# Chapter 7 Random Graphs

**Model A** $G(n,p)$ or $G(n,P=p)$, $0 \le p \le 1$.

The probability of the existence of an edge (independently) is $p$ and the graph induced by using existent edges is $G_p$.

• A graph $H$ with $m$ edges occurs : probability $=_{def}\ p^{\|H\|} \cdot q^{N-\|H\|}$ where $N = \binom{n}{2}$, denoted by $P_p(G_p = H) = p^{\|H\|} \cdot q^{N-\|H\|}$, $q = 1-p$.

• The probability of a graph with $m$ edges is $\binom{N}{m} P^m \cdot q^{N-m}$.

• We regularly use $p = \frac{1}{2}$.

**Model B** $G(n,M)$

We assume that the probability of an $M$ edges graph $G_M$ is equal. Therefore, an $M$ edges graph $H$ occurs with prpbability $\frac{1}{\binom{N}{M}}$, denoted by $P(G_M = H) = \binom{N}{M}^{-1}$ where $N = \binom{m}{2}$.

**Model C** $\tilde{\mathrm{G}}$ Sequence of Random graphs $G_0 \le G_1 \le \cdots \le G_N$

$\tilde{\mathrm{G}} = (e_1, e_2, \cdots, e_N)$ where $e_t \in E(G_t) \setminus E(G_{t-1})$ and $G_t$ has exactly $t$ edges. The probability of $\tilde{\mathrm{G}}$ is $\frac{1}{N!}$.

§ 7.1. **Basic Notions**

**Definition 7.1.1** (Discrete Probabilistic Space) , D.P.S.

A D.P.S. is an ordered pair $(S, f)$ where $S$ is a countable set and $f : S \to \mathbb{R}$ satisfying (i) $0 \leq f(x) \leq 1$ and (ii) $\Sigma_{x \in S} f(x) = 1$.

(Note) A countable set is either a finite set or an infinite set which has the same cardinality as $\mathbb{N}$.

**Definition 7.1.2** (The probability of an event $A \subseteq S$)

Let $(S, f)$ be a D.P.S.. Then the probability of $A \subseteq S$ is $P(A) = \Sigma_{x \in A} f(x)$.

**Definition 7.1.3** (Independent events)

If $P(A \cap B) = P(A)P(B)$, then $A$ and $B$ are independent events.

**Definition 7.1.4** (Random variables)

Let $(S, f)$ be a D.P.S.. Then $\mathbb{X} : S \to \mathbb{R}$ is a random variable where we use $\mathbb{X} = k$ to denote an event.

$$K = \{x \in S \mid \mathbb{X}(x) = k\}.$$

e.g. Let $S = [1, 6]^2$ and $f(x, y) = \frac{1}{36}$ for each $(x, y) \in [1, 6]^2$. $\mathbb{X}((x, y)) = x + y$, $k = 7 \Rightarrow (\mathbb{X} = 7) = \{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\}$.

**Definition 7.1.5** (Expectation)

Let $\mathbb{X}$ be a random variable. Then the expectation of $\mathbb{X}$, $\mathbb{E}(X) = \Sigma_k k \cdot p(\mathbb{X} = k)$. ( We define $P(\mathbb{X} = h) = 0$ if $h$ is not in the image of $\mathbb{X} : S \to \mathbb{R}$.)

e.g. (Continued), $\mathbb{X} = 7$.

$$\begin{aligned}
\mathbb{E}(X) = {} & 2 \cdot \frac{1}{36} + 3 \cdot \frac{1}{18} + 4 \cdot \frac{1}{12} + 5 \cdot \frac{1}{9} + 6 \cdot \frac{5}{36} + 7 \cdot \frac{1}{6} \\
& + 12 \cdot \frac{1}{36} + 11 \cdot \frac{1}{18} + 10 \cdot \frac{1}{12} + 9 \cdot \frac{1}{9} + 8 \cdot \frac{5}{36} \\
& = 14 \cdot (\frac{1}{36} + \frac{1}{18} + \frac{1}{12} + \frac{1}{9} + \frac{5}{36} + \frac{1}{12}) \\
& = 14 \cdot \frac{1 + 2 + 3 + 4 + 5 + 3}{36} = 7.
\end{aligned}$$

**Lemma 7.1.1** (Pigeon-hole principle of Expectation)

Let $\mathbb{X}$ be a random variable of a D.P.S. Then, there exists a $y \in S$ such that $\mathbb{X}(y) \geq \mathbb{E}(\mathbb{X})$.

**Lemma 7.1.2** (Linear Property of Expectation)

Let $X, X_1, \cdots, X_m$ be random variables such that $X = \Sigma_{i=1}^{m} X_i$. Then, $\mathbb{E}(X) = \Sigma_{i=1}^{m}\mathbb{E}(X_i)$.

