2020 FALL MAS583 HW8

Fanchen Bu [20194185]

December 1, 2020

1 Problem 1

1.1 Question

Prove that for every $\epsilon > 0$ there is some $n_0 = n_0(\epsilon)$ so that for every $n > n_0$ there is a graph on n vertices containing every graph on $k \leq (2-\epsilon)\log_2 n$ vertices as an induced subgraph.

1.2 Answer

Suffice it to show that it holds when $k = (2-\epsilon)\log_2 n$. Given n_0, let k_0 be the real such that $k_0 = (2-\epsilon)\log_2 n_0$. To solve the problem, we can equivalently show that for every $\epsilon > 0$, there is some $k_0 = k_0(\epsilon)$ so that for every integer $k > k_0$, there is a graph on $n = n(k) = \lceil 2^{k/(2-\epsilon)} \rceil$ vertices containing every graph on k vertices as an induced subgraph.
First, let I_k be the number of non-isomorphic graphs on k vertices, it's easy to see that

$$\binom{n}{k} \geq I_k \geq 2^{\binom{k}{2}}/k!,$$

which gives

$$n \geq 2^{(k-1)/2}.$$

Therefore, what we want to prove is tight.
WLOG, we assume that k is even and sufficiently large, put $m = k2^{k/2}$, and consider the random graph $G(m, 1/2)$ with vertex set $[m]$. Let H be a fixed graph with vertex set $\{x_1, ..., x_k\}$. Given G on $[m]$, we let $Y_k = Y_k(G)$ be the random variable denoting the number of k-subsets $\{y_1, ..., y_k\}$ of $[m]$ with $y_1 < ... < y_k$, where $x_i \to y_i$ gives an isomorphism between H and the induced subgraph of G on $\{y_1, ..., y_k\}$. Clearly, we have

$$\mathbb{E}[Y_k] \geq \binom{m}{k}/2^{\binom{k}{2}} \geq \frac{1}{2k}(\frac{em}{k})^k/2^{\binom{k}{2}} > e^k,$$

with large k. On the other hand, we let $Y_k^* = Y_k^*(G)$ specifically denote the number of K_k in G, then we have

$$\mathbb{E}(m,k) \coloneqq \mathbb{E}[Y_k^*] = \binom{m}{k}/2^{\binom{k}{2}}.$$

The second moment of Y_k^* is the sum of the probabilities of ordered pairs of K_k. The probability of two K_k with l points in common is $2^{-2\binom{k}{2}+\binom{l}{2}}$, and one can choose $\binom{m}{k}\binom{k}{l}\binom{m-k}{k-l}$ such ordered pairs, which gives the second moment of Y_k^* is

$$\mathbb{E}[(Y_k^*)^2] = \sum_{l=0}^{k} \binom{m}{k}\binom{k}{l}\binom{m-k}{k-l} 2^{-2\binom{k}{2}+\binom{l}{2}},$$

with

$$\mathbb{E}^2[Y_k^*] = \sum_{l=0}^{k} \binom{m}{k}\binom{k}{l}\binom{m-k}{k-l} 2^{-2\binom{k}{2}}$$

we have

$$\mathrm{Var}[Y_k^*] = \sum_{l=2}^{k} \binom{m}{k}\binom{k}{l}\binom{m-k}{k-l} 2^{-2\binom{k}{2}}(2^{\binom{l}{2}}-1),$$

which follows

$$\mathrm{Var}/\mathbb{E}^2 = \sum_{l=2}^{k} \frac{\binom{k}{l}\binom{m-k}{k-l}}{\binom{m}{k}}(2^{\binom{l}{2}-1}) = \sum_{l=2}^{k} F_l.$$

With some routine calculations, we get for large m (which is guaranteed by large k) and $3 \le l \le k-1$,

$$F_l < F_3 + F_{k-1},$$

which gives

$$\begin{aligned}
\mathrm{Var}/\mathbb{E}^2 &< F_2 + F_k + k(F_3 + F_{k-1}) \\
&< \frac{k^4}{2m^2} + \frac{1}{\mathbb{E}(m,k)} + k(\frac{7k^6}{6m^3}) + \frac{km}{2^{m-1}\mathbb{E}(m,k)} \\
&< k^4/m^2 + 2/\mathbb{E}(m,k).
\end{aligned}$$

