

Approximate DP

Def. We say that a mechanism M is (ϵ, δ) -DP with $\epsilon \geq 0$ & $\delta \in [0, 1]$ if

$$\mathbb{E}_{\epsilon} (M_O \| M_{O'}) \leq \delta \quad \forall D \sim D'$$

This relaxation was first introduced by Dwork, Kenthapadi, McSherry, Mironov,
& Naor

in the paper "Our data: Privacy via distributed noise generation", TCC 2006

This definition provides marginally weaker privacy guarantee, but allows us to add significantly less noise to achieve it.

Remark:

1- clearly, setting $\delta=0$ in the above definition, we recover

$$\text{pure DP:} \quad (\Sigma, 0)\text{-DP} \equiv \Sigma\text{-DP}$$

2- Using the expression of HS divergence, we can write the more standard definition of $(\Sigma, \delta)\text{-DP}$:

$$M \text{ is } (\Sigma, \delta)\text{-DP} \Leftrightarrow M_D(A) \leq e^{\Sigma} M_{D'}(A) + \delta \quad \forall A \in \mathcal{A}$$

3- A necessary condition:

$$\text{Let } Z_{D,D'} := \log \frac{p_{Z_D}(Z)}{p_{Z_{D'}}(Z)} \quad \text{where } Z \sim M_D$$

then:

$$\Pr(|Z_{D,D'}| \leq \epsilon) \geq 1 - \delta \quad \forall D \sim D' \Rightarrow M \text{ is } (\epsilon, \delta)\text{-DP}$$

Note that:

$$\Pr(|Z_{D,D'}| \leq \epsilon) = 1 \quad \forall D \sim D' \Leftrightarrow M \text{ is } \epsilon\text{-DP}$$

4- Definition of approximate DP implies:

There might exist a set E such that $M_D(E) = 0$

but $M_{D'}(E) > 0$. The probability of all such

events $\leq \delta$. If the mechanism's output happens to

take value in these sets, then for sure D' couldn't have

been the dataset; so in this case, we have no privacy!

Since for all such events, Privacy is blatantly violated, we say that approximately $\mathcal{D}_p$ may cause catastrophic privacy violation, but it happens with prob $\leq \delta$.

This gives us a general recipe for proving a mechanism is (ϵ, δ) - $\mathcal{D}_p$.

step 1: Find the collection of all events for which

$$\frac{p_{\mathcal{D}_D}(u)}{p_{\mathcal{D}_B}(u)} > e^\epsilon \quad (\text{Bad event})$$

Note that if x takes values in the complement of this set, then $\frac{f_{2D}(x)}{f_{2D'}(x)} \leq e^\varepsilon$.

Step 2: Show the probability of Bad events $\leq \delta$.

Step 1 + Step 2 $\Rightarrow$ (ε, δ) -DP. Why?

$$\text{let } \text{Bad}_{D,D'} := \left\{ x: \frac{f_{2D}(x)}{f_{2D'}(x)} > e^\varepsilon \right\}$$

For any set A :

$$\begin{aligned} M_D(A) &= M_D(A \cap \text{Bad}_{D,D'}) + M_D(A \cap \text{Bad}_{D,D'}^c) \\ &\leq M_D(A \cap \text{Bad}_{D,D'}^c) + M_D(\text{Bad}_{D,D'}) \end{aligned}$$

↗ complement

By definition of Bad events

$$\begin{aligned} & \stackrel{\text{Step 2}}{\leq} M_D(A \cap \text{Bad}_{D,D'}^c) + \delta \\ & \rightarrow \leq e^\Sigma M_{D'}(A \cap \text{Bad}_{D,D'}^c) + \delta \end{aligned}$$

$$\leq e^\Sigma M_{D'}(A) + \delta \quad \Rightarrow \quad M \text{ is } (\Sigma, \delta)\text{-DP}$$

Thus, Steps 1 & 2 are sufficient to prove $(\Sigma, \delta)\text{-DP}$.

How to choose δ ?

We mentioned that δ quantifies the catastrophic failure in privacy! This catastrophe occurs with a quite small probability.

But how small is "quite small"?

Example. (Name-and-Shame)

Let D be a dataset of n individuals together with a sensitive

data $D = \{ (i, x_i) : i = 1, 2, \dots, n \}$

$\uparrow$ individual index $\nwarrow$ sensitive data

Mechanism M iterates over all entries & release each entry as is w.p. δ or does nothing w.p. $(1-\delta)$. Thus, the output is $(Y_1, \dots, Y_n)$ where

$$Y_i = \begin{cases} U(x_i) & \text{w.p. } \delta \\ \perp & \text{w.p. } (1-\delta) \end{cases}$$

$\uparrow$ output nothing

Claim. This mechanism is $(0, \delta)$ -DP.

Proof. Consider datasets $D \sim D'$ that differ in i th entry & let $Z_D = (Y_1, \dots, Y_n)$ &

$$Z_{D'} = (Y'_1, \dots, Y'_n).$$

Let $B_i \sim \text{Bernoulli}(\delta)$. Note that $\Pr(Y_j \neq Y'_j) = 0$ for $j \neq i$ &

$$\Pr(Y_i \neq Y'_i) = \Pr(B_i = 1) = \delta$$

$\downarrow$
 $Z_D = Z_{D'}$ under $B_i = 0$

$$\Pr(Z_D \in E) = \Pr(Z_D \in E \cap B_i = 0) + \Pr(Z_D \in E \cap B_i = 1) = \Pr(Z_D \in E \cap B_i = 0)$$

$$+ \delta \Pr(Z_D \in E \mid B_i = 1) \leq \Pr(Z_{D'} \in E) + \delta$$

However, this mechanism potentially reveals sensitive data for several individuals, leading to considerable privacy violation:

Probability that at least one individual's sensitive gets released = $1 - (1-\delta)^n$

← probability that at least one individual has no privacy!

this mechanism is intuitively private only if $1 - (1-\delta)^n$ is very small:

$$\underbrace{1 - (1-\delta)^n}_{\approx n \cdot \delta} \ll 1$$

$$\Rightarrow n \cdot \delta \ll 1$$

Thus, we want δ to be much smaller than $\frac{1}{n}$.

Usually:

$$\delta \approx \frac{1}{n^{1+\delta}} \quad \text{for instance} \quad \frac{1}{n^{1.1}}$$

In practice:

$$\delta \approx 10^{-7} - 10^{-5} \quad \text{as the size of datasets are typically } 10^5 - 10^7.$$