**Definition 7.1.6** (Indicstor Random Variable)

An indicator random variable is a random variable $X$ such that $\mathbb{X} : S \to \{0, 1\}$ (instead if $\mathbb{R}$).
(Note) A random variable $\mathbb{X}$ can be written as a sum of $|G|$ indicator random variables

$$x_v = \begin{cases} 1, & \text{if } v \in \mathbb{X} \text{ , and} \\ 0, & \text{otherwise} \end{cases}$$

**Theorem 7.1.3** If $\binom{n}{k} \cdot 2^{1-\binom{k}{2}} \leq 1$, then $R(k, k) > n$. Thus, $R(k, k) > \lfloor 2^{\frac{k}{2}} \rfloor$ for all $k \geq 3$.

**Proof.** Consider a random red-blue coloring of the edges of $K_n$. For a fixed set $T$ of $k$ vertices, let $A_T$ be the event that $< T >$ is monochromatic. Hence, $P(A_T) = (\frac{1}{2})^{\binom{k}{2}} \cdot 2 = 2^{1-\binom{k}{2}}$. Since there are $\binom{n}{k}$ possible sets for $T$, the probability that at least one of the events $A_T$ occurs is $\binom{n}{k} \cdot 2^{1-\binom{k}{2}}$. By assumption, $\binom{n}{k} \cdot 2^{1-\binom{k}{2}} < 1$. This implies that no event $A_T$ occurs is of positive probability, i.e., there exists a red-blue coloring such that no monochromatic $K_k$ exists, we have $R(k, k) > n$. Now, if we take $n = \lfloor 2^{\binom{k}{2}} \rfloor$,

$$\begin{aligned} &\binom{n}{k} 2^{1-\binom{k}{2}} \quad \text{where } 1 - \frac{k(k-1)}{2} = 1 - \frac{k^2}{2} + \frac{k}{2} \\ &< \frac{n^k}{k!} \cdot \frac{2^{1+\frac{k}{2}}}{2^{\frac{k^2}{2}}} \\ &\leq \frac{(2^{\frac{k}{2}})^k}{k!} \cdot \frac{2^{1+\frac{k}{2}}}{2^{\frac{k^2}{2}}} < 1. (k \geq 3) \end{aligned}$$

Hence, $R(k, k) > \lfloor 2^{\frac{k}{2}} \rfloor$ for all $k \geq 3$. This concludes the proof. □

**Theorem 7.1.4** (Szele, 1943)

There exists a tournament $T_n$ such that $T_n$ has at least $\frac{n!}{2^{n-1}}$ Hamiltonian paths.

**Proof.** There are $n!$ possible Hamiltonian (undirected) paths and the probability of a undirected Hamiltonian path (H. path) is a directed H. path is $\frac{1}{2^{n-1}}$. Therefore, $\mathbb{E}(X) = n! \cdot \frac{1}{2^{n-1}}$. This concludes the proof. □

(Note) How many are they ? In 1990, Alon proved that the number of H. paths is at most $\frac{n!}{(2-o(1))^n}$.

**Theorem 7.1.5** $\alpha(G) \geq \Sigma_{v \in V(G)} \frac{1}{1+deg_G(v)}$.

**Proof.** (Greedy Algorithm) In a set of $deg_G(v) + 1$ vertices we can select one vertex. This concludes the proof by selecting an independent set one vertex at a time. □

**Proof.** (Random idea) Use $1, 2, ..., |G|$ to label the vertices of the set $V(G)$ randomly call it $\varphi$. Let $v_0 \in S$ (independent set) if $\varphi(v_0) = min\{\varphi(x) | x \in \mathbb{N}[v_0]\}$. So, the probability is $\frac{1}{1+deg_G(v_0)}$ and the expectation value is $\Sigma_{v \in V(G)} \frac{1}{1+deg_G(v)}$. □

**Definition 7.1.7** (Dominating Set)

Let $S \subseteq V(G)$. Then, $S$ is said to be a dominating set of $G$ if for all $v \in V(G)$§, $v$ is adjacent to a vertex in $S$.
$min\{|S| | S$ is a dominating set of $G\} = D(G)$ is the domination number of $G$.

**Theorem 7.1.6** (Alon, 1990)

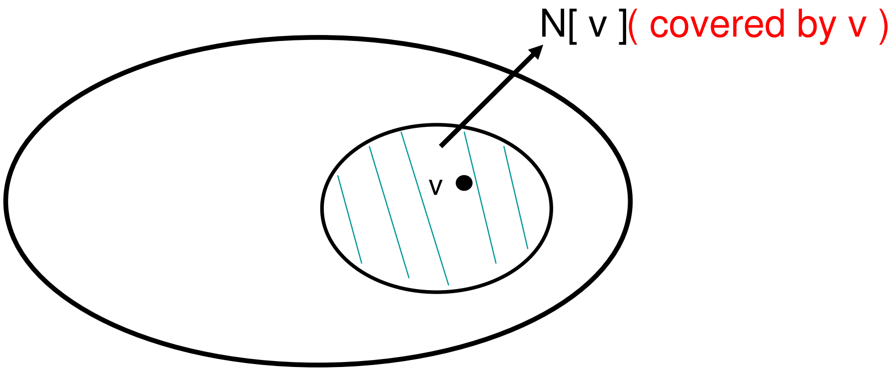
Let $|G| = n$. Then $D(G) \leq \frac{n(1+ln(\delta(G)+1))}{\delta(G)+1}$.

