EN.601.774 Theory of Replicable ML	Spring 2025
Lecture 7	
Instructor: *Jess Sorrell*	Scribe: *Jess Sorrell*

Let's consider the problem of determining the bias of a coin. Given a coin that we're promised comes up heads with probability either $1/2+\tau$ or $1/2-\tau$, how many independent flips do we need to observe to correctly guess the bias (except with probability δ)?

Notice this is a statistical query. $\mathcal{X} = \{\text{heads} : 1, \text{tails} : 0\}$, $\phi(x) = x$, and $\mathbb{E}_D[\phi] = \Pr_D[\text{heads}]$. So without replicability, we know we only need $O(\frac{\log(1/\delta)}{\tau^2})$ (we compute $\mathbb{E}_S[\phi]$. If it's $> 1/2$, we guess "heads" bias, otherwise "tails").

Theorem 0.1. *Let $\tau < 1/4$, and $\rho, \delta < 1/16$. Let $\mathcal{A}$ be an algorithm that correctly solves the coin problem except with probability δ (over the internal randomness r and choice of sample S), and such that*

$$\Pr_{S_1,S_2\sim D^m}[\mathcal{A}(S_1;r) \neq \mathcal{A}(S_2;r)] \leq \rho,$$

even if $\Pr_D[heads] \in (1/2-\tau, 1/2+\tau)$. Then $m \in \Omega(\frac{1}{\tau^2\rho^2})$.

Proof. This proof follows in 3 parts:

1. Show that there must exist a random string r^* such that $\mathcal{A}$ is accurate and replicable with high probability over $S \sim D_p$, once we fix $\mathcal{A}(\cdot, r^*)$.

2. Show that there must be some probability p^* such that $\mathcal{A}(\cdot, r^*)$ guesses heads with probability 1/2 over $S \sim D_{p^*}$. Furthermore, show that the probability $\mathcal{A}(\cdot, r^*)$ guesses heads can't change too quickly in a $O(1/\sqrt{m})$ interval around p^*.

3. Argue that $\mathcal{A}(\cdot, r^*)$ can't be replicable when it's guessing heads with probability near 1/2, and so the region in which we're guessing heads with probability near 1/2 can't be too large. We said this interval has width $O(1/\sqrt{m})$, and so m *must* be large.

Step 1. Let $D_{-\tau}$ denote a coin with bias $1/2-\tau$, let $D_{+\tau}$ denote a coin with bias $1/2+\tau$, and let D_p denote a coin with bias p. Assume we have an algorithm $\mathcal{A}(S;r)$ of sample complexity m that satisfies the above correctness guarantee. That is

- if $S \sim D^m_{-\tau}$, $\Pr_{S\sim D_{-\tau},r}[\mathcal{A}(S;r) \text{ wrong}] \leq \delta$.
- if $S \sim D^m_{+\tau}$, $\Pr_{S\sim D_{+\tau},r}[\mathcal{A}(S;r) \text{ wrong}] \geq 1-\delta$.

Let $p \in [0,1]$ denote the bias of a coin. Since $\mathcal{A}$ is ρ-reproducible, $\mathcal{A}$ is ρ-reproducible for any distribution on p. In particular, pick $p \sim \mathcal{U}([1/2-\tau, 1/2+\tau])$. By Markov's inequality, each of the following is true with probability at least 3/4 over choice of r:

- $\Pr_{S\sim D_{-\tau}}[\mathcal{A}(S;r) \text{ wrong }] \leq 4\delta$.

- $\Pr_{S\sim D_{+\tau}}[\mathcal{A}(S;r) \text{ wrong }] \geq 1-4\delta$
- When $p \sim \mathcal{U}([1/2-\tau,\, 1/2+\tau])$ uniformly, and then $S_1, S_2 \sim D_p$,

$$\Pr_{S_1,S_2}[\mathcal{A}(S_1;r) = \mathcal{A}(S_2;r)] \geq 1-4\rho.$$

To see how this follows from Markov, we'll work out the first case step by step: Let X be the random variable $X = \Pr_{S\sim D_{-\tau}}[\mathcal{A}(S;r) \text{ guesses heads}]$. Then

$$\mathbb{E}_r[X] = \mathbb{E}_r[\Pr_{S\sim D_{-\tau}}[\mathcal{A}(S;r) \text{ wrong}]] \leq \delta$$

so Markov tells us that

$$\begin{aligned}
&\Pr_r[\Pr_{S\sim D_{-\tau}}[\mathcal{A}(S;r) \text{ wrong}] \geq 4\delta] \\
&\quad \leq \Pr_r[\Pr_{S\sim D_{-\tau}}[\mathcal{A}(S;r) \text{ wrong}] \geq 4\,\mathbb{E}_r[X]] \\
&\quad = \Pr_r[\Pr_{S\sim D_{-\tau}}[\mathcal{A}(S;r) \text{ wrong}] \geq 4\,\mathbb{E}_r[\Pr_{S\sim D_{-\tau}}[\mathcal{A}(S;r) \text{ wrong}]]] \\
&\quad \leq \frac{1}{4}.
\end{aligned}$$

By a union bound over these three cases, we see that there must exist an r^* such that once we fix the algorithm to run with that randomness r^*, all three cases above hold.

