (1 - \frac{m}{n})p_3) + (\frac{m}{n}(e^{\varepsilon} - 1))(\frac{m}{n}p_2 + (1 - \frac{m}{n})p_3)}{\frac{m}{n}p_2 + (1 - \frac{m}{n})p_3} \\
&= \frac{(1 + \frac{m}{n}(e^{\varepsilon} - 1))(\frac{m}{n}p_2 + (1 - \frac{m}{n})p_3)}{\frac{m}{n}p_2 + (1 - \frac{m}{n})p_3} \\
&\leq \frac{e^{\frac{m}{n}(e^{\varepsilon} - 1)}(\frac{m}{n}p_2 + (1 - \frac{m}{n})p_3)}{\frac{m}{n}p_2 + (1 - \frac{m}{n})p_3} \\
&= e^{\frac{m}{n}(e^{\varepsilon} - 1)}
\end{aligned}$$

Therefore, M' satisfies

$$\frac{Pr[M'(x) \in S] - \frac{m}{n}\delta}{Pr[M'(x') \in S]} \leq e^{\frac{m}{n}(e^{\varepsilon}-1)}$$

which is $((e^{\varepsilon}-1) \cdot m/n, \delta m/n)$-DP by definition. □