**Proof.** Let $S$ be a subset of $V(G)$ with the probability of each vertex $p =_{def} \frac{ln(\delta(G)+1)}{\delta(G)+1}$. Let $T = \{x | x \notin S, N(x) \cap S = \phi\}$. Then $S \cap T$ is a dominating set of $G$. By assumption, $E(|S|) = n \cdot p$ and $E(|T|) \leq n \cdot (1-p)^{\delta(G)+1}$. Note here that $|S|$ and $|T|$ are random variables. Since $(1-p)^{\delta(G)+1} \leq e^{-p(\delta(G)+1)}$, $E(|S| + |T|) \leq np + ne^{-p(\delta(G)+1)} = n(p + \frac{1}{\delta(G)+1}) = n(\frac{1+ln(\delta(G)+1)}{\delta(G)+1})$. This implies that there exists a dominating set with at most $n(\frac{1+ln(\delta(G)+1)}{\delta(G)+1})$ vertices. □

**Deterministic Algorithm** (idea)

Choose the vertices jof a dominating set one by one !

Step: A vertex that covers the maximum number of vertices which are not covered yet is picked.




$r(1-\frac{\delta+1}{n})$
$r(1-\frac{\delta+1}{n})^2 \ldots$
$r(1-\frac{\delta+1}{n})^{\frac{n}{\delta+1}ln(\delta+1)}$
$ln(\delta+1)(1-x)^{\frac{1}{x}}$ where $x=\frac{\delta+1}{n}$.

**Theorem 7.1.7** If $|G|=n$ and $\|G\|=\frac{nd}{2}$, $d \geq 1$. Then $\alpha(G) \geq \frac{n}{2d}$.

**Proof.** Let $S \subseteq V(G)$ be a random subset defined by

$P[v \in S]=p$. Let $X=|S|$.

For each $e=\{v_i, v_j\} \in E(G)$, let $Y_e$ be the indicator random variable for the event $\{v_i, v_j\} \subseteq S$ and $Y=\Sigma_{e \in E} Y_e$. Now, $E(Y_e)=P[v_i, v_j \in S]=p^2$ and thus $E(Y)=\frac{nd}{2}p^2$. Since $E(X)=np$, $E(X-Y)=np-\frac{nd}{2}p^2=np(1-\frac{d}{2}p)$, $p=\frac{1}{d}$ gives the maximum. Hence, $E(X-Y)=\frac{n}{2d}$.

Thus, there exists a specific $S$ for which $|S|-\|S\| \geq \frac{n}{2d}$. Now, select one vertex from each edge of $S$ and delete it to obtain a set $S^*$ with at least $\frac{n}{2d}$ vertices. Since all edges are gone, $S^*$ is an independent set. $\square$

§ 7.2. **Almost all graphs**

**Definition 7.2.1** We use $n-$th space $G^n$ to denote the distribution of graphs of order $n$. Let $q_n$ be the probability of the existence of "Property $Q$".

**Definition 7.2.2** If $lim_{n\to\infty} q_n = 1$, then we say "$Q$" almost always holds or in this case, almost all graphs have property "$Q$".

**Theorem 7.2.1** (Gilbert,1959)

Let $p$ be a constant such that $0 < p \leq 1$. Then, almost all graphs are connected.

**Proof.** If $G$ is not connected, then there exists a subset $S \subseteq V(G)$ such that $< S, V(G)\S >= \phi$. This implies that the probability $q_n$ of the existence of disconnected graphs of order $n$ satisfies

$0 \leq q_n \leq \Sigma_{k=1}^{\lfloor \frac{n}{2} \rfloor} \binom{n}{k} (1-p)^{k(n-k)} \cdot p^x$

(here $p$ is the probability of edges and $x$ is fixed).

$$\begin{aligned} \text{Hence } \; 0 \leq q_n &\leq \Sigma_{k=1}^{\lfloor \frac{n}{2} \rfloor} (n^k)(1-p)^{k(n-k)} \leq \Sigma_{k=1}^{\lfloor \frac{n}{2} \rfloor} (n(1-p)^{n-k})^k \\ &\leq \Sigma_{k=1}^{\lfloor \frac{n}{2} \rfloor} (n(1-p)^{\frac{n}{2}})^k \\ &< \frac{x}{1-x} \text{ where } x = n(1-p)^{\frac{n}{2}}. \text{ But } lim_{n\to\infty} x = 0 \end{aligned}$$

This implies that $lim_{n\to\infty} q_n = 0$. □

In what follows, let $\mathbb{X}$ be an Integer-Valued Random Variable.

**Lemma 7.2.2** (Markov's Inequality)

Let $p_k = P(\mathbb{X} = k)$, $k \geq 0$. Then, $P(\mathbb{X} \geq t) \leq \frac{E(X)}{t}$. Moreover, if $E(X) \to 0$, then $P(X = 0) \to 1$.

