EXISTENCE OF SMALL SUBGRAPH IN THE RANDOM GRAPH $\mathcal{G}(n,p)$

TONY DENG

Abstract. This expository paper discusses the existence of small subgraphs in the random graph $\mathcal{G}(n,p)$. Given a balanced graph H we present the proof of a result from [10] that determines the threshold for $\mathcal{G}(n,p)$ to contain H as a subgraph. Furthermore, if H is strictly balanced, we include proofs of results from [2] and [8] that determine the asymptotic probability that $\mathcal{G}(n,p)$ contains H at the threshold.

1. Introduction

A binomial random graph $\mathcal{G}(n,p)$ is a graph on the vertex set $[n] = \{1,2,...,n\}$, determined randomly, where each edge is present independently with probability $p = p(n)$. Formally, a random graph $\mathcal{G}(n,p)$ is a finite probability space over the graphs on $[n]$, where for any graph G with $|E(G)| = e_G$, $\mathbb{P}(\mathcal{G}(n,p) = G) = p^{e_G}(1-p)^{\binom{n}{2}-e_G}$. The subject of random graphs was introduced by Erdős and Rényi in the paper *On the evolution of Random Graphs* [5] and the model $\mathcal{G}(n,p)$ was introduced by Gilbert [7]. Random graphs were used to prove many graph-theoretic results, such as the existence of graphs with arbitrarily high girth and high chromatic numbers (see [4]), and the subject is of great research interest in its own right.
Let H be a graph with v_H vertices and e_H edges, and let $p = p(n) \in [0,1]$ be a function of n. This paper investigates the asymptotic probability that $\mathcal{G}(n,p)$ does not contain H as a subgraph. The next definition imposes another condition on H.

Definition 1.1. *We define the density of H as $\rho(H) = e_H/v_H$. We say that H is balanced if for all subgraphs $K \subseteq H$, we have $\rho(K) \leq \rho(H)$. We say that H is strictly balanced if for all proper subgraphs $K \subsetneq H$, we have $\rho(K) < \rho(H)$.*

This paper's main objective is to present the next two Theorems and their proofs. The following theorem was proven by Erdös and Rényi in [5]. The proof we give in this paper is due to Ruciński and Vince [12].

Theorem 1.2. *Let H be a balanced graph. Then*

$$\lim_{n\to\infty} \mathbb{P}(H \subseteq \mathcal{G}(n,p)) = \begin{cases} 0 & \text{if } pn^{1/\rho(H)} \to 0 \\ 1 & \text{if } pn^{1/\rho(H)} \to \infty \end{cases}$$

The next Theorem is due to Bollobás [2] and Karoński and Ruciński [10]. Let $aut(H)$ denote the number of automorphisms of the graph H.

Theorem 1.3. *Let H be a strictly balanced graph and $pn^{1/\rho(H)} \to c$ for some arbitrary $c > 0$. Then,*

$$\lim_{n\to\infty} \mathbb{P}(H \subseteq \mathcal{G}(n,p)) = e^{-\lambda}$$

where $\lambda = c^{v_H}/aut(H)$.

Heuristically, if p is small, then $\mathcal{G}(n,p)$ will have a small number of edges, and is more likely to be H-free. We will characterize this phenomenon by defining the notion of a graph property.

Definition 1.4. *A graph property $\mathscr{P}$ is a set of graphs on $[n]$ that is closed under isomorphism. A property is increasing if given two graphs G_1, G_2 on $[n]$ with $G_1 \subseteq G_2$, $G_1 \in \mathscr{P}$ implies $G_2 \in \mathscr{P}$. A property is decreasing if given $G_1 \subseteq G_2$, $G_2 \in \mathscr{P}$ implies $G_1 \in \mathscr{P}$. The complement of a property $\mathscr{P}$ is the set of graphs on $[n]$ that does not lie in $\mathscr{P}$. If $\mathscr{P}$ is increasing, its complement $\overline{\mathscr{P}}$ is decreasing. A property $\mathscr{P}$ is non-trivial if $\mathscr{P} \neq \emptyset$ and $\overline{\mathscr{P}} \neq \emptyset$.*

Lemma 1.5. *Let $\mathscr{P}$ be an increasing property and $0 < p_1 < p_2 < 1$. Then,*

$$\mathbb{P}(\mathcal{G}(n,p_1) \in \mathscr{P}) \leq \mathbb{P}(\mathcal{G}(n,p_2) \in \mathscr{P})$$