Now we conclude that

$$\Pr[Y_k = 0] \le \Pr[Y_k^* = 0] < \mathrm{Var}/\mathbb{E}^2 < k^4/m^2 + 2/\mathbb{E}(m,k) < \frac{2k^2}{2^k},$$

for k that is sufficiently large. Now we consider random graph $G = G(n, 1/2)$ with $n = k^2 2^{k/2}$ on vertex set $[n]$ and we partition $[n]$ into k sets of m vertices each. Given a fixed graph H on k vertices, each part on m vertices contains no isomorphism of H as an induced subgraph with probability at most $2k^2/2^k$. As we have k such disjoint parts, the probability that the whole G contains no isomorphism of H as an induced subgraph with probability at most $(2k^2/2^k)^k$, as the number of choices of H is at most $2^{\binom{k}{2}}$, we have the probability that G contains at least one isomorphism of every graph on k vertices is at least

$$1 - 2^{\binom{k}{2}}(2k^2/2^k)^k,$$

which goes to 1 as k goes to infinity. Besides, for any given $\epsilon > 0$, we can find a $k_0 = k_0(\epsilon)$ such that when $k > k_0$, $n = k^2 2^{k/2} \le 2^{k/(2-\epsilon)}$, completing the proof.

1.3 More thoughts

The previous answer definitely overkills the problem as it's proved that with high probability, what we want holds, rather than just with positive probability. Therefore, I want to find a simpler solution in the framework of the Poisson paradigm.
First we consider the case of K_k, similarly with the triangle-freeness problem, we consider $G = G(n, 1/2)$ on vertex set $[n]$, let $H = K_k$, and let $\{A_i, i \in I\}$ be the set of all labeled copies of H on vertex set $[n]$ (i.e., in K_n), and let B_i be the event that $A_i \subset G$, in the sense that the induced subgraph of G on $V(A_i)$ is A_i. Furthermore, let

$$X = \sum_{i \in I} X_i,$$

where X_i is the indicator random variable of B_i. Clearly, X is just the Y_k used in the previous answer, so we know

$$\mu = \mathbb{E}[X] = \binom{n}{k}/2^{\binom{k}{2}}$$

and

$$M = \prod_{i \in I} \Pr[\overline{B_i}] = (1 - 2^{-k})^{\binom{n}{k}}$$

and Δ is just $2\mathbb{E}[(Y_k^*)^2]$, which means

$$\Delta = 2 \sum_{l=0}^{k} \binom{n}{k} \binom{k}{l} \binom{n-k}{k-l} 2^{-2\binom{k}{2}+\binom{l}{2}}.$$

By the extended Janson inequality, we have

$$\Pr[\bigwedge_{i \in I} \overline{B_i}] = \Pr[G(n,p) \text{ contains no } H] \leq e^{-\frac{\mu^2}{2\Delta}},$$

which can be clearly upper bounded by $\exp(-\sum_{l=2}^{k} F_l^{-1})$. And the remaining processes are similar with these in the previous answer. Therefore, we actually get a better concentration with the extended Janson inequality, just as mentioned in the lecture notes about Theorem 8.3.

To make the proof more clear, we need to show that the previous process works for any other graph H on k vertices too. Actually, if K_k contains $a = a(H)$ copies of H (in our previous case, $a = a(K_k) = 1$), then we have

$$\mu = \mathbb{E}[X] = a\binom{n}{k}/2^{\binom{k}{2}}.$$

For Δ, we claim that

$$\Delta \leq a^2 2 \sum_{l=0}^{k} \binom{n}{k} \binom{k}{l} \binom{n-k}{k-l} 2^{-2\binom{k}{2}+\binom{l}{2}},$$

which implies that

$$\Pr[\bigwedge_{i \in I} \overline{B_i}] = \Pr[G(n,p) \text{ contains no } H] \leq e^{-\frac{\mu^2}{2\Delta}},$$

is no worse upper bounded than the case of $H = K_k$. To see this, for any ordered pair of two k-subsets of $[n]$ that intersect, which contributes a term in the Δ for $H = K_k$, now the pair contributes a^2 terms, however, the probability of each is at most the same as the case of K_k (the same or 0 actually) because two related B_i can happen together only if the intersection part matches both cases, which completes the proof.

