NOTES ON INTRODUCTION TO RANDOM GRAPHS

ZHIDAN LI

Abstract. This is the personal learning note on Frieze and Karoński's book *Introduction to Random Graphs* [FK23]. It might be a revised version of the contents discussed in the reading seminar hold in Spring 2024.

Contents

1. Mathematical Symbols and Technique Tools

Before we start our discussion on random graphs, it is of great necessity to state some mathematical symbols and technique tools for completeness.

1.1. Inequalities.

For binomial inequalities, the most important tool we use is Stirling's approximation: for every $n \in \mathbb{N}_{>0}$,

$$\sqrt{2\pi n}\left(\frac{n}{e}\right)^n e^{\frac{1}{12n+1}} < n! < \sqrt{2\pi n}\left(\frac{n}{e}\right)^n e^{\frac{1}{12n}}.$$

Lemma 1.1. *The following identities and inequalities hold*

(1) For all $n, k \in \mathbb{N}_{>0}$, $k \le n$,

$$\binom{n}{k} \le \left(\frac{en}{k}\right)^k.$$

(2) For $n \in \mathbb{N}_{>0}$ and $k = o(n)$, it holds that

$$\binom{n}{k} \approx \frac{n^k}{k!}.$$

(3) For all $n, k \in \mathbb{N}_{>0}$,

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

1.2. **Probabilistic methods.** The most common tools we use are the moment methods, especially the *first-moment method (the Markov inequality)* and *the second-moment method (the Chebyshev inequality).*

Lemma 1.2 (The Markov Inequality)**.** *Let X be a non-negative random variable. Then for all $t > 0$,*

$$\mathbf{Pr}\left[X \geq t\right] \leq \frac{\mathbf{E}\left[X\right]}{t}.$$

Theorem 1.3 (The First-Moment Method)**.** *Let X be a non-negative integer-valued random variable. Then*

$$\mathbf{Pr}\left[X > 0\right] \leq \mathbf{E}\left[X\right].$$

Lemma 1.4 (The Chebyshev Inequality)**.** *Let X be a random variable with finite mean and finite variance. Then for $t > 0$, it holds that*

$$\mathbf{Pr}\left[\left|X - \mathbf{E}\left[X\right]\right| \geq t\right] \leq \frac{\mathbf{Var}\left(X\right)}{t^2}.$$

Theorem 1.5 (The Second-Moment Method)**.** *Let X be a non-negative integer-valued random variable. Then*

$$\mathbf{Pr}\left[X = 0\right] \leq \frac{\mathbf{Var}\left(X\right)}{\mathbf{E}\left[X\right]^2}. \tag{1}$$

Furthermore, it holds that

$$\mathbf{Pr}\left[X = 0\right] \leq \frac{\mathbf{Var}\left(X\right)}{\mathbf{E}\left[X^2\right]}. \tag{2}$$

Proof. The first inequality is easy to show by Lemma 1.4. For the second one, note that

$$X = X \cdot \mathbb{1}\left[X \geq 1\right].$$

Then by the Cauchy-Schwarz inequality,

$$\mathbf{E}\left[X\right]^2 = \left(\mathbf{E}\left[X \cdot \mathbb{1}\left[X \geq 1\right]\right]\right)^2 \leq \mathbf{E}\left[X^2\right]\mathbf{Pr}\left[X \geq 1\right].$$

□

1.2.1. *Poisson approximation.* Now we introduce the Poisson approximation which is useful in many aspects of probability.

Theorem 1.6. *Let $S_n = \sum_{i=1}^n I_i$ be a sequence of random variables, $n \geq 1$ and let $B_k^{(n)} = \mathbf{E}\left[\binom{S_n}{k}\right]$. Suppose that there exists $\lambda \geq 0$ such that for every fixed $k \geq 1$,*

$$\lim_{n\to\infty} B_k^{(n)} = \frac{\lambda^k}{k!}.$$

Then for every $j \geq 0$,

$$\lim_{n\to\infty} \mathbf{Pr}\left[S_n = j\right] = e^{-\lambda}\frac{\lambda^j}{j!}.$$

2. Basic Models of Random Graphs

Before we begin all studies on properties, firstly we introduce the models that we usually take into account.

Let $\mathscr{G}_{n,m}$ be the collection of all graphs $G = (V, E)$ with $|V| = n$ and $|E| = m$. For convenience, we assume that $V = \{1, \dots, n\}$. To ensure that $\mathscr{G}_{n,m}$ is well-defined, always suppose that $0 \le m \le \binom{n}{2}$. For every $G \in \mathscr{G}_{n,m}$, we equip it with probability

$$\mathbb{P}(G) = \binom{\binom{n}{2}}{m}^{-1}.$$

It's easy to note that following the probability, we draw a graph with n vertices and m edges uniformly at random. We denote this random graph by $\mathcal{G}_{n,m} = (V = [n], E_{n,m})$ and call it a *uniform random graph.*

Another random graph model we consider is similar. Given a real $p \in [0, 1]$. For $0 \le m \le \binom{n}{2}$ and every graph $G = (V, E)$ with $|V| = n$ and $|E| = m$, we assign to G the probability

$$\mathbb{P}(G) = p^m(1-p)^{\binom{n}{2}-m}.$$

We denote this random graph by $\mathcal{G}_{n,p} = (V = [n], E_{n,p})$ and call it an *Erdős-Rényi random graph.*

The two models are strongly related to each other.

Lemma 2.1. *A random graph $\mathcal{G}_{n,p}$ given that the number of its edge is m, is equally likely to be one of the graph $G \sim \mathcal{G}_{n,m}$.*

Proof. For every $G = (V, E)$ with $|E| = m$, simply we can observe that

$$\{\mathcal{G}_{n,p} = G\} \subseteq \{|E_{n,p}| = m\}.$$

Then by calculation,

$$\begin{aligned}
\mathbf{Pr}[\mathcal{G}_{n,p} = G \mid |E_{n,p}| = m] &= \frac{\mathbf{Pr}[\mathcal{G}_{n,p} = G \wedge |E_{n,p}| = m]}{\mathbf{Pr}[|E_{n,p}| = m]} \\
&= \frac{p^m(1-p)^{\binom{n}{2}-m}}{p^m(1-p)^{\binom{n}{2}-m}\binom{\binom{n}{2}}{m}} \\
&= \binom{\binom{n}{2}}{m}^{-1} \\
&= \mathbf{Pr}[\mathcal{G}_{n,m} = G].
\end{aligned}$$

□

Intuitively, the two random graphs perform a similar fashion when m is closed to the expected number of the edges of $\mathcal{G}_{n,p}$, *i.e.*,

$$m = \binom{n}{2}p = (1 + o(1))\frac{n^2 p}{2}$$

or

$$p = \frac{m}{\binom{n}{2}} = (1 + o(1))\frac{2m}{n^2}.$$

To generate the random graphs, we usually apply a coupling technique. Suppose that $p_1 < p$ and p_2 is defined by

$$1 - p = (1 - p_1)(1 - p_2).$$

Now we independently draw $\mathcal{G}(n, p_1)$ and $\mathcal{G}(n, p_2)$, and let $\mathcal{G}_{n,p} = \mathcal{G}(n, p_1) \cup \mathcal{G}(n, p_2)$. So when we write

$$\mathcal{G}(n, p_1) \subseteq \mathcal{G}_{n,p},$$

it means that the two graphs are coupled so that $\mathcal{G}_{n,p}$ is obtained from $\mathcal{G}(n, p_1)$ by the method described above.

To introduce a similar coupling process for $\mathcal{G}_{n,m}$, firstly consider $m_1 < m$. Then let

$$\mathcal{G}_{n,m} = \mathcal{G}(n, m_1) \cup \mathcal{H}$$

where $\mathcal{H}$ is a random graph with exactly $m_2 = m - m_1$ edges uniformly generated from $\binom{[n]}{2} \setminus E_{n,m_1}$.

Pseudo-random graphs. Besides the 'real' random graph models, the following two models will be taken into account.

- **Model A:** Let $\mathbf{x} = (x_1, \ldots, x_{2m})$ be chosen uniformly at random from $[n]^{2m}$.
- **Model B:** Let $\mathbf{x} = (x_1, \ldots, x_{2m})$ be chosen uniformly at random from $\binom{[n]}{2}^m$.

For $X \in \{A, B\}$, we construct the random graph $\mathcal{G}_{n,m}^{(X)}$ with the vertex set $[n]$ and edge set $E_m = \{(x_{2i-1}, x_{2i}) : i = 1, \ldots, n$ Note that the graph might be a multi-graph. To generate the simple graph $\mathcal{G}_{n,m}^{(X,-)}$ with m^- edges, we remove all self-loops and multiple edges. It can be seen that conditional the value of m^-, the simple graphs generated by the above two models are distributed the same as $\mathcal{G}_{n,m}$.

Also, it holds that, by symmetry for every $G_1 \in \mathscr{G}_{n,m}$ and $G_2 \in \mathscr{G}_{n,m}$,

$$\mathbf{Pr}\left[\mathcal{G}_{n,m}^{(X)} = G_1 \;\middle|\; \mathcal{G}_{n,m}^{(X)} \text{ is simple}\right] = \mathbf{Pr}\left[\mathcal{G}_{n,m}^{(X)} = G_2 \;\middle|\; \mathcal{G}_{n,m}^{(X)} \text{ is simple}\right]$$

for $X \in \{A, B\}$.

When $m = cn$ with constant parameter $c > 0$, it holds that

$$\mathbf{Pr}\left[\mathcal{G}_{n,m}^{(X)} \text{ is simple}\right] \geq \binom{\binom{n}{2}}{m} \frac{m! 2^m}{n^{2m}} \geq (1 - o(1)) \exp\left(-c^2 - c\right).$$

Then we know that

$$\mathbf{Pr}\left[\mathcal{G}_{n,m} \in \mathscr{P}\right] = \mathbf{Pr}\left[\mathcal{G}_{n,m}^{(X)} \in \mathscr{P} \;\middle|\; \mathcal{G}_{n,m}^{(X)} \text{ is simple}\right] \leq (1 + o(1)) e^{c^2 + c} \mathbf{Pr}\left[\mathcal{G}_{n,m}^{(X)} \in \mathscr{P}\right].$$

Then to show the random graph does not satisfy some graph property, when $m = O(n)$, it is feasible to turn to the pseudo-random graph models.

2.1. **Results on random graph properties.** Now we consider the property of graphs.

Definition 2.2 (Graph Property)**.** Fix a vertex set $V = [n]$. A *graph property* $\mathscr{P}$ is a collection of graphs $G = (V, E)$ where $E \subseteq \binom{[n]}{2}$.

