Week 5

Announcements:

- HW1 graded, also HW1 solutions posted
- HW2 posted (shorter/easier? than HW1) DUE OCT 15
- Next Tests: Test 2 Oct 23
  Test 3 Nov 20
  Final Dec 18?

Today: Randomized Algorithms

Randomness : Probability Theory Basics

Finite sample space S

Probability Distribution over S

An Event is a subspace $A \subseteq S$.

$$\Pr_{\mathcal{D}}[A] = \sum_{s \in A} \Pr[s]$$

Let A, B be 2 events $A, B \subseteq S$.

$$\Pr[A \wedge B] = \sum_{s \in A \cap B} \Pr[s] = \Pr[A] \cdot \Pr[B|A] = \Pr[B] \cdot \Pr[A|B]$$

where $$\Pr[A|B] = \frac{\sum_{s \in A \cap B} \Pr[s]}{\sum_{s \in B} \Pr[s]} = \frac{\Pr[A \wedge B]}{\Pr[B]}$$

A, B are independent if $\Pr[A \wedge B] = \Pr[A] \cdot \Pr[B]$

Example

Sample space: all 3-bit strings $S = \{000, 001, \ldots, 111\}$

Distribution $\mathcal{D}$ over S: pick a random string in S, so $\forall s \in S \quad Pr[s] = \frac{1}{8}$

Let $A \subseteq S$ be event that 1st bit of string is 0

$A = \{000, 001, 010, 011\}$

Let $B \subseteq S$ be event that 2nd bit is 1

$B = \{010, 110, 011, 111\}$

then $Pr[A \wedge B] = Pr[A] \cdot Pr[B|A] = \frac{1}{2} \cdot \frac{1}{2}$

A and B here are independent

Let $C \subseteq S$ be event that s has exactly one '1'

$C = \{001, 010, 100\}$

Then $Pr[A \wedge C] = \underbrace{Pr[C]}_{3/8} \cdot Pr[A|C] = \frac{3}{8} \cdot \frac{1}{3} = \frac{1}{8}$

↖ A, C not independent

Random Variables

given sample space S, distrib $\mathcal{D}$

A **random variable** is a function $X : S \to \mathbb{R}$

Ex 1 $S = \{$finite set of people, $\mathcal{D}$ = uniform$\}$

R.V. X : person $\to$ height

Ex 2. $S = \{$6 sides of die, unif dist.$\}$

RV : X : value $\to$ (value)^2

Expectation of a r.v. $\mathbb{E}[X] = \sum_{s \in S} X(s) \cdot \Pr_{s \sim \mathcal{D}}[s]$

($\mathbb{E}[X]$ is a r.v.)

Ex 1 $\mathbb{E}[X]$ = avg height of people in S

Ex 2 $\mathbb{E}[S] = \frac{1}{6} 1^2 + \frac{1}{6} 2^2 + \frac{1}{6} 3^2 + \frac{1}{6} 4^2 + \frac{1}{6} 5^2 + \frac{1}{6} 6^2$

Linearity of Expectation Let $X_1, \ldots, X_N$ be r.v.'s over $\mathcal{D}$

Then $\mathbb{E}[X_1 + \cdots + X_N] = \mathbb{E}[X_1] + \cdots + \mathbb{E}[X_N]$

Example S = all n-bit strings, $\mathcal{D}$ = unif dist.

Let $X(s) = \#$ 1's in s

$$\mathbb{E}[X] = \sum_{i=0}^{n} i \cdot \frac{\binom{n}{i}}{2^n}$$

↑ defn of expectation

Better way: Let $X_i(s) = \begin{cases} 1 & \text{if } s_i = 1 \\ 0 & \text{ow} \end{cases}$

Then $X(s) = X_1(s) + \cdots + X_n(s)$,

So by linearity of exp. $\mathbb{E}[X] = \mathbb{E}[X_1 + \cdots + X_n] = \mathbb{E}[X_1] + \cdots + \mathbb{E}[X_n]$

$= n \cdot \frac{1}{2}$

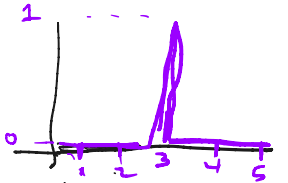
Variance: $\text{Var}[X] = \mathbb{E}[(X-\mu)^2]$ where $\mu = \mathbb{E}[X]$ = expected (average value of X

Variance is called a "second order" statistic.

