Section 5: The Gaussian Mechanism, Synthetic Data, and the Exponential Mechanism

CS 208 Applied Privacy for Data Science, Spring 2022

March 1, 2022

1 Agenda

- Discuss any questions about problem sets.
- Explain differences in definitions of DP (e.g., pure vs. approximate vs. zero-concentrated).
- Re-introduce the Gaussian Mechanism.
- Generating synthetic data via histograms.
- Review the exponential mechanism.

2 Differential Privacy Review

Definition 2.1 ((ϵ, δ)-Differential Privacy)**.** A randomized mechanism M is (ϵ, δ)**-differentially private** if, for all databases x, x' differing on one row, for all queries q, and for all sets $T \subseteq \text{Range}(M)$:

$$\Pr[M(x,q) \in T] \leq e^{\epsilon} \cdot \Pr[M(x',q) \in T] + \delta,$$

where $\delta \in [0,1]$.

We will require that δ be cryptographically negligible. e.g., $\delta \leq n^{-\omega(1)}$.

An equivalent way to define (ϵ, δ)-differential privacy is via the *smoothed max divergence*, as shown below

Definition 2.2 ((ϵ, δ)-Differential Privacy)**.** A randomized mechanism M is (ϵ, δ)**-differentially private** if, for all databases x, x' differing on one row, for all queries q, and for all sets $T \subseteq \text{Range}(M)$:

$$D_{\infty}^{\delta}(M(x,q) \,\|\, M(x',q)) \leq \epsilon,$$

where $\delta \in [0,1]$ and $D_{\infty}^{\delta}(Y \,\|\, Z) = \max_{T \subseteq \text{Supp}(Y) : \Pr[Y \in T] \geq \delta} \left[\ln \frac{\Pr[Y \in T] - \delta}{\Pr[Z \in T]} \right]$.

Definition 2.3 (Privacy Loss)**.** The quantity below is the privacy loss incurred by observing an output $r \sim M(x)$. For neighboring databases x, x', the privacy loss is defined as

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

The privacy loss can be positive or negative depending on whether the output is more likely under x or x'; however, we will mostly care about the absolute value of the privacy loss.

3 z-CDP

Divergence measures (e.g., Renyi or smoothed max divergence) generalize ways to measure closeness of distributons.

Definition 3.1 (Renyi-divergence)**.** Let P and Q be probability distributions on $\mathcal{X}$. For $\alpha \in (1, \infty)$, the *Renyi divergence* of order α between P and Q is

$$D_\alpha(P \,\|\, Q) = \frac{1}{\alpha - 1} \log \left(E_{x \sim P} \left[(P(x)/Q(x))^{\alpha - 1} \right] \right),$$

where $P(\cdot)$ and $Q(\cdot)$ are the probability mass/density functions of P and Q, respectively.

Using Renyi divergence, we can define zCDP.

Definition 3.2 (ρ-zCDP (Zero-Concentrated Differential Privacy) (Bun-Steinke, 2016))**.** A randomized mechanism M is *ρ-zCDP* if for all databases $x \sim x'$ differing on one row, for all queries q, and for all $\alpha \in (1, \infty)$:

$$D_\alpha(M(x, q) \,\|\, M(x, q)) \leq \rho \cdot \alpha,$$

where $D_\alpha(M(x, q) \,\|\, M(x, q))$ is the Renyi divergence of order α between the distribution of $M(x, q)$ and the distribution of $M(x', q)$.

As α goes to ∞, Definition 3.2 reduces to pure differential privacy (Definition 2.1 with $\delta = 0$).

3.1 Interpreting the three definitions of DP

ϵ-DP tells us that the privacy loss is within ϵ for *every* output of the mechanism M:

$$\Pr[\mathcal{L} > \epsilon] = 0.$$

(ϵ, δ)-DP is a relaxation of $(\epsilon, 0)$-DP such that the privacy loss $\mathcal{L}$ is guaranteed to be bounded by ϵ with probability at least $1 - \delta$:

$$\Pr[\mathcal{L} > \epsilon] \leq \delta$$

ρ-zCDP, an even further relaxation, requires strong tail bounds for the privacy loss random variable. The definition implies that for all $\lambda > 0$,

$$\Pr[\mathcal{L} > \lambda + \rho] \leq \exp\left(-\frac{\lambda^2}{4\rho} \right)$$

Intuitively, zCDP requires that the privacy loss follow a distribution similar to the Gaussian distribution, and is concentrated around 0. zCDP is useful for deploying and composing the Gaussian Mechanism.