Lemma 2.3. *Let $\mathscr{P}$ be any graph property and $p = m/\binom{n}{2}$ where $m = m(n) \to \infty$ and $\binom{n}{2} - m \to \infty$ as $n \to \infty$. Then for sufficiently large n,*

$$\mathbf{Pr}\left[\mathcal{G}_{n,m} \in \mathscr{P}\right] \leq 10 m^{1/2} \mathbf{Pr}\left[\mathcal{G}_{n,p} \in \mathscr{P}\right].$$

Proof. By the law of total probability,

$$\begin{aligned}
\mathbf{Pr}\left[\mathcal{G}_{n,p} \in \mathscr{P}\right] &= \sum_{k=0}^{\binom{n}{2}} \mathbf{Pr}\left[\mathcal{G}_{n,p} \in \mathscr{P} \mid |E_{n,p} = k|\right] \mathbf{Pr}\left[|E_{n,p}| = k\right] \\
&= \sum_{k=0}^{\binom{n}{k}} \mathbf{Pr}\left[\mathcal{G}(n,k) \in \mathscr{P}\right] \mathbf{Pr}\left[|E_{n,p}| = k\right] \\
&\geq \mathbf{Pr}\left[\mathcal{G}_{n,m} \in \mathscr{P}\right] \mathbf{Pr}\left[|E_{n,p}| = m\right]
\end{aligned}$$

where the second equality holds by Lemma 2.1. Now it suffices to estimate the term $\mathbf{Pr}\left[|E_{n,p}| = m\right]$. By definition,

$$\mathbf{Pr}\left[|E_{n,p}| = m\right] = \binom{\binom{n}{2}}{m} p^m (1 - p)^{\binom{n}{2} - m}.$$

By Stirling's formula,

$$k! = (1 + o(1)) \sqrt{2\pi k} \frac{k^k}{e^k}.$$

Then when $m = m(n) \to \infty$ and $\binom{n}{2} - m \to \infty$ as $n \to \infty$,

$$\begin{aligned}\mathbf{Pr}\left[|E_{n,p}| = m\right] &= (1 + o(1))\sqrt{\frac{\binom{n}{2}}{2\pi m(\binom{n}{2} - m)}} \\ &\geq \frac{1}{10\sqrt{m}}.\end{aligned}$$

Putting it into the above inequality we conclude the lemma. □

When the property $\mathscr{P}$ is so called *monotone increasing*, the result of Lemma 2.3 can be tightened.

Definition 2.4 (Monotone Increasing Graph Property)**.** A graph property $\mathscr{P}$ is said to be *monotone increasing* if $G \in \mathscr{P}$ implies $G + e \in \mathscr{P}$. Furthermore, it is said to be *non-trivial* if the empty graph $\varnothing \notin \mathscr{P}$ and the complete graph $K_n \in \mathscr{P}$.

Remark 2.5. From the view of coupling, if $\mathscr{P}$ is monotone increasing, then whenever $p \leq p'$ or $m < m'$, if $\mathcal{G}_{n,p} \in \mathscr{P}$ or $\mathcal{G}_{n,m} \in \mathscr{P}$, then

$$\mathcal{G}(n, p') \in \mathscr{P}, \quad \mathcal{G}(n, m_1) \in \mathscr{P}.$$

Lemma 2.6. *Let $\mathscr{P}$ be a monotone increasing graph property. Given integers $n, m > 0$, fix $p = \frac{m}{N}$ where $N = \binom{n}{2}$. Then for large n and $p = o(1)$ such that $Np, \frac{N(1-p)}{\sqrt{Np}} \to \infty$ as $n \to \infty$,*

$$\mathbf{Pr}\left[\mathcal{G}_{n,m} \in \mathscr{P}\right] \leq 3\mathbf{Pr}\left[\mathcal{G}_{n,p} \in \mathscr{P}\right].$$

Proof. Since $\mathscr{P}$ is monotone increasing, we know

$$\mathbf{Pr}\left[\mathcal{G}_{n,p} \in \mathscr{P}\right] \geq \sum_{k=m}^{N} \mathbf{Pr}\left[\mathcal{G}(n,k) \in \mathscr{P}\right] \mathbf{Pr}\left[|E_{n,p}| = k\right].$$

By Remark 2.5, for $m \leq k \leq N$,

$$\mathbf{Pr}\left[\mathcal{G}(n,k)\right] \geq \mathbf{Pr}\left[\mathcal{G}_{n,m} \in \mathscr{P}\right].$$

Then we know

$$\mathbf{Pr}\left[\mathcal{G}_{n,p} \in \mathscr{P}\right] \geq \mathbf{Pr}\left[\mathcal{G}_{n,m} \in \mathscr{P}\right] \sum_{k=m}^{N} u_k$$

where

$$u_k = \binom{N}{k} p^k (1-p)^{N-k}.$$

Using Stirling's formula, we know

$$u_m = \frac{1 + o(1)}{(2\pi m)^{1/2}}.$$

For $0 \leq k - m \leq m^{1/2}$, we know

$$\frac{u_{k+1}}{u_k} = \frac{(N-k)p}{(k+1)(1-p)} \geq \exp\left(-\frac{k-m}{N-k} - \frac{m-k+1}{m}\right).$$

Then it follows that for $0 \leq t \leq m^{1/2}$,

$$u_{m+t} \geq \frac{\exp\left(-\frac{t^2}{2m} - o(1)\right)}{(2\pi m)^{1/2}}.$$

Then we know

$$\sum_{k=m}^{N} u_k \geq \sum_{t=0}^{m^{1/2}} u_{m+t} \geq \frac{1 - o(1)}{(2\pi)^{1/2}} \int_0^1 e^{-x^2/2}\,\mathrm{d}x \geq \frac{1}{3}.$$

This concludes our lemma. □

Lemmas 2.3 and 2.6 show us that if we want to prove $\mathbf{Pr}\left[\mathcal{G}_{n,m} \in \mathscr{P}\right] \to 0$, it suffices to show $\mathbf{Pr}\left[\mathcal{G}_{n,p} \in \mathscr{P}\right] \to 0$. In most cases, $\mathbf{Pr}\left[\mathcal{G}_{n,p} \in \mathscr{P}\right]$ is much easier to compute.

To get rid of the limit between m and p, we have the following asymptotic version.

Theorem 2.7 ([Łuc90])**.** *Let $0 \le p_0 \le 1$ be a real, $s(n) = n\sqrt{p(1-p)} \to \infty$, and $\omega(n) \to \infty$ arbitrary slowly as $n \to \infty$.*

(1) Suppose that $\mathscr{P}$ is a graph property such that $\mathbf{Pr}\left[\mathcal{G}_{n,m} \in \mathscr{P}\right] \to p_0$ for all

$$m \in \left[\binom{n}{2}p - \omega(n)s(n), \binom{n}{2}p + \omega(n)s(n)\right].$$

Then $\mathbf{Pr}\left[\mathcal{G}_{n,p} \in \mathscr{P}\right] \to p_0$ as $n \to \infty$.

(2) Let $p_- = p - \omega(n)s(n)/n^2$ and $p_+ = p + \omega(n)s(n)/n^2$. Suppose that $\mathscr{P}$ is a monotone increasing graph property such that $\mathbf{Pr}\left[\mathcal{G}(n,p_-)\right] \to p_0$ and $\mathbf{Pr}\left[\mathcal{G}(n,p_+)\right] \to p_0$. Then $\mathbf{Pr}\left[\mathcal{G}_{n,m} \in \mathscr{P}\right] \to p_0$ for $m = \left\lfloor \binom{n}{2}p \right\rfloor$.

2.2. **Thresholds and sharp thresholds.** One of the most important observations is that, for a monotone increasing graph property, there might exist a 'threshold'.

Definition 2.8 (Thresholds for $\mathcal{G}_{n,m}$)**.** A function $m^* = m^*(n)$ is called a *threshold* for a monotone increasing property $\mathscr{P}$ in the random graph $\mathcal{G}_{n,m}$ if

$$\lim_{n\to\infty} \mathbf{Pr}\left[\mathcal{G}_{n,m} \in \mathscr{P}\right] = \begin{cases} 0 & m/m^* \to 0, \\ 1 & m/m^* \to \infty. \end{cases}$$

Definition 2.9 (Thresholds for $\mathcal{G}_{n,p}$)**.** A function $p^* = p^*(n)$ is called a *threshold* for a monotone increasing property $\mathscr{P}$ in the random graph $\mathcal{G}_{n,p}$ if

$$\lim_{n\to\infty} \mathbf{Pr}\left[\mathcal{G}_{n,p} \in \mathscr{P}\right] = \begin{cases} 0 & p/p^* \to 0, \\ 1 & p/p^* \to \infty. \end{cases}$$

Remark 2.10. The threshold is not unique since any function which differs from $m^*(n)$ (or $p^*(n)$) by only a constant factor is also a threshold.

Theorem 2.11. *Every non-trivial monotone graph property has a threshold.*

Proof. Without loss of generality, we assume that $\mathscr{P}$ is monotone increasing. Given $0 < \varepsilon < 1$, we define $p(\varepsilon)$ by

$$\mathbf{Pr}\left[\mathcal{G}_{n,p(\varepsilon)} \in \mathscr{P}\right] = \varepsilon.$$

Before the proof, firstly we argue that $p(\varepsilon)$ exists. Note that, for every $0 \le p \le 1$,

$$\mathbf{Pr}\left[\mathcal{G}_{n,p} \in \mathscr{P}\right] = \sum_{G\in\mathscr{P}} p^{|E(G)|}(1-p)^{N-|E(G)|}$$

is a polynomial increasing from 0 to 1. Then we know $p(\varepsilon)$ exists.

Now we will show $p(1/2)$ is a threshold for $\mathscr{P}$. Let $G_1, \ldots, G_k$ be k independent copies of $\mathcal{G}_{n,p}$. Then the graph $G = G_1 \cup \ldots \cup G_k$ is distributed as $\mathcal{G}_{n,1-(1-p)^k}$. Note that $1-(1-p)^k \le kp$. By the coupling argument,

$$\mathcal{G}_{n,1-(1-p)^k} \subseteq \mathcal{G}_{n,kp}.$$

And so, $\mathcal{G}_{n,kp} \notin \mathscr{P}$ implies $G_1, \ldots, G_k \notin \mathscr{P}$ (by monotonicity). Hence,

$$\mathbf{Pr}\left[\mathcal{G}_{n,kp} \notin \mathscr{P}\right] \le \mathbf{Pr}\left[\mathcal{G}_{n,p} \notin \mathscr{P}\right]^k.$$

Then, for any $\omega(n) \to \infty$ arbitrarily slowly as $n \to \infty$ and $\omega(n) \ll \log\log n$, we know

$$\mathbf{Pr}\left[\mathcal{G}_{n,\omega(n)p(1/2)} \notin \mathscr{P}\right] \le 2^{-\omega} = o(1).$$

On the other hand, for $p = p(1/2)/\omega(n)$, we know

$$\mathbf{Pr}\left[\mathcal{G}_{n,p(1/2)/\omega(n)} \notin \mathscr{P}\right] \ge 2^{-1/\omega} = 1 - o(1).$$

$\square$

By observation, there exists a more subtle threshold for some monotone graph properties.

