W1D1
16/11/14

Probabilistic Methods in Combinatorics

Second Moment Method

$\sigma^2 = Var(X) = \mathbb{E}((X-\mathbb{E}(X))^2)$, $\sigma = \sqrt{Var X}$ the standard deviation of X. Chebyshev – $\mathbb{P}(|X-\mathbb{E}(X)| \geq \lambda\sigma) < \frac{1}{\lambda^2}$.

Proof: $\mathbb{P}(|X-\mathbb{E}(X)| \geq \lambda\sigma) = \mathbb{P}((X-\mathbb{E}(X))^2 \geq \lambda^2 \cdot \sigma^2) \overset{Markov}{\leq} \frac{1}{\lambda^2}$.

Computing σ is (mostly) possible.

How to compute? By lin. of expectation $Var(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$.

Suppose $X = X_1 + \dots + X_n$, then $Var X = \sum_{i=1}^{n} Var(X_i) + \sum_{i \neq j}^{n} Cov(X_i, X_j)$, where $Cov(Y,Z) = \mathbb{E}(YZ) - \mathbb{E}(Y)\mathbb{E}(Z)$. If Y, Z indep. then $Cov(Y,Z) = 0$. So, if $X_1, \dots X_n$ (pairwise) indep. then $Var X = \sum_{i=1}^{n} Var X_i$.

Markov $\mathbb{P}(X > 0) \leq \mathbb{E}(X)$ where $X \geq 0$ r.v. in $\mathbb{N}$.

In general $(\mathbb{E}(X_n) \to \infty) \not\Rightarrow (\mathbb{P}(X_n > 0) \not\to 0)$ (–1,

Prop: $\mathbb{P}(X=0) \leq \frac{Var(X)}{\mathbb{E}(X)^2}$. Proof – $\mathbb{P}(X=0) \leq \mathbb{P}(|X-\mathbb{E}(X)| \leq \mathbb{E}(X)) =$ $\mathbb{P}(|X-\mathbb{E}(X)| \leq \sigma\frac{\mathbb{E}(X)}{\sigma}) \leq \frac{\sigma^2}{\mathbb{E}(X)^2}$. In particular if $Var(X_n) = o(\mathbb{E}[X_n]^2)$ then $\mathbb{P}(X_n = 0) \to 1$. In general, $\forall \varepsilon$ $\mathbb{P}([(1-\varepsilon)\mathbb{E}(X_n)] \leq X_n \leq [(1+\varepsilon)\mathbb{E}(X_n)]) \to 1$. Since the prob. of complementary event is $\frac{Var(X_n)}{\varepsilon^2 \mathbb{E}(X_n)^2} \to 0$.

Cor. Weak Law of Large Numbers – $X_1, \dots$ i.i.d. r.v.'s s.t. $\mathbb{E}(X_1^2) < \infty$. Then $\frac{X_1 + \dots + X_n}{n} \xrightarrow{\mathbb{P}} \mathbb{E}(X_1)$ (that is, $\mathbb{P}(|\frac{X_1 + \dots + X_n}{n} - \mathbb{E}(X_1)| < \varepsilon) \to 0$ for any $\varepsilon > 0$). In fact, $\mathbb{E}(X_1) < \infty$ is sufficient.

Def A set $\{x_1, \dots x_n\} \in \mathbb{Z}$ has distinct sums if all sums $\sum_{i \in I} x_i$ are different for different $I \subseteq [n]$.

Obs. For each k $\exists$ A set with distinct sums of size k, say $\{1, 2, \dots, 2^{k-1}\}$, but 2^{k-1} is pretty large.

def. $f(n) = \max_k \{[1, \dots k]$ has a size-n subset with distinct sums$\}$.

$f(k) \geq 1 + \lfloor \log_2 k \rfloor$. Since we have $2^{f(n)}$ sets with max. sum $nf(n)$, so $f(n) \leq \log_2 n + \log_2 \log_2 n + O(1)$. We'll try and get a sharper estimate.

Thm. $f(n) \le \log_2 n + \frac{1}{2}\log_2\log_2 n + O(1)$.

Open Question (Erdős) $f(n) \overset{?}{\le} \log_2(n) + O(1)$ - prove or disprove.

Pf. Using the 2^{nd} moment method. Let $k=f(n)$ and $1\le x_1,\dots x_k \le n$ a set with distinct sums. Take $\varepsilon_1,\dots,\varepsilon_k$ uniform i.i.d on $\{0,1\}$ - and look at $X=\sum_{i=1}^{n} \varepsilon_i x_i$.

