The Second moment method

General idea: Sometimes computing $\mathbb{E}[X]$ is not enough, we need to prove that with high probability, X is close to $\mathbb{E}[X]$. We then use concentration bounds.

1) Chebyshev's inequality

For every $t > 0$, for every a.s. non constant r.v. X,

$$\mathbb{P}\left[|X - \mathbb{E}[X]| > t\right] \leq \frac{\mathrm{Var}(X)}{t^2}.$$

Recall that $\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{Var}(X_i) + \sum_{i \neq j} \mathrm{Cov}(X_i, X_j)$

where $\mathrm{Cov}(X_i, X_j) = \mathbb{E}[X_i X_j] - \mathbb{E}[X_i]\mathbb{E}[X_j]$.

Application: "Almost all integers n have close to $\ln \ln n$ prime factors"

$\nu(n) = \#$ of prime factors of n

[Hardy & Ramanujan 1920]. Let $\omega(n) \to +\infty$ arbitrarily slowly. Then the number of x in $[n]$ such that

$$|\nu(x) - \ln \ln x| > \omega(n)\sqrt{\ln \ln n} \quad \text{is } o(n).$$

Proof by Turán 1934, basis of probabilistic method in number theory

Let x be randomly chosen from $[n]$.

For p prime, set $X_p = \begin{cases} 1 & \text{if } p \mid x \\ 0 & \text{otherwise} \end{cases}$

Let $M = n^{1/10}$ and $X = \sum_{p \leq M} X_p$.

No $x \leq n$ can have more than 10 prime factors larger than M, so $\nu(x) - 10 \leq X(x) \leq \nu(x)$.

$$\mathbb{E}[X_p] = \frac{\lfloor n/p \rfloor}{n} = \frac{1}{p} + O\left(\frac{1}{n}\right).$$

- $\mathbb{E}[X] \underset{\text{L.E.}}{=} \sum_{p \in M} \left(\frac{1}{p} + O\left(\frac{1}{n}\right)\right) = \ln \ln n + O(1).$

(Stirling approx + Abel summation. $A(t) = \sum_{0}^{t} a_n$

$\forall\, \phi \in \mathcal{C}^1[x, y],\ \sum_{x < n \leq y} a_n \phi(n) = A(y)\phi(y) - A(x)\phi(x) - \int_x^y A(u)\phi'(u)\,du$)

- $\mathrm{Var}[X] = \sum_{p < M} \mathrm{Var}[X_p] + \sum_{\substack{p \neq q \\ \leq M}} \underbrace{\mathrm{Cov}[X_p, X_q]}_{= \mathbb{E}[X_p X_q] - \mathbb{E}[X_p]\mathbb{E}[X_q]}$

$$\mathrm{Var}[X_p] = \mathbb{E}[X_p^2] - \mathbb{E}[X_p]^2$$

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

$$= \frac{1}{p}\left(1 - \frac{1}{p}\right) + O\left(\frac{1}{n}\right).$$

So $\sum_{p \leq M} \mathrm{Var}(X_p) = \left(\sum_{p \leq M} \frac{1}{p}\right) + O(1)$

$$= \ln \ln n + O(1)$$

$X_p X_q = 1 \iff p \mid x,\ q \mid x \iff pq \mid x$

so $\mathrm{Cov}(X_p, X_q) = \mathbb{E}[X_p X_q] - \mathbb{E}[X_p]\mathbb{E}[X_q]$

$$= \frac{\lfloor n/pq \rfloor}{n} - \frac{\lfloor n/p \rfloor}{n}\frac{\lfloor n/q \rfloor}{n} \leq \frac{1}{pq} - \left(\frac{1}{p} - \frac{1}{n}\right)\left(\frac{1}{q} - \frac{1}{n}\right)$$

$$\leq \frac{1}{n}\left(\frac{1}{p} + \frac{1}{q}\right)$$

So $\sum_{p \neq q} \mathrm{Cov}(X_p, X_q) \leq \frac{1}{n} \sum_{p \neq q} \left(\frac{1}{p} + \frac{1}{q}\right)$

$$\leq \frac{2M}{n} \sum \frac{1}{p}$$

$$\leq O\left(n^{9/10} \ln\ln n\right) = o(1).$$

Hence, $\mathrm{Var}[X] = \ln\ln(n) + O(1)$.

