CS 6815 Pseudorandomness and Combinatorial Constructions **Fall 2019**

Lecture 2: September 3

Lecturer: Eshan Chattopadhyay *Scribe: Renee Mirka*

2.1 Introduction

Today's lecture provided a proof of the Chernoff bound, a statement of Hoeffding's bound, and a discussion of randomness efficient ways of reducing errors in randomized algorithms.

2.2 Chernoff Bound

Theorem 2.1 *Let $X_1, X_2, \ldots, X_n$ be i.i.d. $\{0,1\}$ r.v.s and $X = \sum_{i=1}^{n} X_i$. Let $\mu = \mathbb{E}[X]$. Then for any $0 < \delta < 1$, $Pr[|X - \mu| > \delta\mu] \leq 2 \cdot exp(\frac{-\mu\delta^2}{3})$.*

Proof: Recall by Markov's inequality that for any $t \geq 0$, $\Pr[e^{sX} \geq e^{st}] \leq \dfrac{\mathbb{E}[e^{sX}]}{e^{st}}$. Furthermore,

$$\mathbb{E}[e^{sX}] = \mathbb{E}[e^{s\sum_{i=1}^{n} X_i}] = \mathbb{E}[\prod_{i=1}^{n} e^{sX_i}] = \prod_{i=1}^{n} \mathbb{E}[e^{sX_i}] = (\mathbb{E}[e^{sX_1}])^n$$

where the third and fourth equalities are true due to the independence and identical distribution, respectively, of the X_i. Now, let $X_1 = 1$ with probability p and $X_1 = 0$ with probability $1-p$ for some $p \in [0,1]$. Then $\mathbb{E}[e^{sX_1}] = pe^s + (1-p) = 1 + p(e^s - 1) \leq exp(p(e^s - 1))$ (using $e^y \geq 1 + y$ for all $y \in \mathbb{R}$). Substituting this into our initial bound from Markov's inequality, we see $\Pr[e^{sX} \geq e^{st}] \leq \dfrac{exp(np(e^s-1))}{e^{st}} = \dfrac{exp(\mu(e^s-1))}{e^{st}}$ since $\mu = \mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X_i] = np$ by linearity of expectation. Now, let $t = (1+\delta)\mu$ and $s = \log(1+\delta)$. Since $\Pr[X \geq (1+\delta)\mu] = \Pr[e^{sX} \geq exp(s(1+\delta)\mu)]$, we see $\Pr[X \geq (1+\delta)\mu] \leq \dfrac{e^{\mu\delta}}{(1+\delta)^{(1+\delta)\mu}} = \left(\dfrac{e^{\delta}}{(1+\delta)^{1+\delta}}\right)^{\mu}$. From this point, one can use exponential approximations and algebraic manipulations to match the formula as stated. ∎

2.3 Hoeffding's Bound

Let $X_1, \ldots, X_n$ be independent r.v.s where X_i is supported on $[a_i, b_i]$ and $X = \sum_{i=1}^{n} X_i$. Then for any $t > 0$, $\Pr[|X - \mathbb{E}[X]| > t] \leq 2 \cdot exp\left(-\dfrac{2t^2}{\sum_{i=1}^{n}(b_i - a_i)^2}\right)$.

2.4 Randomness Efficient Ways of Reducing Errors in Randomized Algorithms

We begin this discussion with a definition of pairwise (2-wise) independence.

Definition 2.2 *$\{X_1, \ldots, X_n\}$ are 2-wise independent if for all $i \neq j$, X_i, X_j are independent random variables.*

Example 2.3 *Consider $X_1, X_2, X_3 \in \{0,1\}$. Let $Pr[X_1 = 1] = Pr[X_1 = 0] = 1/2$ and similarly for X_2. Let $X_3 = X_1 \oplus X_2$. Then, these random variables are 2-wise independent but not i.i.d.*

2.4.1 Constructing 2-wise Independent r.v.s

First fix a finite field $\mathbb{F}_p$ where p is prime, then sample a, b independently and uniformly on $\mathbb{F}_p$. For each $i \in \{0, \ldots, p-1\}$ define $X_i = ai + b$. Here $'+'$ is the field operation (mod p).

