Introduction to Combinatorics — Lecture 15

Probabilistic Method (Graphs)

Definition 15.1 (Random graph with edge probability)**.** $G(n,p)$ or $G(n, P = p)$, where $0 \le p \le 1$. The probability of the existence of an edge (independently) is p and the graph induced by using existent edges is G_p.

Definition 15.2 (Discrete Probabilistic Space, D.P.S.)**.** A D.P.S. is an ordered paired pair (S, f) where S is countable set and $f : S \to \mathbb{R}$ satisfying (i) $0 \le f(x) \le 1$ and
(ii) $\sum_{x \in S} f(x) = 1$.

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

Definition 15.3. Let (S, f) be a D.P.S.. Then the probability of an event $A \subseteq S$ is $P(A) = \sum_{x \in A} f(x)$.

Definition 15.4 (Independent event)**.** If $P(A \cap B) = P(A)P(B)$, then A and B are independent events.

Definition 15.5 (Random variables)**.** Let (S, f) be a D.P.S.. Then $\mathbb{X} : S \to \mathbb{R}$ is a random variable where we use $(\mathbb{X} = k) := \{x \in S \mid \mathbb{X}(x) = k\}$ to denote an event.

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

Definition 15.6 (Expectation)**.** Let $\mathbb{X}$ be a random variable. Then the expectation of $\mathbb{X}$, $\mathbb{E}(\mathbb{X}) = \sum_k k \cdot P(\mathbb{X} = k)$. (We define $P(\mathbb{X} = h) = 0$ if h is not in the image of $\mathbb{X} : S \to \mathbb{R}$.)

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

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

Lemma 15.1 (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 15.2 (Linear Property of Expectation)**.** *Let $X,\ X_1,\ ...,\ X_m$ be random variables such that $X = \sum_{i=1}^{m} X_i$. Then, $\mathbb{E}(X) = \sum_{i=1}^{m} \mathbb{E}(X_i)$.*

Definition 15.7 (Indicator Random Variable)**.** An indicator random variable for the event $A \subseteq S,\ I[A]$, is a random variable $\mathbb{X}$ such that $\mathbb{X} : S \to \{0, 1\}$ (instead of $\mathbb{R}$).

Remark. A random variable $\mathbb{X}$ can be written as a sum of $|S|$ indicator random variables for an event $A \subseteq S$,

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

Here are some examples of probabilistic method.

Theorem 15.3. *If $\binom{n}{k} \cdot 2^{1-\binom{k}{2}} < 1$, then $R(k, k) > n$. Thus,*

$$R(k, k) > \lfloor 2^{\frac{k}{2}} \rfloor,\ \forall\ k \geq 3.$$

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 $\langle T \rangle_{K_n}$ is nomochromatic. 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. Thus, we have $R(k,k) > n$.

Now, if we take $n = \lfloor 2^{\binom{k}{2}} \rfloor$, then

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

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

□

Theorem 15.4 (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 is a directed Hamiltonian path is $\frac{1}{2^{n-1}}$. Therefore, $\mathbb{E}(X) = n! \cdot \frac{1}{2^{n-1}}$. This concludes the proof.

□

Theorem 15.5. $\alpha(G) \geq \sum_{v \in V(G)} \frac{1}{1 + deg_G(v)}$.

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 this bijection φ. Let $v_0 \in S$ (an independent set) if $\varphi(v_0) = \min\{\varphi(x) \mid x \in$

$N(v_0)$ (neighbor of v_0)$\}$. So, the probability is $\dfrac{1}{1+deg_G(v_0)}$ and the expectation value is $\displaystyle\sum_{v\in V(G)} \frac{1}{1+deg_G(v)}$.

□

Theorem 15.6. *If $|G| = n$ and $||G|| = \dfrac{nd}{2}$, $d \geq 1$, then $\alpha(G) \geq \dfrac{n}{2d}$.*

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 $Y = \displaystyle\sum_{e\in E} Y_e$. Now, $\mathbb{E}(Y_e) = P[v_i, v_j \in S] = p^2$ and thus $\mathbb{E}(Y) = \dfrac{nd}{2} \cdot p^2$. Since $\mathbb{E}(X) = np$, $\mathbb{E}(X-Y) = np - \dfrac{nd}{2}p^2 = np(1-\dfrac{d}{2}p)$, $p = \dfrac{1}{d}$ gives the maximum. Hence, $\mathbb{E}(X-Y) = \dfrac{n}{2d}$.

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

□

Definition 15.8. 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 15.9. If $\displaystyle\lim_{n\to\infty} q_n = 1$, then we say "Property Q" almost always holds or in this case, almost all graphs have "Property Q".