**Proof.** $E(X) = \Sigma_{k\geq 0} k p_k \geq \Sigma_{k\geq t} k p_k \geq t\Sigma_{k\geq t} p_k = tP(\mathbb{X} \geq t)$. □

Review

In proving that almost all graphs in $G^p$ are connected, we may assume that the random variable $\mathbb{X} = 0$ when $G$ is connected and $\mathbb{X} = 1$ for the situation that $G$ is disconnected.

**Theorem 7.2.3** Let $0 < p \leq 1$. Then almost all graphs are of diameter 2.

**Proof.** Let $\mathbb{X}_{i,j}$ be the indicator random variable such that

$$\mathbb{X}_{i,j} = \begin{cases} 1, & \text{if } v_i \text{ and } v_j \text{ do not have a common neighbor; and} \\ 0, & \text{otherwise.} \end{cases}$$

Note that the probability of "$v_i$ and $v_j$ do not have a common neighbor" is equal to $(1-p^2)^{n-2}$, hence $P(\mathbb{X}_{i,j} = 1) = (1-p^2)^{n-2}$. By def. $\mathbb{X} = \Sigma_{i \neq j} \mathbb{E}(\mathbb{X}_{i,j}) = \binom{n}{2} \cdot (1-p^2)^{n-2}$.
Since $lim_{n\to\infty} \binom{n}{2}(1-p)^{n-2} = 0$, $\mathbb{E}(\mathbb{X}) \to 0$. This implies that $P(\mathbb{X} = 0) \to 1$, i.e., almost every pair of distinct vertices $v_i$ and $v_j$ have a common neighbor. This concludes the proof. □

Remark
(1) Here, $p$ is a constant !
(2) If $G^p$ has $n$ vertices, then $G^p$ has $p \cdot \binom{n}{2}$ edges.
(3) $p$ may be different as $n$ changes ; this is the idea of $p(n)$ which will be introduced later.

**Definition 7.2.3** (Monotonic property)
Let $Q$ be a property. If $G$ has property $Q$ and for each edge $e \in E(\bar{G})$ $\bar{G}$ also has the property $Q$, then $G$ is a monotonic property.
e.g. If $Q$ is the property "diameter 2", then $Q$ is a monotonic property.
Clearly, planarity is not a monotonic property.

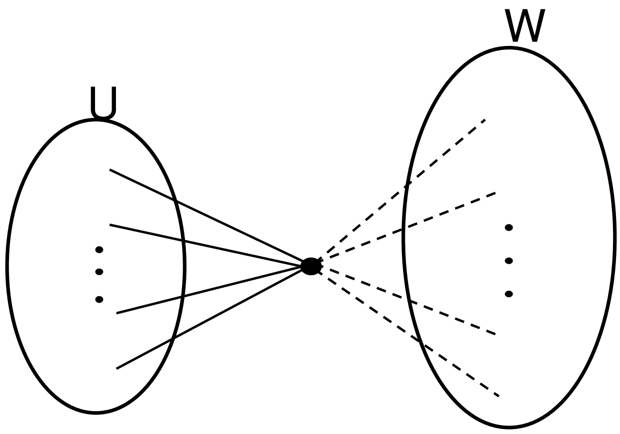
**Definition 7.2.4** Let $Q$ be a monotonic property and $t(n)$ be a function of $n$ such that
(i) $\frac{p(n)}{t(n)} \to 0 \Rightarrow$ Almost all graphs in $G^p$ has no property $Q$, and
(ii) $\frac{p(n)}{t(n)} \to \infty \Rightarrow$ Almost all graphs $G^p$ has property $Q$.
Then $t(n)$ is called a threshold probability function of $Q$.

**Theorem 7.2.4** For every constant $p \in (0,1)$ and every graph $H$, almost all graphs $G^p$ contains an induced copy of $H$.

**Proof.** Let $H$ be given and $|H| = k$. Let $U$ be a set of $k$ (fixed) vertices of $G$ then $< U >_G \cong H$ with a certain probability $r > 0$. ($r$ depends on $p$, not $n$ ?) Now, $G$ contains a collection of $\lfloor \frac{n}{k} \rfloor$ disjoint sets $U_i$ of size "$k$". So, the probability that none of $< U_i >_G$ is isomorphic to $H$ is $(1 - r^{\lfloor \frac{n}{k} \rfloor})$. Hence, $P[H \not\leq G] \leq (1-r)^{\lfloor \frac{n}{k} \rfloor}$ which is going to "0" as $n \to \infty$. □

**Theorem 7.2.5** For every constant $p \in (0,1)$ and $i,j \in \mathbb{N}$, almost all graphs $G^p$ has the property $P_{i,j}$ where $P_{i,j}$ is the property that for any disjoint vertex sets $U$ and $W$ with $|U| \leq i$ and $|W| \leq j$, a vertex $v \notin U \cup W$ that is adjacent to all the vertices of $U$ but to none of vertices in $W$.