Step 2. Want to show that $\Pr_{S\sim D_p^m}[\mathcal{A}(S;r^*) = 1] \in \Theta(1)$ for $p \in I$, where $|I| \in \Omega(\frac{1}{\sqrt{m}})$ and $I \subset (\frac{1}{2}-\tau, \frac{1}{2}+\tau)$.

If we can, then we know that for all $p \in I$

$$\Pr_{S_1,S_2\sim D_p^m}[\mathcal{A}(S_1;r^*) \neq \mathcal{A}(S_2;r^*) \mid p \in I] \in \Theta(1)$$

But we showed that when $p \sim \mathcal{U}([1/2-\tau \text{ and } 1/2+\tau])$ uniformly, and then $S_1, S_2 \sim D_p$,

$$\Pr_{S_1,S_2}[\mathcal{A}(S_1;r^*) \neq \mathcal{A}(S_2;r^*)] < 4\rho.$$

Then

$$\begin{aligned}
4\rho &> \Pr_{S_1,S_2\sim D_p}[\mathcal{A}(S_1;r^*) \neq \mathcal{A}(S_2;r^*)] \\
&= \Pr_{S_1,S_2\sim D_p}[\mathcal{A}(S_1;r^*) \neq \mathcal{A}(S_2;r^*) \mid p \in I] \cdot \Pr[p \in I] \\
&\qquad + \Pr_{S_1,S_2\sim D_p}[\mathcal{A}(S_1;r^*) \neq \mathcal{A}(S_2;r^*) \mid p \notin I] \cdot \Pr[p \notin I] \\
&\geq \Pr_{S_1,S_2\sim D_p}[\mathcal{A}(S_1;r^*) \neq \mathcal{A}(S_2;r^*) \mid p \in I] \cdot \Pr[p \in I] \\
&\in \Theta(\Pr[p \in I]) \\
&\in \Theta(\tfrac{1}{\tau\sqrt{m}}) \Rightarrow m > \tfrac{1}{\tau^2\rho^2}
\end{aligned}$$

Define some shorthand notation:
The probability that $\mathcal{A}$ guesses "heads" when j of its m flips come up heads:

$$a_j = \Pr_{S\sim D_p^m}[\mathcal{A}(S;r^*) = 1 | \sum_{x\in S} x = j]$$

The probability that $\mathcal{A}$ guesses "heads" when given a sample S of m flips from D_p^m

$$H(p) = \Pr_{S\sim D_p^m}[\mathcal{A}(S;r^*) = 1]$$

Note that

$$\begin{aligned}
\mathrm{H}(p) &= \Pr_{S\sim D_p^m}[\mathcal{A}(S;r^*) = 1] \\
&= \sum_j a_j \cdot \Pr_S[\sum_{x\in S} x = j] \\
&= \sum_j a_j \binom{m}{j} p^j (1-p)^{m-j}.
\end{aligned}$$

Things we know from accuracy and assuming $\delta < 1/16$:

- $\mathrm{H}(1/2 - \tau) < 4\delta < 1/4$
- $\mathrm{H}(1/2 + \tau) > 1 - 4\delta > 3/4$

This is a continuous and differentiable function, and so there must be some $p^* \in (1/2 - \tau, 1/2 + \tau)$ with $\mathrm{H}(p^*) = 1/2$. We also know that, because we assumed $\tau < 1/4$, that $(\frac{1}{2} - \tau, \frac{1}{2} + \tau) \in (1/4, 3/4)$. So we'll bound the derivative in this interval.

Now we take the derivative of H with respect to p

$$
\begin{aligned}
\mathrm{H}'(p) &= \sum_j a_j \binom{m}{j} (j p^{j-1}(1-p)^{m-j} - (m-j)p^j(1-p)^{m-j-1}) && \text{product rule} \\
&= \sum_j a_j \binom{m}{j} p^j (1-p)^{m-j} (\frac{j}{p} - \frac{m-j}{1-p}) && \text{factor out } p^j(1-p)^{m-j} \\
&= \sum_j a_j \binom{m}{j} p^j (1-p)^{m-j} \frac{j-mp}{p(1-p)} && \text{collect terms} \\
&\le \sum_j \binom{m}{j} p^j (1-p)^{m-j} \frac{j-mp}{p(1-p)} && a_j \le 1 \\
&= \sum_j \Pr_S[\sum_{x \in S} x = j] \cdot \frac{j-mp}{p(1-p)} \\
&= \mathbb{E}_j \left[\frac{j-mp}{p(1-p)} \right] \\
&\le \mathbb{E}_j [6(j-mp)] && \tfrac{1}{4} < p < \tfrac{3}{4},\ \text{so } p(1-p) \ge \tfrac{1}{6} \\
&\le 6\, \mathbb{E}_j [|j-mp|] \\
&\in O(\sqrt{m}) && \mathbb{E}_j[|j-mp|] \le \sqrt{\mathsf{Var}(j)}
\end{aligned}
$$

□