- By Chebyshev's inequality,

$$\mathbb{P}\left[|X - \ln\ln n| > \lambda\sqrt{\ln\ln n}\right] \leq \lambda^{-2} + o(1)$$

As $|X - \nu| \leq 10$, the same holds for ν ∎

Actually [Erdős and Kac 1940] showed that ν behaves like a normal distribution with mean & variance $\ln\ln n$:

$$\forall \lambda \in \mathbb{R}, \lim_{n \to +\infty} \frac{1}{n} \#\left\{x \in [n] : \nu(x) \geq \ln\ln(x) + \lambda\sqrt{\ln\ln n}\right\} = \int_{\lambda}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt.$$

2) Chernoff Bound

Let $X \sim Bin(n,p)$. For any $0 \leq t \leq np$,

$$\mathbb{P}(|X - \mathbb{E}[X]| > t) \leq 2e^{-\frac{t^2}{3np}}.$$

Markov's inequality only gave $\mathbb{P}(|X-\mathbb{E}[X]|>t) < \frac{1}{1+\frac{t}{\mathbb{E}[X]}}$

Application: Hajós' conjecture

Hadwiger conjecture 1943: For any t, $\chi(G) \geq t \Rightarrow G$ contains a K_t-minor

G contains a K_t-subdivision $\Rightarrow$ G contains a K_t-minor.

Hajós' conjecture 1961: For any t, $\chi(G) \geq t \rightarrow G$ contains a K_t-subdivision.

Disproved by Caitlin in 1979

Proof by [Erdős, Fajtlowicz 1981]

We consider $G \sim G(n, \frac{1}{2})$.

Step 1: W.h.p, $\chi(G) \geq \frac{n}{2\log_2 n}$

Step 2: W.h.p. G has no K_ℓ-subdivision where $\ell = \lceil 8\sqrt{n} \rceil$.

1. we show that $\alpha(G) \leq 2\log_2 n$ w.h.p

Let X be the number of stable sets of size $\lceil 2\log_2 n \rceil$.

$$\mathbb{E}[X] = \binom{n}{k} 2^{-\binom{k}{2}} \qquad \text{where } k = 2\lceil \log_2 n \rceil.$$

$$\leq \frac{n!}{k!(n-k)!} \, 2^{-\frac{k(k-1)}{2}}$$

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

$$\leq \frac{n}{k!} < \frac{1}{n} \quad \text{for } n \text{ sufficiently large.}$$

By $\boxed{\text{Stirling } n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n}$

Alternatively,

$$\boxed{\text{when } k = o(n),\ \binom{n}{k} \sim \left(\frac{ne}{k}\right)^k \cdot (2\pi n)^{-1/2} \exp\left(-\frac{k^2}{2n}(1+o(1))\right)}$$

So by Markov's inequality, $\mathbb{P}[X>0] < \frac{1}{n}$.

Hence, $\mathbb{P}\left[\chi(G) < \frac{n}{2\log_2 n}\right] < \frac{1}{n}$.

2. $U \subseteq V(G)$ is a $\frac{3}{4}$-clique if $|E(G[U])| \geq \frac{3}{4}\binom{|U|}{2}$.

Let $l = \lceil 8\sqrt{n} \rceil$, we show that with high probability, G has no $\frac{3}{4}$-clique of size l.

