Privacy

1 Privacy Loss

Definition 1.1. Let Y and Z be two random variables. The privacy loss random variables $\mathcal{L}_{Y||Z}$ is distributed by drawing $t \sim \mathrm{Law}(Y)$, and outputting $\log\left(\frac{\mathbb{P}[Y=t]}{\mathbb{P}[Z=t]}\right)$. If the support of Y and Z are not equal, then the privacy loss random variable is undefined.

2 $\varepsilon - \delta$ DP

2.1 4 ways to see δ

Proposition 2.1. Let P and Q be two probability distributions on $\mathcal{Y}$ such that the privacy loss distribution $\mathrm{PrivLoss}(P||Q)$ is well-defined. Fix $\varepsilon \geq 0$ and define

$$\delta := \sup_{S \subset \mathcal{Y}} P(S) - e^{\varepsilon} Q(S). \tag{2.1}$$

Then

$$\begin{aligned}
\delta &= \underset{Z \sim \mathrm{PrivLoss}(P||Q)}{\mathbb{P}}[Z > \varepsilon] - e^{\varepsilon} \cdot \underset{Z' \sim \mathrm{PrivLoss}(Q||P)}{\mathbb{P}}[-Z' > \varepsilon] \\
&= \underset{Z \sim \mathrm{PrivLoss}(P||Q)}{\mathbb{E}}[\max\{0, 1 - \exp(\varepsilon - Z\}] \\
&= \int_{\varepsilon}^{\infty} e^{\varepsilon - z} \underset{Z \sim \mathrm{PrivLoss}(P||Q)}{\mathbb{P}}[Z > z] \, \mathrm{d}z \\
&\leq \underset{Z \sim \mathrm{PrivLoss}(P||Q)}{\mathbb{P}}[Z > \varepsilon].
\end{aligned}$$

2.2 Moment difference bound

Let X and Y be a random variable supported on $[-\Delta, \Delta]$ satisfying $\mathbb{P}[X \in S] \leq e^{\varepsilon} \mathbb{P}[Y \in S] + \delta$ for all measurable S and vice versa. Then

$$\mathbb{E}[X] - \mathbb{E}[Y] \leq (e^{\varepsilon} - 1)\, \mathbb{E}[|Y|] + 2\delta\Delta \tag{2.2}$$

3 zCDP

Definition 3.1. A randomised mechanism $M : \mathcal{X}^n \to \mathcal{Y}$ is (ξ, ρ)-zero-concentrated differentially private if, for all $x, x' \in \mathcal{X}^n$ differing on a single entry and all $\alpha \in (1, \infty)$,

$$\mathbb{E}[e^{(\alpha-1)Z}] \leq e^{(\alpha-1)(\xi + \rho\alpha)}, \tag{3.1}$$

where $Z = \mathrm{PrivLoss}(M(x)||M(x'))$ is the privacy loss random variable.

3.1 Key properties

1. Pure ε-DP imples $\frac{1}{2}\varepsilon^2$-zCDP
2. The composition of k independent $\frac{1}{2}\varepsilon^2$-zCDP algorithms satisfies $\frac{1}{2}\varepsilon^2 k$-zCDP.
3. $\frac{1}{2}\varepsilon^2 k$-zCDP implies approximate (ε', δ)-DP with $\delta \in (0,1)$ arbitrary and $\varepsilon' = \varepsilon \cdot \sqrt{2k \log(1/\delta)} + \frac{1}{2}\varepsilon^2 k$.

4 Approximate Rényi Differential Privacy

> Rényi differential privacy was introduced by Minorov and was motivated by analyzing privacy amplification by subsampling interleaved with composition, which arises in differentially private deep learning
>
> \- Thomas Steinke

Defintion (RDP). An algorithm M is said to be (λ, ε)-RDP with $\lambda \geq 1$ and $\varepsilon \geq 0$, if for any adjacent inputs x, x'

$$D_\lambda(M(x)||M(x')) := \frac{1}{\lambda - 1} \log \mathop{\mathbb{E}}_{Y \leftarrow M(x)} \left[\left(\frac{\mathbb{P}[M(x) = Y]}{\mathbb{P}[M(x') = Y]} \right)^{\lambda - 1} \right] \leq \varepsilon \tag{4.1}$$

Tip: The ε should be thought of as a function $\varepsilon(\lambda)$, rather than a single number.