Proof. Since $0 < 1-p_2 < 1-p_1 < 1$, we can choose $p \in (0,1)$ so that $(1-p)(1-p_1) = (1-p_2)$. Consider two independent random graphs, $\mathcal{G}(n,p), \mathcal{G}(n,p_1)$ on $[n]$. Then, an edge is not included in $\mathcal{G}(n,p_2)$ if it is not included in both $\mathcal{G}(n,p)$ and $\mathcal{G}(n,p_1)$. By independence, we know $\mathcal{G}(n,p_2)$ has the same distribution as $\mathcal{G}(n,p) \cup \mathcal{G}(n,p_1)$. Thus, for any fixed graph G on $[n]$, we have

$$\mathbb{P}(\mathcal{G}(n,p_2) = G) = \mathbb{P}(\mathcal{G}(n,p_1) \cup \mathcal{G}(n,p) = G)$$

and

$$\mathbb{P}(\mathcal{G}(n,p_2) \in \mathscr{P}) = \mathbb{P}(\mathcal{G}(n,p_1) \cup \mathcal{G}(n,p) \in \mathscr{P}) \geq \mathbb{P}(\mathcal{G}(n,p_1) \in \mathscr{P})$$

since $\mathcal{G}(n,p_1) \cup \mathcal{G}(n,p) \supseteq \mathcal{G}(n,p_1)$ and the property $\mathscr{P}$ is increasing. $\square$

In this proof, we know $\mathcal{G}(n,p) \cup \mathcal{G}(n,p_1)$ has the same distribution as $\mathcal{G}(n,p_2)$. The argument used is known as the coupling argument. It follows from this proof that if $0 < p_1 < p_2 < 1$ and $\mathscr{P}$ is decreasing, then $P(\mathcal{G}(n,p_1) \in \mathscr{P}) \geq P(\mathcal{G}(n,p_2) \in \mathscr{P})$. Since the property $\mathscr{P} = \{G : G \text{ contains } H \text{ as a subgraph}\}$ is increasing, it is reasonable to guess that if p is large, $\mathbb{P}(H \subseteq G) \to 1$ and if p is small, $\mathbb{P}(H \subseteq G) \to 0$. The following definition and theorem formalize this intuition.

Definition 1.6. *A function $p^* = p^*(n)$ is a threshold for an increasing property $\mathscr{P}$ if for any $p = p(n)$*

$$\lim_{n\to\infty} \mathbb{P}(\mathcal{G}(n,p) \in \mathscr{P}) = \begin{cases} 0 & \text{if } p/p^* \to 0 \\ 1 & \text{if } p/p^* \to \infty \end{cases}$$

The next theorem establishes the existence of a threshold. It was proved by Bollobas and Thompson [3].

Theorem 1.7. *Every non-trivial increasing property has a threshold.*

Proof. We use the coupling argument in Lemma 1.5. Let k be a positive integer, and let $\mathcal{G}_1(n,p), ..., \mathcal{G}_k(n,p)$ be k independent random graphs on $[n]$. The probability an edge e in K_n is not included in all of the graphs $\mathcal{G}_1(n,p), ..., \mathcal{G}_k(n,p)$ equals $(1-p)^k$, and it is the same as the probability $e \notin \mathcal{G}(n, 1-(1-p)^k)$. Therefore, the distribution of $\mathcal{G}_1(n,p) \cup ... \cup \mathcal{G}_k(n,p)$ is the same as the distribution of $\mathcal{G}(n, 1-(1-p)^k)$ by the same argument in Lemma 1.5. By the binomial theorem, $1-(1-p)^k \leq kp$. Therefore,

$$[\mathbb{P}(\mathcal{G}(n,p) \notin \mathscr{P})]^k = \mathbb{P}(\mathcal{G}(n, 1-(1-p)^k) \notin \mathscr{P}) \geq \mathbb{P}(\mathcal{G}(n,kp) \notin \mathscr{P}) \tag{1}$$

It follows that

$$\mathbb{P}(\mathcal{G}(n,p/k) \notin \mathscr{P}) \geq [\mathbb{P}(\mathcal{G}(n,p) \notin \mathscr{P})]^{1/k} \tag{2}$$

We know

$$\mathbb{P}(\mathcal{G}(n,p) \in \mathscr{P}) = \sum_{G \in \mathscr{P}} p^{e_G}(1-p)^{\binom{n}{2}-e_G}$$

is a polynomial on p. Since it is continuous and non-decreasing, it is strictly increasing, and it maps $[0,1] \to [0,1]$. Hence, there exists p^* such that $\mathbb{P}(G(n,p^*) \in \mathscr{P}) = 1/2$. We will prove that p^* is a threshold.
Suppose $p/p^* \to 0$. Let $\omega = \lfloor p^*/p \rfloor$. Then, $\omega \to \infty$ as $n \to \infty$. We know

$$\begin{aligned}
\mathbb{P}(\mathcal{G}(n,p) \in \mathscr{P})) &\leq \mathbb{P}(\mathcal{G}(n,p^*/\omega) \in \mathscr{P}) \qquad \text{as } p \leq p^*/\omega \\
&= 1 - \mathbb{P}(\mathcal{G}(n,p^*/\omega) \notin \mathscr{P}) \\
&\leq 1 - \mathbb{P}(\mathcal{G}(n,p^*) \notin \mathscr{P})^{1/\omega} \qquad \text{by (2)} \\
&= 1 - 2^{-1/\omega} \\
&\to 0
\end{aligned}$$

