Random graphs and thresholds

March 26, 2026

Exercise 1.

Show that for every $\delta > 0$, a series of n independent coin flips contains k consecutive heads

- with probability $o(1)$ if $k \leq (1+\delta)\log_2 n$,
- with probability $1 - o(1)$ if $k \leq (1-\delta)\log_2 n$.

Solution 1.

Let $X_1, \dots X_n$ be n independent Bernoulli variable of parameter $1/2$. Let H_k be the number of sequences of k consecutive 1 in $(X_1, \dots X_n)$. That is, $H_k = \sum_{i=1}^{n-k} X_{i+0} \cdots X_{i+k-1}$. We have $\mathbb{E}(H_k) = (n-k+1)2^{-k}$ by linearity of expectation. We also have

$$
\begin{aligned}
\mathrm{Var}(H_k) &= \sum_{i=1}^{n-k+1} \mathrm{Var}(X_{i+0} \cdots X_{i+k-1}) + \sum_{j=1}^{k-1} \sum_{i=1}^{n-k-j} \mathrm{Cov}(X_{i+0} \cdots X_{i+k-1}, X_{i+j+0} \cdots X_{i+j+k-1}) \\
&\leq n2^{-k}(1-2^{-k}) + n\sum_{j=1}^{k-1}(2^{-(2k-j)} - 2^{-2k}) \\
&\leq n2^{-k} + n2^{-2k}\sum_{j=1}^{k-1}(2^j - 1) \\
&\leq n2^{-(k-1)}
\end{aligned}
$$

- If $k \leq (1+\delta)\log_2 n$, then $\mathbb{E}(H_k) \leq n^{-\delta}$ so by Markov's inequality, $\mathbb{E}(H_k) = o(1)$ and by Markov's inequality, $\mathbb{P}(H_k = 0) \leq 1 - o(1)$
- If $k \geq (1-\delta)\log_2 n$, then $\mathbb{E}(H_k) \geq n^{\delta}/2 \to \infty$ and $\mathrm{Var}(H_k) \leq n^{\delta}$.

$$
\begin{aligned}
\mathbb{P}(H_k \leq \mathbb{E}(H_k)/2) &\leq \frac{\mathrm{Var}(H_k)}{(\mathbb{E}(H_k)/2)^2} && \text{By Chebyshev's inequality} \\
&\leq 4n^{-\delta}
\end{aligned}
$$

So $\mathbb{P}(H_k > 0) = 1 - o(1)$.

Exercise 2.

Let $S_{n,p}$ be a random subset of $[n]$ such that each element of $[n]$ is selected independently with probability p. For any fixed $k \geq 3$, determine the threshold at which $S_{n,p}$ contains an arithmetic progression of length k. Is it a a coarse or a sharp threshold?

Solution 2.

Let X count the number of arithmetic progressions of length k in $S_{n,p}$. We have $\mathbb{E}[X] = p^k \cdot \mathrm{AP}(k,n)$ where $\mathrm{AP}(k,n)$ counts the number of arithmetic progressions in $[n]$. We have $\mathrm{AP}(k,n) = \sum_{i=1}^{\lfloor n/k \rfloor}(n-ki) \sim \frac{n^2}{2k}$.

If $p \leq (1-\delta)\left(\frac{2k}{n^2}\right)^{1/k}$, then $\mathbb{E}(X) \lesssim (1-\delta)^k$. So by Markov's inequality, $S_{n,p}$ has with high probability no arithmetic progression of length k if $p \ll \left(\frac{2k}{n^2}\right)^{1/k}$ and the threshold is coarse.

I haven't finished typing this correction.

Exercise 3. Threshold for cycles

What is the threshold for $G(n,p)$ to contain a cycle? Is it coarse or sharp?

Solution 3.

