Intro. & Moment.

Sunday, March 30, 2025 2:01 PM

Plan of this course

Asymptotics in Prob. Theory

① Basic tools

② Random Matrix Theory

③ Interacting Particle Systems
Voter / Contact / Exclusion

Reminder: random variable X

Markov ineq: for $X \geq 0$, $\mathbb{E}|X| < \infty$, $\mathbb{P}[X \geq b] \leq \frac{\mathbb{E}X}{b}$,

Chebyshev's ineq: for $\mathbb{E}X^2 < +\infty$, $\mathbb{P}[|X - \mathbb{E}X| > \beta] \leq \frac{\mathrm{Var}[X]}{\beta^2}$

$(\mathrm{Var}[X] = \mathbb{E}X^2 - (\mathbb{E}X)^2 = \mathbb{E}(X - \mathbb{E}X)^2)$.

Example 1. Coupon Collector

$(X_i)_{i \in \mathbb{N}}$ i.i.d. uniform random in $\{1, \cdots, n\}$

smallest T_n s.t. $\{X_1, \cdots, X_T\}$ contains all of $1, \cdots, n$?

$$\mathbb{P}\left[\left|T_n - n\left(1 + \tfrac{1}{2} + \tfrac{1}{3} + \cdots + \tfrac{1}{n}\right)\right| \geq \varepsilon n \log n\right] \to 0 \quad \text{as } n \to \infty$$

Proof. $T_n = \tau_1 + \tau_2 + \cdots + \tau_n$

τ_i: time to get the i-th new number — $\mathrm{Geo}(p)$

$\tau_1 = 1$

$\tau_2 = \mathrm{Geo}(p_2)$ $\quad (\mathbb{P}[\tau_2 = k] = (1-p_2)^{k-1} p_2)$

with $p_2 = \frac{n-1}{n}$

$\tau_3 = \mathrm{Geo}(p_3)$ $\quad p_3 = \frac{n-2}{n}$ $\quad [\mathbb{E}\,\mathrm{Geo}(p) = \frac{1}{p}]$

$\vdots$

$\mathbb{E}T_n = \mathbb{E}\tau_1 + \mathbb{E}\tau_2 + \cdots + \mathbb{E}\tau_n = n(1 + \frac{1}{2} + \cdots + \frac{1}{n})$

$\mathrm{Var}(T_n) = \mathrm{Var}(\tau_1) + \mathrm{Var}(\tau_2) + \cdots + \mathrm{Var}(\tau_n)$

$[\mathrm{Var}(\mathrm{Geo}(p)) = \frac{1-p}{p^2}]$

$= 0 + \frac{n}{(n-1)^2} + \frac{2n}{(n-2)^2} + \cdots + \frac{n^2}{1^2} < 2n^2$

$$\Rightarrow \mathbb{P}\left[\left|T_n - n(1 + \tfrac{1}{2} + \cdots + \tfrac{1}{n})\right| \geq \varepsilon n \log n\right] \leq \frac{2n^2}{(\varepsilon n \log n)^2} \to 0 \quad \text{as } n \to \infty$$

Example 2. Random permutation

$(\sigma_1, \sigma_2, \cdots, \sigma_n) \in S_n$

$L_n :=$ length of longest increasing subsequence

Then L_n is of order $\sqrt{n}$

Upper bound $\mathbb{P}[L_n \geq l] \leq \underbrace{\binom{n}{l}}_{\#\text{ len. } l \text{ subseq}} \cdot \underbrace{\frac{1}{l!}}_{\text{Prob of subseq. increasing}} \leq \frac{n^l}{(l!)^2} \leq \left(\frac{e\sqrt{n}}{l}\right)^{2l}$

for $l = (e+\varepsilon)\sqrt{n}$, any $\varepsilon > 0$

$\mathbb{P}[L_n \geq l] \to 0$ as $n \to \infty$. (Moreover, $\mathbb{E}[L_n] \leq l + \underbrace{n\mathbb{P}[L_n \geq l]}_{\to 0}$, so $\limsup_{n\to\infty} \mathbb{E}[L_n] \leq e$)

Lower bound D_n: length of longest decreasing subsequence

Claim $L_n D_n \geq n$. (Erdős–Szekeres)

$\Rightarrow \mathbb{E}L_n = \frac{1}{2}\mathbb{E}(L_n + D_n) \geq \mathbb{E}\sqrt{L_n D_n} \geq \sqrt{n}$

proof $(L_n^{(k)}, D_n^{(k)})$ length of LIS / LDS ending at k

then $(L_n^{(k)}, D_n^{(k)}) \neq (L_n^{(j)}, D_n^{(j)})$ for $j \neq k$

Baik–Deift–Johansson: $L_n = 2\sqrt{n} + n^{\frac{1}{6}} \cdot \underline{\text{Tracy-Widom}_2}$

Example 3 Erdős–Rényi graph. $G(n,p)$

n vertices, any two connected with probability p.