4 The Gaussian Mechanism

The *Gaussian Mechanism* adds noise from a Gaussian distribution to a query in order to satisfy (ϵ, δ)-DP, ρ-zCDP, or other DP variants. *It does not satisfy pure differential privacy!*

For any database $x \in \mathcal{X}^n$ and query q, the Gaussian Mechanism is $q(x) + \mathcal{N}(0, \sigma^2)$ for some $\sigma > 0$. Let GS_q be the global sensitivity of the query q. Then, we have the following theorems about how the parameters from the Gaussian mechanism relate to the privacy guarantees it satisfies.

Theorem 4.1. *Let $\epsilon \in (0,1)$. For any $c^2 > 2\ln(1.25/\delta)$, the Gaussian Mechanism is (ϵ,δ)-differentially private if $\sigma \geq cGS_q/\epsilon$.*

Theorem 4.2. *The Gaussian Mechanism is ρ-zCDP if $\sigma^2 \geq (GS_q)^2/(2\rho)$.*

Exercise 4.3. Suppose that I have n items in the database x. I wish to compute the mean of x via the Gaussian Mechanism. Suppose that for all $i \in [n]$, $x_i \in [-b,b]$ and we need to satisfy (ϵ,δ)-differential privacy via the Gaussian mechanism.

Is it possible to always get to within error of $t > 0$ with high probability? What are the necessary conditions? What if we use zCDP to implement the Gaussian Mechanism? (Hint: use the Gaussian tail bound.)

Solution. To satisfy (ϵ,δ)-differential privacy via the Gaussian Mechanism, we could do the following:

$$M(x,q) = q(x) + \mathcal{N}(0,\sigma^2), \quad \sigma^2 = \frac{2}{\epsilon^2}\ln\left(\frac{1.25}{\delta}\right)\frac{(2b)^2}{n^2},$$

since for all $i \in [n]$, $x_i \in [-b,b]$, the global sensitivity of q is $(2b)/n$.

To satisfy ρ-zCDP via the Gaussian Mechanism, we can do the following:

$$M(x,q) = q(x) + \mathcal{N}(0,\sigma^2), \quad \sigma^2 = \frac{2b^2}{\rho n^2}.$$

First, we recall the Gaussian tail bound (Claim 4.4) which is helpful for our solution.

Claim 4.4 (Gaussian Tail Bound). *Let Z be a standard normal random variable with mean 0 and variance 1. Then for every $t > 0$,*

$$\Pr[Z > t] \leq \exp(-t^2/2).$$

Since the x_i's are constants, the only source of randomness is from noise to ensure differential privacy so we can use the following to get a high probability bound:

$$\Pr\left[\frac{|M(x,q) - q(x)|}{\sigma} > t\right] \leq \exp\left(-t^2/2\right),$$

via the Gaussian tail bound.

So we need n to be large enough to always get to within error $t > 0$ with high probability.

Theorem 4.5 (Advanced Composition). *For all $\epsilon, \delta, \delta' \geq 0$, the class of (ϵ,δ)-differentially private algorithms satisfies $(\epsilon', k\delta + \delta')$-differential privacy under k-fold adaptive composition for:*

$$\epsilon' = \epsilon\sqrt{2k\ln(1/\delta')} + k\epsilon(e^\epsilon - 1).$$

Exercise 4.6. I, again, want to compute the mean of x where $i \in [n]$, $x_i \in [-b,b]$. This time I decide to add noise to the individual points x_i via the Gaussian Mechanism to satisfy (ϵ,δ)-DP. Is this better or worse than adding noise to $\bar{x}$? Why or why not?

Solution. If we add noise to the individual points x_i using advanced composition, the noise variance for the sum will be on the order of $b^2 \ln(n/\delta)/\epsilon^2$. This is clearly worse than adding noise to $\bar{x}$ with variance on the order of $b^2 \ln(n/\delta)/\epsilon^2 n^2$. Intuitively, this makes sense because we are releasing much more information if we estimate the x_i's separately.

Exercise 4.7. Suppose for all $i \in [n]$, $x_i \in [-b, b]$ for some $b > 0$. I wish to compute $f(x) = \sum_{i=1}^n x_i^2$ while satisfying zCDP. How can I use the Gaussian mechanism to accomplish this task?

Solution. The global sensitivity of f is b^2, so you would add Gaussian noise with $\sigma^2 = b^4/(2\rho)$ to satisfy ρ-zCDP.