Properties

Let P, Q be probability distributions over $\mathcal{Y}$ with a common sigma-algebra such that P is absolutely continuous with respect to Q.

1. **Postprocessing (a.k.a. data processing inequality) & non-negativity:** Let $f : \mathcal{Y} \to \mathcal{Z}$ be a measurable function. Let $f(P)$ denote the distribution on $\mathcal{Z}$ obtained by applying f to a sample from P; define $f(Q)$ similarly. Then

$$0 \leq D_\alpha(f(P)\|f(Q)) \leq D_\alpha(P\|Q) \quad \text{for all } \alpha \in [1, \infty].$$

2. **Composition:** If $P = P' \times P''$ and $Q = Q' \times Q''$ are product distributions, then

$$D_\alpha(P\|Q) = D_\alpha(P'\|Q') + D_\alpha(P''\|Q'') \quad \text{for all } \alpha \in [1, \infty].$$

More generally, suppose P and Q are distributions on $\mathcal{Y} = \mathcal{Y}' \times \mathcal{Y}''$. Let P' and Q' be the marginal distributions on $\mathcal{Y}'$ induced by P and Q respectively. For $y' \in \mathcal{Y}'$, let $P''_{y'}$ and $Q''_{y'}$ be the conditional distributions on $\mathcal{Y}''$ induced by P and Q respectively. That is, we can generate a sample $Y = (Y', Y'') \leftarrow P$ by first sampling $Y' \leftarrow P'$ and then sampling $Y'' \leftarrow P''_{Y'}$, and similarly for Q. Then

$$D_\alpha(P\|Q) \leq D_\alpha(P'\|Q') + \sup_{y' \in \mathcal{Y}'} D_\alpha(P''_{y'}\|Q''_{y'}) \quad \text{for all } \alpha \in [1, \infty].$$

3. **Monotonicity:** For all $1 \leq \alpha \leq \alpha' \leq \infty$,

$$D_\alpha(P\|Q) \leq D_{\alpha'}(P\|Q).$$

4. **Gaussian divergence:** For all $\mu, \mu' \in \mathbb{R}$ with $\sigma > 0$ and all $\alpha \in [1, \infty)$,

$$D_\alpha(\mathcal{N}(\mu, \sigma^2)\|\mathcal{N}(\mu', \sigma^2)) = \alpha \cdot \frac{(\mu - \mu')^2}{2\sigma^2}.$$

5. **Pure DP to Concentrated DP:** For all $\alpha \in [1, \infty)$,

$$D_\alpha(P\|Q) \leq \frac{\alpha}{8} \cdot (D_\infty(P\|Q) + D_\infty(Q\|P))^2.$$

6. **Quasi-convexity:** Let P' and Q' be probability distributions over $\mathcal{Y}$ such that P' is absolutely continuous with respect to Q'. For $s \in [0,1]$, let $(1-s)\cdot P + s\cdot P'$ denote the convex combination of the distributions P and P' with weighting s. For all $\alpha \in (1,\infty)$ and all $s \in [0,1]$,

$$\begin{aligned} &D_\alpha\left((1-s)\cdot P + s\cdot P' \parallel (1-s)\cdot Q + s\cdot Q'\right) \\ &\leq \frac{1}{\alpha-1}\log\left((1-s)\cdot\exp\left((\alpha-1)D_\alpha(P\|Q)\right) + s\cdot\exp\left((\alpha-1)D_\alpha(P'\|Q')\right)\right) \\ &\leq \max\left\{D_\alpha(P\|Q), D_\alpha(P'\|Q')\right\}, \end{aligned}$$

and

$$D_1\left((1-s)\cdot P + s\cdot P' \parallel (1-s)\cdot Q + s\cdot Q'\right) \leq (1-s)\cdot D_1(P\|Q) + s\cdot D_1(P'\|Q').$$