We prove that $1/n$ is a coarse threshold for containing a cycle. Let X be the number of cycles in $G(n,p)$. We have $\mathbb{E}[X] = \sum_{k=3}^{n} \binom{n}{k} \frac{(k-1)!}{2} p^k = \sum_{k=3}^{n} \frac{n! p^k}{2k(n-k)!} \leq \frac{(np)^k}{k}$. Hence, $E[X] = o(1)$ if $p \lesssim (1-\delta)/n$.

Let $p \gtrsim (1+\delta)/n$. By Chernoff's bound, the number of edges in $G(n,p)$ is at least n with probability $1 - o(1)$ so $G(n,p)$ contains a cycle asymptotically almost surely.

Exercise 4. Poisson limit

The k^{th} *moment* of a random variable X is $\mathbb{E}[X^k]$. A random variable X with finite moments is *determined by its moments* if for every random variable Y such that $\mathbb{E}[X^k] = \mathbb{E}[Y^k]$ for every k, then $X \overset{d}{=} Y$. In particular every random variable X such that there exists $C > 0$ such that $|\mathbb{E}[X^k]| \leq C^k k!$ for every k, is determined by its moments.

Let X be a random variable determined by its moments and let $(X_n)_{n\geq 0}$ be a sequence of random variables with finite moments such that for every k, $\mathbb{E}[X_n^k] \to \mathbb{E}[X^k]$. Then X_n converges in distribution to X. Alternatively, all the previous definition hold by replacing moment by the k^{th} *factorial moment* $\mathbb{E}[X \cdot (X-1) \cdots (X-k+1)] = k! \cdot \mathbb{E}[\binom{X}{k}]$.

In particular Poisson distributions are determined by their moments and:

Theorem 1 *Let X_n be a sequence of random variables with finite moments, such that $\mathbb{E}[\binom{X_n}{k}] \to \lambda^k k!$ for every k. Then X_n converges in distribution to a random Poisson variable of parameter λ.*

Let $p(n) \sim c/n$ for some fixed constant $c > 0$. Let X_n be the random variable counting the number of triangles in $G(n,p)$.

1. Determine the asymptotics of $\mathbb{E}\left[\binom{X_n}{k}\right]$.
2. Let $\lambda > 0$ and $Y_n \sim \text{Bin}(n, \lambda/n)$. Determine the asymptotics of of $\mathbb{E}\left[\binom{Y_n}{k}\right]$.
3. Let $m \in \mathbb{N}$, compute the limit of $\mathbb{P}(X_n = m)$ (we admit the fact that Y_n converges in distribution to the Poisson distribution of parameter λ). What is the asymptotic probability that $G(n,p)$ contains a triangle ?

Solution 4.

1. We have

$$\begin{aligned}\mathbb{E}\left[\binom{X_n}{k}\right] &= \sum_{T_1,\dots T_k \text{ distinct triangles}} \mathbb{P}(T_1,\dots T_k \text{ are triangles of } G(n,p)) \\ &= \sum_{T_1,\dots T_k \text{ distinct triangles}} p^{e(T_1\cup\dots T_k)}\end{aligned}$$

Let $T_1 \dots T_k$ be vertex-disjoint triangles. We have $e(T_1 \cup \dots T_k) = 3k$ and there exists $\frac{1}{k!}\binom{n}{3} \cdot \binom{n-3}{3} \cdots \binom{n-3k+3}{3} \sim \left(\frac{c^3}{6}\right)^k \cdot \frac{1}{k!}$ such choices of triangles. Let $T_1 \dots T_k$ be distinct triangles such that some T_i shares at least one edge with some other triangle. Then $e(T_1 \cup \dots T_k) \leq 3k-1$ and the number of choices of such triangle is $O(n^{3(k-1)+1})$. Thus,

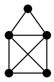
$$\mathbb{E}\left[\binom{X_n}{k}\right] \sim \left(\frac{c^3}{6}\right)^k \cdot \frac{1}{k!}$$

2. We have

$$\begin{aligned}\mathbb{E}\left[\binom{Y_n}{k}\right] &= \sum_{i_1\dots i_k \text{ distinct integers}} \mathbb{P}(\text{we tossed 1 for every integer } i_1,\dots i_k) \\ &= \binom{n}{k} p^k \sim \frac{\lambda^k}{k!}\end{aligned}$$

3. X_n converges in distribution towards a Poisson distribution of parameter $\lambda = c^3/6$. Hence $\mathbb{P}(X_n = m) \to \frac{\lambda^m e^{-\lambda}}{m!}$. Hence $\mathbb{P}(G(n,p)$ contains a triangle$) \to 1 - e^{-c^3/6}$.

