EN.601.774 Theory of Replicable ML	Spring 2025
Lecture 19	
Instructor: *Jess Sorrell*	Scribe: *Jess Sorrell*

1 Accuracy of Composed Stable Mechanisms

Last time we finished proving this wonderful theorem:

Theorem 1.1. *Bassily et al. [2016] Let $\varepsilon \in [\sqrt{\frac{12}{n}}, \frac{1}{8}]$ and $\delta \leq \frac{\varepsilon}{16}$. Let $\mathcal{M} : \mathcal{X}^m \to \mathcal{Q}$ be an (ε, δ)-private algorithm, where $\mathcal{Q}$ is the class of all queries such that $|q(S) - q(S')| \leq \frac{1}{m}$ for $|S| = m$. Then for any distribution D on $\mathcal{X}$:*

$$\Pr_{S \sim D^m q \leftarrow \mathcal{M}(S)}[|q(S) - q(D)| \geq 6\varepsilon] \leq \max\{\tfrac{4\delta}{\varepsilon}, e^{\frac{-\varepsilon^2 m}{8}}\}$$

We now want to argue that we can ensure all queries are sufficiently accurate as well!

Theorem 1.2. *Let $\mathcal{M}$ be the k-fold sequential adaptive composition of k mechanisms for answering queries. Let $(a_1, a_2, \ldots, a_k) \leftarrow \mathcal{M}(S)$ for $|S| = m$. Suppose that $\mathcal{M}(S)$ is (ε, δ)-DP and the* empirical error *of each a_j is smaller than α except with probability β. That is, for all $j \in [k]$:*

$$\Pr_{S \sim D^m, \mathcal{M}}[|a_j - \phi_j(S)| \geq \alpha] < \beta.$$

Then for every distribution D, we have

$$\Pr_{S,\mathcal{M}}[\max_{j=1}^{k} |a_j - \phi_j(D)| \geq 6\varepsilon + \alpha] \leq \beta k + \max\{\tfrac{4\delta}{\varepsilon}, e^{\frac{-\varepsilon^2 m}{8}}\}.$$

Proof.

$$\max_{j=1}^{k} |a_j - \phi_j(D)| \leq \max_{j=[k]} |a_j - \phi_j(S)| + |\phi_j(S) - \phi_j(D)|$$

We have from assumption that $|a_j - \phi_j(S)| < \alpha$ except with probability at most β. Union bounding over k queries then gives us

$$\Pr_{S \sim D^m, \mathcal{M}}[\max_{j \in [k]} |a_j - \phi_j(S)| \geq \alpha] \leq \beta k.$$

What about the second term? How do we bound $|\phi_j(S) - \phi_j(D)|$ for the worst query? Note that we can't just use a standard Chernoff-Hoeffding bound, because the data is no longer independent of the query. We'll use a monitor argument again (but simpler this time). This time our monitor will look at all of the queries ϕ_j produced by $\mathcal{M}$ and select the one that overfits the most. That is,

$$Monitor_D(\phi_1, a_1, \phi_2, a_2, \ldots, \phi_k, a_k) = \phi_{j^*}$$

where $j^* = \text{argmax}_{j\in[k]} |a_j - \phi(D)|$. Note that the Monitor algorithm is just a post-processing of an (ε,δ)-DP algorithm. Therefore the algorithm $Monitor \circ \mathcal{M}(S)$ is also (ε,δ)-DP. We just finished showing (ε,δ)-DP algorithms outputting statistical queries generalize, so we have

$$\Pr_{S\sim D^m,\mathcal{M}}[|\phi_{j^*}(S) - \phi_{j^*}(D)| \geq 6\varepsilon] \leq \max\{\tfrac{4\delta}{\varepsilon}, e^{\frac{-\varepsilon^2 m}{8}}\}$$

Putting it all together, we have

$$\begin{aligned}\Pr_{S,\mathcal{M}}[\max_{j=1}^{k} |a_j - \phi_j(D)| \geq 6\varepsilon + \alpha] &\leq \Pr_{S,\mathcal{M}}[\max_{j=1}^{k} |a_j - \phi_j(S)| + |\phi_j(S) - \phi_j(D)| \geq 6\varepsilon + \alpha] \\ &\leq \beta k + \max\{\tfrac{4\delta}{\varepsilon}, e^{\frac{-\varepsilon^2 m}{8}}\}\end{aligned}$$

□

Example: Laplace Mechanism. Recall the Laplace distribution $Lap(\varepsilon,\mu)$ has PDF $f(x) = \frac{1}{2\varepsilon}e^{\frac{-|x-\mu|}{\varepsilon}}$. We'll write $Lap(\varepsilon) = Lap(\varepsilon, 0)$.