7. **Triangle-like inequality (a.k.a. group privacy):** Let R be a distribution on $\mathcal{Y}$ and assume that Q is absolutely continuous with respect to R. For all $1 < \alpha < \alpha' < \infty$,

$$D_\alpha(P\|R) \leq \frac{\alpha'}{\alpha'-1}\cdot D_{\alpha'\cdot\frac{\alpha'-1}{\alpha'-\alpha}}(P\|Q) + D_{\alpha'}(Q\|R).$$

In particular, if $D_\alpha(P\|Q) \leq \rho_1\cdot\alpha$ and $D_\alpha(Q\|R) \leq \rho_2\cdot\alpha$ for all $\alpha \in (1,\infty)$, then

$$D_\alpha(P\|R) \leq \left(\sqrt{\rho_1}+\sqrt{\rho_2}\right)^2\cdot\alpha \quad \text{for all } \alpha \in (1,\infty).$$

8. **Conversion to approximate DP:** For all measurable $S \subset \mathcal{Y}$, all $\alpha \in (1,\infty)$, and all $\tilde{\varepsilon} \geq D_\alpha(P\|Q)$,

$$\begin{aligned} P(S) &\leq e^{\tilde{\varepsilon}}\cdot Q(S) + e^{-(\alpha-1)(\tilde{\varepsilon}-D_\alpha(P\|Q))}\cdot\frac{1}{\alpha}\left(1-\frac{1}{\alpha}\right)^{\alpha-1} \\ &\leq e^{\tilde{\varepsilon}}\cdot Q(S) + e^{-(\alpha-1)(\tilde{\varepsilon}-D_\alpha(P\|Q))}. \end{aligned}$$

Definition 4.1 (Approximate RDP). A randomized algorithm $M : \mathcal{X}^n \to \mathcal{Y}$ is δ-approximately (λ,ε)-Rényi differentially private if, for all neighboring pairs of inputs $x, x' \in \mathcal{X}^n$,

$$D_\lambda^\delta(M(x)||M(x')) \leq \varepsilon.$$

4.1 Properties

1. (ε,δ)-DP is equivalent to δ-approximate (∞,δ)-RDP.
2. (ε,δ)-DP implies δ-approximate $(\lambda, \frac{1}{2}\varepsilon^2\delta)$-RDP for all $\lambda \in (1,\infty)$.
3. δ-approximate (λ,ε)-RDP implies $(\hat{\varepsilon},\hat{\delta})$-DP for

$$\hat{\delta} = \delta + \frac{\exp((\lambda-1)(\hat{\varepsilon}-\varepsilon))}{\lambda}\cdot\left(1-\frac{1}{\lambda}\right)^{\lambda-1}. \tag{4.2}$$

4. δ-approximate (λ,ε)-RDP is closed under postprocessing.
5. If M_1 is δ_1-approximately (λ,ε_1)-RDP and M_2 is δ_2-approximately (λ,ε_2)-RDP, then their composition is $(\delta_1+\delta_2)$-approximately $(\lambda,\varepsilon_1+\varepsilon_2)$-RDP.

5 Composition

5.1 Advanced composition (ε, δ)

Theorem 5.1. If each mechanism m_i is in a k-fold adaptive composition $m_1, \ldots, m_k$ satisfies ε-differential privacy, then for any $\delta' \geq 0$, the entire k-fold adaptive composition satisfies (ε', δ')-differential privacy, where

$$\varepsilon' = \varepsilon\sqrt{2k\log(1/\delta')} + k\varepsilon(e^{\varepsilon} - 1) \tag{5.1}$$