Definition 2.12 (Sharp Thresholds for $\mathcal{G}_{n,m}$)**.** A function $m^* = m^*(n)$ is called a *sharp threshold* for a monotone increasing property $\mathscr{P}$ in the random graph $\mathcal{G}_{n,m}$ if for every $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbf{Pr}\left[\mathcal{G}_{n,m} \in \mathscr{P}\right] = \begin{cases} 0 & m/m^* \leq 1-\varepsilon, \\ 1 & m/m^* \geq 1+\varepsilon. \end{cases}$$

Definition 2.13 (Sharp Thresholds for $\mathcal{G}_{n,p}$)**.** A function $p^* = p^*(n)$ is called a *sharp threshold* for a monotone increasing property $\mathscr{P}$ in the random graph $\mathcal{G}_{n,p}$ if for every $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbf{Pr}\left[\mathcal{G}_{n,p} \in \mathscr{P}\right] = \begin{cases} 0 & p/p^* \leq 1-\varepsilon, \\ 1 & p/p^* \geq 1+\varepsilon. \end{cases}$$

To illustrate Definitions 2.8 and 2.9 more precisely, we state the following simple example. We deal with the graph $\mathcal{G}_{n,p}$ and the property

$$\mathscr{P} = \{G = (V(G), E(G)) \mid V(G) = n, E(G) \neq \varnothing\}. \tag{3}$$

Now we will show $p^* = 1/n^2$ is a threshold.

Theorem 2.14. *Let $\mathscr{P}$ be the graph property defined as (3). Then*

$$\lim_{n\to\infty} \mathbf{Pr}\left[\mathcal{G}_{n,p} \in \mathscr{P}\right] = \begin{cases} 0 & p \ll n^{-2}, \\ 1 & p \gg n^{-2}. \end{cases}$$

Proof. Let X be the number of edges in $\mathcal{G}_{n,p}$. By the definition of the random model, it holds that

$$\mathbf{E}\left[X\right] = \binom{n}{2}p, \quad \mathbf{Var}\left(X\right) = \binom{n}{2}p(1-p) = (1-p)\mathbf{E}\left[X\right].$$

By the Markov inequality, it holds that

$$\mathbf{Pr}\left[X > 0\right] \leq \mathbf{E}\left[X\right] \leq \frac{n^2}{2}p.$$

When $p \ll n^{-2}$, it holds that $\lim_{n\to\infty} \mathbf{Pr}\left[X > 0\right] = 0$. Thus we conclude the first part of the theorem.

To show the second result, we consider the concentration of the random variable X. By the Chebyshev inequality,

$$\mathbf{Pr}\left[X > 0\right] \geq 1 - \frac{\mathbf{Var}\left(X\right)}{\mathbf{E}\left[X\right]^2} = 1 - \frac{1-p}{\mathbf{E}\left[X\right]}.$$

When $p \gg n^{-2}$, it holds that $\frac{1-p}{\mathbf{E}[X]} \to 0$ and we know $\lim_{n\to\infty} \mathbf{Pr}\left[X > 0\right] = 1$. $\square$

Now we consider the degree of a fixed vertex $v \in V$ in random graphs. By definition, it is easy to show:

$$\mathbf{Pr}_{\mathcal{G}_{n,p}}\left[\deg(v) = d\right] = \binom{n-1}{d} p^d (1-p)^{n-1-d}.$$

and for the model $\mathcal{G}_{n,m}$,

$$\mathbf{Pr}_{\mathcal{G}_{n,m}}\left[\deg(v) = d\right] = \frac{\binom{n-1}{d}\binom{\binom{n-1}{2}}{m-d}}{\binom{\binom{n}{2}}{m}}.$$

Let $\mathscr{P}$ be the graph property such that the graph contains an isolated vertex, *i.e.*,

$$\mathscr{P} := \{G = (V(G), E(G)) \mid \exists v \in V(G), \deg(v) = 0\}.$$

Now we show $m = \frac{1}{2}n\log n$ is a sharp threshold for $\mathscr{P}$ in $\mathcal{G}_{n,m}$.

Lemma 2.15. *Let $\mathscr{P}$ be the property defined as above, and $m = \frac{1}{2}n(\log n + \omega(n))$. Then*

$$\lim_{n\to\infty} \mathbf{Pr}\left[\mathcal{G}_{n,m} \in \mathscr{P}\right] = \begin{cases} 1 & \omega(n) \to -\infty, \\ 0 & \omega(n) \to \infty. \end{cases}$$

Proof. We define a random variable X as the number of isolated vertices in $\mathcal{G}_{n,m}$, and for every $v \in V$, we define a random variable I_v to denote whether v is isolated. Then

$$X = \sum_{v\in V} I_v$$

and for each $v \in V$,

$$\begin{aligned} \mathbf{E}\left[I_v\right] &= \mathbf{Pr}\left[I_v = 1\right] \\ &= \binom{\binom{n-1}{2}}{m} \Big/ \binom{\binom{n}{2}}{m} \\ &= \prod_{i=0}^{m-1} \left(\frac{\frac{(n-1)(n-2)}{2} - i}{\frac{n(n-1)}{2} - i} \right) \\ &= \left(\frac{n-2}{n}\right)^m \prod_{i=0}^{m-1} \left(1 - \frac{4i}{n(n-1)(n-2) - 2i(n-2)}\right). \end{aligned}$$

Thus we obtain

$$\begin{aligned} \mathbf{E}\left[X\right] &= \sum_{v\in V} \mathbf{E}\left[I_v\right] \\ &= n\left(\frac{n-2}{n}\right)^m \prod_{i=0}^{m-1} \left(1 - \frac{4i}{n(n-1)(n-2) - 2i(n-2)}\right). \end{aligned}$$

To bound the product, notice that, if $0 \le x_0, \ldots, x_{m-1} \le 1$, it holds that

$$n\left(1 - \sum_{i=0}^{m-1} x_i\right) \le n \prod_{i=0}^{m-1} (1 - x_i) \le n.$$

Thus we obtain that, if we assume that $\omega(n) = o(\log n)$,

$$n\left(\frac{n-2}{n}\right)^m \prod_{i=0}^{m-1} \left(1 - \frac{4i}{n(n-1)(n-2) - 2i(n-2)}\right) \le n\left(1 - \frac{2}{n}\right)^m \le e^{-\omega(n)}.$$

When $\omega(n) \to \infty$, we know $\mathbf{E}\left[X\right] \to 0$ and by the first-moment method, we know $X = 0$ with high probability.

For the counterpart, note that

$$\prod_{i=0}^{m-1} \left(1 - \frac{4i}{n(n-1)(n-2) - 2i(n-2)}\right) \ge 1 - \frac{4}{n-2} \sum_{i=0}^{m-1} \frac{i}{n(n-1) - 2i} = 1 - O\left(\frac{(\log n)^2}{n}\right).$$

Then it holds that

$$\mathbf{E}\left[X\right] = (1 - o(1))n\left(\frac{n-2}{n}\right)^m \ge (1 - o(1))ne^{-\frac{2m}{n-2}} \ge (1 - o(1))e^{-\omega(n)} \to \infty.$$

Also, to show the concentration of X, we compute the second moment of X. By calculation,

$$\begin{aligned}\mathbf{E}\left[X^2\right] &= \mathbf{E}\left[\left(\sum_{v\in V} I_v\right)^2\right] \\ &= \sum_{u,v\in V} \mathbf{Pr}\left[I_u = I_v = 1\right] \\ &= n(n-1)\binom{\binom{n-2}{2}}{m} / \binom{\binom{n}{2}}{m} + \mathbf{E}\left[X\right] \\ &\le (1+o(1))\mathbf{E}\left[X\right]^2 + \mathbf{E}\left[X\right].\end{aligned}$$

Then we know

$$\mathbf{Pr}\left[X > 0\right] \ge \frac{\mathbf{E}\left[X\right]^2}{\mathbf{E}\left[X^2\right]} \ge \frac{1}{1 + o(1) + \mathbf{E}\left[X\right]^{-1}} = 1 - o(1)$$

whenever $\omega(n) \to -\infty$. □

At the end of the part, we show a more complicated example.

Theorem 2.16. *If $m/n \to \infty$, then with high probability the random graph $\mathcal{G}_{n,m}$ contains a triangle.*

Proof. It is easy to observe that the property is monotone increasing. Then it suffices to show that, when p satisfies some regular requirements, the random graph $\mathcal{G}_{n,p}$ contains at least one triangle with high probability.

By coupling method, it suffices to show the case $\omega := np \le \log n$. Let Z be the random variable denoting the number of triangles in $\mathcal{G}_{n,p}$. Then

$$\mathbf{E}\left[Z\right] = \binom{n}{3} p^3 \ge \frac{(1-o(1))\omega^3}{6} \to \infty.$$

For the second moment, let $T_1, \ldots, T_M$ be the triangles of the complete graph K_n where $M = \binom{n}{3}$. Then,

$$\begin{aligned}\mathbf{E}\left[Z^2\right] &= \sum_{i,j=1}^{M} \mathbf{Pr}\left[T_i, T_j \in \mathcal{G}_{n,p}\right] \\ &= \sum_{i=1}^{M} \mathbf{Pr}\left[T_i \in \mathcal{G}_{n,p}\right] \sum_{j=1}^{M} \mathbf{Pr}\left[T_j \in \mathcal{G}_{n,p} \mid T_i \in \mathcal{G}_{n,p}\right] \\ &= M\mathbf{Pr}\left[T_1 \in \mathcal{G}_{n,p}\right] \sum_{j=1}^{M} \mathbf{Pr}\left[T_j \in \mathcal{G}_{n,p} \mid T_1 \in \mathcal{G}_{n,p}\right] \\ &= \mathbf{E}\left[Z\right] \sum_{j=1}^{M} \mathbf{Pr}\left[T_j \in \mathcal{G}_{n,p} \mid T_1 \in \mathcal{G}_{n,p}\right].\end{aligned}$$

Separating the summation according to the number of edges T_1, T_j share, we obtain

$$\begin{aligned}\sum_{j=1}^{M} \mathbf{Pr}\left[T_j \in \mathcal{G}_{n,p} \mid T_1 \in \mathcal{G}_{n,p}\right] &= 1 + 3(n-3)p^2 + \left(\binom{n}{3} - 3n + 8\right) p^3 \\ &\le 1 + \frac{3\omega^2}{n} + \mathbf{E}\left[Z\right].\end{aligned}$$

Then we know

$$\mathbf{Var}\left(Z\right) \le \mathbf{E}\left[Z\right]\left(1 + \frac{3\omega^2}{n} + \mathbf{E}\left[Z\right]\right) - \mathbf{E}\left[Z\right]^2 \le 2\mathbf{E}\left[Z\right].$$

By the Chebyshev inequality, we conclude

$$\mathbf{Pr}[Z=0] \leq \frac{\mathbf{Var}(Z)}{\mathbf{E}[Z]^2} \leq \frac{2}{\mathbf{E}[Z]} = o(1).$$

The result then comes immediately. □

3. Evolution of Random Graphs

Now we view the random graph model as an evolution process of a graph sequence:

$$G_0 = ([n], \varnothing) \subseteq G_1 \subseteq \ldots \subseteq G_N = K_n$$

where for $m \geq 1$, G_m is generated from G_{m-1} by uniformly and independently adding a remaining edge. Then we know G_m and $\mathcal{G}_{n,m}$ share the same distribution.

