Lecture #2

- First and Second Moment Methods
- Probabilistic Method
- Random Graphs (Erdös-Renyi Random Graph Model).

Specific Ideas

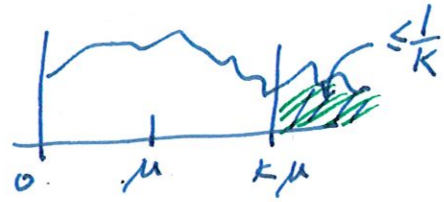
① Markov's Inequality

"first moment" mean

If X is a random variable (and non-negative) with $\mathbb{E}X = \mu$

then $Pr(X \geq K\mu) \leq 1/K$. $\forall K \geq 0$

$\Leftrightarrow \; Pr(X \geq \lambda) \leq \frac{E(X)}{\lambda} \quad \forall \lambda \geq 0$



"At most half the people can be at least twice the average".

"concentration of measure"

② Chebyshev's Inequality

X is r.v. (not necc. non-neg anymore) with mean $\mathbb{E}X = \mu$

And variance $\mathbb{E}[(X-\mu)^2] = \sigma^2$

σ = standard deviation

σ^2 = variance "second moment"

then

$$Pr(|X-\mu| \geq K\sigma) \leq \frac{1}{K^2}.$$

Probabilistic method

Want to show that some object exists.

Set up a random experiment such that

Pr(get object at the end) > 0

$\Rightarrow$ the object must exist

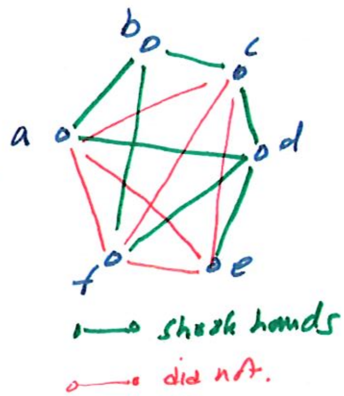
Thm 1: At any party you invite 6 people. Some handshakes happen.

Then either $\exists$ 3 people who all shook hands with each other

or $\exists$ 3 —— none of whom shook hands with any other.

Rephrase I: $\exists$ a green Δ or a red Δ.

in a 2 coloring of the complete graph K_6.



Pf: sps vtx a has green degree ≥ 3, say b, c, d

if a green edge between any pair {b,c}, {c,d}, {b,d} $\Rightarrow$ green Δ.

else no green edge $\Rightarrow$ {b, c, d} is a red Δ.

case of a having red degree ≥ 3 same.

$\boxtimes$

Drop all red ~~eg~~ edges. (No edge $\equiv$ red edge, edge = green edge)

Rephrase II :

[Every graph on 6 vertices has either a clique of size 3

or an independent set of size 3.]

Ramsey-type theorem.

Thm 2: Every graph G on n vertices has either

- an clique of size s or
- an independent set of size t

if $2^{s+t} \le n+1$

Pf: by induction on $s+t$.

if $s=1$ or $t=1$, trivial

$s=2$ or $t=2$, easy: $\begin{cases} \text{a graph has an edge (then has } K_2 = K_s) \\ \text{or else is independent set.} \end{cases}$

similar for $t=2$

Now suppose true for $s+t = k-1$.

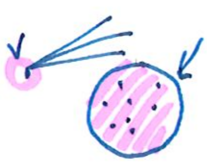
to prove for $s+t = k$, consider graph on $n \geq 2^k - 1$ vertices.

Pick v. either has $\geq \frac{n-1}{2} = 2^{k-1} - 1$ neighbors.

$\Rightarrow$ recurse. has either clique of size $s-1$ ← combine with v to get K_s clique

or I.S of size t

else is not adjacent to $\geq 2^{k-1} - 1$ vertices.



Find clique of size s or IS of size $t-1$. ← extend with v to get IS of size t. ☑

— x —

Corollary: Every graph on n vertices has either

clique of size $\frac{1}{2}\log_2 n$

or

I.S of size $\frac{1}{2}\log_2 n$

Can we do better?

No!

(up to constants at least)

How to prove

Need magical graph that has no large clique — of size $2\log n$

and no large IS too! — of size $2\log n$.

— x —

Use the probabilistic Method!

actually a random variable of "type" graph.

Consider the random graph $G(n,p)$ $\quad n \in \mathbb{Z}_{\geq 1}$, $p \in [0,1]$.

n vertices

each of the $\binom{n}{2}$ possible edges present independently with prob. p.