1.4 References

[1] Bollobás, Béla, and Andrew Thomason. "Graphs which contain all small graphs." European Journal of Combinatorics 2.1 (1981): 13-15.

[2] Bollobás, Béla, and Paul Erdos. "Cliques in random graphs." Mathematical Proceedings of the Cambridge Philosophical Society. Vol. 80. No. 3. Cambridge Univ Press, 1976.

2 Problem 2

2.1 Question

Find a threshold function for the property: $G(n, p)$ contains at least $n/6$ pairwise vertex disjoint triangles.

2.2 Answer

We claim that $p = n^{-2/3}$ is a threshold function for the given property. WLOG, we assume that 6 divides n.

Lemma 2.1 *There exists a positive constant c such that almost surely $G(n, cn^{-2/3})$ contains less than $n/6$ pairwise vertex disjoint triangles.*

Proof. Let $\{A_i : i \in I\}$ denote the set of all distinct labeled copies of K_3 in the complete labeled graph on n vertices. Let B_i be the event that $A_i \subset G(n, p)$, and let X_i be the indicator random variable for B_i. Let $X = \sum_{i \in I} X_i$ be the number of distinct copies of K_3 in $G(n, p)$ (note that here we do not need the copies to be vertex disjoint). Suffice it to show that $X < n/6$ with high probability. It's easy to check that

$$\mathbb{E}[X] = \binom{n}{3} p^3 \leq c^3 n/6,$$

which can be upper bounded by $n/12$ with sufficiently small c, and implies that $\mathbb{E}[X] = \Theta(n)$. For two copies A_i and A_j we say $i \sim j$ if they share one edge, and over all ordered pairs, let

$$\Delta = \sum_{i \sim j} \Pr[B_i \wedge B_j].$$

As $\mathrm{Var}[X] \leq \mathbb{E}[X] + \Delta$, it remains to show that $\Delta = o(\mathbb{E}^2[X])$ before using Chebyshev's Inequality to complete the proof. Actually,

$$\Delta = \binom{n}{5} 5! p^5 \leq c^5 n^{5/3} = o(\mathbb{E}^2[X]),$$

with sufficiently large n, completing the proof. ■

Lemma 2.2 *There exists a positive constant C such that almost surely $G(n, Cn^{-2/3})$ contains at least $n/6$ pairwise vertex disjoint triangles.*

Proof. Consider random graph $G(n/2, Cn^{-2/3})$, with the same notations used in the proof of Lemma 2.1 (except that the number of vertices is $n/2$ now), we have

$$\mu = \mathbb{E}[X] \geq C^3 n/16$$

and

$$\Delta \leq C^5 n^{5/3}/32,$$

with sufficiently large C and n. As $\Delta > \mu$ when C and n are sufficiently large, by the extended Janson inequality, we have

$$\Pr[X = 0] = \Pr[\bigwedge_{i \in I} \overline{B_i}] \leq e^{-\mu^2/2\Delta} \leq e^{-100 n^{1/3}}.$$

Therefore, we conclude that almost surely $G(m, Cn^{-2/3})$ contains a triangle if $m \geq n/2$. From a random graph $G(n,p)$, with probability as least $(1-e^{-100n^{1/3}})^{n/6}$, we can repeatedly do for $i \in [n/6]$,

- Find a triangle in the current graph on the remaining $n-3(i-1)$ vertices
- Remove the 3 vertices and all edges incident to them, now we essentially have $G(n-3i, Cn^{-2/3})$

Finally, we have $n/6$ pairwise vertex disjoint triangles with high probability, completing the proof of both the lemma and the whole problem. ■

2.3 References

Alon, Noga, and Raphael Yuster. "Threshold functions for H-factors." Combinatorics, Probability and Computing 2.2 (1993): 137-144.