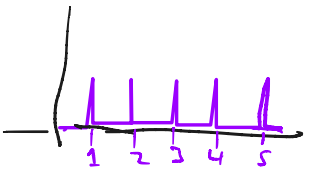
Expectation $\mathbb{E}[X]$ is avg value of X

Variance is a measure of "how much X is spread around its expected value"

Examples:



Low variance



high variance

both have $\mathbb{E}[X] = 3$

Variance = 0

Variance $= (1-3)^2 \cdot \frac{1}{5} + (2-3)^2 \frac{1}{5} + \dots + (5-5)^2 \cdot \frac{1}{5}$

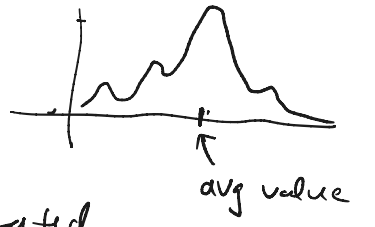
$= \frac{1}{5}[4+1+0+1+0] = 6/5$

Concentration Inequalities / Tail Bounds

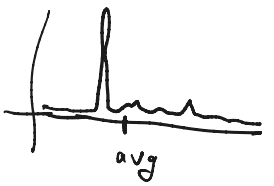
In many applications (esp in complexity theory) we want to bound the probability of an event --

For example in randomized algs we want to bound the event that $A(x,r)$ errs on input x:

show $\Pr_r[A(x,r) \text{ errs}] \leq \frac{1}{n}$ or $\frac{1}{2^n}$



(1) Simplest, most general bound (but often not concentrated enough - concentrated means very close to avg):



Markov Inequality Let $X : S \to \mathbb{R}^+$, $\mathcal{D}$ distrib over S

Then $\forall k \geq 1$ $\quad \Pr_{s \in S}[X(s) \geq k \cdot E[X]] \leq \frac{1}{k}$ ← often we write as $\Pr[X \geq k \cdot E[X]]$

Equivalently: $\Pr[X \geq K] \leq \frac{E[X]}{K}$

[so we leave out Distrib $\mathcal{D}$ over S when it is assumed]

Markov pf

since $\sum_{v} \Pr[X=v] = 1$, if $\Pr[X \geq a] > \frac{E[X]}{a}$

$$\therefore \quad E[X] = \sum_{v<a} \Pr[X=v] \cdot v + \sum_{v \geq a} \Pr[X=v] \cdot v$$

$$> 0 + a \cdot \Pr_{v \geq a} \Pr[X=v] = a \cdot \frac{E[X]}{a} = E[X]$$ Contradiction

② Chebyshev's Inequality: $\Pr[|X-\mu| \geq a] \leq \frac{\mathrm{Var}[X]}{a^2}$

Proof $\Pr[|X-\mu| \geq a] = \Pr[(X-\mu)^2 \geq a^2] \leq \frac{\mathrm{Var}[X]}{a^2}$ by Markov

③ Chernoff / Hoeffding Chernoff Inequality

Much tighter bounds but only holds for RV's that are sums of independent 0/1 valued r.v.'s

($X = X_1 + \ldots + X_m$, and $\forall i$ $Pr[X_i = 1] = p$

Chernoff Ineq Let $X = X_1 + \ldots + X_m$, each X_i indep. 0/1 r.v., $Pr[X_i = 1] = p$. (So $\mathbb{E}[X] = m \cdot p$)

Then $\forall\ 0 \le \gamma \le 1$:

$$Pr[X < (1-\gamma)mp] \le e^{-\frac{1}{2}\gamma^2 mp} \quad \text{and}$$

$$Pr[X \ge (1+\gamma)mp] \le e^{-\frac{1}{3}\gamma^2 mp}$$

Hoeffding Ineq Let $X = X_1 + \ldots + X_m$, $\hat{p} = \frac{1}{m}(X_1 + \ldots + X_m)$.