— x —

Thm 4: $G \sim G(n,\frac{1}{2})$ has no indep. set or clique of size $2\log n \underbrace{+1}_{\text{for simplicity}}$. with high prob.

"G sampled from the distribution described above"

Pf: $\forall\, S \subseteq V$ define indicator random variable

$$X_S = \begin{cases} 1 & \text{if } S \text{ clique in } G \\ 0 & \text{ow.} \end{cases} \iff X_S = \mathbb{1}(S \text{ is a clique in } G)$$

① $\mathbb{E}[X_S] = \left(\frac{1}{2}\right)^{\#\text{edges in } S} = \left(\frac{1}{2}\right)^{\binom{|S|}{2}}$. (independence of each edge) $p=\frac{1}{2}$

if $|S| = 2\log n + 1 = s \Rightarrow \quad \mathbb{E}X_S = \left(\frac{1}{2}\right)^{\frac{s(s-1)}{2}}$

② $\mathbb{E}[\#\text{ of cliques of size } s] = \binom{n}{s}\left(\frac{1}{2}\right)^{\frac{s(s-1)}{2}}$ (linearity of expectations)

$$\mathbb{E}\Big[\sum_{\substack{S\subseteq V\\|S|=s}} X_S\Big] = \sum_{\substack{S\subseteq V\\|S|=s}} \mathbb{E}X_S$$

Let $X = \#$ of cliques of size s $= \sum_{|S|=s} X_S$

<u>Some algebra</u>: $\mathbb{E}X = \binom{n}{s} \cdot 2^{-\frac{s(s-1)}{2}}$

$\leq \frac{n^s}{s!} \cdot \left(2^{-(s-1)/2}\right)^s$

$\left\{ \text{but } \binom{n}{s} \leq \frac{n^s}{s!} \right.$

$\left\{ \text{but } \frac{s-1}{2} = \log_2 n \right.$

$\Rightarrow 2^{-\left(\frac{s-1}{2}\right)} = 2^{-\log n} = \frac{1}{n}$

$= \frac{1}{s!} = o(1)$

goes to 0 as $n \to \infty$

(even for $s=3$, $\leq \frac{1}{3}$)

So $E(X) \leq o(1)$

Markov

$\Rightarrow \Pr(X \neq 0) = \Pr(X \geq 1) \leq E(X) \leq o(1)$.

$\Pr$(G has a t-clique)

<u>Similarly</u>: if $Y_S = \mathbb{1}(S \text{ is indep set})$ and $Y = \sum_{|S|=s} Y_S$

then $\Pr(Y \neq 0) \leq o(1)$.

<u>Final Step</u>: Union Bound. $P(A \cup B) \leq P(A) + P(B)$

$\Pr$(G has a clique of size s or an indep set of size s) $= \Pr(X \neq 0 \text{ or } Y \neq 0)$

$\leq \Pr(X \neq 0) + \Pr(Y \neq 0)$

$= o(1)$. ∎

in fact almost all graphs of this size have this property!

<u>Corollary</u>: $\exists$ a graph on n nodes with no s-clique or s-IS when $s = 2\log n + 1$.

Recap: run some random experiment.

define: for each forbidden ~~object~~ structure S, $X_S = \mathbb{1}(S \text{ exists})$

$$X := \sum_S X_S = \#\text{forbidden objects}$$

Show: $\mathbb{E}X < 1$

Markov (first moment)

Use Markov: $\Pr(X \neq 0) = \Pr(X \geq 1) \leq \mathbb{E}X < 1$

$\Rightarrow \Pr(X=0) > 0 \Rightarrow \exists$ solution with no forbidden structure.

Balls & Bins

n balls

n bins

Each ball thrown into a $\underbrace{\text{uniformly at random}}_{\text{"u.a.r"}}$ bin

Fact: $E[\text{load of bin } \#1] = (\#\text{balls}) \cdot \left(\frac{1}{\#\text{bins}}\right) = \frac{n}{n} = 1$

linearity of expectations

But what is maximum load ~~of~~ (over all bins)

Theorem: with n balls, n bins, maximum load $= O\left(\frac{\log n}{\log \log n}\right)$ whp.

