HW3 out
Due : 11/14

Recap

Total Probability and Expectation

$$\mathbb{P}[B] = \mathbb{P}[A] \cdot \mathbb{P}[B \mid A] + \mathbb{P}[A^c] \cdot \mathbb{P}[B \mid A^c]$$

$$\mathbb{E}[X] = \mathbb{P}[A] \cdot \mathbb{E}[X \mid A] + \mathbb{P}[A^c] \cdot \mathbb{E}[X \mid A^c]$$

Random Variables

Bernoulli (p) : $X = \begin{cases} 1 & \text{wp } p \\ 0 & \text{wp. } 1-p \end{cases}$ $\quad \mathbb{E}[X] = p \quad \mathrm{Var}[X] = p(1-p)$

Binomial (n, p) : $Z = \underbrace{X_1 + \dots + X_n}_{\text{i.i.d Bernoulli }(p)}$ $\quad \mathbb{E}[Z] = n \cdot p \quad \mathrm{Var}[Z] = n \cdot p(1-p)$

Geometric (p) : $Y =$ position of first 1 in $X_1, X_2 \dots$ $\quad \mathbb{E}[Y] = 1/p$

Polynomial Identity Lemma

[Ore '22, Muller '54
Reed '55, Schmidt '76
DeMillo - Lipton '78
Zippel '79, Schwartz '80]

- Let $f(x_1 \ldots x_n)$ be a non-zero polynomial of degree $d \geq 0$.

$$f(x_1 \ldots x_n) = \sum_{i_1 + \ldots + i_n \leq d} c_{i_1 \ldots i_n} \cdot x_1^{i_1} \cdots x_n^{i_n}$$

over field $\mathbb{F}$. Let $S \subseteq \mathbb{F}$ and let $x_1 \ldots x_n$ be chosen independently and uniformly from S. Then,

$$P[f(x_1 \ldots x_n) = 0] \leq \frac{d}{|S|}$$

Proof: (Induction on n)

$n = 1$ $\quad P[f(x_1) = 0] \leq \dfrac{d}{|S|}$

Induction step

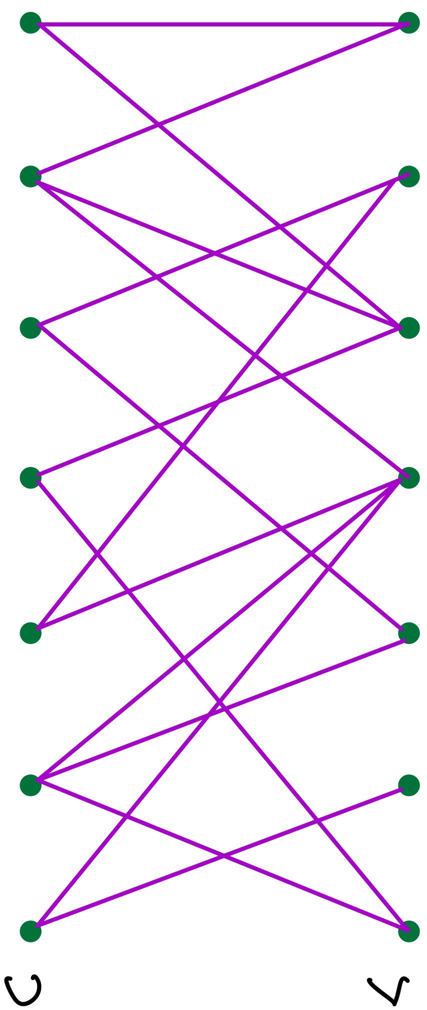
$$f(x_1, \dots, x_n) = \overset{\text{largest power of } x_n}{x_n^k} \cdot \underbrace{g(x_1, \dots, x_{n-1})}_{\deg \le d-k} + h(x_1, \dots, x_n)$$

$$A \equiv \{ g(x_1 \dots x_{n-1}) = 0 \} \qquad B \equiv \{ f(x_1, \dots, x_n) = 0 \}$$

$$\mathbb{P}[B] = \underbrace{\mathbb{P}[A]}_{\le \frac{d-k}{|S|} \text{ (induction)}} \cdot \underbrace{\mathbb{P}[B \mid A]}_{\le 1} + \underbrace{\mathbb{P}[A^c]}_{\le 1} \cdot \underbrace{\mathbb{P}[B \mid A^c]}_{\frac{k}{|S|} \text{ (base case)}}$$

$$\le \frac{d-k}{|S|} + \frac{k}{|S|}$$

Application: Testing for bipartite matching



Perfect matching: $E' \subseteq E$ s.t. each vertex has exactly one edge in E' incident to it.

