CS59200-PRS: Pseudorandomness Aug 18, 2025

Lecture 4: δ-biased distributions

Lecturer: Wei Zhan *Scribe: Abhigyan Dutta*

1 δ-biased distribution:

Definition 1. *A probability distribution $p : \{\pm 1\}^n \to \mathbb{R}$ is δ-biased if $|\widehat{p}(S)| \leq 2^{-n}\delta, \forall S \neq \varnothing$.*

Equivalently,

Definition 2. *A probability distribution $p : \{\pm 1\}^n \to \mathbb{R}$ is δ-biased if p δ-fools all characteristic function χ_S, i.e.*

$$\left| \mathop{\mathbb{E}}_{x\sim p} [\chi_S(x)] - \mathop{\mathbb{E}}_{x\sim\{\pm 1\}^n} [\chi_S(x)] \right| \leq \delta.$$

To prove the equivalence, it suffices to note that $\mathop{\mathbb{E}}_{x\sim p} [\chi_S(x)] = 2^n \widehat{p}(S)$, while $\mathop{\mathbb{E}}_{x\sim\{\pm 1\}^n} [\chi_S(x)] = 0$ whenever $S \neq \varnothing$.

1.1 Construction by Alon-Goldreich-Hastad-Peralta

We will use the following property of $\mathbb{F}_{2^l}$: There exists an isomorphism between the additive groups $\pi \colon \mathbb{F}_{2^l} \to \mathbb{F}_2^l$, that is $\pi(a+b) = \pi(a) + \pi(b), \pi(0) = (0, \ldots, 0)$.

Construction: We uniformly draw the seeds $s \in \mathbb{F}_{2^l}$ and $s' \in \mathbb{F}_2^l$, and consider the distribution of $(X_1, \ldots, X_n) \in \mathbb{F}_2^n$ where

$$X_i = \langle \pi(s^i), s' \rangle.$$

Then for every parity function over $T \subseteq [n]$ we have,

$$\sum_{i\in T} X_i = \left\langle \pi\left(\sum_{i\in T} s^i\right), s' \right\rangle.$$

Since s' is uniform in $\mathbb{F}_2^l$, the inner product is uniform over $\{0,1\}$ unless the $\pi\left(\sum_{i\in T} s^i\right) = 0$, so we have

$$\Pr\left[\sum_{i\in T} X_i = 0\right] = \frac{1}{2} + \frac{1}{2}\Pr\left[\pi\left(\sum_{i\in T} s^i\right) = 0\right] = \frac{1}{2} + \frac{1}{2}\Pr\left[\sum_{i\in T} s^i = 0\right] \leq \frac{1}{2}\left(1 + \frac{n}{2^l}\right),$$

where the last inequality is due to $i \leq n-1$ and hence the polynomial $\sum_{i \in T} s^i$ having at most n roots and s being uniformly sampled from a field of size 2^l. Since the distribution is δ-biased when $\left|\Pr\left[\sum_{i \in T} X_i = 0\right] - \Pr\left[\sum_{i \in T} X_i = 1\right]\right| \leq \delta$, we need $\frac{n}{2^l} \leq \delta$, or the seed length $2l = O(\log(n/\delta))$.

What functions are fooled by δ-biased distribution? Recall that p ε-fools f if $\left|\sum_{S \neq \varnothing} \widehat{p}(S)\widehat{f}(S)\right| \leq 2^{-n}\varepsilon$. Since $|\widehat{p}(S)| \leq 2^{-n}\delta$, we immediately have:

Lemma 1. *If $\left|\sum_{S \neq \phi} \widehat{f}(S)\right| \leq c$, f is $c\delta$-fooled by every δ-biased distribution.*

2 Total Variation Distance:

For two distributions on the same ground set $p, q : S \to \mathbb{R}$ we define the total variation distance to be the best amount that an $\{0,1\}$-function on S can distinguish between them:

$$d_{TV}(p,q) = \max_{A\colon S \to \{0,1\}} \left| \mathop{\mathbb{E}}_{x \sim p}[A(x)] - \mathop{\mathbb{E}}_{x \sim q}[A(x)] \right|.$$

We also define the l_1 distance between two distributions p, q as,

$$|p - q|_1 = \sum_{x \in S} |p(x) - q(x)|.$$

Theorem 2.