Mainly focus on sparse case: $p = \frac{d}{n}$

- **Degree distribution**

Given vertex: Binomial $(n-1, d/n)$

$\mathbb{P}(\deg = k) = \binom{n-1}{k}\left(\frac{d}{n}\right)^k \left(1 - \frac{d}{n}\right)^{n-1-k}$

$\to \frac{d^k}{k!} e^{-d}$ as $n \to \infty$

$(\mathrm{Poiss}(d))$

- Max deg: $\dfrac{\max_{v \in V} d_v}{\log n / \log(\log n)} \to 1$ in Prob.

(i.e. $\mathbb{P}\left[\left|\frac{\max_v d_v}{\log n/\log(\log n)} - 1\right| > \varepsilon\right] \to 0$ as $n \to \infty$)

proof. Upper bound:

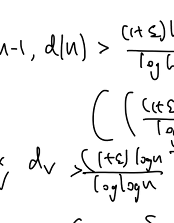
Check: $\mathbb{P}\left[\mathrm{Bino}(n-1, d/n) > \frac{(1+\varepsilon)\log n}{\log\log n}\right] \leq n^{-(1+\varepsilon)+o(1)}$

$\left(\left(\frac{(1+\varepsilon)\log n}{\log\log n}\right)! \approx n^{(1+\varepsilon)+o(1)}\right)$

$\Rightarrow \mathbb{P}\left[\max_{v \in V} d_v > \frac{(1+\varepsilon)\log n}{\log\log n}\right] \leq n^{-\varepsilon + o(1)}$

Lower bound:

Split into 2: S_1 S_2



$\tilde{d}_v := \#\{u \in S_2, (v,u) \in E\}$

$\mathbb{P}\left(\tilde{d}_v = (1-\varepsilon)\frac{\log(n/2)}{\log(\log(n/2))}\right) = \left(\frac{n}{2}\right)^{-(1-\varepsilon)+o(1)}$

by independence:

$\mathbb{P}\left[\max_{v \in S_1} \tilde{d}_v < (1-\varepsilon)\frac{\log(n/2)}{\log\log(\frac{n}{2})}\right] \leq \left(1 - \left(\frac{n}{2}\right)^{-(1-\varepsilon)+o(1)}\right)^{\frac{n}{2}}$

$\leq \exp(-n^{\varepsilon + o(1)})$

- **Cycles** C_k: # of k cycles

$\mathbb{E}C_k = \frac{n(n-1)\cdots(n-k+1)}{2k} \cdot \left(\frac{d}{n}\right)^k \to \frac{d^k}{2k}$ as $n \to \infty$

$C_k \to \mathrm{Poisson}\left(\frac{d^k}{2k}\right)$ (no cycle through a given v) $\Rightarrow$ tree like locally

proof $\mathbb{E}C_k(C_k - 1)$

$= \sum_{\substack{A \neq A' \\ k \text{ cycles}}} \mathbb{P}[A, A' \text{ exists in } G(n, p = \frac{d}{n})]$

$= \underbrace{\sum_{\substack{A \neq A', A \cap A' = \phi \\ k \text{ cycles}}} \cdots}_{\to \left(\frac{d^k}{2k}\right)^2} + \sum_{\substack{A \neq A', A \cap A' \neq \phi \\ k \text{ cycles}}} \cdots$

$\binom{n}{\#V(A \cup A')} \cdot \left(\frac{d}{n}\right)^{\#E(A \cup A')}$

$\approx \frac{n^{\#V(A \cup A')}}{\#V(A \cup A')!} \cdot \left(\frac{d}{n}\right)^{\#E(A \cup A')} \to 0$ as $n \to \infty$

$(\#V(A \cup A') - \#E(A \cup A') \leq -1)$

More generally,

$\mathbb{E}C_k(C_k - 1)\cdots(C_k - a + 1) \to \left(\frac{d^k}{2k}\right)^a$

(For $\mathrm{Poisson}(\lambda)$, $\mathbb{E}X(X-1)\cdots(X-a+1) = \lambda^a$).