Theorem 15.7 (Gilbert, 1959)**.** *Let $0 < p \leq 1$ be a constant. Then, almost all graphs are connected.*

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

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

where x is fixed. Hence,

$$\begin{aligned} 0 \le q_n &\le \sum_{k=1}^{\lfloor \frac{n}{2} \rfloor} n^k \cdot (1-p)^{k(n-k)} \\ &\le \sum_{k=1}^{\lfloor \frac{n}{2} \rfloor} (n(1-p)^{n-k})^k \\ &\le \sum_{k=1}^{\lfloor \frac{n}{2} \rfloor} (n(1-p)^{\frac{n}{2}})^k \\ &< \frac{x}{1-x} \quad \text{where } x = n(1-p)^{\frac{n}{2}}. \end{aligned}$$

But $\lim_{n\to\infty} x = \lim_{n\to\infty} n(1-p)^{\frac{n}{2}} = 0$. This implies that $\lim_{n\to\infty} q_n = 0$.

□

Lemma 15.8 (Markov's Inequality)**.** *Let* $p_k = P(\mathbb{X} = k),\ k \ge 0$. *Then,* $p(\mathbb{X} \ge t) \le \dfrac{\mathbb{E}(\mathbb{X})}{t}$. *Moreover, if* $\mathbb{E}(\mathbb{X}) \to 0$, *then* $P(\mathbb{X} = 0) \to 1$.

Proof.

$$\mathbb{E}(\mathbb{X}) = \sum_{k\ge 0} kp_k \ge \sum_{k \ge t} kp_k \ge t \cdot \sum_{k\ge t} p_k = tP(\mathbb{X} \ge t).$$

□

Theorem 15.9. *Let* $0 < p \le 1$ *be a constant. Then almost all graphs are of diameter* 2.

Proof. Let $\mathbb{X} = \sum_{i \ne j} \mathbb{X}_{i,j}$ where $\mathbb{X}_{i,j}$ is the indicator random variables 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 $P(\mathbb{X}_{i,j} = 1) = (1-p^2)^{n-2}$. Thus, $\mathbb{E}(\mathbb{X}) = \sum_{i\ne j} \mathbb{E}(\mathbb{X}_{i,j}) = \binom{n}{2}(1-p^2)^{n-2}$. Since $\lim_{n\to\infty} \binom{n}{2}(1-p^2)^{n-2} = 0$, $\mathbb{E}(\mathbb{X}) \to 0$. 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.

□

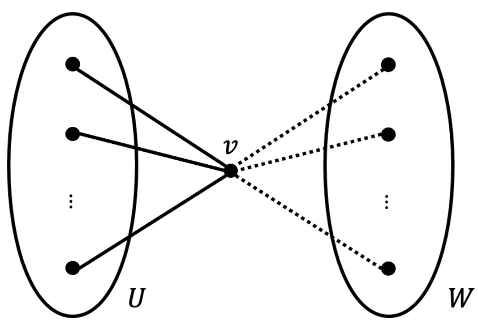
Theorem 15.10. *For every constant $p \in (0, 1)$ and every graph H, almost all graphs G^p contains an induced copy of H.*

Proof. Let H be given and $|H| = k$. Let U be a set of k (fixed) vertices of G. Then, $\langle U \rangle_G \cong H$ 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 $\langle U_i \rangle_G$ is isomorphic to H is $(1-r)^{\lfloor \frac{n}{k} \rfloor}$. Hence, $P[H \not\leq G] \leq (1-r)^{\lfloor \frac{n}{k} \rfloor} \to 0$ as $n \to \infty$.

□

Theorem 15.11. *Let $P_{i,j}$ be the property that for any disjoint vertex sets U and W with $|U| \leq i$ and $|W| \leq j$, there exists at least one vertex $v \notin U \cup W$ that is adjacent to all the vertices of U but to none of the vertices of W. Then, for every constant $p \in (0, 1)$ and $i, j \in \mathbb{N}$, almost all graphs G^p has property $P_{i,j}$.*

Proof. Let $i, j \in \mathbb{N}$ be fixed and $q = 1 - p$. Let U and W be two disjoint vertex sets with $|U| \leq i$ and $|W| \leq j$. The probability that $v \in V(G) \setminus (U \cup W)$ is adjacent to U but not to W is $p^{|U|}q^{|W|} \geq p^i q^j$. Hence, the probability that no suitable v exists for these U and W is $(1 - p^{|U|}q^{|W|})^{n-|U|-|W|} \leq (1 - p^i q^j)^{n-i-j}$. Since the number of $\langle U, W \rangle$ pairs is at most n^{i+j}, the probability that $P_{i,j}$ does not hold is $n^{i+j} \cdot (1-p^i q^j)^{n-i-j} \to 0$ as $n \to \infty$.



□

Corollary 15.12. *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 has property $P_{2,k-1}$, $|G| \geq k + 2$. Let W be an arbitrary set of at most $k - 1$ vertices. Then for all $x, y \in V(G) \setminus W$, either x is adjacent to y or x and y have a common neighbor. ($U = \{x, y\}$) Therefore, W is not a vertex cut. This concludes the proof.

□