Then: $Pr[p - \hat{p} \ge \varepsilon] \le e^{-2m\varepsilon^2}$ and $Pr[\hat{p} - p \ge \varepsilon] \le e^{-2m\varepsilon^2}$

(So $Pr[|p - \hat{p}| \ge \varepsilon] \le e^{-2m\varepsilon^2}$)

p is avg value

$\hat{p}$ is sampled value of $\frac{1}{x}(X_1 + \ldots + X_m)$

Randomized Algorithms

A randomized algorithm $A(x, r)$ has 2 inputs,
x is actual input, and r is a random binary string

[r like "advice in NTM defn but NOW acceptance condition is different]

"Running $A(x, *)$": outputs a distribution over outputs, one for each r.

Note: If A runs in time $T(n)$, then without loss of generality $|r| \leq T(n)$.

Defn (RP) $L \in RP$ if $\exists$ polytime TM $M(x, r)$ [polytime in $|x|=n$]

s.t. $\forall x$:

$$x \in L \Rightarrow \Pr_r[M(x,r) \text{ accepts}] \geq \tfrac{2}{3}$$

$$x \notin L \Rightarrow \Pr_r[M(x,r) \text{ accepts}] = 0$$

} RP algorithm with error $\varepsilon = \frac{1}{3}$

one sided error!

Defn (RP) $L \in RP$ if $\exists$ polytime TM $M(x,r)$ [polytime in $|x|=n$]

s.t. $\forall x$:

$x \in L \Rightarrow \Pr_r[M(x,r) \text{ accepts}] \geq \frac{2}{3}$

$x \notin L \Rightarrow \Pr_r[M(x,r) \text{ accepts}] = 0$

one-sided error!

Defn (coRP) $L \in RP$ if $\exists$ polytime TM $M(x,r)$ [polytime in $|x|=n$]

s.t. $\forall x$:

$x \notin L \Rightarrow \Pr_r[M(x,r) \text{ accepts}] \geq \frac{2}{3}$

$x \in L \Rightarrow \Pr_r[M(x,r) \text{ accepts}] = 0$

one-sided error!

equivalently $L \in coRP$ iff $\bar{L} \in RP$

Defn [BPP] $L \in BPP$ if $\exists$ polyTIME TM $M(x,r)$ [polytime in $|x|=n$]

s.t. $\forall x$:

$x \in L \Rightarrow \Pr_r[M(x,r) \text{ accepts}] \geq \frac{2}{3}$

$x \notin L \Rightarrow \Pr_r[M(x,r) \text{ accepts}] \leq \frac{1}{3}$

2 sided error!

BPP alg with (2-sided) error $\varepsilon = \frac{1}{3}$

Error Amplification

Theorem 1 If A is an RP algorithm for L with error $\varepsilon = \frac{1}{3}$ then for all polynomials $p(n) > 0$ there is another RP algorithm A' for L with error $\left(\frac{1}{2}\right)^{p(n)}$ ← easier

Theorem 2 Same as Theorem 1 but with RP replaced by BPP

Proof (in HW2)

Idea for RP: Run $A(x, *)$ on fresh random strings $r_1 \dots r_{p(n)}$ $p(n)$ times
Accept x iff $A(x, r_i)$ accepts $\forall r_i$

Idea for BPP: Run $A(x, *)$ on fresh random strings $r_1 \dots r_{p(n)}$
Accept x iff majority answer of $\{A(x, r_1) \dots A(x, r_{p(n)})\}$ is 1

RP amplification:

Let $A \in RP$ accept L (with error ε : $x \in L \Rightarrow Pr[A(x,r) \text{ accepts}] \geq 1-\varepsilon$)

A' on x:

Run $A(x,\cdot)$ k times each independently. reject ↓ accept ↑

Let outcomes be $\{o_1, o_2, \ldots, o_k\}$, $o_i \in \{0, 1\}$

Accept x if $\exists o_i$ s.t. $o_i = 1$; Reject x ow.