$$d_{TV}(p,q) = \frac{1}{2}|p - q|_1.$$

Proof. By optimizing the function $A\colon S \to \{0,1\}$, we have

$$\begin{aligned} d_{TV}(p,q) &= \max_{A\colon S \to \{0,1\}} \left| \mathop{\mathbb{E}}_{x \sim p}[A(x)] - \mathop{\mathbb{E}}_{x \sim q}[A(x)] \right| \\ &= \max_{A\colon S \to \{0,1\}} \left| \sum_{x \in S} A(x)(p(x) - q(x)) \right| \\ &= \max\left(\sum_{\substack{x \in S \\ p(x) \geq q(x)}} (p(x) - q(x)), \sum_{\substack{x \in S \\ p(x) < q(x)}} (q(x) - p(x)) \right). \end{aligned}$$

The two sums in the last line are actually equal, since

$$\sum_{\substack{x \in S \\ p(x) \geq q(x)}} (p(x) - q(x)) = \sum_{\substack{x \in S \\ p(x) \geq q(x)}} p(x) + \sum_{\substack{x \in S \\ p(x) < q(x)}} q(x) - 1 = \sum_{\substack{x \in S \\ p(x) < q(x)}} (q(x) - p(x)).$$

And their sum is exactly $|p - q|_1$. Therefore $d_{TV}(p,q) = \frac{1}{2}|p - q|_1$. □

The range of the distinguishing function can also vary. For instance, when A is allowed to take value in the interval $[0,1]$, from the proof above we see that A id still optimized when taking $\{0,1\}$ values. Therefore,

$$\max_{A\colon S\to[0,1]} \left| \mathop{\mathbb{E}}_{x\sim p}[A(x)] - \mathop{\mathbb{E}}_{x\sim q}[A(x)] \right| = d_{TV}(p,q).$$

If instead, A is allowed to take value in the interval $[-1,1]$, we can think of $A(x) = 2A'(x)-1$ for $A'\colon S\to[0,1]$, and thus

$$\max_{A\colon S\to[-1,1]} \left| \mathop{\mathbb{E}}_{x\sim p}[A(x)] - \mathop{\mathbb{E}}_{x\sim q}[A(x)] \right| = \max_{A'\colon S\to[0,1]} \left| \mathop{\mathbb{E}}_{x\sim p}[2A'(x)-1] - \mathop{\mathbb{E}}_{x\sim q}[2A'(x)-1] \right| = 2d_{TV}(p,q).$$

Theorem 3. *If $p\colon \{\pm1\}^n \to \mathbb{R}$ is δ-biased, then $d_{TV}(p,u) \le 2^{n/2}\delta$, where u represents the uniform distribution.*

Proof. By definition, $d_{TV}(p,u) \le 2^{n/2}\delta$ iff p could $2^{n/2}\delta$-fool every $A\colon \{\pm1\}^n \to \{0,1\}$. We now show $\left|\sum_{S\neq\varnothing} \widehat{A}(S)\right| \le 2^{n/2}$ and the theorem immediately follows from Lemma 1.

We have, by Parseval's identity,

$$\left|\sum_{S\neq\varnothing} \widehat{A}(S)\right| \le \sqrt{2^n-1}\sqrt{\sum_{S\neq\varnothing} \widehat{A}^2(S)} \le \sqrt{2^n}\sqrt{\sum_{S\subseteq[n]} \widehat{A}^2(S)} = \sqrt{2^n}\sqrt{\mathop{\mathbb{E}}_{x\sim\{\pm1\}^n} A^2(x)} \le 2^{n/2},$$

as desired. □

3 ε-almost k-wise uniformity

Definition 3. *Random variables $X_1,\dots,X_n \in S$ are ε-almost k-wise uniform if*

$$d_{TV}\left((X_{i_1},\dots,X_{i_k}), u^k\right) \le \varepsilon$$

for all distinct $i_1,\dots,i_k \in [n]$, where u^k is the uniform distribution over S^k.