Algorithm 1 Laplace mechanism $\mathcal{LM}(\varepsilon, S, \phi)$
Inputs/Parameters:
ε, scale parameter for Laplace distribution
$S = \{x_i\}_{i=1}^m$, dataset

1: $\nu \leftarrow Lap(\frac{1}{m\varepsilon})$
2: **return** $a = \frac{1}{m}\sum_{i=1}^m \phi(x_i) + \nu$

We ultimately want a sample complexity bound for answering k adaptive statistical queries using the Laplace mechanism, such that the answers for all queries have error at most α, and the final query has generalization error α. So we need to do the following:

1. Recall that the Laplace mechanism is $(\varepsilon, 0)$-DP
2. Use our results on the stability of the k-fold adaptive composition of DP mechanisms to say that the sequential adaptive composition of k queries is also DP
3. Use $DP \Rightarrow$ high prob generalization to fix the privacy parameter ε for the laplace mechanism
4. Obtain high probability bounds on the empirical error of the Laplace mechanism with privacy parameter ε
5. Determine sample size necessary to ensure empirical error of all queries is sufficiently small

For (2), recall that we showed the k-fold adaptive composition of $(\varepsilon, 0)$-DP mechanisms is $(\epsilon\sqrt{2k \ln 1/\delta} + k\varepsilon(e^{\varepsilon} - 1), \delta)$-DP, for all δ. Recall that we said if $\varepsilon < \frac{1}{\sqrt{k}}$ and we didn't worry too much about log factors, this is $(\tilde{O}(\varepsilon\sqrt{k}), \delta)$-DP

We want generalization error α, and our $DP \Rightarrow$ h.p. generalization result gives us generalization error bounded by 6ε, so we need $\tilde{O}(\varepsilon\sqrt{k}) \in O(\alpha) \Rightarrow \varepsilon \in O\left(\frac{\alpha}{\sqrt{k}}\right)$. For target failure rate β, we need $\frac{4\delta}{\alpha} < O(\beta)$, and $e^{\frac{-\alpha^2 m}{8}} < O(\beta)$. This means $\delta \in O(\alpha\beta)$ and $m > \frac{\log 1/\beta}{\alpha^2}$.

Now we need empirical error bounds for the Laplace mechanism!

Claim 1.3. *Let $\mathcal{LM}$ be the Laplace mechanism. For any $\alpha, \beta', \varepsilon > 0$, let $m \in O(\frac{\log(1/\beta')}{\varepsilon\alpha})$ and let $S \sim D^m$. Then with probability at least $1 - \beta'$:*

$$|a - \phi(S)| \leq \alpha$$

Proof. Note that $|a - \phi(S)| = \nu \leftarrow Lap(\frac{1}{m\varepsilon})$, so it suffices to get high probability bounds on $|\nu|$.

$$\begin{aligned}
\Pr_{\eta \sim Lap(\frac{1}{m\varepsilon})}[|\eta| \geq \tfrac{t}{m\varepsilon}] &= 2 \Pr_{\eta \sim Lap(\frac{1}{m\varepsilon})}[\eta \geq \tfrac{t}{m\varepsilon}] \\
&= 2\int_{\frac{t}{m\varepsilon}}^{\infty} \tfrac{m\varepsilon}{2} e^{-xm\varepsilon} dx \\
&= \int_{t}^{\infty} e^{-x} dx \\
&= e^{-t}
\end{aligned}$$

□

Therefore, for $m \in O(\frac{\log(1/\beta')}{\alpha\varepsilon})$, we have that

$$\Pr_{\eta \sim Lap(\frac{1}{m\varepsilon})}[|\eta| \geq \alpha] \leq \beta'$$

This brings us to (4). We've already ensured $\max\{\frac{4\delta}{\varepsilon}, e^{\frac{-\varepsilon^2 m}{8}}\} < \beta$, so we just need to ensure that β', the empirical accuracy failure rate for a single query is less than $\frac{\beta}{k}$. This implies that we need $m \in O(\frac{\log k/\beta}{\alpha\varepsilon})$. Substituting our value of ε, we have that $m \in O(\frac{\sqrt{k}\log k/\beta}{\alpha^2})$ to guarantee that except with probability β/k:

$$|a_j - \phi_j(S)| \leq \alpha$$

for any $j \in [k]$, and therefore:

$$\Pr_{S,\mathcal{M}}[\max_{j=1}^{k} |a_j - \phi_j(D)| \geq O(\alpha)] \leq O(\beta).$$

Therefore taking $m \in O(\frac{\sqrt{k}\log k/\beta}{\alpha^2})$ and using privacy parameter $\varepsilon \in O(\frac{\alpha}{\sqrt{k}})$ suffices to obtain generalization error α for the adaptive sequential composition of k statistical queries, except with probability at most β.

References

Raef Bassily, Kobbi Nissim, Adam Smith, Thomas Steinke, Uri Stemmer, and Jonathan Ullman. Algorithmic stability for adaptive data analysis. In *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*, pages 1046–1059, 2016.