In this part, we will focus on the structure of the random graph $\mathcal{G}_{n,m}$ and $\mathcal{G}_{n,p}$. We view the random graph model $\mathcal{G}_{n,m}$ as the above evolution process, and for $\mathcal{G}_{n,p}$, we consider the structure along with the growth of $p = p(n)$ from 0 to 1.

3.1. **Sub-critical phase.** Firstly we focus on sub-critical phases. We set $m = o(n)$ and thus $np = o(1)$. Since most of the properties we consider are monotone, we assume that $\omega = \omega(n)$ is growing more slowly than n.

Theorem 3.1. *If $m \ll n$, then G_m is a forest w.h.p..*

Proof. Assume that $m = n/\omega$, $N = \binom{n}{2}$ and $p = m/N = \frac{2}{\omega(n-1)}$. Let X be the number of cycles in $\mathcal{G}_{n,p}$. Then

$$\begin{aligned}
\mathbf{E}[X] &= \sum_{k=3}^{n} \binom{n}{k} \frac{(k-1)!}{2} p^k \\
&\leq \sum_{k=3}^{n} \frac{n^k}{k!} \frac{(k-1)!}{2} \frac{4}{\omega^k n^k} \\
&\leq \sum_{k=3}^{n} \frac{1}{k} \left(\frac{1}{\omega}\right)^k \\
&= o(\omega^{-3}) \to 0.
\end{aligned}$$

Then we know $\mathbf{Pr}[X>0] \leq \mathbf{E}[X] = o(1)$. Then we conclude this theorem. □

Theorem 3.2. *The function $m^*(n) = n^{1/2}$ is the threshold for the property that a random graph G_m contains a path of length* 2.

Proof. Firstly we consider the case $m \ll n^{1/2}$. Assume that $m = n^{1/2}/\omega$ and $p = m/\binom{n}{2}$. Let X be the number of paths of length 2 in the random graph $\mathcal{G}_{n,p}$. Then by the Markov inequality,

$$\mathbf{Pr}[X>0] \leq \mathbf{E}[X] = 3\binom{n}{3}p^2 \leq O(\omega^{-2}) \to 0$$

as $n \to \infty$. Hence we conclude the negative aspect of the threshold with the fact that the property is monotone increasing.

When $m \gg n^{1/2}$, assume that $m = \omega n^{1/2}$ and $p = m/\binom{n}{2}$. Furthermore, let Path_2 be the collection of all paths of length 2 in the complete K_n. One might apply the second-moment method to obtain the result. However, the direct calculation might be complex.

Now we consider the random variable $\widehat{X}$ denoting the number of isolated paths of length 2. Assume that $m = o(n)$ and then $np = o(1)$. Then we know

$$\mathbf{E}\left[\widehat{X}\right] = 3\binom{n}{3}p^2(1-p)^{3(n-3)+1} \geq 2(1-o(1))\omega^2(1-3np) \to \infty$$

as $n \to \infty$.

Now we compute the second moment of $\widehat{X}$. For a path $P \in \mathsf{Path}_2$, we use $P \subseteq_i \mathcal{G}_{n,p}$ to denote the event P is an isolated path in $\mathcal{G}_{n,p}$. Then,

$$\mathbf{E}\left[\widehat{X}^2\right] = \sum_{P\in\mathsf{Path}_2}\sum_{Q\in\mathsf{Path}_2} \mathbf{Pr}\left[P \subseteq_i \mathcal{G}_{n,p} \wedge Q \subseteq_i \mathcal{G}_{n,p}\right].$$

Observe that, the event $\{P \subseteq_i \mathcal{G}_{n,p} \wedge Q \subseteq_i \mathcal{G}_{n,p}\}$ occurs only when $P = Q$ or P, Q are disjoint. Together with symmetry, we know

$$\begin{aligned}\mathbf{E}\left[\widehat{X}^2\right] &= \mathbf{E}\left[\widehat{X}\right]\left(1 + \sum_{Q\in\mathsf{Path}_2\wedge Q\cap\{1,2,3\}=\varnothing} \mathbf{Pr}\left[Q \subseteq_i \mathcal{G}_{n,p} \mid (1 \leftrightarrow 2 \leftrightarrow 3) \subseteq \mathcal{G}_{n,p}\right]\right)\\ &\le \mathbf{E}\left[\widehat{X}\right]\left(1 + 3\binom{n-3}{3}p^2(1-p)^{3(n-6)+1}\right)\\ &\le \mathbf{E}\left[\widehat{X}\right]\left(1 + (1-p)^{-9}\mathbf{E}\left[\widehat{X}\right]\right).\end{aligned}$$

Then by the second moment method,

$$\mathbf{Pr}\left[\widehat{X} > 0\right] \ge \frac{\mathbf{E}\left[\widehat{X}\right]^2}{\mathbf{E}\left[\widehat{X}^2\right]} \ge \frac{1}{(1-p)^{-9} + \mathbf{E}\left[\widehat{X}\right]^{-1}} \to 0$$

as $n \to \infty$. Then we know when $p = o(1)$ but $\binom{n}{2}p \gg n^{1/2}$, the random graph $\mathcal{G}_{n,p}$ contains a path of length 2 w.h.p.. By coupling method and the fact that the property is monotone increasing, we know that when $m \gg n^{1/2}$, G_m contains a path of length 2 w.h.p.. □

When the number of edges grows, trees occur in the random graph.

Theorem 3.3. *Given a fixed vertex $k \ge 3$, the function $m^*(n) = n^{\frac{k-2}{k-1}}$ is the threshold for the property that a random graph $\mathcal{G}_{n,m}$ contains a tree with $k \ge 3$ vertices.*

Proof. When $m \ll n^{\frac{k-2}{k-1}}$, assume that $m = n^{\frac{k-2}{k-1}}/\omega$. And then

$$p = \frac{m}{\binom{n}{2}} \le \frac{3}{\omega n^{\frac{k}{k-1}}}.$$

Let X_k be the random variable denoting the number of trees with k vertices a random graph $\mathcal{G}_{n,p}$ contains. Then by direct calculation,

$$\mathbf{E}\left[X_k\right] = \binom{n}{k}p^{k-1}k^{k-2}$$

where k^{k-2} comes from Caylay's lemma to count the number of trees formed by $[k]$. Then by the binomial inequality,

$$\begin{aligned}EX_k &\le \frac{n^k e^k}{k^k}k^{k-2}\frac{3^{k-1}}{\omega^{k-1}n^k}\\ &= \frac{e^k 3^{k-1}}{k^2\omega^{k-1}} \to 0\end{aligned}$$

as $n \to \infty$. Then we can prove the negative part by the similar argument we have been familiar with.

When $m \gg n^{\frac{k-2}{k-1}}$, assume that $m = \omega n^{\frac{k-2}{k-1}}$ and $p = m/\binom{n}{2}$. Since the property is monotone increasing, we further suppose that $\omega = o(\log n)$. Now we consider this much stronger property: given a tree T with k vertices, $\mathcal{G}_{n,p}$ contains an isolated copy of T w.h.p.. Let $\widehat{X}_k$ be the number of isolated copies of T in $\mathcal{G}_{n,p}$. For a graph

H, define $\mathrm{aut}(H)$ to be the number of automorphisms of H. Then by direct calculation, the number of copies of T is $\frac{k!}{\mathrm{aut}(T)}$. Then we know

$$\begin{aligned}\mathbf{E}\left[\widehat{X}_k\right] &= \binom{n}{k}\frac{k!}{\mathrm{aut}(T)}p^{k-1}(1-p)^{k(n-k)+\binom{k}{2}-(k-1)} \\ &= \frac{n(n-1)\cdots(n-k+1)}{k!}\frac{k!}{\mathrm{aut}(T)}\frac{(2\omega)^{k-1}n^{k-2}}{n^{k-1}(n-1)^{k-1}}(1-p)^{k(n-k)+\binom{k}{2}-(k-1)} \\ &\geq (1-o(1))\frac{(2\omega)^{k-1}}{\mathrm{aut}(T)} \to \infty\end{aligned}$$

as $n \to \infty$, where the last inequality comes from the fact that

$$(1-p)^{k(n-k)+\binom{k}{2}-(k-1)} \geq 1 - p\left(k(n-k) + \binom{k}{2} - (k-1)\right) = 1 - o(1).$$

To compute the second moment, we apply the similar method we use in Theorem 3.2 and obtain

$$\mathbf{E}\left[\widehat{X}_k^2\right] \geq (1-o(1))\mathbf{E}\left[\widehat{X}_k\right]\left(1 + (1-p)^{-k^2}\mathbf{E}\left[\widehat{X}_k\right]\right).$$

Then we know

$$\mathbf{Pr}\left[\widehat{X}_k > 0\right] \geq \frac{1+o(1)}{1+(1-p)^{-k^2}\mathbf{E}\left[\widehat{X}_k\right]^{-1}} = 1 - o(1).$$

What remains is quite routine and we omit it here. □

It is interesting that when m is exactly at the threshold, which kind of phenomena the random graph will perform?

Theorem 3.4. *Given $m = cn^{\frac{k-2}{k-1}}$ where $c > 0$ is a constant and a fixed k-vertex tree T with $k \geq 3$. Then*

$$\mathbf{Pr}\left[G_m \text{ contains an isolated copy of } T\right] \to 1 - e^{-\lambda}$$

as $n \to \infty$, where $\lambda = \frac{(2c)^{k-1}}{\mathrm{aut}(T)}$. In other words, the number of copies of T is asymptotically distributed as the Poisson distribution with expectation λ.