Suppose $p/p^* \to \infty$. Let $\omega = \lfloor p/p^* \rfloor$. Then, $\omega \to \infty$ and

$$\begin{aligned}
\mathbb{P}(\mathcal{G}(n,p) \in \mathscr{P}) &\geq \mathbb{P}(\mathcal{G}(n,\omega p^*) \in \mathscr{P}) \qquad \text{as } p \geq \omega p^* \\
&= 1 - \mathbb{P}(\mathcal{G}(n,\omega p^*) \notin \mathscr{P}) \\
&\geq 1 - \mathbb{P}(\mathcal{G}(n,p^*) \notin \mathscr{P})^{\omega} \qquad \text{by (1)} \\
&= 1 - 2^{-\omega} \\
&\to 1
\end{aligned}$$

□

If H is a balanced graph, then Theorem 1.2 states that $n^{-1/\rho(H)}$ is the threshold for the property that G contains H as a subgraph. We will prove this threshold in section 3.

2. Inequalities

In this section, we present a toolbox of inequalities used in the proof in section 3. The next theorem is known as the Markov inequality.

Theorem 2.1. *Let X be a non-negative random variable. Then, for all $a > 0$*

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}X}{a}$$

Proof. Let $I_{\{X \geq a\}}$ be the indicator function for the event $\{X \geq a\}$. We know

$$\mathbb{E}X \geq \mathbb{E}XI_{\{X \geq a\}} \geq \mathbb{E}aI_{\{X \geq a\}} = a\mathbb{P}(X \geq a)$$

Hence,

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}X}{a}$$

□

Corollary 2.2. *Let X be a non-negative integer valued random variable. Then,*

$$\mathbb{P}(X > 0) \leq \mathbb{E}X$$

Proof. Since X takes values on non-negative integers,

$$\mathbb{P}(X > 0) = \mathbb{P}(X \geq 1) \leq \mathbb{E}X$$

by Theorem 2.1. $\square$

The next theorem is knowned as the Chebyshev inequality, and the application of the next theorem is known as the second moment method.

Theorem 2.3. *Let X be a random variable and let $\mu = \mathbb{E}X$. Then, for all $\lambda > 0 \in \mathbb{R}$,*

$$\mathbb{P}(|X - \mu| \geq \lambda) \leq \frac{\text{Var}(X)}{\lambda^2}$$

Proof. By Markov inequality, we have

$$\begin{aligned}\mathbb{P}(|X - \mu| \geq \lambda) &= \mathbb{P}((X - \mu)^2 \geq \lambda^2) \\ &\leq \frac{\mathbb{E}(X - \mu)^2}{\lambda^2} \\ &= \frac{\text{Var}(X)}{\lambda^2}\end{aligned}$$

$\square$

Corollary 2.4. *Let X be a random variable and let $\mu = \mathbb{E}X$. Then,*

$$\mathbb{P}(X = 0) \leq \frac{\text{Var}(X)}{\mu^2}$$

Proof. If $X = 0$ then $|X - \mu| \geq \mu$. Substitute λ with μ in Theorem 2.3. $\square$

In the rest of this paper, we will denote $X = X(n)$ as a sequence of random variables. Markov inequality is frequently used to prove $\mathbb{P}(X = 0) \to 1$, and Chebyshev inequality is frequently used to prove $\mathbb{P}(X = 0) \to 0$. In this paper and in many other circumstances, let $t = t(n)$, $X = \sum_1^t X_i$ where $X_i = X_i(n)$ are mostly independent indicator random variables. Let $\mu = \mathbb{E}X = \sum_i \mathbb{E}X_i$. If $\mu \to 0$, then $\mathbb{P}(X = 0) \to 1$. However, if $\mu \to \infty$, that does not show $\mathbb{P}(X = 0) \to 0$.
We know

$$\text{Var}(X) = \sum_i \text{Var}(X_i) + \sum_{i \neq j} \text{Cov}(X_i, X_j)$$