Let U be a set of l vertices and $Y = |E(G[U])|$.

$$Y \sim \mathrm{Bin}\left(\binom{l}{2}, \frac{1}{2}\right).$$

So $$\Pr\left(Y \geq \frac{3}{4}\binom{l}{2}\right) \leq 2e^{-\frac{\binom{l}{2}^2}{16\cdot 3\binom{l}{2}/2}} \leq 2e^{-\binom{l}{2}/24} < 2e^{-\frac{5}{4}n}$$

So $\mathbb{E}\left[\#\,\frac{3}{4}\text{-cliques of size } \ell\right]$

$$\underset{\text{L.E.}}{\leq} \binom{n}{\ell} 2 e^{-\frac{5}{4}n} < 2^n e^{-\frac{5}{4}n} < e^{-\frac{n}{4}}.$$

So by Markov's inequality, $\mathbb{P}(G \text{ has a } \frac{3}{4}\text{-clique of size } \ell) < e^{-\frac{n}{4}}$

By subadditivity of probabilities, as $1 - \frac{1}{n} - e^{-\frac{n}{4}} > 0$ for large n,

G has chromatic number $\geqslant \dfrac{n}{2\log_2 n}$ and no $\frac{3}{4}$-clique of size ℓ with positive probability.

In K_ℓ subdivision on at most 8ℓ vertices, the centres induce a $\frac{3}{4}$ clique of size ℓ. ∎

3) Application: Weierstrass approximation theorem

[Weierstrass 1885]: The polynomials are dense in the space of continuous real functions over a segment:

$$\forall \varepsilon > 0, \ \forall f : [a,b] \to \mathbb{R}, \ \exists p \in \mathbb{R}[X], \ \|p - f\|_\infty < \varepsilon.$$

Proof (Bernstein 1912)

Up to rescaling, let us assume that

- $[a,b] = [0,1]$
- $\|f\|_\infty = 1$ (f is bounded because continuous on a compact)

For $x \in [0,1]$, let $S_n(x) \sim \mathrm{Bin}(n, x)$.

$\frac{S_n(x)}{n}$ concentrates around x, so $\mathbb{E}\left[f\left(\frac{S_n(x)}{n}\right)\right]$ should be close to $\mathbb{E}[f(x)]$.

Let $B_n(x) = \mathbb{E}\left[f\left(\frac{S_n(x)}{n}\right)\right] = \sum_{k=0}^{n} \mathbb{P}\left(S_n(x) = k\right) f\left(\frac{k}{n}\right)$

$$= \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^k f\left(\frac{k}{n}\right)$$

is a polynomial in x.

$f(x) - B_n(x) = \mathbb{E}\left[f(x) - f\left(\frac{S_n(x)}{n}\right)\right]$ by linearity of expectation,

so $|f(x) - B_n(x)| \leq \mathbb{E}\left[\left|f(x) - f\left(\frac{S_n(x)}{n}\right)\right|\right]$

As f is continuous on a compact, f is uniformly continuous:

$\forall \varepsilon > 0, \exists \delta > 0, \forall x, y \in [0,1], \quad |x-y| \leq \delta \to |f(x) - f(y)| \leq \varepsilon.$

So $|f(x) - B_n(x)| \leq \mathbb{E}\left[\left| f(x) - f\left(\frac{S_n(x)}{n}\right) \right| \right]$

$$\leq \varepsilon \cdot \mathbb{P}\left(\left| x - \frac{S_n(x)}{n} \right| \leq \delta \right) + 2 \|f\|_\infty \mathbb{P}\left(\left| x - \frac{S_n(x)}{n} \right| > \delta \right)$$

By Chebyshev's inequality,

$$\mathbb{P}\left(\left| x - \frac{S_n(x)}{n} \right| > \delta \right) = \mathbb{P}\left(\left| nx - S_n(x) \right| > n\delta \right)$$

$$\leq \frac{n x (1-x)}{(n\delta)^2} \leq \frac{1}{4n\delta^2}$$

So for n large enough,

$$|f(x) - B_n(x)| \leq 2\varepsilon \qquad \forall x \in [0,1] \quad \blacksquare$$

Exercise: Prove that with probability $1 - o(1)$, $G(n,p)$ has at most $n^2/8 + n^{3/2}$ edges.