Proof. Let $T_1, \ldots, T_M$ be the copies of T in the complete graph K_n. For every $1 \leq i \leq M$, define the event

$$A_i := \{T_i \text{ occurs isolately in } G_m\}.$$

Suppose that $J \subseteq [M] = \{1, \ldots, M\}$ with $|J| = t$. Let $A_J = \cap_{i\in J}A_i$. If T_i, T_j share a point for some $i, j \in J$, then it holds that $\mathbf{Pr}[A_J] = 0$. Otherwise,

$$\mathbf{Pr}[A_J] = \binom{\binom{n-kt}{2}}{m-(k-1)t}\Big/\binom{\binom{n}{2}}{m}.$$

If say $t \leq \log n$, then we know

$$\binom{n-kt}{2} = N\left(1 - O(ktn^{-1})\right).$$

Then we know

$$\binom{\binom{n-kt}{2}}{m-(k-1)t} = (1+o(1))\frac{N^{m-(k-1)t}\left(1-O(mktn^{-1})\right)}{(m-(k-1)t)!}.$$

Then we know

$$\mathbf{Pr}[A_J] = (1+o(1))m^{(k-1)t}N^{-(k-1)t}.$$

Let Z_T denote the number of components of G_m which are copies of T. Then we know

$$\begin{aligned}\mathbf{E}\left[\binom{Z_T}{t}\right] &= (1+o(1))\frac{1}{t!}\binom{n}{k,\ldots,k}\left(\frac{k!}{\mathrm{aut}(T)}\right)^t\left(\frac{m}{N}\right)^{(k-1)t} \\ &= (1+o(1))\frac{2^{(k-1)t}n^{kt}}{t!(k!)^t}\frac{(k!)^t}{\mathrm{aut}(T)^k}\frac{n^{(k-2)t}c^{(k-1)t}}{n^{2(k-1)t}} \\ &= (1+o(1))\frac{\lambda^t}{t!}.\end{aligned}$$

Then by Theorem 1.6, we conclude the theorem. □

The last ingredient of this part is to show the size of the maximal component in very sparse random graphs.

Theorem 3.5. *If $m = \frac{1}{2}cn$, where $0 < c < 1$ is a constant, then w.h.p. the order of the largest component of a random graph G_m is $O(\log n)$.*

To prove Theorem 3.5, we need the following three lemmas together.

Lemma 3.6. *If $p \le \frac{1}{n} - \frac{\omega}{n^{4/3}}$ where $\omega = \omega(n) \to \infty$, then w.h.p. every component in $\mathcal{G}_{n,p}$ contains at most one cycle.*

Proof. We consider the graph with two cycles. Firstly it can be shown that the number of graphs with exactly two cycles of size k can be bounded by $k^2 \cdot k!$. Then, let X be the number of such graphs in $\mathcal{G}_{n,p}$. It can be shown that

$$\begin{aligned}\mathbf{E}[X] &\le \sum_{k=4}^{n}\binom{n}{k}k^2k!p^{k+1} \\ &\le \sum_{k=4}^{n}\frac{n^k}{k!}k^2k!\left(\frac{1}{n}-\frac{\omega}{n^{4/3}}\right)^{k+1} \\ &\le \sum_{k=4}^{n}k^2n^{-1}\left(1-\omega n^{-1/3}\right)^{k+1} \\ &\le \frac{1}{n}\int_0^\infty x^2(1-\omega n^{-1/3})^{x+1}\,\mathrm{d}x \\ &\le \frac{1}{n}\int_0^\infty x^2\exp\Big(-\omega x n^{-1/3}\Big)\,\mathrm{d}x \\ &= \frac{2}{\omega^3} \\ &= o(1).\end{aligned}$$

Then by the first-moment method, we know $\mathbf{Pr}[X > 0] = o(1)$. □

Remark 3.7. When $p = c/n$ where $0 < c < 1$, then we know

$$\mathbf{Pr}[X > 0] \le \sum_{k=4}^{n} k^2c^{k+1}n^{-1} = O(n^{-1}).$$

Lemma 3.6 means we can only focus on unicyclic components and tree-components. The next lemma will show that the number of vertices on unicyclic components is rather small.

Lemma 3.8. *Let $p = c/n$ where $c \neq 1$ is a constant. Then in $\mathcal{G}_{n,p}$ w.h.p. the number of vertices in components with exactly one cycle is $O(\omega)$ for any growing function $\omega = \omega(n) \to \infty$.*

Proof. Let X_k be the number of vertices on unicyclic components with k vertices. Then

$$\mathbf{E}[X_k] \le \binom{n}{k}k^{k-2}k\binom{k}{2}p^k(1-p)^{k(n-k)+\binom{k}{2}-k}.$$

Firstly we assume that $c < 1$. Applying the inequality

$$\binom{n}{k} \le \frac{n^k}{k!} e^{-\frac{k(k-1)}{2n}},$$

we obtain

$$\begin{aligned}
\mathbf{E}\left[X_k\right] &\le \frac{n^k}{k!} e^{\frac{k(k-1)}{2n}} k^{k+1} \frac{c^k}{n^k} \left(1 - \frac{c}{n}\right)^{k(n-k)+\binom{k}{2}-k} \\
&\le \frac{n^k}{k!} k^{k+1} \frac{c^k}{n^k} \exp\left(-\frac{k(k-1)}{2n} - \frac{ck(n-k)}{n} - \frac{ck(k-1)}{2n} + \frac{ck}{n}\right) \\
&\le \frac{e^k k^{k+1} c^k}{k^k} \exp\left(-ck - \frac{(1-c)k(k-1)}{2n}\right) \\
&\le k\left(ce^{1-c}\right)^k.
\end{aligned}$$

Then by $ce^{1-c} < 1$ for any $c \ne 1$

$$\mathbf{E}\left[\sum_{k=3}^{n} X_k\right] \le \sum_{k=3}^{n} k\left(ce^{1-c}\right)^k = O(1).$$

Thus we know for every $\omega = \omega(n) \to \infty$,

$$\mathbf{Pr}\left[\sum_{k=3}^{n} X_k \ge \omega\right] \le O(\omega^{-1}) \to 0$$

as $n \to \infty$.

For $c > 1$, note that the inequality holds as well when $k = o(n)$ since $e^{k^2/n} = e^{o(k)}$. Then we will prove in the later part that when $c > 1$, w.h.p. there is a unique giant component of size $\Omega(n)$ and all other components are of size $O(\log n)$, and this giant component is not unicyclic. Then we complete the proof for $c > 1$. □

Now we investigate the tree components.

Lemma 3.9. *Let $p = c/n$ where $c \ne 1$ is a constant, $\alpha = c - 1 - \log c > 0$ and $\omega = \omega(n) \to \infty$, $\omega(n) = o(\log\log n)$. Then*

(1) w.h.p. there exists an isolated tree of order

$$k_- := \frac{1}{\alpha}\left(\log n - \frac{5}{2}\log\log n\right) - \omega;$$

(2) w.h.p. there is no isolated tree of order

$$k_+ := \frac{1}{\alpha}\left(\log n - \frac{5}{2}\log\log n\right) + \omega.$$

Proof. Let X_k be the number of isolated trees of order k. Then

$$\mathbf{E}\left[X_k\right] = \binom{n}{k} k^{k-2} p^{k-1} (1-p)^{k(n-k)+\binom{k}{2}-(k-1)}.$$

Suppose that $k = O(\log n)$. Then we know

$$\begin{aligned}
\mathbf{E}\left[X_k\right] &= (1+o(1)) \frac{n^k}{k!} k^{k-2} \frac{c^{k-1}}{n^{k-1}} e^{-ck} \\
&= \frac{1+o(1)}{c\sqrt{2\pi}} n k^{-5/2} \left(ce^{1-c}\right)^k \\
&= \frac{1+o(1)}{c\sqrt{2\pi}} n k^{-5/2} e^{-\alpha k}.
\end{aligned}$$

Let $k = k_-$, we know

$$\mathbf{E}[X_k] = \frac{1+o(1)}{c\sqrt{2\pi}}(k^{-1}\log n)^{5/2}e^{\alpha\omega} \geq Ce^{\alpha\omega}$$

for some constant $C > 0$.

To estimate the second moment of X_k, it is not hard to obtain

$$\mathbf{E}\left[X_k^2\right] \leq \mathbf{E}[X_k]\left(1 + (1-p)^{-k^2}\mathbf{E}[X_k]\right).$$

Then we know

$$\begin{aligned}\mathbf{Var}(X_k) &\leq \mathbf{E}[X_k] + \mathbf{E}[X_k]^2\left((1-p)^{-k^2} - 1\right) \\ &\leq \mathbf{E}[X_k] + 2ck^2\left(\mathbf{E}[X_k]\right)^2/n.\end{aligned}$$

Thus by Chebyshev inequality, for every $\varepsilon > 0$,

$$\mathbf{Pr}\left[|X_k - \mathbf{E}[X_k]| \geq \varepsilon\mathbf{E}[X_k]\right] \leq \frac{1}{\varepsilon^2\mathbf{E}[X_k]} + \frac{2ck^2}{\varepsilon^2 n} = o(1).$$

Thus w.h.p. $X_k \geq Ce^{\alpha\omega}$, leading to the proof of 1.

For the second part, note that for some constant $C_1 > 0$ such that

$$\begin{aligned}\mathbf{E}[X_k] &\leq C_1 k^{-1/2}\left(\frac{ne}{k}\right)^k k^{k-2}\left(1 - \frac{k}{2n}\right)^{k-1}\left(\frac{c}{n}\right)^{k-1} e^{-ck+\frac{ck^2}{2n}} \\ &\leq \frac{2An}{c_k k^{5/2}}\left(c_k e^{1-c_k}\right)^k\end{aligned}$$

where $c_k := c\left(1 - \frac{k}{2n}\right)$.