Pf: S be set of $s = \frac{8 \log n}{\log \log n}$ balls

$X_S = \mathbb{1}(S \text{ goes into the same bin})$

$X = \sum' X_S.$

$$E[X] = \sum_{|S|=s} E(X_S) = \binom{n}{s} \cdot \left(\frac{1}{n}\right)^s \cdot n$$

$\#$ of S. all balls into bin i n bins

$$= n \cdot \binom{n}{s} \cdot \left(\frac{1}{n^s}\right)$$

$$\binom{n}{s} \le \frac{n^s}{s!}$$

$$\le n \cdot \frac{\cancel{n^s}}{s!} \cdot \frac{1}{\cancel{n^s}}$$

But $s! = 1.2.3.\ldots.s$

$$\ge \underbrace{(1\cdot s)(2(s-1))(3\cdot(s-3))\ldots}_{s/2 \text{ terms}} \ge s^{s/2}$$

$$\le \frac{n}{s^{s/2}}.$$

But $s = \frac{8\log n}{\log\log n}$ say. $\Rightarrow s^{s/2} \ge n^2$ (algebra)

~~please try and if you have~~

$$s^{s/2} \ge \left(\sqrt{\log n}\right)^{\frac{4\log n}{\log\log n}}$$

$$= \left(2^{\frac{1}{2}\log\log n}\right)^{\frac{4\log n}{\log\log n}}$$

$$= 2^{2\log n} = n^2$$

$$\le \frac{1}{n}$$

$\Rightarrow Pr(\exists S$ of size s that falls in same bin$)$

$= Pr[X \ge 1] \le E[X] \le 1/n.$

$\Rightarrow Pr(\text{maxload} \le s)$ w.p. $1 - 1/n$

$\blacksquare$

— X —

Good: used 1st Moment method to show.

(a) $G(n, p=1/2)$ does not have a clique of size $2\log n + 1$ whp.

(b) no bin has $\ge \frac{8\log n}{\log\log n}$ balls whp.

Are these results tight? YES!

How to prove? SECOND MOMENT METHOD!

Second Moment Method

Collection $\mathcal{C}$ of good items. We generate a random graph say.

$X_S = \mathbb{1}$(random graph contains item S from $\mathcal{C}$)

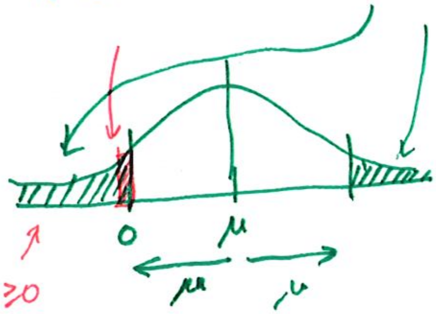
$X = \sum_{S \in \mathcal{C}} X_S$ = ~~#copies~~ good items we contain in G

Want to show $P(X \neq 0)$ with high probability $1-o(1)$

i.e. $P(X=0)$ with tiny probability $o(1)$.

$$Pr(X=0) \le Pr(|X-\mu| \ge \mu) \le \frac{\text{Variance}(X)}{\mu^2} \qquad (\mu = EX)$$

Chebyshev's Ineq.



btw $X \ge 0$

so

$Pr(X \le 0) = Pr(X=0)$

here

How does variance(X) relate to $(E(X))^2$

$E(X-\mu)^2 \qquad \mu^2$

$= E(X^2) - E(X)^2 \qquad E(X)^2$

$$\boxed{\text{to show that } Pr(X=0) = o(1) \quad \text{want that } E(X^2) = (1+o(1))(E(X))^2} \Leftrightarrow \boxed{\text{want } Var(X) = o(1) \cdot E(X)}$$

Recall: $E(X^2) \ge E(X)^2$ by Jensen's Inequality.

~~also~~:

Cliques in
(Random Graphs later or in HWs).

Let's do it for balls and bins

Thm: $\boxed{\text{The max load for n-balls n-bins is } \Omega\left(\frac{\log n}{\log\log n}\right) \text{ whp.}}$

$s = \dfrac{\log n}{3 \lg\lg n}$

OK: ~~Again, $X_S = \mathbb{1}$(all balls in S fall in~~

This time $X_i = \mathbb{1}$(some set of $\geq s$ balls falls into i). for bin i

$E(X_i) \geq \Pr(\text{Bin } i \text{ has exactly } s \text{ balls})$

$= \sum_{|S|=s} \Pr(\text{Bin } i \text{ gets exactly the balls in } S)$

$= \binom{n}{s} \cdot \left(\frac{1}{n}\right)^s \cdot \left(1-\frac{1}{n}\right)^{n-s}$

$\geq \left(\frac{n}{s}\right)^s \cdot \frac{1}{n^s} \cdot \frac{1}{4} = \frac{1}{4s^s} \geq \frac{1}{n^{1/3}}$

for our choice of s.