Exercise 4.8. Let

$$X = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \cdots & \cdots \\ 1 & x_n \end{pmatrix},$$

for any $x_1, \dots, x_n \in [-b, b]$. Assume that n is public information. How would you compute $X^T X$ while satisfying ρ-zCDP for any $\rho > 0$? It is known that $X^T X$ is always PSD. If $(x_1, \dots, x_n)^T$ is not a constant vector, X is full-rank and $X^T X$ is positive definite. With noise addition to satisfy differential privacy, would $X^T X$ always remain PSD? Why or why not?

Solution. $X^T X$ is PSD since for all $z \in \mathbb{R}^2$, $z^T X^T X z = \|Xz\|^2 \geq 0$.

Compute

$$X^T X = \begin{pmatrix} 1 & n\bar{x} \\ n\bar{x} & \sum_{i=1}^n x_i^2 \end{pmatrix}.$$

Assuming that n is public, you can split your budget to $\rho_0 = \rho/2$ and then add Gaussian noise to the following terms: $\bar{x}, \sum_{i=1}^n x_i^2$.

Note that adding noise to $\bar{x}, \sum_{i=1}^n x_i^2$ could result in $X^T X$ having negative eigenvalues, resulting in non-PSD matrices.

Exercise 4.9. Write some python code to compute the mean via the Gaussian Mechanism.

5 Histograms

For answering k queries, the $\sqrt{k}$ error growth is inherent for arbitrary, worst-case sets of counting queries. For specific sets of queries, we may be able to do much better. Consider the trivial scenario of the same query being asked many times. Here, we can compute just one noisy answer, adding noise $\text{Lap}(1/\epsilon)$, and give the same answer for all the queries.

5.1 Laplace Histograms

Now consider a more interesting example: a mechanism that releases a histogram of the database, where each bin count corresponds to one query. Since the bins are disjoint, the mechanism only needs to add noise of $\text{Lap}(2/\epsilon)$ (with no dependence on k) to each query.

Theorem 5.1 (Arbitrary counting queries via the Laplace histogram)**.** *For every set Q of counting queries on data universe $\mathcal{X}$, $|Q| = k, n \in N$, and $\epsilon > 0$, there is an ϵ-differentially private mechanism $M : \mathcal{X}^n \to \mathcal{R}^Q$ such that on every dataset $x \in \mathcal{X}^n$, with high probability $M(x)$ answers all the queries to within error*

$$O\left(\frac{\sqrt{|\mathcal{X}| \log(k)}}{\epsilon n}\right)$$

Note that the dependence on $k = |Q|$ has improved from $\sqrt{k}$ obtained by advanced composition to $\sqrt{\log(k)}$. However, we have introduced a large dependence on $|\mathcal{X}|$. Thus, we must choose between using the Laplace histogram and using advanced composition based on the relative sizes of $|Q|$ and $|\mathcal{X}|$.

5.2 Synthetic Data via DP Histograms

We can use histograms to create synthetic data as follows.

1. Create singleton bins $B_y = \{y\}$ for each $y \in \mathcal{X}$
2. Construct a DP histogram $(a_1, \ldots, a_{|\mathcal{X}|}) \leftarrow M_{\text{hist}}(x)$, where $a_i \approx \{\#i : x_i = y\}$.
3. Output a synthetic dataset that has a_y copies of each element $y \in \mathcal{X}$.

As we saw earlier, the mechanism only has to add noise of $\text{Lap}(2/\epsilon)$ to each bin to achieve $(\epsilon, 0)$-differential privacy. However, there are a few problems with this approach. First, the a_y's may be negative. This can be solved by clamping, but that causes the results to be biased. Second, since the mechanism has to release an a_y for every element in the data universe, this method will be highly inefficient when $\mathcal{Y}$ is large.

Exercise 5.2. Suppose I have n items generated from the following distribution $\mathcal{N}(0, 1)$. I have a histogram with 11 bins with the left-most bin at -1 and right most bin at 1. For $n = 100, 1000, 10000$, with high probability, what is the maximum number of items that will reside in the first 2 bins and the last 2 bins? Suppose each bin is made to satisfy $(\epsilon, 0)$-differential privacy via Laplace noise. How would your answer change? What if we increase the number of bins from 11 to 1001?

6 Exponential Mechanism

The exponential mechanism is used when adding noise directly to the result of the query makes the value meaningless. This is often the case when queries have a non-numeric range.