**Proof.** The probability that $v \in V(G) \backslash (U \cup W)$ is adjacent to $U$ but not to $W$ is $p^{|U|} q^{|W|} \geq p^i q^j$. $|U| \leq i$, $|W| \leq j$. □




Hence, the probability that no suitable $v$ exists for these $U$ and $W$ is $(1 - p^{|U|} g^{|W|})^{n-|U|-|W|} \leq (1 - p^i g^j)^{n-i-j}$ where $(1 - p^i g^j)$ enlarge, and $n - i - j$ diminish, for $n \geq i + j$. Since the number of $< U, W >$ pairs is at most $n^{i+j}$, the total probability is $n^{i+j}(1 - p^i q^j)^{n-i-j} \to 0$ ($i, j$ are constants!) as $n \to \infty$. □.

**Corollary 7.2.6** For every constant $p \in (0,1)$ and $k \in \mathbb{N}$, almost all graphs are k-connected.

**Proof.** Let $i = 2$ and $j = k - 1$. Since almost all graphs $G$ has property $P_{k-1}$, $|G| \geq k + 2$. Let W be an arbitrary set of at most $k - 1$ vertices. Then $\forall x, y \in V(G) \backslash W$, then either $x$ is adjacent to $y$ or $x$ and $y$ have a common neighbor. Therefore, $W$ is not a vertex cut. This concludes the proof. □

**Lemma 7.2.7**. For all integers $n, k$ with $n \geq k \geq 2$, the probability that $G^p$ has a set of $k$ independent vertices is at most $P[\alpha(G) \geq k] \leq \binom{n}{k} \cdot g^{\binom{k}{2}}$.

**Proof.** (Exercise) $\square$

**Theorem 7.2.8** For every $p \in (0,1)$ and $\varepsilon > 0$, almost every graph $G$ has chromatic number

$$\mathbb{X}(G) > \frac{log(\frac{1}{1-p})}{2+\varepsilon} \cdot \frac{n}{logn}.$$

**Proof.** Since

$$\begin{aligned} P[\alpha \geq k] &\leq \binom{n}{k}(1-p)^{\binom{k}{2}} \\ &\leq n^k(1-p)^{\binom{k}{2}} \\ &= q^{k\frac{logn}{logq}+\frac{1}{2}k(k-1)} \\ &= q^{\frac{k}{2}[-\frac{2logn}{log(\frac{1}{q})}+k-1]}. \end{aligned}$$

Let $k =_{def} (2+\varepsilon)\frac{logn}{log(\frac{1}{q})}$.
Then $q^{\frac{k}{2}[\varepsilon\cdot\frac{logn}{log\frac{1}{q}}-1]} = q^{(1+\frac{\varepsilon}{2})\frac{logn}{log\frac{1}{q}}(\varepsilon\frac{logn}{log\frac{1}{q}}-1)} \to 0$.
This implies that in almost every graph no $k$ vertices are the same color. Hence $\mathcal{X}(G) > \frac{n}{k} = \frac{log\frac{1}{1-p}}{2+\varepsilon} \cdot \frac{n}{logn}$. $\square$

We remark have that the upper bound of $\mathcal{X}(G)$ was obtained in 1998 by *Bollobaś*.

**Theorem 7.2.9** $\mathcal{X}(G) \leq \frac{log\frac{1}{q}}{2-\varepsilon} \cdot \frac{n}{logn}$.

$$p = p(n) = \begin{cases} n^{-2} \to & \text{No edges !} \\ \sqrt{n}n^{-2} \to & \text{G has a nontrivial component which grows like a tree.} \\ n^{-1} \to & \text{Contains a cycle.} \\ \frac{logn}{n} \to & \text{connected.} \\ (1+\varepsilon)(logn)n^{-1} \to & \text{has hamiltonian cycles !} \end{cases}$$

From the above results, it seems that as long as $p$ is a constant in $(0,1)$, no matter $p = 0.1$ or $p = 0.9$, we always have the same conclusion. But, in fact, if we allow $p = p(n)$, then we will see the difference, see the following table.
**Lemma 7.2.10** Let $k > 0$ be an integer and let $p = p(n) \geq (6klnn)n^{-1}$ for large $n$. Then $lim_{n\to\infty}P(\alpha \geq \frac{1}{2}n/k) = 0$.

Table :

| $p(n)$ | graphs |
|---|---|
| $n^{-2}$ | No edges. |
| $\sqrt{n}n^{-2}$ | G has a nontrivial component which grows like a tree. |
| $n^{-1}$ | Contains a cycle. |
| $\frac{\log n}{n}$ | Connected, |
| $(1+\varepsilon)(\log n)n^{-1}$ | has hamiltonian cycles ! |