- $\binom{n}{s} \geq \left(\frac{n}{s}\right)^s$
- $\left(1-\frac{1}{n}\right)^{n-s} \geq \left(1-\frac{1}{n}\right)^n \geq \frac{1}{4}$

for $n \geq 2$

def $X = \sum_i X_i$
$=$ #bins with $\geq s$ balls.

$\Rightarrow E[\#\text{bins with at least } s \text{ balls}] \geq n \cdot 1/n^{1/3} = n^{2/3}$

$E(X) = \sum_i E(X_i)$

linearity of Expectation

$\boxed{\text{Great: So we expect } n^{2/3} \text{ bins to have many balls. in Expectation}}$

But can we infer that some bin will have many balls whp?

← with high prob
ie. w.p. $1-o(1)$?

Not yet!

(E.g maybe w.p. $1/2$, each bin has 1 ball
w.p. $1/2$, $2n^{2/3}$ bins have lots of balls) ← not possible, as we will see, but we need more work...

To use the second moment idea, we say:—

$$\Pr(X=0) \le \Pr(|X-EX| \ge EX) \overset{\text{Chebyshev.}}{\le} \frac{Var(X)}{(EX)^2} \le \frac{Var(X)}{n^{4/3}}.$$

<u>What is Var(X)?</u> Remember $X = \sum_{i=1}^{m} X_i$

$$Var(X) = \sum_i Var(X_i) + \sum_{i \ne j} \text{Covariance}(X_i, X_j)$$

$Cov(X_i, X_j)$
$= E[(X_i - \mu_i)(X_j - \mu_j)]$

$Var(X_i) = Cov(X_i, X_i)$
$= E[(X_i - \mu_i)^2]$

<u>Fact 1</u> In this experiment

$$Cov(X_i, X_j) \le 0$$

"X_i and X_j are negatively correlated: if $X_i = 1$ then X_j is less likely to be 1". (Proof later.)

<u>So</u>: $Var(X) \le \sum_i Var(X_i) \overset{\text{all bins identically distributed.}}{=} n \cdot Var(X_i)$

<u>Fact 2</u>: Since $X_i \in \{0,1\}$, $Var(X_i) \le E(X_i)$

<u>Pf</u>: $Var(X_i) = E(X_i^2) - E(X_i)^2 \le E(X_i^2) = E(X_i)$ ▨ since $1^2 = 1$, $0^2 = 0$.

$$\Rightarrow \Pr(X=0) \cancel{\frac{Var(X)}{(EX)^2}} \le \frac{Var(X)}{(EX)^2} \le \frac{E(X)}{(EX)^2} = \frac{1}{EX} \le \frac{1}{n^{2/3}}$$

$\Rightarrow$ with prob $1 - 1/n^{2/3}$, some bin has $\ge \dfrac{\log n}{3 \log \log n}$ balls ☺

Great: Similarly you can show (HW) that

whp, the graph $G_{n,1/2}$ contains a clique of size $(2-\varepsilon)\log n$.

[Diff: cannot just use negative correlation, need to do some work to control $E[X_S X_T]$]
(not true)

— x —

Wrap up: actually, tighter analysis can show that

#balls in most loaded bin is $(1 \pm o(1)) \cdot \frac{\ln n}{\ln \ln n}$ with high prob $1-o(1)$.

natural log $\log_e n$.

[See notes on webpage from MU book]

— x —

Cool: today saw

① Markov's Ineq : $\Pr(X \geq \lambda) \leq \frac{E(X)}{\lambda}$ $\boxed{\text{if } X \geq 0.}$

② Chebyshev : $\Pr(|X-\mu| \geq \lambda) \leq \frac{Var(X)}{\lambda^2}$ for any random var.

③ First moment method : if $E(X) = o(1)$ $X = $ #bad things

then $\underbrace{\Pr(X=0)}_{\text{zero bad things happen.}} \geq 1 - E(X) = \underbrace{1-o(1)}_{\text{whp}}$.

④ Second moment method : $X = $ #good things

if $\frac{Var(X)}{E(X)^2} = o(1)$ then

$\Pr(X=0) \leq \frac{Var(X)}{(EX)^2} = o(1) \Rightarrow \Pr(X \geq 1) \geq \overbrace{1-o(1)}^{\text{whp}}$.

at least one good thing happens