When $c < 1$ and $k \geq k_+$, we know $c_k e^{1-c_k} \leq ce^{1-c}$ and $c_k \geq \frac{1}{2}c$. Then we know

$$\begin{aligned}\sum_{k=k_+}^{n}\mathbf{E}[X_k] &\leq \frac{4C_1 n}{c}\sum_{k=k_+}^{n}\frac{(ce^{1-c})^k}{k^{5/2}} \\ &\leq \frac{4C_1 n}{ck_+^{5/2}}\sum_{k=k_+}^{\infty}e^{-\alpha k} \\ &= \frac{4C_1 ne^{-\alpha k_+}}{ck_+^{5/2}(1-e^{-\alpha})} \\ &= \frac{(4C_1 + o(1))e^{-\alpha\omega}\alpha^{5/2}}{c(1-e^{-\alpha})} = o(1).\end{aligned}$$

When $c > 1$, we use the following two inequalities: when $k \leq \frac{n}{\log n}$, $c_k e^{1-c_k} = e^{-\alpha - O(1/\log n)}$ while $c_k \geq c/2$ and $c_k e^{1-c_k} \leq 1$ for $k > \frac{n}{\log n}$. Then we know

$$\sum_{k=k_+}^{n}\mathbf{E}[X_k] \leq \frac{4C_1 n}{ck_+^{5/2}}\sum_{k=k_+}^{n/\log n}e^{-(\alpha + O(1/\log n))k} + \frac{20C_1 n}{c}\sum_{k=n/\log n}^{n}\frac{1}{k^{5/2}} = o(1).$$

□

Now we prove the useful identity. For any $c > 0$, suppose $x = x(c) \in (0,1]$ such that

$$x(c) = \begin{cases} c & c \leq 1 \\ \text{the solution of } xe^{-x} = ce^{-c} & c > 1. \end{cases}$$

Lemma 3.10. *If $c > 0$ is a constant, and $x = x(c)$ is given above. Then*

$$\frac{1}{x}\sum_{k=1}^{\infty}\frac{k^{k-1}}{k!}(ce^{-c})^k = 1.$$

Proof. We prove the case $c < 1$. The case $c > 1$ can be solved by the observation $ce^{-c} = xe^{-x}$. To prove $x = 1$, note that the function $f(c) = \frac{1}{x}\sum_{k=1}^{\infty}\frac{k^{k-1}}{k!}(ce^{-c})^k$ is continuous.

Let $p = c/n$. Let X be the number of vertices in $\mathcal{G}_{n,p}$ that lie in non-tree components. Let X_k be the number of isolated trees of order k. Note that

$$n = \sum_{k=1}^{n} kX_k + X,$$

meaning that

$$n = \sum_{k=1}^{n} k\mathbf{E}\left[X_k\right] + \mathbf{E}\left[X\right].$$

Then we know $\mathbf{E}\left[X\right] = O(1)$, and if $k < k_+$,

$$\mathbf{E}\left[X_k\right] = (1+o(1))\frac{n}{ck!}k^{k-2}(ce^{-c})^k.$$

Then we know that

$$\begin{aligned} n &= o(n) + \frac{n}{c}\sum_{k=1}^{k_+}\frac{k^{k-1}}{k!}(ce^{-c})^k \\ &= o(n) + \frac{n}{c}\sum_{k=1}^{\infty}\frac{k^{k-1}}{k!}(ce^{-c})^k. \end{aligned}$$

Then we prove the identity when $c < 1$. □

3.2. **Super-critical phase.** The structure of a random graph $\mathcal{G}_{n,m}$ changes dramatically when $m = \frac{1}{2}cn$ where $c > 1$ is a constant.

Theorem 3.11. *Suppose that $m = \frac{1}{2}cn$, $c > 1$. Then w.h.p. G_m consists of a unique giant components with $\left(1 - \frac{x}{c} + o(1)\right)n$ vertices and $\left(1 - \frac{x^2}{c^2} + o(1)\right)\frac{cn}{2}$ edges, where x is the unique solution in $(0, 1)$ with $xe^{-x} = ce^{-c}$. The remaining components are of order at most $O(\log n)$.*

Proof. Suppose that Z_k is the number of components of order k in $\mathcal{G}_{n,p}$. To bound $\mathbf{E}\left[Z_k\right]$, we could bound the number of trees of order k.

$$\begin{aligned} \mathbf{E}\left[Z_k\right] &\le \binom{n}{k}k^{k-2}p^{k-1}(1-p)^{k(n-k)} \\ &\le \frac{n^k}{k!}k^{k-2}\left(\frac{c}{n}\right)^{k-1}\left(1-\frac{c}{n}\right)^{k(n-k)} \\ &\le C\frac{n^k e^k}{k^k}k^{k-2}\frac{c^{k-1}}{n^{k-1}}e^{-ck+ck^2/n} \\ &\le \frac{Cn}{ck^2}\left(ce^{1-c+ck/n}\right)^k \end{aligned}$$

Define the two quantities $\beta_0 = \beta_0(c), \beta_1 = \beta_1(c)$ as

$$ce^{1-c+c\beta_1} < 1, \quad \left(ce^{1-c+o(1)}\right)^{\beta_0 \log n} < \frac{1}{n^2}.$$

Then we know that w.h.p. there is no component of order $k \in [\beta_0 \log n, \beta_1 n]$.

Now we consider the number of vertices on small components of order $[1, \beta_0 \log n]$. Recall the setting of $\alpha = c - 1 - \log c$. Firstly assume that $1 \le k \le k_0$, where $k_0 = \frac{1}{2\alpha} \log n$. Then

$$\begin{aligned}\mathbf{E}\left[\sum_{k=1}^{k_0} kX_k\right] &= (1+o(1))\frac{n}{c}\sum_{k=1}^{k_0}\frac{k^{k-1}}{k!}(ce^{-c})^k \\ &= (1+o(1))\frac{n}{c}\sum_{k=1}^{\infty}\frac{k^{k-1}}{k!}(ce^{-c})^k\end{aligned}$$

where we use the fact that $\frac{k^{k-1}}{k!} < e^k$ and $ce^{-c} < e^{-1}$ for $c \neq 1$ to extend the summation from k_0 to ∞.

Note that, the probability $\sum_{k=1}^{k_0} kX_k$ deviates from its mean by more than $1 \pm O(\log n^{-1})$ is at most $o(1)$. Then we know that

$$\sum_{k=1}^{k_0} kX_k \approx (1+o(1))\frac{nx}{c}$$

where $x \in (0,1)$ is the unique solution such that $xe^{-x} = ce^{-c}$.

Now consider $k_0 < k \le \beta_0 \log n$.

$$\begin{aligned}\mathbf{E}\left[\sum_{k=k_0+1}^{\beta_0 \log n} kX_k\right] &\le \frac{n}{c}\sum_{k=k_0+1}^{\beta_0 \log n}\left(ce^{1-c+ck/n}\right)^k \\ &= O\left(n(ce^{1-c})^{k_0}\right) \\ &= O\left(n^{1/2+o(1)}\right).\end{aligned}$$

So by the Markov inequality, w.h.p.,

$$\sum_{k=k_0+1}^{\beta_0 \log n} kX_k = o(n).$$

Now, we consider the non-tree components. Let Y_k be the number of non-tree components of k vertices where $1 \le k \le \beta_0 \log n$.

$$\begin{aligned}\mathbf{E}\left[\sum_{k=1}^{\beta_0 \log n} kY_k\right] &\le \sum_{k=1}^{\beta_0 \log n}\binom{n}{k}k^{k-1}\binom{k}{2}\frac{c^k}{n^k}\left(1-\frac{c}{n}\right)^{k(n-k)} \\ &= \sum_{k=1}^{\beta_0 \log n} k\left(ce^{1-c+ck/n}\right)^k \\ &= O(1).\end{aligned}$$

Then by the Markov inequality, w.h.p.,

$$\sum_{k=1}^{\beta_0 \log n} kY_k = o(n).$$

Combining the above, we know that w.h.p. there are $\frac{nx}{c}$ vertices on components of order not more than $\beta_0 \log n$.

Now we show the uniqueness of the giant component. We prove the argument by coupling random graphs. let

$$c_1 = c - \frac{\log n}{n}, p_1 = \frac{c_1}{n}, p_2 = 1 - \frac{1-p}{1-p_1}.$$

Then by the coupling argument, it holds that

$$\mathcal{G}_{n,p} = \mathcal{G}_{n,p_1} \cup \mathcal{G}_{n,p_2}.$$

Additionally, we observe that

$$p_2 = 1 - \frac{1-p}{1-p_1} = \frac{p-p_1}{1-p_1} = \frac{\log n}{n^2 - \log n} \geq \frac{\log n}{n^2}.$$

Let $x_1 \in (0,1)$ such that $x_1 e^{-x_1} = c_1 e^{-c_1}$. Since $c_1 = c - o(1)$, we know that $c_1 \approx c$. Then similar argument indicates that, w.h.p. there are no components of order in $[\beta_0 \log n, \beta_1 n]$ in $\mathcal{G}_{n,p_1}$.

Suppose that $C_1, \ldots, C_\ell$ are maximal giant components in $\mathcal{G}_{n,p_1}$, *i.e.*, $|C_i| > \beta_1 n$ for every $i \in [\ell]$ and $\ell \leq 1/\beta_1$. By coupling argument, we view $\mathcal{G}_{n,p}$ as a graph induced based on $\mathcal{G}_{n,p_1}$ by adding edges according to $\mathcal{G}_{n,p_2}$ independently of $\mathcal{G}_{n,p_1}$. Then,

$$\begin{aligned}
\mathbf{Pr}\left[\exists i, j \in [\ell], C_i, C_j \text{ are not linked}\right] &\leq \binom{\ell}{2}(1-p_2)^{(\beta_1 n)^2} \\
&\leq \ell^2 e^{-\beta_1^2 \log n} \\
&= o(1).
\end{aligned}$$

Thus we conclude that $\mathcal{G}_{n,p}$ has an unique giant components of size $\left(1 - \frac{x}{c} + o(1)\right) n$.

The last ingredient of the theorem is to estimate the edges in the giant component. Now we consider the evolution of random graphs

$$G_0 = \varnothing, G_1, \ldots, G_m = \mathcal{G}_{n,m}.$$

Denote by C_0 the giant component in $\mathcal{G}_{n,m}$. Note that G_{m-1} is distributed as $\mathcal{G}_{n,m-1}$. Then we know G_{m-1} has a unique giant components C_{-1} and other components are of size $O(\log n)$. Then,

$$\mathbf{Pr}\left[e \notin C_0 \mid |C_{-1}| \approx (1-x/c)n\right] = \mathbf{Pr}\left[e \cap C_{-1} = \varnothing \mid (1-x/c)n\right] \approx \left(\frac{x}{c}\right)^2.$$

This means that the expected number of edges in the giant component is $\left(1 - \frac{x}{c} + o(1)\right) n$. What remains to do is to show the concentration. Fix $i \neq j \leq m$. Let C_{-2} denote the unique giant component of $G_m - \{e_i, e_j\}$. Then we know that

$$\mathbf{Pr}\left[e_i, e_j \subseteq C_0\right] = (1+o(1))\mathbf{Pr}\left[e_i \in C_0\right]\mathbf{Pr}\left[e_j \in C_0\right].$$

Then by the Chebyshev inequality, we can easily show the concentration of the number of edges in the giant component. □

For more dense random graphs, the structure of them is much more complex. However, there are some simple structures as well.

Theorem 3.12. *Let $\omega = \omega(n) \to \infty$ as $n \to \infty$ be some slowly growing function. If $m \geq \omega n$ but $m \leq n(\log n - \omega)/2$, then G_m is disconnected and all components are trees w.h.p. except the giant one.*

3.2.1. *Cores.* Given a positive integer $k > 1$, the k-core of a graph $G = (V, E)$ is the largest subset $S \subseteq V$ such that the minimum degree $\delta_S := \min_{v \in S} \deg_{G[S]}(v)$ in the induced subgraph $G[S]$ is at least k.