Theorem 5.2. For $j \in [k]$, let $M_j \in \mathcal{X}^n \times \mathcal{Y}_{i-1} \to \mathcal{Y}_i$ be randomized algorithms. Suppose M_j is $(\varepsilon_j, \delta_j)$-DP for each $j \in [k]$. For $j \in [k]$, inductively define $M_{1\ldots j} : \mathcal{X}^n \to \mathcal{Y}_j$ by $M_{1\ldots j}(x) = M_j(x, M_{1\ldots(j-1)}(x))$, where each algorithm is run independently and $M_{1\ldots 0} = y$ for some fixed $y_0 \in \mathcal{Y}_0$. Then $M_{1\ldots k}$ is (ε, δ)-DP for any $\delta > \sum_{j=1}^{k} \delta_j$ with

$$\varepsilon = \min\left\{\sum_{j=1}^{k}\varepsilon_j, \frac{1}{2}\sum_{j=1}^{k}\varepsilon_j^2 + \sqrt{2\log(1/\delta')\sum_{k=1}^{k}\varepsilon_j^2}\right\} \tag{5.2}$$

6 Joint Differential Privacy

Definition 6.1. For $\varepsilon, \delta \geq 0$, a randomized algorithm $\mathcal{M} : \mathbb{N}^{\mathcal{X}} \to \mathcal{Y}^N$ is (ε, δ)-joint differentially private if for every possible pair of $z, z' \in \mathcal{X}$, for every $i \in [N]$, and for every subset of possible outputs $E \subseteq \mathcal{Y}^{N-1}$, we have

$$\mathbb{P}_{\mathcal{M}}[\mathcal{M}(z \cup D_{-z})_{-i} \in E] \leq e^{\varepsilon}\mathbb{P}_{\mathcal{M}}[\mathcal{M}(z' \cup D_{-z})_{-i} \in E] + \delta \tag{6.1}$$

where $\mathcal{M}_{-i}$ denotes the output of $\mathcal{M}$ that excludes the ith dimension.

7 Lower Bound Tools

7.1 Query moment w.r.t binary data

Theorem 7.1 (Correlation–Variance Dichotomy). Let $f : \{0,1\}^d \to [0,1]$ be an arbitrary function. Let $P \in [0,1]$ be uniformly random and, conditioned on P, let $X_1, X_2, \ldots X_n$ be independent with $\mathbb{E}[X_i] = P$ for each $i \in [n]$. Then

$$\underbrace{\mathop{\mathbb{E}}_{X,P}\left[(f(X) - P)\cdot\sum_{i=1}^{n}(X_i - P)\right]}_{\text{일종의 total correlation}} + \mathop{\mathbb{E}}_{P}\left[\mathop{\mathbb{E}}_{X}\left[f(X) - \overline{X}\right]^2\right] \geq \frac{1}{12} \tag{7.1}$$

8 Decomposition

8.1 Basic decomposition

Let P and Q be probability distributions over $\mathcal{Y}$. Fix $\varepsilon, \delta \geq 0$. Suppose that, for all measurable $S \subset \mathcal{Y}$, we have $P(S) \leq e^{\varepsilon}Q(S) + \delta$ and vice versa.
Then there exist $\delta' \in [0, \delta]$ and distributions P', Q', P'' and Q'' over $\mathcal{Y}$ such that the following three properties are all satisfied.

1. We can express P and Q as convex combinations:

$$P = (1 - \delta')P' + \delta' P''$$
$$Q = (1 - \delta')Q' + \delta' Q''$$

2. Second, for all measurable $S \subset \mathcal{Y}$, we have $e^{-\varepsilon} P'(S) \leq Q'(S) \leq e^{\varepsilon} P'(S)$

3. There exists measurable $S, T \subset \mathcal{Y}$ such that $P''(S) = 1, Q''(T) = 1, \forall S' \subset S\ P(S') \geq Q(S')$, and $\forall T' \subset T\ Q(T') \geq P(T')$