We write $i \sim j$ if X_i, X_j are not independent and $i \neq j$. If $i \nsim j$ and $i \neq j$, then $\text{Cov}(X_i, X_j) = 0$. If $\mathbb{P}(X_i = 1) = p_i$, then $\text{Var}(X_i) = p_i(1 - p_i) \leq p_i = \mathbb{P}(X_i = 1)$ and if $i \sim j$, then $\text{Cov}(X_i, X_j) = \mathbb{E}[X_i X_j] - \mathbb{E}[X_i]\mathbb{E}[X_j] \leq \mathbb{E}[X_i X_j] = \mathbb{P}(X_i X_j = 1)$. Therefore, we know

$$\text{Var}(X) \leq \sum_i \mathbb{P}(X_i = 1) + \sum_{i \sim j} \mathbb{P}(X_i X_j = 1) = \mu + \sum_{i \sim j} \mathbb{P}(X_i X_j = 1)$$

Let $\Delta = \sum_{i \sim j} \mathbb{P}(X_i X_j = 1)$. Then, if $\mu \to \infty$ and $\Delta = o(\mu^2)$, then by Corollary 2.4, we have

$$\mathbb{P}(X = 0) \leq \frac{\mu + \Delta}{\mu^2} \to 0$$

The next two inequalities are used to study the edge cases where $\mathbb{E}X \to c$ for some constant c. In this case, we want to show

$$\mathbb{P}(X=0) \approx \prod_i \mathbb{P}(X_i = 0) = \prod_i (1 - \mathbb{E}X_i) \approx e^{-\sum_i \mathbb{E}X_i} = e^{-\mu}$$

given that the dependency between the X_is is weak, $\mathbb{P}(X_i)$ is small, and the number of i is large. Such a model is sometimes referred to as the Poisson Paradigm.

We know a graph property is a subset of $\{(i,j) : i < j \in [n]\}$. To develop the next two inequalities, we generalize from the graph case. To simplify, we consider a property $\mathscr{P}$ on $[n]$, which is a subset of $\mathcal{P}([n])$. We say a property $\mathscr{P}$ is increasing if given $A, B \subseteq [n], A \subseteq B$, $A \in \mathscr{P}$ implies $B \in \mathscr{P}$, and it is decreasing if given $A \subseteq B$, $B \in \mathscr{P}$ implies $A \in \mathscr{P}$. We know an intersection of increasing properties is increasing, and an intersection of decreasing properties is decreasing.

We consider a random subset $S \subseteq [n]$ where each element $k \in [n]$ is included in S independently with probability p. We refer to p as the parameter. For any fixed subset $A \subseteq [n]$, we have $\mathbb{P}(S = A) = p^{|A|}(1-p)^{n-|A|}$. Given a property $\mathscr{P}$ on $[n]$, we define

$$\mathbb{P}(S \subseteq \mathscr{P}) = \sum_{A \in \mathscr{P}} \mathbb{P}(S = A)$$

The next result is called a correlation inequality and it is due to Harris [8] and Kleitman [11].

Theorem 2.5. *Let $\mathscr{P}_1, \mathscr{P}_2$ be two increasing properties on $[n]$ and let S be a random set on $[n]$ with parameter p. Then*

$$\mathbb{P}(S \in \mathscr{P}_1 \cap \mathscr{P}_2) \geq \mathbb{P}(S \in \mathscr{P}_1)(S \in \mathscr{P}_2)$$

Proof. We will prove it by induction on n. In the base case where $n = 1$, the only increasing properties on $[1]$ are $\emptyset, \{\{1\}\}, \{\emptyset, \{1\}\}$, and the statement holds trivially.

For $n \geq 2$, let $\mathscr{P}$ be a property on $[n]$. Define

$$\mathscr{P}^- = \{A \subseteq [n-1] : A \in \mathscr{P}\} \qquad \mathscr{P}^+ = \{A \subseteq [n-1] : A \cup \{n\} \in \mathscr{P}\}$$

Let S^* denote the random subset on $[n-1]$ with parameter p. Then,

$$\mathbb{P}(S \in \mathscr{P}) = (1-p)\mathbb{P}(S^* \in \mathscr{P}^-) + p\mathbb{P}(S^* \in \mathscr{P}^+)$$