Example: Consider the randomized approximate MAX-CUT algorithm, where we want to derandomize the algorithm by fooling the function $A(G,x) = \frac{1}{|E|}\sum_{(i,j)\in E} \mathbb{1}_{x_i\neq x_j}$. We showed how to do it with pairwise uniformity, and the function was fooled exactly (without any error in expectation). But since an $1/n^2$ error is allowed, almost pairwise uniformity also suffices. If $x_1,\cdots x_n$ are ε-almost pairwise uniform then,

$$\left|\mathop{\mathbb{E}}_{x}[A(G,x)] - \mathop{\mathbb{E}}_{r\sim\{0,1\}^n}[A(G,r)]\right| \le \frac{1}{|E|}\sum_{(i,j)\in E} d_{TV}((x_i,x_j),(r_i,r_j)) \le \varepsilon,$$

and $\varepsilon = 1/n^2$ suffices to find a cut of size $|E|/2$.

We can also define ε-almost k-wise independence:

Definition 4. *Random variables $X_1, \ldots, X_n \in S$ are ε-almost k-wise independent, for all distinct $i_1, \ldots i_k$, there exists a mutually independent distribution $(Y_{i_1}, \ldots, Y_{i_k})$ on S^k such that $d_{TV}((X_{i_1}, \ldots, X_{i_k}), (Y_{i_1}, \ldots, Y_{i_k})) \leq \varepsilon$.*

3.1 Fourier Analysis

There is no simple Fourier characterization of almost k-wise uniformity, but we have a necessary condition in terms of Fourier coefficients:

Theorem 4. *If $p\colon \{\pm 1\}^n \to \mathbb{R}$ is ε-almost k-wise uniform then $|\widehat{p}(S)| \leq 2^{1-n}\varepsilon, \forall 1 \leq |S| \leq k$.*

Proof.

$$|\widehat{p}(S)| = \left| \mathop{\mathbb{E}}_{x\sim p}[\chi_S(x)] \right| \cdot 2^{-n} = \left| \mathop{\mathbb{E}}_{x\sim p}[\chi_S(x)] - \mathop{\mathbb{E}}_{x\sim \{0,1\}^n}[\chi_S(x)] \right| \cdot 2^{-n} \leq 2\varepsilon \cdot 2^{-n},$$

where the last inequality follows from the d_{TV} definition of ε-almost k-wise uniform distribution and the fact that $\chi_S(x) \in \{\pm 1\}$. □

3.2 Construction from δ-biased distributions

Theorem 5 (Vazirani's XOR Lemma)**.** *If $p : \{\pm 1\}^n \to \mathbb{R}$ is δ-biased, then p is also $2^{k/2}\delta$-almost k-wise uniform for every $k \leq n$.*

Proof. It suffices to prove that, for all $A\colon \{\pm 1\}^n \to \{0, 1\}$ that depends only on k coordinates $\{i_1, \ldots, i_k\}$ of the input x, p $2^{k/2}\delta$ fools A. We only need to bound $|\sum_{S\neq\varnothing} \widehat{A}(S)|$.

Claim 6. *If $S \not\subseteq \{i_1, \ldots, i_k\}$ then $\widehat{A}(S) = 0$.*

To prove the claim, notice that on the coordinates $\{i_1, \ldots, i_k\}$ we can already perform Fourier expansion of A and write A as a polynomial in $x_{i_1}, \ldots, x_{i_k}$. Since Fourier expansion is unique, it would the same polynomial when we perform Fourier expansion of A on the entire n coordinates.

Thus we have, again by Parseval's identity

$$\left| \sum_{S\neq\varnothing} \widehat{A}(S) \right| \leq \sqrt{2^k} \sqrt{\sum_{S\subseteq\{i_1,\cdots,i_k\}} \widehat{A}^2(S)} \leq \sqrt{2^k} \sqrt{\mathop{\mathbb{E}}_{x\sim\{\pm 1\}^n} A^2(x)} \leq 2^{k/2}.$$

Now we can use Lemma 1 to conclude our theorem. □

Since $\varepsilon = 2^{k/2}\delta$, the construction of the δ-biased distribution gives a construction of ε-almost k-wise uniform distribution with seed length $O(\log(n/\delta)) = O(k + \log n + \log(1/\varepsilon))$. There are several better constructions, most notably by Naor-Naor, that has seed length $O(k + \log\log n + \log(1/\varepsilon))$:

- Joseph Naor and Moni Naor. *Small-bias probability spaces: efficient constructions and applications.*