More generally,

$C_3, C_4, \cdots \to$ independent $\mathrm{Poi}\left(\frac{d^3}{6}\right), \mathrm{Poi}\left(\frac{d^4}{8}\right), \ldots$

Random d-regular graph

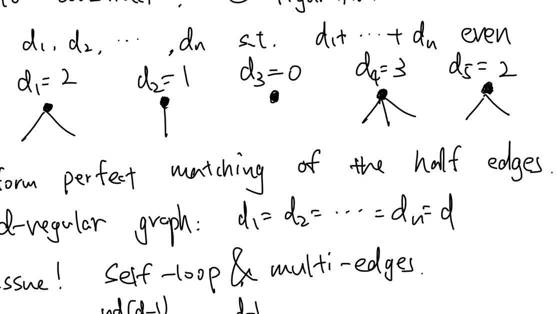
All d-regular graphs with n labeled vertices (nd even)

choose one uniformly at random

How to construct? Configuration model (Bollobás)

Take $d_1, d_2, \cdots, d_n$ s.t. $d_1 + \cdots + d_n$ even

$d_1 = 2$ $\quad d_2 = 1$ $\quad d_3 = 0$ $\quad d_4 = 3$ $\quad d_5 = 2$



Uniform perfect matching of the half edges.

random d-regular graph: $d_1 = d_2 = \cdots = d_n = d$

Issue! Self-loop & multi-edges.

$\mathbb{E}\underbrace{C_1}_{\#\text{ cycle of length 1, i.e. self loop}} = \frac{nd(d-1)}{2(nd-1)} \to \frac{d-1}{2}$

$\mathbb{E}C_2 = \frac{n}{2} \cdot \frac{d(d-1)}{2} \frac{(d-1)}{(nd-3)} \to \frac{(d-1)^2}{4}$

Higher moments $\Rightarrow$ $C_1, C_2 \to \mathrm{Poi}\left(\frac{d-1}{2}\right), \mathrm{Poi}\left(\frac{(d-1)^2}{4}\right)$, independent

$\Rightarrow \underbrace{\mathbb{P}(C_1 = C_2 = 0)}_{\text{(simplicity)}} \to \exp\left(-\frac{d-1}{2} - \frac{(d-1)^2}{4}\right) > 0$

From construction: $G \sim$ config with $d_1 = d_2 = \cdots = d_n = d$, conditional on simple

$= G$ random d-reg of n vertices

- For fixed $d \geq 3$

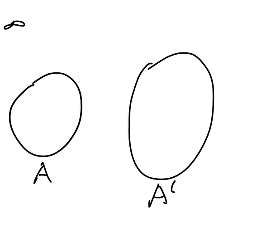
$\mathbb{P}[G \sim \text{random } d\text{-reg is connected}] \to 1$ as $n \to \infty$

proof Suffices to show:

$\mathbb{P}[G \sim \text{config is disconnected}] \to 0$ as $n \to \infty$

This is bounded by

$\frac{1}{2}\sum_{A \subseteq \{1, \cdots, n\}} \mathbb{P}[\text{no edge between } A, A^c]$



$= \frac{1}{2}\sum_{A \subseteq \{1, \cdots, n\}} \frac{M_{d|A|} M_{d(n-|A|)}}{M_{dn}}$ $\qquad M_m$: # perfect matching of size m

$= \sum_{k=1}^{n/2} \binom{n}{k} \frac{M_{dk} M_{d(n-k)}}{M_{dn}}$ $\qquad M_m = \frac{m!}{2^{m/2}(m/2)!} \approx \sqrt{2}\left(\frac{m}{e}\right)^{m/2}$

$\approx \sum_{k=1}^{n/2} \sqrt{\frac{n}{2\pi k(n-k)}} \frac{n^n}{k^k (n-k)^{n-k}} \sqrt{2} \cdot \frac{k^{dk/2}(n-k)^{d(n-k)/2}}{n^{dn/2}}$ $\qquad (m! \approx \sqrt{2\pi m}(m/e)^m$, Stirling)

$\leq C \cdot \sum_k \frac{1}{\sqrt{k}} \left(\frac{k}{n}\right)^{(\frac{d-2}{2})k} \to 0$ as $n \to \infty$

As a comparison:

For Erdős–Rényi $G(n,p)$

If $p = p_n \ll \frac{\log n}{n}$ (i.e. $\frac{p_n}{\log n/n} \to 0$ as $n \to \infty$)

$\mathbb{P}[G(n, p_n) \text{ connected}] \to 0$ as $n \to \infty$

proof X_n: # of isolated points

$\mathbb{E}X_n = n(1-p)^{n-1}$

$\mathbb{E}X_n(X_n - 1) = n(n-1)(1-p)^{2n-3}$

$\mathrm{Var}X_n = n(n-1)(1-p)^{2n-3} + n(1-p)^{n-1} - n^2(1-p)^{2n-2}$

$= n^2 p(1-p)^{2n-3} + n((1-p)^{n-1} - (1-p)^{2n-3})$

Using Chebyshev's ineq.