Theorem 3.13. *Suppose that $c > 1$ and $x \in (0,1)$ is the unique solution of the equation $xe^{-x} = ce^{-c}$. Then w.h.p. the 2-core C_2 of the random graph $\mathcal{G}_{n,p}$ with parameter $p = c/n$ has $(1-x)\left(1 - \frac{c}{x} + o(1)\right) n$ vertices and $\left(1 - \frac{x}{c} + o(1)\right)^2 \frac{cn}{2}$ edges.*

The proof of this theorem follows the routine and we omit it here.

4. Degrees in Random Graphs

In this section, we show the distribution of degrees of vertices in random graphs. We consider two different cases.

4.1. **Degrees of sparse random graphs.** Firstly we consider the sparse case. Let X_0 be the number of isolated vertices. Then

$$\mathbf{E}[X_0] = n(1-p)^{n-1}.$$

As $n \to \infty$,

$$\mathbf{E}[X_0] \to \begin{cases} \infty & np - \log n \to -\infty \\ e^{-c} & np - \log n \to c < \infty \\ 0 & np - \log n \to \infty \end{cases}.$$

We have the following theorem for the distribution of degrees.

Theorem 4.1. *Let X_0 be the random variable denoting the number of isolated vertices in $\mathcal{G}_{n,p}$. Then as $n \to \infty$,*

1. $\widehat{X_0} = (X_0 - \mathbf{E}[X_0])/(\mathbf{Var}(X_0))^{1/2} \xrightarrow{D} \mathcal{N}(0,1)$ *if $n^2 p \to \infty$ and $np - \log n \to -\infty$.*
2. $X_0 \xrightarrow{D} \mathrm{Poisson}(e^{-c})$ *if $np - \log n \to c$, $c < \infty$.*
3. $X_0 \xrightarrow{D} 0$ *if $np - \log n \to \infty$.*

If we work further on the case of the number of vertices with a specific degree $d \in \mathbb{N}$, the following theorem will come.

Theorem 4.2. *Let $X_d = X_{n,d}$ be the random variable denoting the number of vertices in $\mathcal{G}_{n,p}$ with degree $d > 0$. Then as $n \to \infty$,*

1. $X_d \xrightarrow{D} 0$ *if $p \ll n^{-(d+1)/d}$.*
2. $X_d \xrightarrow{D} \mathrm{Poisson}(c^d/d!)$ *if $p \approx cn^{-(d+1)/d}$, $c < \infty$.*
3. $\widehat{X_d} := (X_d - \mathbf{E}[X_d])/\mathbf{Var}(X_d)^{1/2} \xrightarrow{D} \mathcal{N}(0,1)$ *if $p \gg n^{-(d+1)/d}$ but $pn - \log n - d\log\log n \to -\infty$.*
4. $X_d \to \mathrm{Poisson}(e^{-c}/d!)$ *if $pn - \log n - d\log\log n \to c$, $c \in \mathbb{R}$.*
5. $X_d \to 0$ *if $pn - \log n - d\log\log n \to \infty$.*

The next theorem shows the concentration of X_d around its expectation when $p = c/n$.

Theorem 4.3. *Let $p = c/n$ where c is a constant. Let X_d denote the number of vertices of degree d in $\mathcal{G}_{n,p}$. Then for $d = O(1)$, w.h.p.*

$$X_d \approx \frac{c^d e^{-c}}{d!} n.$$

We end this part with the behavior of the maximum degree in a sparse random graph.

Theorem 4.4. *Let $\Delta(\mathcal{G}_{n,p})$ ($\delta(\mathcal{G}_{n,p})$) be the maximum (minimum) degree of vertices of $\mathcal{G}_{n,p}$.*

1. *If $p = c/n$ for some constant $c > 0$, then w.h.p.,*
$$\Delta(\mathcal{G}_{n,p}) \approx \frac{\log n}{\log\log n}.$$
2. *If $p = \omega \log n$ where $\omega \to \infty$, then w.h.p.,*
$$\delta(\mathcal{G}_{n,p}) \approx \Delta(\mathcal{G}_{n,p}) \approx np.$$

4.2. **Degrees of dense random graphs.** When p is a constant, the degrees are well concentrated.

Theorem 4.5. *Let*

$$d_{\pm} = (n-1)p + (1 \pm \varepsilon)\sqrt{2(n-1)p(1-p)\log n}.$$

If p is a constant and $\varepsilon > 0$ is a small constant, then w.h.p.

1. $d_- \le \Delta(\mathcal{G}_{n,p}) \le d_+$.
2. *There is a unique vertex of maximum degree.*

More precisely, we have the following strengthened version of Theorem 4.5.

Theorem 4.6. *Let $\varepsilon = 1/10$ and p be a constant. Let*

$$d_{\pm} = (n-1)p + (1 \pm \varepsilon)\sqrt{2(n-1)p(1-p)\log n}.$$

Then w.h.p.,

(1) $\Delta(\mathcal{G}_{n,p}) \le d_+$.
(2) There are $\Omega(n^{2\varepsilon(1-\varepsilon)})$ *vertices of degree at least* d_-.
(3) There exists no $u \ne v$ *such that* $\deg(u), \deg(v) \ge d_-$ *and* $|\deg(u) - \deg(v)| \le 10$.

Application to graph isomorphism. Now we introduce an application of results on degrees in random graphs to graph isomorphism. Given two graph $G = (V_G, E_G)$ and $H = (V_H, E_H)$, our goal is to determine whether $G \cong H$. The first step is to give a labeling algorithm.

Algorithm 1: Graph Labeling Label

Input: a graph $G = (V, E)$ and a parameter $L > 0$.
Output: a boolean signal **flag** $\in \{\mathsf{True}, \mathsf{False}\}$ to determine whether the algorithm succeeds and a labeling $\{v_1, \dots, v_n\}$ of V.

Relabel V as $\{v_i\}_{i=1}^n$ so they satisfy

$$\deg(v_1) \ge \deg(v_2) \ge \cdots \ge \deg(v_n).$$

if $\exists i < L, \deg(v_i) = \deg(v_{i+1})$ **then**
  flag $\leftarrow$ False;
  return flag and $\{v_i\}_{i=1}^n$;
foreach $i > L$ **do**
  $X_i \leftarrow \{j \in [L] : \{v_i, v_j\} \in E\}$;
relabel $\{v_{L+1}, \dots, v_n\}$ so that

$$X_{L+1} \succ X_{L+2} \succ \cdots \succ X_n$$

where $\succ$ denotes the lexicographic order;
if $\exists i < n, X_i = X_{i+1}$ **then**
  flag $\leftarrow$ False;
  return flag and $\{v_i\}_{i=1}^n$;
flag $\leftarrow$ True;
return flag and $\{v_i\}_{i=1}^n$;

Now we are able to give Algorithm 2 for graph isomorphism.

Algorithm 2: Graph Isomorphism Isomorphism

Input: two graphs $G = (V_G, E_G)$ and $H = (V_H, E_H)$ and a parameter $L > 0$.
Output: a boolean signal to determine whether $G \cong H$.

$\{\mathbf{flag}_G, \{v_i\}_{i=1}^n\} \leftarrow$ Label(G, L);
$\{\mathbf{flag}_H, \{w_i\}_{i=1}^n\} \leftarrow$ Label(H, L);
if $\mathbf{flag}_G =$ False $\vee$ $\mathbf{flag}_H =$ False **then**
  return False;
if $v_i \mapsto w_i$ *is an isomorphism* **then**
  return True;
else
  return False;

Usually the parameter L is not easy to choose. However, when the two underlying graphs G, H are dense random graphs, the algorithm works well w.h.p..

Theorem 4.7. *Let p be a fixed constant and $\rho = p^2 + q^2$. Then Algorithm 2 works well on $\mathcal{G}_{n,p}$ w.h.p..*

Proof. The first step is to show that Algorithm 1 succeeds w.h.p. on $\mathcal{G}_{n,p}$. It suffices to show that when $i \ne j > L$, w.h.p. $X_i \ne X_j$.

Fix $i \neq j$, $i, j > L$. Let $\widetilde{G} := \mathcal{G}_{n,p} \setminus \{v_i, v_j\}$. By Theorem 4.6, w.h.p. the L largest degree vertices of $\widetilde{G}$ and $\mathcal{G}_{n,p}$ coincide. So w.h.p. we compute X_i and X_j from $\widetilde{G}$ w.h.p. which are independent of the edge incident with v_i and v_j. Then we know

$$\mathbf{Pr}\left[\mathbf{flag} = \mathsf{False}\right] \leq o(1) + \binom{n}{2}\left(p^2 + (1-p)^2\right)^L = o(1).$$

Finally, by Algorithm 1, w.h.p. $\mathcal{G}_{n,p}$ only has exactly one automorphism, meaning that Algorithm 2 works well w.h.p.. □

Application to edge colorings. For the typical edge-coloring problem, the famous Vizing's lemma has told us that for every graph $G = (V, E)$,

$$\Delta(G) \leq \chi_E(G) \leq \Delta(G) + 1$$

where $\chi_E(G)$ denotes the minimum number of colors which can form a proper edge coloring of G. It is often a hard problem to determine the exact number of $\chi_E(G)$. However, when the graph has a unique vertex of maximum degree, it can be shown that $\chi_E(G) = \Delta(G)$. By Theorem 4.5, w.h.p. $\chi_E(\mathcal{G}_{n,p}) = \Delta(\mathcal{G}_{n,p})$.

5. Connectivity of Random Graphs

In this section, we look at the behavior of the connectivity in random graphs. It is interesting to see when the random graph is connected.

5.1. Connectivity.

Firstly we state the following result for the connectivity of $\mathcal{G}_{n,m}$.

Theorem 5.1. *Let $m = \frac{1}{2}n(\log n + c_n)$. Then*

$$\lim_{n\to\infty} \mathbf{Pr}\left[\mathcal{G}_{n,m} \text{ is connected}\right] = \begin{cases} 0 & c_n \to -\infty, \\ e^{-e^{-c}} & c_n \to c, \\ 1 & c_n \to \infty. \end{cases}$$

Comparing Theorem 5.1 with Theorem 4.1, we can see the following interesting phenomenon.

Proposition 5.2. *Consider the evolution process $\{G_m\}$. Let*

$$m_1^* := \min\left\{m : \ \delta(G_m) \geq 1\right\}$$

and

$$m_c := \min\left\{m : \ G_m \text{ is connected}\right\}.$$

Then w.h.p. $m_1^* = m_c$.

This proposition is quite impressive. On one hand, it tells us the picture through the whole evolution process. On the other hand, it emphasizes the importance of vertex degrees in random graphs.