Similarly, define $\mathscr{P}_1^-, \mathscr{P}_1^+, \mathscr{P}_2^-, \mathscr{P}_2^+$. Note that all these four properties on $[n-1]$ are increasing, and $(\mathscr{P}_1 \cap \mathscr{P}_2)^- = \mathscr{P}_1^- \cap \mathscr{P}_2^-$ and $(\mathscr{P}_1 \cap \mathscr{P}_2)^+ = \mathscr{P}_1^+ \cap \mathscr{P}_2^+$. Moreover, $\mathscr{P}_1^- \subseteq \mathscr{P}_1^+$ and $\mathscr{P}_2^- \subseteq \mathscr{P}_2^+$ because $\mathscr{P}_1, \mathscr{P}_2$ are increasing. Therefore,

$$\begin{aligned} P(S \in \mathscr{P}_1 \cap \mathscr{P}_2) &= (1-p)\mathbb{P}(S^* \in (\mathscr{P}_1 \cap \mathscr{P}_2)^-) + p\mathbb{P}(S^* \in (\mathscr{P}_1 \cap \mathscr{P}_2)^+) \\ &= (1-p)\mathbb{P}(S^* \in \mathscr{P}_1^- \cap \mathscr{P}_2^-) + p\mathbb{P}(S^* \in \mathscr{P}_1^+ \cap \mathscr{P}_2^+) \\ &\geq (1-p)\mathbb{P}(S^* \in \mathscr{P}_1^-)\mathbb{P}(S^* \in \mathscr{P}_2^-) + p\mathbb{P}(S^* \in \mathscr{P}_1^+)\mathbb{P}(S^* \in \mathscr{P}_2^+) \end{aligned}$$

where the last inequality is due to the inductive hypothesis. Let $a = \mathbb{P}(S^* \in \mathscr{P}_1^-)$, $b = \mathbb{P}(S^* \in \mathscr{P}_2^-), c = \mathbb{P}(S^* \in \mathscr{P}_1^+), d = \mathbb{P}(S^* \in \mathscr{P}_2^+)$. We want to show

$$(1-p)ab + pcd \geq ((1-p)a + pc)((1-p)b + pd) = \mathbb{P}(S \in \mathscr{P}_1)\mathbb{P}(S \in \mathscr{P}_2)$$

Since

$$(1-p)ab + pcd - ((1-p)a + pc)((1-p)b + pd) = p(1-p)(a-c)(b-d)$$

and $a < c, b < d$, the proof follows. $\square$

The following corollary can be proven using Theorem 2.5 and basic algebra. We will omit the proof here.

Corollary 2.6. *(a). If $\mathscr{P}_1, \mathscr{P}_2$ are decreasing, then*

$$\mathbb{P}(S \in \mathscr{P}_1 \cap \mathscr{P}_2) \geq \mathbb{P}(S \in \mathscr{P}_1)\mathbb{P}(S \in \mathscr{P}_2)$$

(b). If $\mathscr{P}_1$ is increasing and $\mathscr{P}_2$ is decreasing, then

$$\mathbb{P}(S \in \mathscr{P}_1 \cap \mathscr{P}_2) \leq \mathbb{P}(S \in \mathscr{P}_1)\mathbb{P}(S \in \mathscr{P}_2)$$

(c). If $\mathscr{P}_1, ..., \mathscr{P}_k$ are all increasing/decreasing, then

$$\mathbb{P}\left(S \in \bigcap_{i=1}^{k} \mathscr{P}_i\right) \geq \prod_{i=1}^{k} \mathbb{P}(S \in \mathscr{P}_i)$$

Consider the random set $S \subseteq [n]$ with parameter p. Let $\{A_i\}_{i=1}^k$ be a collection of subsets of $[n]$. Let $\mathscr{P}_i = \{A \subseteq [n] : A \supseteq A_i\}$. Clearly, $\mathscr{P}_i$ is increasing. Moreover, $\mathscr{P}_i \cap \mathscr{P}_j = \{A : A \supseteq A_i \cup A_j\}$. Consider the indicator random variable X_i where $X_i(S) = 1$ iff $S \supseteq A_i$. Then, $\mathbb{P}(X_i = 1) = p^{|A_i|}$. Notice that if $A_i \cap A_j = \emptyset$, then X_i and X_j are independent. We say $i \sim j$ if $i \neq j$ and $A_i \cap A_j \neq \emptyset$ and we define $\sum_i \mathbb{E}\left[X_i\right] = \mu$ and $\sum_{i \sim j} \mathbb{E}[X_i X_j] = \Delta$.

The next theorem is known as the Janson Inequality and is due to Janson, Luczak and Ruciński [9].

Theorem 2.7. *Consider the setup above. Then*

$$\mathbb{P}\left(S \in \bigcap \overline{\mathscr{P}_i}\right) \leq e^{-\mu + \Delta/2}$$

Proof. Let B_i be the event $\mathscr{P}_i$ in the probability space $\mathcal{P}([n])$ and we say $\mathbb{P}(B_i) = \mathbb{P}(S \in \mathscr{P}_i)$. We abbreviate $B_i B_j$ for $B_i \cap B_j$. Let $r_i = \mathbb{P}(B_i | \overline{B}_1 ... \overline{B}_{i-1})$.