Error prob of A' on x:

1. $x \in L$: $\text{Prob}[A'(x) \text{ rejects}] \leq \varepsilon^k$ (all o_i's must be 0 to reject)

2. $x \notin L$: $\text{Prob}[A'(x) \text{ accepts}] = 0$ so A' still never errs on x's not in L

RP amplification contd

So if $x \in L$ A' accepts x wp $\geq (1 - \varepsilon^k)$

$\varepsilon = \frac{1}{3} \Rightarrow 1 - \left(\frac{1}{3}\right)^k$

Runtime of A' on $x \approx$ [Runtime of $A(x, \cdot)$] $\times k$.

$\therefore$ can drive error on yes inputs to be exponentially small in $|x|$ (pick $k = n = |x|$) and still run in polytime

BPP amplification

$A'(x)$: Run $A(x,*)$ k times each independently. reject ↓ accept ↑

Let outcomes be $\{o_1, o_2, \ldots, o_k\}$, $o_i \in \{0, 1\}$

If $\underbrace{o_1 + o_2 + \ldots + o_k}_{\text{counts \# of accepts}} \geq \frac{k}{2} \Rightarrow$ accept

ow fail

Analysis : use Chernoff / Hoeffding: Fix x, $|x| = n$. Fix length of random string r.

$o_1 \ldots\ldots o_k$ }

$S = \{$all possible $r \in \{0,1\}^{poly(n)}\}$

random variables : $o_i(r) \in \{0,1\}$

Analysis: use Chernoff / Hoeffding: Fix x, $|x| = n$. Fix length of random string r.

$S = \{\text{all possible } r \in \{0,1\}^{\text{poly}(n)}\}$ $\mathcal{D}$: uniform over S

Random vars: $\{O_1, O_2, \ldots O_k\}$ $\quad O_i(r) \in \{0,1\}$ $\quad O_i(r) \in \{0,1\}$

XGL: Then $\Pr[O_i = 1] = \frac{2}{3}$, each O_i a/l indep r.v's

$O = O_1 + O_2 + \ldots + O_k$, $\mathbb{E}[O] = k \cdot \underbrace{\frac{2}{3}}_{p}$

want to show: $\Pr[O \le \frac{k}{2}]$ is small

$$= \Pr[O \le (1 - \tfrac{1}{4}) k \cdot \tfrac{2}{3}]$$

$$\le \exp\left(-\tfrac{1}{2}\left(\tfrac{1}{4}\right)^2 kp\right)$$

$$\sim \left[\frac{1}{e}\right]^{kp/32} = \left[\frac{1}{e}^{k \cdot 2/32}\right]$$

$\frac{k}{2} = (1-\gamma) kp$

$= (1-\gamma) k \frac{2}{3}$

$\Rightarrow \frac{3}{4} = 1 - \gamma$

$\Rightarrow \gamma = \frac{1}{4}$

exponentially small in k

Analysis: use Chernoff / Hoeffding: Fix x, $|x| = n$. Fix length of random string r.

$S = \{\text{all possible } r \in \{0,1\}^{\text{poly}(n)}\}$ $\mathcal{D}$: uniform over S

Random vars: $\{\theta_1, \theta_2, \ldots \theta_k\}$ $\theta_i(r) \in \{0,1\}$ $\theta_i(r) \in \{0,1\}$

$x \in L$: Then $\Pr[\theta_i = 1] = 2/3$, each θ_i d/i indep r.v's

$\theta = \theta_1 + \theta_2 + \ldots + \theta_k$, $\mathbb{E}[\theta] = k \cdot \underbrace{\frac{2}{3}}_{p}$

want to show: $\Pr[\theta \le \frac{k}{2}]$ is small

$$= \Pr[\theta \le (1 - \tfrac{1}{4}) k \cdot \tfrac{2}{3}]$$

$$\le \exp\left(-\tfrac{1}{2}\left(\tfrac{1}{4}\right)^2 kp\right)$$

$$\sim \left[\frac{1}{e}\right]^{kp/32} = \left[\frac{1}{e}^{k \cdot 2/32}\right]$$