$\mathbb{P}[X_n = 0] \leq \frac{\mathrm{Var}X_n}{(\mathbb{E}X_n)^2} = \frac{p}{1-p} + \frac{1}{n(1-p)^{n-1}} - \frac{1}{n(1-p)} \to 0$

If $p = p_n \gg \frac{\log n}{n}$ (i.e. $\frac{p_n}{\log n/n} \to \infty$ as $n \to \infty$)

$\mathbb{P}[G(n, p_n) \text{ connected}] \to 1$ as $n \to \infty$

proof $\mathbb{P}[\text{disconnected}] \leq \sum_{k=1}^{n/2} \binom{n}{k} \cdot (1-p)^{k(n-k)}$

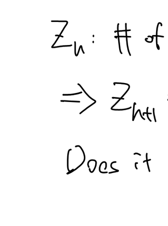
$\leq \sum_{k=1}^{n/2} n^k (1-p)^{nk/2} = \sum_{k=1}^{n/2} \left(n(1-p)^{n/2}\right)^k$

as $p \gg \frac{\log n}{n}$, $n(1-p)^{n/2} \to 0$

$\Rightarrow \mathbb{P}[\text{disconnected}] \to 0$ as $n \to \infty$

Galton–Watson Branching Process

Offspring distribution: X



Each vertex has independent number of offspring with law X

Let $\mu = \mathbb{E}X < \infty$

Z_n: # of vertex at level n

$\Rightarrow Z_{n+1} = \sum_{i=1}^{Z_n} X_i \Rightarrow \mathbb{E}Z_{n+1} = \mu\mathbb{E}Z_n = \mu^{n+1}$

Does it die out? $p = \mathbb{P}[\text{infinite tree}] > 0$ iff $\mu > 1$.

(Reveal one-by-one / breadth first search $\Rightarrow$ random walk with step $X - 1$)

(Also $p = 1 - \mathbb{E}(1-p)^X$)

Local Weak Limit

for rooted graphs, $(G, v) \simeq (G', v')$

if $\exists$ bijection $\psi: V \to V'$, $\psi(v) = v'$

$(u,w) \in E$ iff $(\psi(u), \psi(w)) \in E'$

For a sequence of graphs $G_1, G_2, \cdots, G_n, \cdots$

let I_n be uniformly chosen from G_n

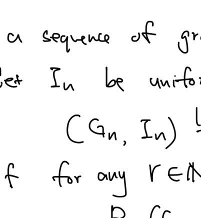
$(G_n, I_n) \xrightarrow{\text{lwc}} (G_\infty, I_\infty)$ ← a random rooted graph

if for any $r \in \mathbb{N}$

$B_r(G_n, I_n) \xrightarrow[\text{weakly}]{} B_r(G_\infty, I_\infty)$

$(\mathbb{P}[B_r(G_n, I_n) = (G, v)] \to \mathbb{P}[B_r(G_\infty, I_\infty) = (G, v)]$ $\forall (G, v))$

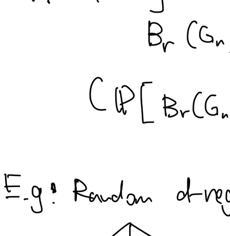
E.g.: Random d-regular graph:



$\mathbb{P}[\text{cycle in } B_r(G_n, I_n)] \to 0$, $\forall r \in \mathbb{N}$

$\Rightarrow$ lwc to infinite d-reg. tree

- Erdős–Rényi: $G(n, \frac{d}{n})$



$\mathbb{P}[\text{cycle in } B_r(G_n, I_n)] \to 0$

$\Rightarrow$ lwc to Galton–Watson $\mathrm{Poiss}(d)$.

- config model with each $d_i \sim Y$? (Still tree like)

Root distribution: Y

for $v \neq$ root

$\mathbb{P}[\deg(v) = k]$

$= \frac{(\#\text{ vertex of deg } k) \cdot k}{\#\text{ half edges}} \approx \frac{\mathbb{P}[Y = k] \cdot k}{\mathbb{E}Y}$

size biased distribution Y^*

$\Rightarrow$ lwc to Galton–Watson tree, root offspring Y

other offspring $Y^* - 1$

(For $Y \sim \mathrm{Pois}(\lambda)$, $Y^* - 1 \sim \mathrm{Pois}(\lambda)$)