$$\begin{aligned}
\mathbb{P}(\overline{B}_1 ... \overline{B}_k) &= \mathbb{P}(\overline{B}_1)\mathbb{P}(\overline{B}_2|\overline{B}_1)...\mathbb{P}(\overline{B}_k|\overline{B}_1...\overline{B}_{k-1}) \\
&= (1 - r_1)(1 - r_2)...(1 - r_k) \\
&\leq e^{-r_1 + ... + r_k}
\end{aligned}$$

We will to show that $r_i \geq \mathbb{P}(B_i) - \sum_{j \sim i, j < i} \mathbb{P}(B_i B_j)$. Then, we have $\sum_i r_i \geq \mu - \Delta/2$. Fix i, let $D_1 = \bigcap_{j:j<i, j \sim i} \overline{B}_j$, $D_2 = \bigcap_{j:j<i, j \nsim i} \overline{B}_j$. Then,

$$\begin{aligned}
r_i &= \mathbb{P}(B_i | D_1 D_2) && (3) \\
&= \frac{\mathbb{P}(B_i D_1 D_2)}{\mathbb{P}(D_1 D_2)} && (4) \\
&\geq \frac{\mathbb{P}(B_i D_1 D_2)}{\mathbb{P}(D_2)} && (5) \\
&= \frac{\mathbb{P}(B_i D_2)}{\mathbb{P}(D_2)} - \frac{\mathbb{P}(B_i \overline{D}_1 D_2)}{D_2} && (6) \\
&= \mathbb{P}(B_i) - \frac{\mathbb{P}(B_i \overline{D}_1 D_2)}{\mathbb{P}(D_2)} && (7) \\
&\geq \mathbb{P}(B_i) - \mathbb{P}(B_i \overline{D}_1) && (8)
\end{aligned}$$

The equation (7) is due to the fact that $\bigcup_{j\sim i, j\sim i} A_j$ is disjoint from A_i, hence B_i and D_2 are independent. The equation (8) is due to the correlation inequality, since $B_i, \overline{D}_1$ are increasing and D_2 is decreasing. Finally,

$$\mathbb{P}(B_i\overline{D}_1) = \mathbb{P}\left(B_i \cap \bigcup_{j<i, j\sim i} B_j\right) \leq \sum_{j<i, j\sim i} \mathbb{P}(B_iB_j)$$

by the subadditivity of probability measure. □

Both the correlation inequality and the Janson inequality are statements that bound the probability of certain properties on $[n]$. Since a random graph can be relabeled to a random subset of $[\binom{n}{2}]$, the correlation inequality and the Janson inequality apply to $\mathcal{G}(n,p)$ as well. Moreover, a subset A_i in the setup of the Janson inequality can be viewed as a subgraph of K_n. In many cases, if we want to apply Janson inequality, we want to show μ is constant or approaches a constant value, and $\Delta = o(1)$. If $\mathbb{P}(B_i) = o(1)$ is the same for each i, then in this case,

$$\left(\prod_i (1 - \mathbb{P}(B_i))\right) = e^{-\mu + o(1)} \leq \mathbb{P}(\overline{B}_1 ... \overline{B}_k) \leq e^{-\mu + o(1)}$$

Then, we know $\mathbb{P}(\overline{B}_1 ... \overline{B}_k) \to e^{-\mu}$.

3. Thresholds

In this section, we prove Theorem 1.2 and 1.3. We will be using the first and the second moment method as well as the correlation inequality and the Janson inequality. We will first use a lemma for counting the number of copies of H in K_{v_H}.

Lemma 3.1. *The number of copies of H in K_{v_H} is $v_H!/aut(H)$.*

Proof. Let $n = v_H$ and let $\mathcal{H}$ be the set of all copies of H in K_n. For each $\sigma \in S_n$,, $H \in \mathcal{H}$, define H_σ to be the graph on $[n]$ with the edge set $\{(\sigma(i), \sigma(j)) : (i,j) \in E(H)\}$. Then, S_n defines a group action on $\mathcal{H}$. Fix $H_0 \in \mathcal{H}$. The orbit $\mathcal{O}$ of H_0 is $\mathcal{H}$ and the stabilizer of H_0, denoted S, is the group of automorphisms of H_0. Therefore, by the orbit stabilizer theorem,

$$n! = |S_n| = |\mathcal{O}||S| = |\mathcal{H}|aut(H)$$

□

The next lemma calculates the expected number of copies of H in the random graph $\mathcal{G}(n,p)$.