$\frac{k}{2} = (1-\gamma) kp$

$= (1-\gamma) k \frac{2}{3}$

$\Rightarrow \frac{3}{4} = 1 - \gamma$

$\Rightarrow \gamma = \frac{1}{4}$

exponentially small in k

exercise: Do other case when $x \notin L$ and runtime of A'

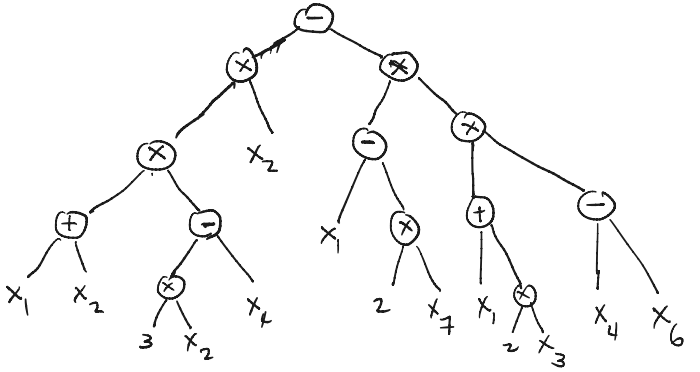
An example where Randomness seems to help:

Polynomial Identity Testing ("PIT"):

Input is an n-variable algebraic formula/expression $p(x_1 \ldots x_n)$

Example $p(x_1 \ldots x_n) = \left[(x_1 + x_2)(3x_2 - x_4)\right]x_2 - \left[x_1 - 2x_7\right]\cdot\left[(x_1 + 2x_3)(x_4 - x_6)\right]$

$p(x_1 \ldots x_n) =$



accept iff $p(x_1, \ldots, x_n) \equiv 0$ ← as a formal polynomial after cancellation

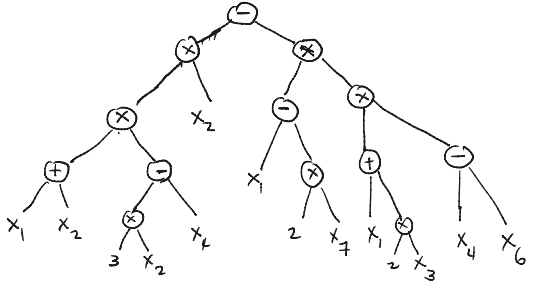
Polynomial Identity Testing ("PIT")

Example $p(x_1 \ldots x_n) = \left[(x_1+x_2)(3x_2-x_4)\right]x_2 - \left[x_1-2x_7\right]\cdot\left[(x_1+2x_3)(x_4-x_6)\right]$

operations:
binary $+, -, \times$

Input: vars x_i or an integer

$p(x_1 \ldots x_n) =$



accept iff $p(x_1, \ldots, x_n) \equiv 1$ ← as a formal polynomial after cancellation

Obvious algorithm: Expand $p(x_1 \ldots x_n)$ to write as sum of monomials, and check if $\equiv 0$

Ex. $q(x_1, x_2) = -x_1^2 - x_2^2(4x_2+2) + (x_1-2x_2)(x_1+2x_2+1) = 1$

PIT - Better Randomized Algorithm

PIT $[p(x_1 \ldots x_n)]$:

Plug in random values $a_1, \ldots, a_n$ for input vars and check if $p(a_1 \ldots a_n) = 0$

Note: can't just try one deterministically chosen value

A on input $p(x_1 \ldots x_n)$:

$|p|$ = # operations = size of algebraic formula p

Let $m = 2^{|p|}$, $M = 2^m$

Pick n random integers $a_1 \ldots a_n$ from $\{1, \ldots, M\}$

Evaluate $p(a_1 \ldots a_n)$. Accept iff $p(a_1 \ldots a_n) = 0$

(will continue next time)