$$\underbrace{A}_{\text{Tutte Matrix}} = \begin{bmatrix} & & \\ & \cdot & \\ & & \end{bmatrix} \longrightarrow \begin{cases} x_{ij} \text{ (variable)} & \text{if } (i,j) \in E \\ 0 & \text{otherwise} \end{cases}$$

Ex: G has perfect matching iff $\mathrm{Det}(A) \not\equiv 0$

Check using polynomial identity lemma

The probabilistic method [Alon & Spencer]

- Let $X : \Omega \to \mathbb{R}$ be a random variable.

  If $\mathbb{E}[X] \geq c$ then $\exists\, \omega \in \Omega$ s.t. $X(\omega) \geq c$

<u>Proof:</u> If not,

$$\mathbb{E}[X] = \sum \nu(\omega) X(\omega) < \left(\sum \nu(\omega)\right) \cdot c = c$$

To show $\exists\, \omega \in \Omega$ s.t. $X(\omega) \geq c$, it suffices

to prove $\mathbb{E}[X] \geq c$ for any measure ν of our choice.

Example: Satisfying 3-SAT clauses

Max 3-SAT : $\varphi \equiv C_1, C_2, \ldots C_m$ clauses in n vars

$$C_i = (l_{i_1} \vee l_{i_2} \vee l_{i_3})$$

$\hookrightarrow$ Can be a variable x_j or neg. $\bar{x}_j$.

e.g. $(x_1 \vee \bar{x}_9 \vee x_2), (x_5 \vee \bar{x}_2 \vee x_3), (\bar{x}_4 \vee \bar{x}_1 \vee \bar{x}_3)$

Goal : Assign $x_1 \ldots x_n \in \{\text{True}, \text{False}\}$

To satisfy maximum no. of $C_1 \ldots C_m$

- $\exists$ choice of $x_1, \ldots, x_n$ satisfying $\geq \frac{7m}{8}$ clauses

Proof: Assign $x_1 \ldots x_n$ i.i.d $\in \{\text{True}, \text{False}\}$

Z = # Satisfied clauses $\quad \mathbb{E}[Z] = \frac{7m}{8}$

Finding an assignment : Method of conditional expectations

$$\frac{7m}{8} \leq \mathbb{E}[Z] = \underbrace{P[x_1 = \text{true}]}_{1/2} \cdot \mathbb{E}[Z \mid x_1 = \text{true}] + \underbrace{P[x_1 = \text{false}]}_{1/2} \cdot \mathbb{E}[Z \mid x_1 = \text{false}]$$

- $\exists\, b_1 \in \{\text{true}, \text{false}\}$ s.t. $\mathbb{E}[Z \mid x_1 = b_1] \geq \frac{7m}{8}$.

  Can also be found efficiently

Proof: $\frac{1}{2} \cdot P[Z \mid x_1 = \text{true}] + \frac{1}{2} \cdot P[Z \mid x_1 = \text{false}] \geq \frac{7m}{8}$

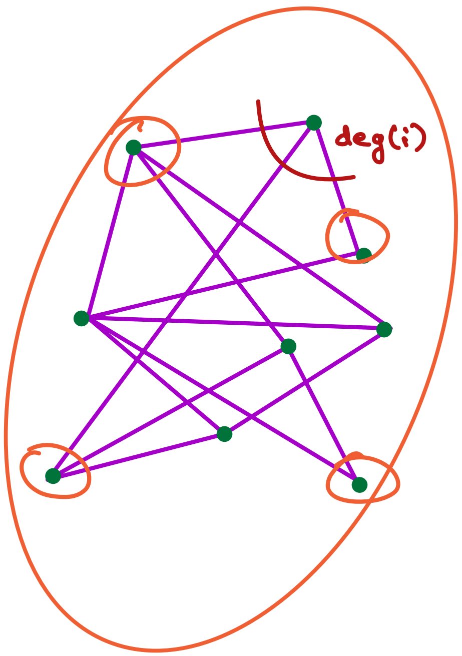
$\therefore$ one of the two above $\geq 7m/8$.