$\mu = \mathbb{E}(X) = \frac{\sum^{k} x_i}{2}$, $\sigma^2 = Var\, X = \sum_{i=1}^{n} Var(\varepsilon_i x_i) = \frac{\sum^{k} x_i^2}{4} \le \frac{kn^2}{4}$. By Chebyshev (for some $\lambda>1$) $\mathbb{P}(|X-\mu|\le \lambda\sigma) \ge 1-\lambda^{-2}$. This prob. is $\frac{1}{2^k}$ times $2\lambda\sigma+1$. We get $2^k \le \frac{\lambda\sqrt{k}\,n+1}{1-\lambda^{-2}}$. For any $\lambda>1$ we get $k\le \log_2 n + \frac{1}{2}\log_2 \log_2 n + O(1)$.

Let $\nu(n) = \#$ different prime divisors of n.

Thm. If $\nu(n) \to \infty$ (arbitrarily slowly). Then $|\{x\in\{1,\dots,n\}: |\nu(x)-\log\log n| > \omega(n)\sqrt{\log\log n}\}| = o(n)$.

also n

Pf. Pick $x \sim U[1,\dots,n]$. for each p, $X_p = \mathbb{1}[p|x]$. Let $X = \sum_{p\in P} X_p = \sum_{\substack{p\in P\\ p\le n^{0.1}}} X_p$ (Note $|\nu(x)-X(x)|\le 10$). $\mathbb{E}X_p = \frac{1}{n}\lfloor\frac{n}{p}\rfloor = \frac{1}{p} - O(\frac{1}{n}) \Rightarrow$

$\mathbb{E}(X) = \sum_{p<n^{0.1}} \frac{1}{p} - O(\frac{1}{n})$ $(= \log\log n \pm O(1))$. It's enough to show the same order of magnitude for $Var\, X$ to get the thm via Chebyshev. $Var(X) = \sum_{\substack{p\in P\\ p<n^{0.1}}} Var(X_p) + \sum_{\substack{q\ne p\\ q\in P\\ q<n^{0.1}}} Cov(X_p,X_q)$. $Var(X_p) =$

$\mathbb{E}[X_p^2] - \mathbb{E}[X_p]^2 = \mathbb{E}[X_p] - \mathbb{E}[X_p]^2 = \frac{1}{p}(1-\frac{1}{p}) \pm O(\frac{1}{n})$.

$Cov(X_p,X_q) = \frac{1}{n}\lfloor\frac{n}{pq}\rfloor - \frac{1}{n^2}\lfloor\frac{n}{p}\rfloor\lfloor\frac{n}{q}\rfloor \le \frac{1}{pq} - (\frac{1}{p}-\frac{1}{n})(\frac{1}{q}-\frac{1}{n}) = \frac{1}{n}(\frac{1}{p}+\frac{1}{q}-\frac{1}{n}) \le \frac{1}{n}(\frac{1}{q}+\frac{1}{p})$. Therefore $Var(X) \le \sum \frac{1}{p} + O(1) + \frac{1}{n}\sum \frac{1}{p}+\frac{1}{q} \le \log\log n + O(1) + \underbrace{\frac{2n^{0.1}}{n}\sum \frac{1}{p}}$ $= \log\log n + O(1)$.

Random Graphs and Threshhold Functions

Take $G(n,p)$. $\mathcal{P}$ a property - a collection of graphs closed under isomorphism. $\mathcal{P}$ Montone (inc.) $\Leftrightarrow$ $H\subseteq G\in\mathcal{P} \Leftarrow H\in\mathcal{P}$.

Suppose $\mathcal{P}$ non-trivial monotone property.

Add edges at a random order ($p_0 \nearrow 1$) and find p when $\mathbb{P}(G\in\mathcal{P})$ changes from being near 0 to near 1.
for $\mathcal{P}$
A function $r:\mathbb{N}\to[0,1]$ is a <u>threshold function</u> iff

(1) $p(n) << r(n) \Rightarrow G(n,p)\notin\mathcal{P}$ a.a.s.

(2) $p(n) >> r(n) \to G(n,p)\in\mathcal{P}$ a.a.s.

Every such $\mathcal{P}$ admits a theshold function.

Let's start with some <u>examples</u>:

$\mathcal{P}$ – $\omega(G)\geq 4$. Thershold of $n^{-2/3}$.

1) suppose $p(n) << n^{-2/3}$. Let $X=\#$ copies of K_4, then

$X=\sum_{S\in\binom{[n]}{4}} X_S$. $\mathbb{E}(X)=\sum \mathbb{E}(X_S)\le n^4p^6 << 1$.

$X_S \to \mathbb{1}_S$ induces K_4: 1 with prob. p^6

By Markov, $\mathbb{P}(X=0)\to 1$.