Exercise 5.

Let H be the following graph:  Using an ad hoc proof, what is the threshold for having H as a subgraph?

Exercise 6.

Prove that there exists $c > 0$ such that $G(n, n^{-1/2})$ has asymptotically almost surely at least $cn^{3/2}$ edge-disjoint triangles.

Solution 6.

Let $H(n,p)$ be the graph whose vertices are the triangles of of $G(n,p)$ with an edge between two triangles if they share an edge. Let X be the number of vertices in $H(n,p)$ and Y be the number of edges. We have $\mathbb{E}[X] = \binom{n}{3}p^3 \sim n^{3/2}/6$ by linearity of expectation and

$$\begin{aligned}\mathrm{Var}(X) &= \sum_T \mathrm{Var}(X_T) + \sum_{T_1,T_2 \text{ sharing an edge}} \mathrm{Cov}(X_{T_1}, X_{T_2}) \\ &= \binom{n}{3}p^3(1-p^3) + \binom{n}{3}\frac{3(n-3)}{2}\cdot(p^5 - p^6) \\ &\sim \frac{n^{3/2}}{6} + \frac{n^4}{4}n^{-5/2} \le \frac{5}{12}n^{3/2}\end{aligned}$$

Hence by Chebyshev's inequality, $\mathbb{P}(|X - \mathbb{E}(X)| \ge \frac{n^{3/2}}{4}) = O(n^{-3/2})$.

We have $\mathbb{E}[Y] = \binom{n}{3}\frac{3(n-3)}{2}p^5 \sim n^{3/2}/12$ and

$$\begin{aligned}\mathrm{Var}(Y) &= \sum_{T_1,T_2 \text{ sharing an edge}} \mathrm{Var}(Y_{T_1T_2}) + \sum_{T_1,..T_4 \text{ sharing some edges}} \mathrm{Cov}(X_{T_1T_2}, X_{T_3T_4}) \\ &\le \binom{n}{3}\frac{3(n-3)}{2}p^5 + O(n^6p^9) \\ &\le O(n^{3/2})\end{aligned}$$

To bound the covariances, one needs to distinguish cases depending on number of intersections. So by Chebyshev $\mathbb{P}(|Y - \mathbb{E}(Y)| \ge \frac{n^{3/2}}{2}0) = O(n^{-3/2})$. Hence asymptotically almost surely, $H(n,p)$ has constant average degree and thus a linear size independent set.

★ Exercise 7. Threshold for connectedness

Let $p = \frac{\ln(n)+c_n}{n}$. The goal of this exercise is to determine the connectedness threshold of $G(n,p)$.

1. Prove that $\mathbb{P}[G(n,p)$ has no isolated vertex$] \to \begin{cases} 0 & \text{if } c_n \to -\infty \text{ arbitrarily slowly} \\ 1 & \text{if } c_n \to \infty \text{ arbitrarily slowly} \end{cases}$

2. By considering X_k count the number of connected components of size exactly k in $G(n,p)$, show that $\mathbb{P}[G(n,p)$is connected$] = 1 - o(1)$ if $c_n \to \infty$ arbitrarily slowly

★ Solution 7.