**Proof.** For $n \geq r \geq 2$,
$P(\alpha \geq r) \leq \binom{n}{r} \cdot (1-p)^{\binom{n}{r}} = \binom{n}{r} q^{\frac{r(r-1)}{2}} \leq n^r q^{\frac{r(r-1)}{2}} = (nq^{\binom{r-1}{2}})^r$
Since $1 - x \leq e^{-x}$ for all positive integers $x$,

$$
\begin{aligned}
(nq^{\frac{r-1}{2}})^r &\leq (n(e^{-p})^{\frac{r-1}{2}})^r \\
&= (ne^{-p(r-1)/2})^r \\
&= (ne^{-pr/2+p/2})^r \\
&\leq (n \cdot e^{-\frac{3}{2}lnn+p/2})^r \\
&\leq n \cdot n^{-\frac{3}{2}} \cdot e^{p/2} \\
&\leq n^{-\frac{1}{2}} \cdot e^{1/2} \rightarrow_{n\rightarrow\infty} 0.
\end{aligned}
$$

$$
\begin{aligned}
\therefore lim_{n\rightarrow\infty} P(\alpha \geq \frac{1}{2}n/k) &= lim_{n\rightarrow\infty} P(\alpha \geq r) wherer =_{def} \lceil \frac{1}{2}n/k \rceil \\
&= 0
\end{aligned}
$$

□

**Theorem 7.2.11** For every integer $k \geq 3$, there exists a triangle-free graph $G$ such that $\mathcal{X}(G) = k$.

**Theorem7.2.12** (*Erdös*, 1959)

For every integer $k$ there exists a graph $H$ with $g(H) > k$ and $\mathcal{X}(H) > k$.

**Proof.** Assume that $k \geq 3$, fix $\varepsilon$ with $0 < \varepsilon < \frac{1}{k}$, and let $p =_{def} n^{\varepsilon-1}$. Note that $n^{\varepsilon-1} = n^{\varepsilon} \cdot n^{-1} \geq (6klnn)n^{-1}$ for large $m$.
( $n^{\frac{1}{10}}/6klnn \approx n^{\frac{1}{10}}/lnn \rightarrow_{n\rightarrow\infty} \frac{\frac{1}{10}\cdot n^{-\frac{9}{10}}}{\frac{1}{n}} \rightarrow_{n\rightarrow\infty} \frac{n^{\frac{1}{10}}}{10} \rightarrow \infty$.)

Let the random variable $X(G)$ denote the number of short cycles in a random graph $G^p$, i.e., its number of cycles of length at most $k$. Then $E(X) = \Sigma_{k'=3}^{k} \frac{n^{k'}}{2k'} p^{k'}$, this is by the fact that each $k'$-cycle has expectation value $1 \cdot p^{k'}$ and there are ($\frac{n^{k'}}{k'} \leq \binom{n}{k'} \leq n^{k'}$) $\frac{n^{k'}}{2k'}$ cycles.

Hence,
$$\begin{aligned} E(X) &= \frac{1}{2}\Sigma_{k'=3}^{k} \frac{n^{k'}}{k'} p^{k'} \\ &\leq \frac{1}{2}\Sigma_{i=3}^{k} n^i p^i \\ &= \frac{1}{2}\Sigma_{i=3}^{k} (np)^i, n \cdot n^{\varepsilon-1} = n^{\varepsilon} \geq 1. \\ &\leq \frac{1}{2}\Sigma_{i=3}^{k} (np)^k \leq \frac{1}{2}(k-2)n^k p^k. \end{aligned}$$

Now, consider
$$\begin{aligned} p(X \geq \frac{n}{2}) &\leq \mathbb{E}(X)/\frac{n}{2} \\ &\leq (k-2)n^{k-1}p^k \\ &= (k-2)n^{k-1}n^{(\varepsilon-1)k} \\ &= (k-2)n^{k\varepsilon-1} \rightarrow_{n\rightarrow\infty} 0 \end{aligned}$$

(By assumption $k\varepsilon < 1$.)
This implies taht $lim_{n\rightarrow\infty} P(X \geq \frac{n}{2}) = 0$.
Let $n$ be large enough that $P(X \geq \frac{n}{2}) < \frac{1}{2}$ and $P(\alpha \geq \frac{n}{2k}) < \frac{1}{2}$. (By Lemma 10).

Then, there exists a graph $G$ of order $n$ with fewer than $\frac{n}{2}$ short cycles and $\alpha(G) < \frac{n}{2k}$. Now, from each cycle we delete one vertex and obtain a new graph $H$. Then $g(H) > k$, moreover, $|H| \geq \frac{n}{2}$. Now, $\mathcal{X}(H) \geq \frac{|H|}{\alpha(H)} \geq \frac{\frac{n}{2}}{\alpha(G)} > k$. We conclude the proof. □

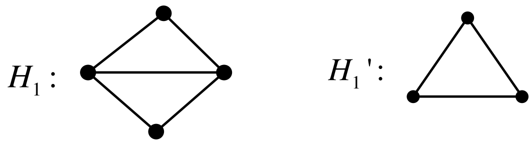
§3. **Threshold Functions**

Review

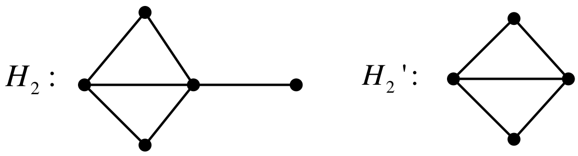
$$d(G) =_{def} \frac{1}{|V(G)|}\Sigma_{v\in V(G)} deg_G(v)$$
$$\varepsilon(G) =_{def} \frac{1}{2} d(G).$$

**Definition7.3.1** A graph $H$ is called balanced if $\varepsilon(H') \leq \varepsilon(H)$ for all subgraphs $H'$ of $H$. Clearly, a regular graph is balanced.