Given some arbitrary range $\mathcal{R}$, the exponential mechanism is defined with respect to some utility function $u : \mathbb{N}^{|X|} \times \mathcal{R} \to \mathbb{R}$, which maps database/output pairs to utility scores. It is important that this range is independent of the database. The utility score has global sensitivity

$$GS_u = \max_{r \in \mathcal{R}} \max_{x \sim x'} |u(x, r) - u(x', r)|$$

Intuitively, we will output each possible $r \in \mathcal{R}$ with probability proportional to $\exp(\epsilon u(x,r)/GS_u)$, which yields privacy loss of

$$\ln \frac{\exp(\epsilon u(x,r)/GS_u)}{\exp(\epsilon u(x',r)/GS_u)} = \frac{\epsilon(u(x,r) - u(x',r))}{GS_u} \leq \epsilon$$

Note that since an additional person in the database can cause the utilities of some elements $r \in \mathcal{R}$ to increase and others to decrease, we actually add a factor of 2 to the global sensitivity.

Definition 6.1 (Exponential Mechanism)**.** The exponential mechanism $M_E(x,u,\mathcal{R})$ selects and outputs an element $r \in \mathcal{R}$ with probability proportional to $\exp(\frac{\epsilon u(x,r)}{2GS_u})$.

Theorem 6.2. *The exponential mechanism is* $(\epsilon, 0)$*-differentially private.*

Proof. For clarity, we assume the range $\mathcal{R}$ of the outputs is finite, but this is not necessary. First, we can quantify the probability that $M(x,u,\mathcal{R})$ outputs $r \in \mathcal{R}$ as

$$\Pr[M(x,u,\mathcal{R}) = r] = \frac{\exp(\epsilon u(x,r)/2GS_u)}{\sum_{r' \in \mathcal{R}} \exp(\epsilon u(x,r')/2GS_u)}.$$

We can write a similar quantity for the x', the neighboring database. Now, we take the ratio of the two to compute the privacy loss.

$$\begin{aligned}\frac{\Pr[M(x,u,\mathcal{R}) = r]}{\Pr[M(x',u,\mathcal{R}) = r]} &= \frac{\frac{\exp(\epsilon u(x,r)/2GS_u)}{\sum_{r' \in \mathcal{R}} \exp(\epsilon u(x,r')/2GS_u)}}{\frac{\exp(\epsilon u(x',r)/2GS_u)}{\sum_{r' \in \mathcal{R}} \exp(\epsilon u(x',r')/2GS_u)}} \\ &= \frac{\exp(\epsilon u(x,r)/2GS_u)}{\exp(\epsilon u(x',r)/2GS_u)} \cdot \frac{\sum_{r' \in \mathcal{R}} \exp(\epsilon u(x',r')/2GS_u)}{\sum_{r' \in \mathcal{R}} \exp(\epsilon u(x,r')/2GS_u)}.\end{aligned}$$

The first term can be re-written as

$$\frac{\exp(\epsilon u(x,r)/2GS_u)}{\exp(\epsilon u(x',r)/2GS_u)} = \exp\left(\frac{\epsilon(u(x,r) - u(x',r))}{2GS_u}\right) \leq \exp\left(\frac{\epsilon}{2}\right).$$

The second term can be re-written as the following. Note that we have re-written the numerator in terms of x using the observation that $u(x',r') \leq u(x,r') + GS_u$. This allows us to cancel out the sums in the numerator and denominator.

$$\begin{aligned}\frac{\sum_{r' \in \mathcal{R}} \exp(\epsilon u(x',r')/2GS_u)}{\sum_{r' \in \mathcal{R}} \exp(\epsilon u(x,r')/2GS_u)} &\leq \frac{\sum_{r' \in \mathcal{R}} \exp(\epsilon(u(x,r') + GS_u)/2GS_u)}{\sum_{r' \in \mathcal{R}} \exp(\epsilon u(x,r')/2GS_u)} \\ &= \exp\left(\frac{\epsilon \cdot GS_u}{2 \cdot GS_u}\right) \frac{\sum_{r' \in \mathcal{R}} \exp(\epsilon u(x,r')/2GS_u)}{\sum_{r' \in \mathcal{R}} \exp(\epsilon u(x,r')/2GS_u)} \\ &= \exp\left(\frac{\epsilon}{2}\right).\end{aligned}$$

Putting it all together, we have that the ratio of the probabilities is bounded by $\exp(\epsilon)$.

$$\frac{\Pr[M(x,u,\mathcal{R}) = r]}{\Pr[M(x',u,\mathcal{R}) = r]} \leq \exp\left(\frac{\epsilon}{2}\right) \exp\left(\frac{\epsilon}{2}\right) = \exp(\epsilon).$$

The opposite is true by symmetry. □

References

[1] Dwork, Cynthia, and Aaron Roth. ”The algorithmic foundations of differential privacy.” Foundations and Trends in Theoretical Computer Science 9.34 (2014): 211-407.