Random Graphs — Assignment 2

Yuval Filmus

December 4, 2019

Question 1. Let X_k be the number of k-cliques in $G(n, 1/2)$. Recall that

$$N_k := \mathbb{E}[X_k] = \binom{n}{k} 2^{-\binom{k}{2}}$$

and

$$\frac{\mathbb{E}[X_k^2]}{\mathbb{E}[X_k]^2} = \sum_{\ell=0}^{k} \frac{\binom{k}{\ell}\binom{n-k}{k-\ell}}{\binom{n}{k}} 2^{\binom{\ell}{2}} =: \sum_{\ell=0}^{k} J_\ell.$$

Let k_0 be the maximal k such that $N_k \geq 1$, and recall that $k_0 = 2\log_2 n - 2\log_2 \log_2 n + O(1)$. The goal of this exercise is to show that with high probability, $G(n, 1/2)$ contains a k-clique, for $k = k_0 - 1$.

(a) Calculate $J_{\ell+1}/J_\ell$.

(b) Calculate $(J_{\ell+2}/J_{\ell+1})/(J_{\ell+1}/J_\ell)$.

(c) Show that there exists N_1 such that if $n \geq N_1$ and $0 \leq \ell \leq \frac{1}{2}\log_2 n$ then $J_{\ell+1} < J_\ell$.

(d) Show that there exists N_2 such that if $n \geq N_2$ and $k > \ell \geq \frac{3}{2}\log_2 n$ then $J_{\ell+1} > J_\ell$.

(e) Show that there exists N_3 such that if $n \geq N_3$ then the sequence $J_{\ell+1}/J_\ell$ is increasing for $\frac{1}{2}\log_2 n \leq \ell \leq \frac{3}{2}\log_2 n$.

(f) Deduce that for $N \geq \max(N_1, N_2, N_3)$ the sequence $J_0, \ldots, J_k$ is unimodal (decreasing and then increasing, with perhaps two identical values at the middle).

(g) Deduce that for $N \geq \max(N_1, N_2, N_3)$ the maximum of $J_1, \ldots, J_k$ is attained at one of the endpoints.

(h) Conclude that $\mathbb{E}[X_k^2]/\mathbb{E}[X_k]^2 = 1 + o(1)$, and so with high probability $G(n, 1/2)$ contains a k-clique.

Question 2. In this exercise, we will empirically explore algorithms for finding cliques in $G(n, 1/2)$.

(a) Consider the heuristic which constructs a clique by iteratively picking an arbitrary vertex which is connected to all vertices chosen so far. How large a clique does this heuristic find, when $n = 1000$?

(b) Modify this heuristic by always picking a vertex of maximal degree. Does this improve on the original heuristic when $n = 1000$?

(c) Optional: Come up with a better heuristic.

In both cases, I suggest performing at least 10^4 experiments. In each experiment, generate a $G(n, 1/2)$ random graph, run the heuristic, and record the result. After running all experiments, report the average, using two significant digits.

Due to speed considerations, I recommend using a compiled language like C/C++/Java rather than an interepreted language such as Python/Matlab.

Question 3. Let $p_{n,k}$ be the probability that the heuristic in Question 2(a) constructs a clique of size k when run on $G(n, 1/2)$.

(a) Write a recurrence relation for $p_{n,k}$.

(b) Compute the expected size of the clique when $n = 1000$ exactly (but display the result as a decimal). You can use a computer algebra system such as Mathematica, Maple or Sage, or a library such as `libgmp`, which supports multi-precision integer or floating point arithmetic.

Question 4 (Bonus)**.** Let $H := \rhd\!\!-$. In the last assignment we calculated the probability that $G(n, c/n)$ contains no copy of H. Now we calculate the entire distribution:

$$p_k := \lim_{n\to\infty} \Pr[G(n, c/n) \text{ contains exactly } k \text{ copies of } H].$$

Say that a triangle has *type* $t = (a, b, c)$ if $0 \le a \le b \le c$ and the degrees of vertices in the triangle are $2 + a, 2 + b, 2 + c$ (that is, the triangle has a edges dangling from one vertex, b edges from another, and c from the remaining vertex). Let $|t| = a + b + c$.

Fix an arbitrary ordering on types. For a vector $\tau = (t_1, \ldots, t_m)$ of non-decreasing types, define

$$q_\tau = \lim_{n\to\infty} \Pr[G(n, c/n) \text{ contains exactly } m \text{ triangles, of types } t_1, \ldots, t_m].$$

Let $|\tau| = \sum_{i=1}^m |t_i|$.

(a) Show that with high probability, any two copies of H in $G(n, c/n)$ are either disjoint or share a triangle.

(b) Show that

$$p_k = \sum_{|\tau|=k} q_\tau.$$

(Hint: use part (a) and modify the argument of Question 3 in Assignment 1.)

(c) Calculate q_τ.

(d) Let τ be a vector not including any isolated triangles. Calculate z_τ, which is the sum of $q_{\tau'}$ over all vectors τ' obtained from τ by adding an arbitrary number of isolated triangles.

(e) Calculate p_0, p_1, p_2.