Corollary 8.1. Let P and Q be probability distribution over $\mathcal{Y}$. Fix ε, δ. Suppose that for all measurable $S \subset \mathcal{Y}$, we have $P(S) \leq e^{\varepsilon} Q(S) + \delta$ and $Q(S) \leq e^{\varepsilon} P(S) + \delta$. Then there exist distributions A, B, P'', and Q'' over $\mathcal{Y}$ such that

$$P = (1-\delta)\frac{e^{\varepsilon}}{e^{\varepsilon}+1}A + (1-\delta)\frac{1}{e^{\varepsilon}+1}B + \delta P'',$$
$$Q = (1-\delta)\frac{e^{\varepsilon}}{e^{\varepsilon}+1}B + (1-\delta)\frac{1}{e^{\varepsilon}+1}A + \delta Q''$$

Interpretation: All (ε, δ) DP distributions can be represented as a postprocessing of the (ε, δ) randomized response with the postprocessing F such that $F(0, \perp) = A, F(1, \perp) = B, F(0, \top) = P''$ and $F(1, \top) = Q''$ subsubsectionBayesian version

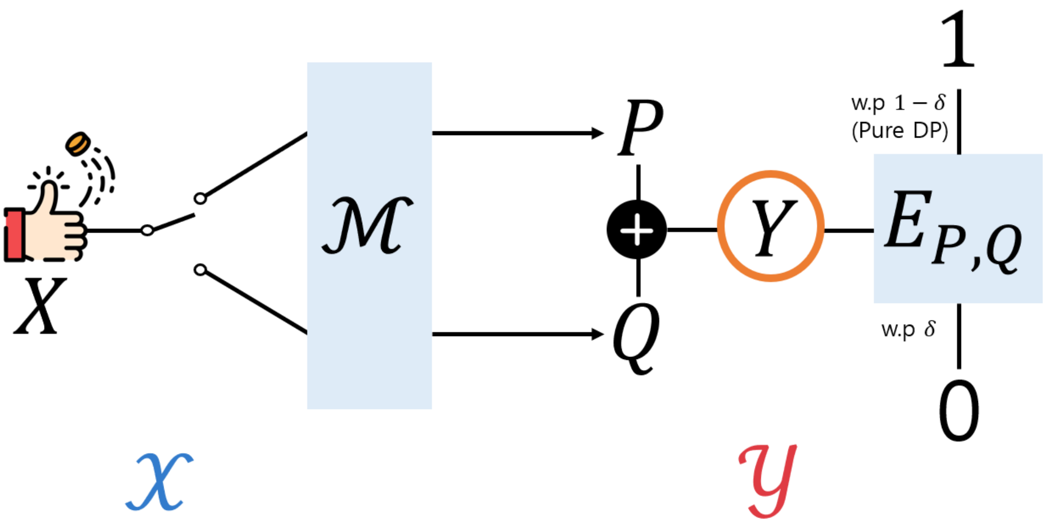

Figure 1: Visualization of the Bayesian version decomposition

Question

Suppose we observe a sample from either P or Q and we have a prior on these two possibilities, what is the posterior distribution of possibilities? We need to account for the event with δ where things "fail" arbitrarily.

Let P and Q be probability distributions over $\mathcal{Y}$. Fix $\varepsilon, \delta \geq 0$. Suppose that, for all measurable $S \subset \mathcal{Y}$, we have $P(S) \leq e^{\varepsilon} Q(S) + \delta$ and vice versa.
Then there exists a randomized function $E_{P,Q} : \mathcal{Y} \to \{0, 1\}$ with the following properties:

1. Fix $p \in [0, 1]$ and suppose $X \sim \text{Bernoulli}(p)$. If $X = 1$, sample $Y \sim P$ else $Y \sim Q$. Then for all $Y \in \mathcal{Y}$, we have

$$\underset{\substack{X \sim \text{Bernoulli(p)} \\ Y \sim XP + (1-X)Q}}{\mathbb{P}} [X = 1 \wedge E_{P,Q}(Y) = 1 | Y = y] \leq \frac{p}{p + (1-p)e^{-\varepsilon}}$$

2. Under each hypothesis $Y \sim P$ and $Y \sim Q$, the expected value $E_{P,Q}(Y)$ is equal or greater than $1 - \delta$.