Lemma 3.2. *Let X be the number of copies of H in $\mathcal{G}(n,p)$. Then,*

$$\mathbb{E}X = \frac{n_{(v_H)}}{aut(H)} p^{e_H}$$

Proof. We first determine the number of copies of H in K_n. There are $\binom{n}{v_H}$ ways to choose v_H vertices from $[n]$. Among these vertices, there are $v_H!/aut(H)$ copies of H. Therefore, there are a total of $(n)_{(v_H)}/aut(H)$ copies of H in K_n. We label them with $H_1, H_2, ..., H_t$ where $t = (n)_{(v_H)}/aut(H)$. For each $i \in [t]$, let X_i be the indicator random variable for the event $H_i \subseteq \mathcal{G}(n,p)$, and $X = \sum_i X_i$. Then, $\mathbb{P}(X_i = 1) = p^{e_H}$, and we have

$$\mathbb{E}X = \sum_i \mathbb{E}X_i = \frac{n_{(v_H)}}{aut(H)} p^{e_H}$$

$\square$

The next lemma gives an upper bound for the constant Δ.

Lemma 3.3. *If H is a balanced graph, then*

$$\Delta = O\left(\sum_{j=2}^{v_H} n^{2v_H-j} p^{2e_H-j\rho(H)}\right)$$

If H is strictly balanced, then

$$\Delta = O\left(\sum_{j=2}^{v_H} n^{2v_H-j} p^{2e_H-j\rho(H)+1}\right)$$

Proof. Again, let $t = (n)_{(v_H)}/aut(H)$, $H_1, ..., H_t$ be the copies of H in K_n, $X_1, ..., X_t$ be the indicator random variables. For $i \neq j \in [k]$, we denote $i \sim_k j$ if $|V(H_i) \cap V(H_j)| = k$, and $i \sim j$ if $i \sim_k j$ for some $2 \leq k \leq v_H$. Suppose $i \sim_k j$. We first choose the vertices of $V(H_i) \cup V(H_j)$ in $\binom{n}{2v_H-j} = \mathcal{O}(n^{2v_H-j})$ ways. Then, we partition the vertex set into three parts, $V(H_i)/V(H_j), V(H_i) \cap V(H_j), V(H_j)/V(H_i)$ in $(2v_H-j)!/(v_H-j)!^2 j! = O(1)$ ways. Finally, we choose the edges of H_i and H_j in at most $(v_H!/aut(H))^2 = O(1)$ ways. If H is balanced, then suppose $K = H_i \cap H_j$. We know $K \subsetneq H$ and $e_K \leq j\rho(H)$. Therefore, $H_i \cup H_j$ will have $2e_H - e_K \geq 2e_H - j\rho(H)$ edges. We have

$$\mathbb{P}(X_i X_j) \leq p^{2e_H-j\rho(H)}$$

and

$$\Delta = O\left(\sum_{j=2}^{v_H} n^{2v_H-j} p^{2e_H-j\rho(H)}\right)$$

If H is strictly balanced, then $K \subsetneq H$ is a proper subgraph, and $e_K \leq j\rho(H) - 1$. Hence,

$$\Delta = O\left(\sum_{j=2}^{v_H} n^{2v_H-j} p^{2e_H-j\rho(H)+1}\right)$$

$\square$

Now, we are able to prove Theorem 1.2.

Proof of Theorem 1.2. Suppose H is balanced. If $pn^{1/\rho(H)} \to 0$, then $n^{v_H} p^{e_H} \to 0$ and $\mathbb{E}X \to 0$. Therefore, $\mathbb{P}(X = 0) \to 1$. Suppose $pn^{1/\rho(H)} \to \infty$. Then, $n^{v_H} p^{e_H} \to \infty$ and $\mathbb{E}X = \Theta(n^{v_H} p^{e_H}) \to \infty$. Moreover,

$$\Delta/(\mathbb{E}X)^2 = \mathcal{O}\left(\sum_{j=2}^{v_H} n^{-j} p^{-j\rho(H)}\right) = o(1)$$

By applying the second moment method, we know

$$\mathbb{P}(X = 0) \leq \Delta/(\mathbb{E}X)^2 \to 0$$

so we finished the proof. $\square$

Proof of Theorem 1.3. Suppose H is strictly balanced and $n^{1/\rho(H)}p \to c$. We know

$$\mu = \frac{n^{v_H}p^{e_H}}{aut(H)}(1+o(1)) \to \frac{c^{v_H}}{aut(H)} = \lambda$$

Moreover,

$$\Delta = O\left(\sum_{j=2}^{v_H} n^{2v_H - j}e^{2e_H - j\rho(H)+1}\right) = \mathcal{O}(p) = o(1)$$

Finally, let $t = n_{(v_H)}!/aut(H)$, and let $X_1, ..., X_t$ denote the indicator random variable for $H_i \subseteq \mathcal{G}(n,p)$, we know the event $\{H_i \subseteq \mathcal{G}(n,p)\}$ is increasing. By the correlation inequality and the Janson inequality, we have

$$\prod_i \mathbb{P}(X_i = 0) \leq \mathbb{P}(X = 0) \leq e^{-\mu + \Delta/2}$$

and

$$\prod_i \mathbb{P}(X_i = 0) = \left(1 - \frac{\mu}{t}\right)^t = e^{-\mu}(1+o(1)) = e^{-\lambda}(1+o(1))$$