5.2. k-connectivity.

Now we look at k-connectivity of a random graph. A graph is said to be k-connected if it remains connected after removing arbitrary k vertices. Inspired by Proposition 5.2, the following result holds.

Theorem 5.3. *Let $m = \frac{1}{2}n(\log n + (k-1)\log n + c_n)$. Then*

$$\lim_{n\to\infty} \mathbf{Pr}\left[\mathcal{G}_{n,m} \text{ is } k\text{-connected}\right] = \begin{cases} 0 & c_n \to -\infty, \\ e^{-\frac{e^{-c}}{(k-1)!}} & c_n \to c, \\ 1 & c_n \to \infty. \end{cases}$$

6. Small Subgraphs in Random Graphs

In this section, we put our sight on the existence of small subgraphs in random graphs. That is to say, given a graph $H = (V_H, E_H)$, we investigate the occurrence of H in $\mathcal{G}_{n,m}$ or $\mathcal{G}_{n,p}$, as the growth of m/n and p. Before all, we define some quantities and useful facts here. Let $v_H := |V_H|$ be the number of vertices in H and let $e_H := |E_H|$ be the number of edges in H. Then, define the density of H as

$$d(H) = \frac{e_H}{v_H}.$$

Note that $2d(H) = \alpha(H)$ where $\alpha(H)$ is the average degree of H. Moreover, define the quantity $m(H)$ as

$$m(H) := \sup_{K \subseteq H} d(K).$$

Intuitively, $m(H)$ measures how 'dense' a part of H could be. If $d(H) = m(H)$, we say that the graph H is *balanced*. Moreover, if for every $K \subsetneqq H$, it holds that $d(K) < d(H)$, then we say that the graph H is *strictly balanced*.

Recall that $\mathrm{aut}(H)$ means the number of automorphisms of H. The following fact computes the copies of H in the complete graph K_n.

Fact 6.1. *The number of copies of H in a complete graph K_n is*

$$\binom{n}{v_H} \frac{v_H!}{\mathrm{aut}(H)}.$$

6.1. Threshold of occurrence.

At the beginning of this section, we state the threshold result of the occurrence of H in random graphs.

Theorem 6.2. *Fix a graph H with $e_H > 0$. Then*

$$\lim_{n\to\infty} \mathbf{Pr}\left[H \subseteq \mathcal{G}_{n,p}\right] = \begin{cases} 0 & pn^{1/m(H)} \to 0, \\ 1 & pn^{1/m(H)} \to \infty. \end{cases}$$

Proof. The first result is not hard to verify. To show the second result, we apply the second-moment method. Set $p = \omega n^{-1/m(H)}$. Denote by $H_1, H_2, \ldots, H_t$ all copies of H in complete graph K_n. Then by Fact 6.1,

$$t = \binom{n}{v_H} \frac{v_H!}{\mathrm{aut}(H)}.$$

For every $i = 1, \ldots, t$, define the indicator I_i as

$$I_i := \mathbb{1}\left[H \subseteq \mathcal{G}_{n,p}\right].$$

Let X_H be the number of H occurring in $\mathcal{G}_{n,p}$, *i.e.*, $X_H = \sum_{i=1}^t I_i$. Then by direct calculation,

$$\begin{aligned}
\mathbf{Var}\left(X_H\right) &= \sum_{i=1}^t \sum_{j=1}^t \mathbf{Cov}\left(I_i, I_j\right) \\
&= \sum_{i=1}^t \sum_{j=1}^t \left(\mathbf{Pr}\left[I_i = I_j = 1\right] - \mathbf{Pr}\left[I_i = 1\right] \mathbf{Pr}\left[I_j = 1\right]\right) \\
&= \sum_{i=1}^t \sum_{j=1}^t \left(\mathbf{Pr}\left[I_i = I_j = 1\right] - p^{2e_H}\right).
\end{aligned}$$

Note that, when H_i and H_j are disjoint, $\mathbf{Cov}(I_i, I_j) = 0$. When H_i and H_j intersect, we consider the intersection of them.

$$\begin{aligned}\mathbf{Var}(X_H) &= O\left(\sum_{K\subseteq H, e_K>0} n^{2v_H - v_K}\left(p^{2e_H-e_K} - p^{2e_H}\right)\right)\\ &= O\left(n^{2v_H}p^{2e_H}\sum_{K\subseteq H, e_K>0} n^{-v_K}p^{-e_K}\right).\end{aligned}$$

On the other hand,

$$\mathbf{E}[X_H] = \binom{n}{v_H}\frac{v_H!}{\mathrm{aut}(H)}p^{e_H} = \Omega\left(n^{v_H}p^{e_H}\right).$$

Then by the second-moment method,

$$\begin{aligned}\mathbf{Pr}[X_H = 0] &\le \frac{\mathbf{Var}(X_H)}{(\mathbf{E}[X_H])^2}\\ &= O\left(\sum_{K\subseteq H, e_K>0}\left(\frac{1}{\omega n^{1/d(K)-1/m(H)}}\right)^{e_K}\right)\\ &= o(1).\end{aligned}$$

Hence w.h.p. the random graph $\mathcal{G}_{n,p}$ contains a copy of H when $pn^{1/m(H)} \to \infty$. □

6.2. **Asymptotic distributions.** When p is relatively large, the asymptotic distribution of the occurrence of H is interesting. However, for general cases, the distribution is seemingly impossible to compute, when p is at the threshold. The situation, however, changes when we only consider the strictly balanced subgraphs.

Theorem 6.3. *Let H be a strictly balanced subgraph and $np^{m(H)} \to c, c > 0$. Define a random variable X_H to denote the copies of H in the random graph $\mathcal{G}_{n,p}$. Then as $n \to \infty$,*

$$X_H \sim \mathrm{Poisson}(\lambda), \quad \text{where } \lambda = \frac{c^{v_H}}{\mathrm{aut}(H)}.$$

Proof. Similarly, let $H_1, \ldots, H_t$ denote all copies of H in the complete graph K_n, and for every $i \in [t]$, define the indicator I_i as

$$I_i := \mathbb{1}[H_i \subseteq \mathcal{G}_{n,p}].$$

Then $X_H = \sum_{i=1}^t I_i$. To investigate the distribution of X_H, by the Poisson approximation, we compute:

$$\mathbf{E}[(X_H)_k] := \mathbf{E}[X_H(X_H-1)\cdots(X_H-k+1)].$$

By definition,

$$\begin{aligned}\mathbf{E}[(X_H)_k] &= \sum_{i_1, i_2, \ldots, i_k}\mathbf{Pr}[I_{i_1} = I_{i_2} = \cdots = I_{i_k} = 1]\\ &= S_k + \overline{S_k},\end{aligned}$$

where the summation takes over all distinct k indices of $\{1, \ldots, k\}$, and S_k and $\overline{S_k}$ denote the partial sums taken over all elements which are respectively pairwise disjoint or not. For S_k, we have that

$$\begin{aligned}S_k &= \binom{n}{v_H, v_H, \ldots, v_H}\left(\frac{v_H!}{\mathrm{aut}(H)}p^{e_H}\right)^k\\ &\approx \left(\frac{n^{v_H}p^{e_H}}{\mathrm{aut}(H)}\right)^k\\ &= \left(\frac{c^{v_H}}{\mathrm{aut}(H)}\right)^k\end{aligned}$$

by the assumption that H is strictly balanced.

On the other hand, we will show that $\overline{S_k} \to 0$. Consider the following family of graphs

$$\mathcal{F}_k := \{F = H_{i_1} \cup \cdots \cup H_{i_k} \,:\, H_{i_1}, \ldots, H_{i_k} \text{ are not pairwise disjoint}\}.$$

For the sake of simplicity, we assume that the graphs in $\mathcal{F}_k$ are mutually non-isomorphic. Firstly we prove the following claim.

Claim. *For every $F \in \mathcal{F}_k$, it holds that $d(F) > m(H)$.*

Assuming the claim, for every $F \in \mathcal{F}_k$, let C_F be the number of sequences such that $H_{i_1}, \ldots, H_{i_k}$ are pairwise disjoint and

$$\bigcup_{j=1}^{k} H_{i_j} \cong F.$$

Then we conclude that

$$\overline{S_k} = \sum_{F \in \mathcal{F}_k} \binom{n}{v_F} C_F p^{e_F} = O(n^{v_F} p^{e_F}) = o(1).$$

Thus we conclude that for every $k \in \mathbb{N}$,

$$\mathbf{E}\left[(X_H)_k\right] \to \left(\frac{c^{v_H}}{\mathrm{aut}(H)}\right)^k$$

and by Theorem 1.6, we obtain the asymptotic distribution.

What remains to do is to prove the claim. Define the function f over the domain of all graphs as

$$f(F) := m(H)v_F - e_F.$$

Then it is equivalent to show that $f(F) < 0$ for every $F \in \mathcal{F}_k$. Observe that for each pair of graphs (F_1, F_2),

$$f(F_1 \cup F_2) = f(F_1) + f(F_2) - f(F_1 \cap F_2).$$

We prove the claim by hypothesis induction on $k \geq 2$. For $k = 2$, by the assumption that H is strictly balanced, it holds that for every $K \subsetneq H$, $f(K) > 0$. Then for every $F = H_{i_1} \cup H_{i_2}$, it holds that

$$f(F) = -f(H_{i_1} \cap H_{i_2}) < 0.$$

Thus we prove the case $k = 2$. For arbitrary $k \geq 3$, let $F' = \bigcup_{j=1}^{k-1} H_{i_j}$ and $K = F' \cap H_{i_k}$. Then it holds that $f(F') < 0$ by induction and $f(K) > 0$ by the fact that $K \subsetneq H$. Thus we know

$$f(F) = f(F') + f(H_{i_k}) - f(K) = f(F') - f(K) < 0.$$

Then we conclude the claim. □

When p is beyond the threshold, regardless of whether or not H is strictly balanced, the following result holds.

Theorem 6.4. *Let H be a non-empty graph. If $np^{m(H)} \to \infty$ and $n^2(1-p) \to \infty$, then as $n \to \infty$,*

$$\frac{X_H - \mathbf{E}\left[X_H\right]}{\left(\mathbf{Var}\left(X_H\right)\right)^{1/2}} \xrightarrow{D} \mathcal{N}(0,1).$$

References

[FK23] Alan Frieze and Michał Karoński. *Introduction to Random Graphs.* the edited version of the book published in 2015, 2023. 1

[Łuc90] Tomasz Łuczak. On the Equivalence of Two Basic Models of Random Graphs. In Michał Karoński, Jerzy Jaworski, and Andrzej Rucinski, editors, *Proceedings of Random Graphs'87*, pages 151–158, Chichester, 1990. Wiley. 6