MCS 549 – Foundations of Data Science
Fall 2024
Problem Set 2

Lev Reyzin

Due: 11/1/24 at the beginning of class

Instructions: Atop your problem set, please write your name and list your collaborators.

Problems

1. Find the threshold for $p(n)$ for the existence of 4-cliques in $G(n, p(n))$. Prove your answer correct.

2. The example at the end of Section 8.1.1 in the book computes that if the degrees in $G(n, \frac{1}{n})$ were independent, there would be a vertex of degree

$$d = \Omega \left(\frac{\log n}{\log \log n} \right)$$

with constant positive probability. However, the degrees are not independent. Show how to overcome this difficulty and reach the same conclusion.

3. Show that in $G(n, 1/2)$ there are almost surely are no cliques of size greater than or equal to $2 \log_2 n$. Then, use the second moment method to show that in $G(n, 1/2)$, almost surely there are cliques of size $(2 - \varepsilon) \log_2 n$ (for any constant $\varepsilon > 0$).

4. What is the expected number of simple paths of length 3, $\log n$, $\sqrt{n}$, and $n - 1$ in $G(n, \frac{d}{n})$?