1. Let X count the number of isolated vertices in $G(n,p)$. By linearity of expectation, we have

$$\begin{aligned}\mathbb{E}(X) &= n(1-p)^{n-1} \\ &= \exp\left(\ln(n) - (n-1)\frac{\ln(n)+c_n}{n} + O\left(\left(\frac{\ln(n)+c_n}{n}\right)^2\right)\right) \qquad \text{because } 1 - y = e^{-y+O(y^2)} \\ &= (1-o(1))\exp(-c_n)\end{aligned}$$

Hence, if $c_n \to \infty$, then $\mathbb{P}[G(n,p)$ has no isolated vertex$] = \mathbb{P}[X < 1] = o(1)$ by Markov's inequality. Conversely, let X_v be the indicator variable that v is an isolated vertex. We have $X = \sum_{v \in V(G)} X_v$ and $\mathrm{Var}(X) = \sum_{v \in V} \mathrm{Var}(X_v) + \sum_{u \neq v \in V} \mathrm{Cov}(X_v, X_u)$. We have $\mathrm{Var}(X_v) = (1-p)^{n-1}(1-(1-p)^{n-1})$ and $\mathrm{Cov}(X_u, X_v) = p(1-p)^{2n-3}$ and for $\lambda \in (0,1)$,

$$\begin{aligned}
\mathbb{P}(X \leq (1-\lambda)\mathbb{E}[X]) &\leq \frac{\mathrm{Var}(X)}{\lambda^2 \mathbb{E}[X]^2} && \text{by Chebyshev's inequality} \\
&\leq \sum_v \frac{1-(1-p)^{n-1}}{\lambda^2 n(1-p)^{n-1}} + \frac{p}{2\lambda^2(1-p)} = o(1)
\end{aligned}$$

2. Let $Y = \sum_{k=1}^n X_k$ and $X = \sum_{k \leq n/2} X_k$. If $Y \geq 2$ then one of the connected components of $G(n,p)$ must have size at most $n/2$, so $X \geq 1$. Thus,

$$\begin{aligned}
\mathbb{P}[Y \geq 2] &\leq \mathbb{P}[X \leq 1] \\
&\leq \mathbb{E}[X] && \text{by Markov's inequality} \\
&\leq \sum_{k=1}^{\lfloor n/2 \rfloor} \mathbb{E}[X_k] && \text{by linearity of Expectation}
\end{aligned}$$

We have

$$\begin{aligned}
\mathbb{E}[X_k] &\leq \binom{n}{k}(1-p)^{k(n-k)} \\
&\leq \left(\frac{ne}{k}\right)^k e^{-pk(n-k)} && \text{By Stirling approximation formula} \\
&\leq \left(\frac{\exp(\ln(n) - p(n-k) + 1)}{k}\right)^k
\end{aligned}$$

Without loss of generality, we assume that $c_n = o(\ln(n))$. We distinguish two cases. If $k \in [n^{3/4}, n/2]$, then

$$\begin{aligned}
\mathbb{E}[X_k] &\leq \left(\frac{\exp((\ln(n) + c_n)/2 + 1}{n^{3/4}}\right)^k \\
&\leq o(1/n)
\end{aligned}$$

If $2 \leq k \leq n^{3/4}$, then

$$\begin{aligned}
\mathbb{E}[X_k] &\leq \left(\frac{\exp(\ln(n) - (\ln(n) + c_n)(1 + o(1)))}{k}\right)^k \\
&\leq \left(\frac{\exp(-c_n(1 + o(1)))}{k}\right)^k = o(2^{-k})
\end{aligned}$$

Combining these bounds and the fact that $E(X_1) = o(1)$ is the number of isolated vertices,

$$\begin{aligned}
\mathbb{P}[Y \geq 2] &\leq o(1) + \sum_{k=2}^{n^{3/4}} \mathbb{E}[X_k] + \sum_{k=n^{3/4}}^{n/2} \mathbb{E}[X_k] \\
&\leq o(1) + o(\sum_{k=2}^{n^{3/4}} 2^{-k}) + n \cdot o(1/n) \\
&\leq o(1)
\end{aligned}$$

So $G(n,p)$ is connected with high probability.