Since

$$e^{-\mu+\Delta/2} = e^{-\mu+o(1)} = e^{-\lambda}((1+o(1))$$

by the squeeze theorem, we know $\mathbb{P}(X=0) \to e^{-\lambda}$. ∎

4. Remarks

By the second moment method, we can actually show that when H is balanced, for all $\epsilon > 0$, $\mathbb{P}(|X - \mu| \geq \epsilon\mu) \to 0$. See Chapter 4.3 of [1] for more details. We can also show that if H is strictly balanced, then $X \to \mathrm{Po}(\lambda)$ in distribution, where $\lambda = c^{v_H}/aut(H)$ and $\mathrm{Po}(\lambda)$ is the Poisson distribution with mean λ. See Chapter 5.2 of [6] for more details. The proof uses the principle of Inclusion-Exclusion as opposed to the correlation inequality and the Janson inequality we used here.

5. Acknowledgement

This paper was written for the course Math509: Probabilistic Constructions in Model Theory, taught by Professor Rehana Patel at Wesleyan University.

In the summer of 2024, I conducted research with Professor Collins on the edge intersections between sparse graphs and random perfect matchings on the balanced complete r-partite graph $K_{r\times 2n/r}$. To better understand the literature in this area, I studied *The Probabilistic Method* by Alon and Spencer. This textbook introduced me to a wide range of probabilistic tools used in combinatorics, graph theory, and discrete mathematics.

To further pursue this line of study, I enrolled in Math509. In addition, I explored several textbooks on random graph theory, a field of intersections between probability, combinatorics, and graph theory. Random graph theory is a field closely related to the probabilistic method and directly relevant to my research interests.

For this course, I chose to write a paper on the existence of small subgraphs in the Erdős–Rényi random graph $G(n,p)$. This topic is suitable for a paper of appropriate length and allows me to apply several tools I learned from The Probabilistic Method, including the first and second moment methods, correlation inequalities, and Janson's inequality. The

discussion of the Poisson distribution in this context also connects to my current research, particularly with regard to the Poisson paradigm and the Poisson limit theorem.

Overall, I find the probabilistic method and random graph theory to be fascinating and rapidly evolving areas of mathematics that have seen significant development over the past 50 years.

References

[1] N. Alon and J. H. Spencer. *The probabilistic method*. Wiley Series in Discrete Mathematics and Optimization. John Wiley & Sons, Inc., Hoboken, NJ, fourth edition, 2016.

[2] B. Bollobás. Random graphs. In *Combinatorics (Swansea, 1981)*, volume 52 of *London Math. Soc. Lecture Note Ser.*, pages 80–102. Cambridge Univ. Press, Cambridge-New York, 1981.

[3] B. Bollobás and A. Thomason. Hereditary and monotone properties of graphs. In *The mathematics of Paul Erdős, II*, volume 14 of *Algorithms Combin.*, pages 70–78. Springer, Berlin, 1997.

[4] P. Erd˝ os. Graph theory and probability. *Canadian J. Math.*, 11:34–38, 1959.

[5] P. ERDOS and A. RENYI. On the evolution of random graphs. *Bulletin of the International Statistical Institute*, 38(4):343–347, 1960.

[6] A. Frieze and M. Karoński. *Introduction to random graphs*. Cambridge University Press, Cambridge, 2016.

[7] E. N. Gilbert. Random graphs. *The Annals of mathematical statistics*, 30(4):1141–1144, 1959.

[8] T. E. Harris. A lower bound for the critical probability in a certain percolation process. *Proc. Cambridge Philos. Soc.*, 56:13–20, 1960.

[9] S. Janson, T. Ł uczak, and A. Ruciński. An exponential bound for the probability of nonexistence of a specified subgraph in a random graph. In *Random graphs '87 (Poznań, 1987)*, pages 73–87. Wiley, Chichester, 1990.

[10] M. Karoński and A. Ruciński. On the number of strictly balanced subgraphs of a random graph. In *Graph theory (Łagów, 1981)*, volume 1018 of *Lecture Notes in Math.*, pages 79–83. Springer, Berlin, 1983.

[11] D. J. Kleitman. Families of non-disjoint subsets. *J. Combinatorial Theory*, 1:153–155, 1966.

[12] A. Ruciński and A. Vince. Strongly balanced graphs and random graphs. *Journal of graph theory*, 10(2):251–264, 1986.