So are forest ?




$\varepsilon(H) = \frac{1}{2}\frac{10}{4} = \frac{5}{4}$, $\varepsilon(H') = \frac{1}{2}\frac{6}{3} = 1.$




$\varepsilon(H) = \frac{1}{2}\frac{12}{5} = \frac{6}{5}$, $\varepsilon(H') = \frac{5}{4}$.

$H_1$ is balanced but not $H_2$.

**Definition 7.3.2** We call a real function for a graph property $P$ if the following holds for all $p = p(n)$:

$$lim_{n\to\infty} P(G \in P) = \begin{cases} 0 & \text{if } p(n)/t(n) \to 0 \text{ as } n \to \infty \\ 1 & \text{if } p(n)/t(n) \to \infty \text{ as } n \to \infty \end{cases}$$

Intutively, if the order of $p(n)$ is less than the order of $t(n)$, then almost all graphs do not have the property $P$; and on the other hand, if $o(p(n)) \geq o(t(n))$, then almost all graphs do have the property $P$.

e.g.

(1) $P$: containing a cycle, Threshold function $t(n) = n^{-1}$.

(2) $P$: containing a copy of a tree $T$ of order $k$, Threshold function $t(n) = n^{-k/(k-1)}$.

(3) $P$: containing $K_{k'}$, Threshold function $t(n) = n^{-2/(k-1)}$.

All the three properties mentioned above are "Monotone properties".

**More concepts of probability**

$X$: Random variable.

$\mathbb{E}(X)$: Expectation of $X$, also denoted by $\mu := \mathbb{E}(X)$.

$\sigma^2 = \mathbb{E}((X-\mu)^2)$: variance of $X$.

$\sigma$: Standard deviation.

$\mathbb{E}(X^r)$: $r$-th moment of $X$.

**Lemma 7.3.1**(Second Moment Method)

Let $X$ be a random variable, then $P(X=0) \leq \frac{\mathbb{E}(X^2)-\mathbb{E}(X)^2}{\mathbb{E}(X)^2}$; Moreover, if $\frac{\mathbb{E}(X^2)}{\mathbb{E}(X)^2} \to 1$, then $P(X=0) \to 0$.

**Proof.** Consider $[X-\mathbb{E}(X)]^2$ and value $t^2$, $t>0$, then $P([X-\mathbb{E}(X)]^2 \geq t^2) = P(|X-\mathbb{E}(X)| \geq t) \leq_{(\text{Markov's Inequality})} \frac{\mathbb{E}([X-\mathbb{E}(X)]^2)}{t^2} = \frac{\sigma^2}{t^2}$. Now,

$$\begin{aligned}\frac{\sigma^2}{t^2} &= \frac{\mathbb{E}(X^2 - 2X\mathbb{E}(X) + \mathbb{E}(X)^2)}{t^2} \\ &= \frac{\mathbb{E}(X^2) - 2\mathbb{E}(X)\mathbb{E}(X) + \mathbb{E}(X)^2}{t^2} \qquad (\mathbb{E}(\mathbb{E}(X)) = \mathbb{E}(X)). \\ &= \frac{\mathbb{E}(X^2) - \mathbb{E}(X)^2}{t^2}.\end{aligned}$$

Hence, $P(|X-\mathbb{E}(X)| \geq t) \leq \frac{\mathbb{E}(X^2)-\mathbb{E}(X)^2}{t^2}$.
Let $t = \mathbb{E}(x)$. Then $P(|X-\mathbb{E}(X)| \geq \mathbb{E}(X)) \geq P(X=0)$. This concludes the proof of the first part, and the second part is easy to see. □

**Theorem 7.3.2** In Model $A$, $lnn/n$ is a threshold probability function for the disappearance of isolated vertices, i.e., to ensure $\delta(G) \geq 1$.

**Proof.** Let $X$ be the $\sharp$ of isolated vertices and $X_1, X_2, ..., X_n$ be Indicator r.v.. Then $\mathbb{E}(X) = n \cdot (1-p)^{n-1}$.

Here $\quad p = p(n)$.

$$(1-p)^n = e^{n(ln(1-p))}$$
$$= e^{n(-p-\frac{p^2}{2}-\frac{p^3}{3}-\frac{p^4}{4}-...)}$$
$$= e^{-np} \cdot e^{-np^2(\frac{1}{2}+\frac{p}{3}+\frac{p^2}{4}+...)}$$
$$\backsim e^{-np} \qquad (np^2 \to 0 \Rightarrow p \in o(\frac{1}{\sqrt{n}}))$$
$$\Rightarrow (1-p)^n \backsim e^{-np}, \quad p = c \cdot \frac{lnn}{n}$$
$$\Rightarrow n(1-p)^n \backsim ne^{-clnn} = n^{1-c}.$$

$c > 1, \mathbb{E}(X) \to 0; c < 1, \mathbb{E}(X) \to \infty$. ($c = 1$ is just the threshold function !)