Claim 2.4 *$\{X_i\}_{i=0,\ldots,p-1}$ are 2-wise independent.*

Proof: Note each X_i is uniformly distributed on $\mathbb{F}_p$. For $i \neq j$ and any $\alpha, \beta \in \mathbb{F}_p$, $\Pr_{a,b}[X_i = \alpha,\ X_j = \beta] = \Pr_{a,b}[ai + b = \alpha,\ aj + b = \beta]$. Solving algebraically, this is equal to $\Pr[a = \frac{\alpha-\beta}{i=j}, b = \beta - aj] = \frac{1}{p^2}$. ■

A noteworthy part of this construction, is that we were able to construct p 2-wise independent (on $[p]$) r.v.'s using only $2\lceil \log p \rceil$ bits.

2.4.2 2-wise Independent r.v.s Needed to Reduce Error of Randomized Algorithms

Now let $L \in BPP$ be a language, $x \in L$ and A an algorithm for L using r bits of randomness. We know $\Pr_{y \in \{0,1\}^r}[A(x, y) = 1] \geq 2/3$, but we want to bound the probability even further to $1 - \epsilon$. (We saw in last class using i.i.d. iterations of A requires $O(r \log(1/\epsilon))$ bits.)

Start by letting $Y^1, \ldots, Y^n$ be 2-wise independent r.v.s on $\mathbb{F}_p$ where each Y^i is uniform on $\mathbb{F}_p$ and n is a parameter to be fixed later. Choose any p such that $p > \max\{2^r, n\}$. Define $Z_i = A(x, Y^i)$ ($Z_i \in \{0, 1\}$) and output the majority vote. Then $Z_1, \ldots, Z_n$ are 2-wise independent and $\mathbb{E}[Z_i] \geq 2/3$. Therefore, $Z = \sum_{i=1}^n Z_i$ implies $\mathbb{E}[Z] = \sum_{i=1}^n \mathbb{E}[Z_i] \geq \frac{2}{3}n$. Denote the algorithm that repeats A n times using $Y^1, \ldots, Y^n$ by A'. Then $\Pr[A' \text{ is wrong on } x] = \Pr[Z \leq n/2] \leq \Pr[|Z - \mathbb{E}[Z]| > n/10] \leq 100\dfrac{Var(Z)}{n^2}$ (using $2/3n - n/2 > n/10$ and Chebychev's inequality in the last step).

Claim 2.5 $Var(Z) = \sum_{i=1}^n Var(Z_i)$

Proof:

$$
\begin{aligned}
Var(Z) &= \mathbb{E}[(Z - \mathbb{E}[Z])^2] = \mathbb{E}[(\sum_{i=1}^n (Z_i - \mathbb{E}[Z_i]))^2] \\
&= \sum_{i=1}^n \mathbb{E}[(Z_i - \mathbb{E}[Z_i])^2] + 2\sum_{i<j} (\mathbb{E}[(Z_i - \mathbb{E}[Z_i])]\, \mathbb{E}[(Z_j - \mathbb{E}[Z_j])]) = \sum_{i=1}^n Var(Z_i)
\end{aligned}
$$

since $\mathbb{E}[(Z_i - \mathbb{E}[Z_i])] = 0$. ■

Continuing from where we left off before the claim, $\Pr[A' \text{ is wrong on } x] \leq 100\dfrac{Var(Z)}{n^2} \leq 100\dfrac{Var(Z_1)}{n}$. Furthermore, $Var(Z_1) = \mu \cdot (1 - \mu) \leq 2/9$, so $\Pr[A' \text{ is wrong on } x] = O(1/n)$. Choose $n = O(1/\epsilon)$ for the desired bound.

2.4.3 Error Reduction Table $(2/3 \to (1 - \epsilon))$

Let A be an algorithm for $L \in BPP$ using R bits of randomness and time T.

Error Reduction	Randomness Required	Time Required
By i.i.d. Randomness	$O(R \log(1/\epsilon))$	$O(T \log(1/\epsilon))$
By 2-wise Independent	$2 \cdot R + 2\log(1/\epsilon) + O(1)$	$O(T \cdot 1/\epsilon)$