$\mathbb{E}[Z \mid x_1 = b_1]$ computable using linearity on clauses obtained plugging in $x_1 = b_1$.

Ex: (Induction) Given $b_1 \ldots b_k$ s.t. $\mathbb{E}[Z \mid x_1 = b_1 \ldots x_k = b_k] \geq \frac{7m}{8}$

Can efficiently find b_{k+1} s.t. $\mathbb{E}[Z \mid x_1 = b_1 \ldots x_{k+1} = b_{k+1}] \geq 7m/8$.

Example : Independent Sets



$G = (V, E)$

Independent Set S

$\forall (i,j) \in E$

$i \notin S$ or $j \notin S$

- For any $G = (V, E)$ $\exists$ independent set S s.t. $|S| \geq \sum_{i \in V} \frac{1}{deg(i)+1}$

Proof:

- Arrange in random order
- Select a vertex if no neighbors selected.

$E[|S|] = \sum E[X_i]$ i-th vertex selected.

$\geq \sum_i P[\text{vertex } i \text{ appears before all it's neighbors}]$

$= \sum_i \frac{(deg(i))!}{(deg(i)+1)!} = \sum_i \frac{1}{deg(i)+1}$

Tail Bounds

Probabilistic Method : $\mathbb{E}[Z] \geq c$

$$\Downarrow$$

$$\exists\, \omega \in \Omega \text{ s.t. } \underbrace{Z(\omega) \geq c}_{\text{"Good event" w.p. } > 0}$$

Tail Bounds : $\mathbb{P}[\text{"Bad event"}] \leq \epsilon$

Markov's inequality

- Let $Z : \Omega \to \mathbb{R}_{\geq 0}$ be a non-negative random variable

Then, $\forall t > 0 \quad \mathbb{P}[Z \geq t] \leq \dfrac{\mathbb{E}[Z]}{t}$

Proof: $\mathbb{E}[Z] = \sum_{v} v \cdot \mathbb{P}[Z = v]$

$$= \sum_{v \geq t} v \cdot \mathbb{P}[Z = v] + \underbrace{\sum_{v < t} v \cdot \mathbb{P}[Z = v]}_{\geq 0 \text{ since } Z \geq 0}$$

$$\geq \sum_{v \geq t} t \cdot \mathbb{P}[Z = v] + 0$$

$$= t \cdot \mathbb{P}[Z \geq t]$$

Chebyshev's inequality

- Let $Z : \Omega \longrightarrow \mathbb{R}$ be a real-valued random variable.

  Let $\mu = \mathbb{E}[Z]$. Then, $\mathbb{P}[|Z-\mu| \geq t] \leq \frac{\mathrm{Var}[Z]}{t^2}$

<u>Proof</u>: Let $Y = (Z-\mu)^2$ } non-negative r.v.

$$\mathbb{P}[|Z-\mu| \geq t] = \mathbb{P}[(Z-\mu)^2 \geq t^2]$$

$$= \mathbb{P}[Y \geq t^2]$$

$$\leq \frac{\mathbb{E}[Y]}{t^2} = \frac{\mathrm{Var}[Z]}{t^2}$$

Tail Bounds comparison

Let $Z = \underbrace{X_1 + \cdots + X_n}_{\text{i.i.d. Bernoulli}(1/2)}$

$E[Z] = \frac{n}{2}$ $\qquad$ $Var[Z] = n \cdot p(1-p) = \frac{n}{4}$

	$P[Z \geq \frac{3n}{4}] \leq ?$	$P[Z \geq \frac{n}{2} + \sqrt{n}] \leq ?$
Markov	$\leq \frac{E(Z)}{\frac{3n}{4}} \leq \frac{2}{3}$	
Chebyshev	$\leq \frac{Var[Z]}{(n/4)^2} = \frac{4}{n}$	