We claim that $P(X = 0) \to 0$. By Second Moment Method, it suffices to claim $\mathbb{E}(X^2) \backsim \mathbb{E}(X)^2$, i.e., $\frac{\mathbb{E}(X^2)}{\mathbb{E}(X)^2} \to 1$.

$\mathbb{E}(X^2) = \mathbb{E}((\Sigma_{i=1}^n X_i)^2)$

$= \Sigma_{i=1}^n \mathbb{E}(X_i^2) + \Sigma_{i \neq j} \mathbb{E}(X_i Y_j) = \Sigma_{i=1}^n \mathbb{E}(X) (= \mathbb{E}(X)) + n(n-1) \cdot \mathbb{E}(X_i X_j)$

where $n(n-1) = 2 \cdot \binom{n}{2}$ pairs and $\mathbb{E}(X_i X_j) = (1-p)^{2n-3}$

$\approx \mathbb{E}(X) + n(n-1) \cdot (1-p)^{2n-3} \approx \mathbb{E}(X) + (\mathbb{E}(X))^2$

where $\mathbb{E}(X) \to +\infty$ and $n(n-1) \cdot (1-p)^{2n-3} \sim [n \cdot (1-p)^n]^2$.

□

**Theorem 7.3.3** If $H$ is a balanced graph with $k$ vertices and $l$ edges, then $p = n^{-k/l}$ is a threshold function in Model A for the appearance of $H$ as a subgraph of almost every graph $G^p$.

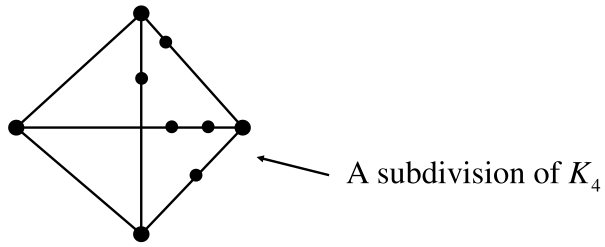
**Proof.** Exercise □

$\varepsilon(H) = \frac{l}{k}$ is also named as the density of $H$.

$\rho(G) = max_{F \subseteq G} \varepsilon(F)$ is called the maximum density.

**Generalized form**

(*) $P = n^{-1/\rho(H)}$ is always a threshold function for the appearance of $H$.




**Definition 7.3.3** Let $\gamma(G)$ denote the largest $r$ such that $G$ contains a subdivision of $K_r$.

**Lemma 7.3.4** For $p = \frac{1}{2}$ and Model A, $mathcalX(G) \geq \frac{n}{2log_2 n}$.

**Proof.** Let $k = \lfloor 2log_2 n \rfloor$. Then
$P(\alpha(G) \geq k) \leq \binom{n}{k}(1-\frac{1}{2})^{\binom{k}{2}} = \binom{n}{k}2^{-\binom{k}{2}} \leq (\frac{en}{k})^k \cdot 2^{-\binom{k}{2}}$
$< (\frac{e\sqrt{2}n}{k2^{k/2}})^k \rightarrow_{n\rightarrow\infty} 0.$ (?)
Since $\mathcal{X}(G)\alpha(G) \geq n$, $\mathcal{X}(G) \geq \frac{n}{2log_2 n}$ (a.s.) □

(Remark here now, $\alpha(G) \leq \lfloor 2log_2 n \rfloor$. )

**Lemma 7.3.5** For $p = \frac{1}{2}$ and Model A, $\gamma(G) \leq \sqrt{6n}$.

**Proof.** Set $r = \lceil \sqrt{6n} \rceil$. Then $n \leq r^2/6$.
There are $\binom{n}{r}$ potential $K_r$ subdivisions, let $X$ induce one. Then $< X >_G$ has **at least** $\binom{r}{2} - (n-r) \geq \binom{r}{2} + r - \frac{r^2}{6} \geq \frac{2}{3}\binom{r}{2}$ edges.
Let $e(X)$ denote the number of edges in $< X >_G$.
The probability for $G$ contains a subdivision of $K_r$ is

$$P(\gamma(G) \geq r) \leq \Sigma_X P(e(X) \geq \frac{2}{3}\binom{r}{2})$$
$$\leq \binom{n}{r} e^{-\frac{1}{18}\binom{r}{2}} \rightarrow_{n\rightarrow\infty} 0$$

(Chernoff Bound $P(\|X\| \geq n(1+\delta)/2) \leq e^{\lceil -n\delta^2/2 \rceil}$, +ake $\delta = \frac{1}{3}$.)
Thus, $\gamma(G) \leq \sqrt{6n}$ almost surely. □

**Theorem 7.3.6** Hajos Conjecture is false, and almost all graps $G^p$, $p = \frac{1}{2}$, are counterexamples.

**Proof.** $\mathcal{X}(G) - \gamma(G) \geq \frac{n}{2log_2 n} - \sqrt{6n} \rightarrow_{n\rightarrow\infty} \infty.$ $\square$