2) When $p(n) >> n^{-2/3}$, then $\mathbb{E}(X) \overset{\to\infty}{>>} 1$. It suffices to show $Var(X)=o(\mathbb{E}(X)^2)$. $Var(X)=\sum_{S\in\binom{[n]}{4}} Var(X_S)+\sum_{\substack{t\in\binom{[n]}{4}\\ t\neq S}} Cov(X_S,X_t)$.

$\to \mathbb{E}(X_S^2)=\mathbb{E}(X_S)=p^6$

X_S,X_t indep. if $|S\cap t|\le 1$. So, $Var(X)=\sum_{S\in\binom{[n]}{4}} Var(X_S) + \sum_{\substack{2\le|S\cap t|\le 3\\ t\in\binom{[n]}{4}}} Cov(X_S,X_t)\le n^4p^6+\sum_{\substack{2\le|S\cap t|\le 3\\ t\in\binom{[n]}{4}}} cov(X_S,X_t)$. We just need to show the 2nd sum is $o(\mathbb{E}(X)^2)$. Fix S then

$$\sum_{\substack{t\in\binom{[n]}{4}\\ 2\le|S\cap t|\le 3}} \underset{p^{e|S\cap t|}-p^{12}}{Cov(X_S,X_t)}\le\sum p^{e(K_S\cup K_t)}=\underbrace{\binom{4}{2}\binom{n-4}{2}p^{11}}_{O(n^2p^{11})}+\underbrace{\binom{4}{3}\binom{n-4}{1}p^{9}}_{O(np^9)}.$$

So $\sum_S\sum_t Cov(X_S,X_t)=O(n^6p^{11})+O(n^5p^9)=p^6n^4\underbrace{[O(n^2p^5+np^3)]}_{o(p^6n^4)}$

We try and generalize

<u>Def</u> H graph with v vertices and e edges. $\frac{e}{v}$ is called the <u>density</u> ρ of H. H is <u>balanced</u> if $\forall G\subsetneq H$ $\rho(G)\le\rho(H)$. H is <u>strictly balanced</u> if $\rho(G)<\rho(H)$.

<u>Thm</u> $n^{-1/\rho(H)}=n^{-\frac{v}{e}}$ is a threshold for "G has a copy of H", Assuming H is balanced.

<u>Remark</u> the balanced assumption is crucial:  happens around $n^{-2/3}$ obviously.

<u>proof</u>: 1) Let $X = \#$ copies of H in $G(n,p)$. $X = \sum_{S \in \binom{[n]}{v}} X_S$ (S has a copy of H). Then $p^e \leq \mathbb{P}(X_S = 1) \leq v!\, p^e$. $\mathbb{E}(X) \leq \binom{n}{v} p^e v! \leq p^e n^v \begin{cases} << 1 & \text{if } p << n^{-1/\rho(H)} \\ >> 1 & \text{if } p >> n^{-\frac{1}{\rho(H)}} \end{cases}$.

2) $\operatorname{Var} X = \sum \operatorname{Var} X_S + \sum \operatorname{cov}(X_S, X_t)$.

Var: $\mathbb{E}(X) = o(\mathbb{E}(X)^2)$

$\sum_{2 \leq |S \cap t| \leq v-1} \operatorname{Cov}(X_S, X_t)$. $\operatorname{Cov}(X_S, X_t) \leq \mathbb{P}(X_S X_t = 1) \leq \sum_{\varphi_S, \varphi_t : V(H) \to V(S), V(t)} \mathbb{P}(\varphi_S(H) \cup \varphi_t(H) \subseteq G(n, p))$

But $H' = \varphi_S(H) \cap \varphi_t(H) \subsetneq H$ (since $S \neq t$). So $p^{e(\varphi_S(H) \cup \varphi_t(H))} = p^{2e - e(\varphi_S(H) \cap \varphi_t(H))}$

$v(H') = |S \cap t|$, $e(H') \leq |S \cap t|\, \rho(H)$. So this is at most $(v!)^2 \left(p^{2e - \frac{e}{v}|S \cap t|}\right)$. $\sum \operatorname{Cov}(X_S, X_t) \leq n^v \sum_{i=2}^{n-1} n^{v-i} p^{2e - i\frac{e}{v}} (v!)^2 \leq$

$\underbrace{n^v p^e}_{O(\mathbb{E}(X))} \sum_{i=2}^{n-1} p^{e(1-\frac{i}{v})} n^{v(1-\frac{i}{v})} (v!)^2 = C \cdot \mathbb{E}(X) \sum_{i=2}^{n-1} \underbrace{(p^e n^v)^{(1-\frac{i}{v})}}_{O(\mathbb{E}(X))} (<1) = o(\mathbb{E}(X)^2)$.