Practical session 6

Exercise 1 For two vertices u and v in a graph G, we denote by $\operatorname{dist}(u,v)$ the length of a shortest path connecting them. The diameter of a graph is defined to be $\operatorname{diam}(G) = \max_{u,v \in V(G)} \operatorname{dist}(u,v)$. Let $G \sim G(n,p)$ for some n and p. Prove the following two claims.

1. If $p \geq \sqrt{\frac{3 \ln n}{n}}$, then $\lim_{n \to \infty} \Pr(\operatorname{diam}(G(n,p)) \leq 2) = 1$.

2. If $p \leq \sqrt{\frac{\ln n}{10n}}$, then $\lim_{n \to \infty} \Pr(\operatorname{diam}(G(n,p)) \leq 2) = 0$.

Solution

For every $1 \leq i < j \leq n$ let $X_{i,j}$ be the indicator random variable for the event "$\operatorname{dist}(i,j) > 2$". By the independence of the appearance of edges in $G(n,p)$, the probability that i and j have no common neighbours is exactly $(1-p^2)^{n-2}$. Indeed, the probability of any given vertex in $[n] \setminus \{i,j\}$ to be connected to at most one of i and j is $1-p^2$. Therefore $\Pr(X_{i,j} = 1) = (1-p) \cdot (1-p^2)^{n-2}$. Let X denote the number of pairs of vertices $1 \leq i < j \leq n$ that are not connected by an edge of G and have no common neighbours, that is, $ij \notin E(G)$ and $N_G(i) \cap N_G(j) = \emptyset$. Then $X = \sum_{1 \leq i < j \leq n} X_{i,j}$.

1. By monotonicity, we may assume that $p = \sqrt{\frac{3 \ln n}{n}}$. Since $\mathbb{E}(X_{i,j}) = (1-p) \cdot (1-p^2)^{n-2}$ for every $1 \leq i < j \leq n$, it follows by the linearity of expectation that

$$\begin{aligned}
\mathbb{E}(X) &= \sum_{1 \leq i < j \leq n} \mathbb{E}(X_{i,j}) \\
&= \binom{n}{2} \cdot (1-p) \cdot (1-p^2)^{n-2} \\
&\leq n^2 \cdot e^{-p} \cdot e^{-p^2(n-2)} \\
&= n^2 \cdot e^{-p^2 n} \cdot e^{2p^2 - p} \\
&\leq n^2 \cdot e^{-3 \ln n} \\
&= n^2 \cdot n^{-3} \\
&= o(1),
\end{aligned}$$

where the first inequality is due to the fact that $1 + x \leq e^x$ for every $x \in \mathbb{R}$, and the second inequality holds since $2p^2 - p \leq 0$ for sufficiently large n. It thus follows by Markov's inequality that

$$\Pr(\operatorname{diam}(G) > 2) = \Pr(X \geq 1) \leq \mathbb{E}(X) = o(1).$$

2. By monotonicity, we may assume that $p = \sqrt{\frac{\ln n}{10n}}$. Similarly to the previous part of this exercise, it follows by the linearity of expectation that

$$\begin{aligned}
\mathbb{E}(X) &= \sum_{1\le i<j\le n} \mathbb{E}(X_{i,j}) \\
&= \binom{n}{2} \cdot (1-p) \cdot (1-p^2)^{n-2} \\
&\ge \frac{n^2}{3} \cdot \frac{1}{2} e^{-p} \cdot \frac{1}{2} e^{-p^2(n-2)} \\
&= \frac{n^2}{12} \cdot e^{-p^2 n} \cdot e^{2p^2 - p} \\
&\ge \frac{n^2}{13} \cdot e^{-\ln n/10} \\
&= \frac{1}{13} \cdot e^{2\ln n - \ln n/10} \\
&= \Omega(n^{1.9}),
\end{aligned}$$

where the first inequality is due to the fact that $1 + x \ge e^x/2$ for every $-1/2 \le x \le 1/2$, and the second inequality holds since $\lim_{n\to\infty} 2p^2 - p = 0$ and e^x is a continuous function. To see as to why $1 + x \ge e^x/2$, observe that by the Taylor expansion of e^x we have that

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots \le 1 + x + \frac{x^2}{2} + \frac{x^2}{2^2} + \frac{x^2}{2^3} + \ldots = 1 + x + x^2 \le 2(1+x).$$

We next show that $\mathrm{Var}(X) = o\left((\mathbb{E}(X))^2\right)$; it will then follow by the second moment method that $X \ge 1$ with high probability. It holds that

$$\mathrm{Var}(X) = \sum_{1\le i<j\le n} \mathrm{Var}(X_{i,j}) + \sum_{\substack{\{i,j\}\ne\{i',j'\} \\ i<j, i'<j'}} \mathrm{Cov}\left(X_{i,j}, X_{i',j'}\right). \tag{1}$$

Since the $X_{i,j}$'s are indicators, it follows that

$$\mathrm{Var}(X_{i,j}) = \mathbb{E}\left(X_{i,j}^2\right) - (\mathbb{E}(X_{i,j}))^2 \le \mathbb{E}\left(X_{i,j}^2\right) = \mathbb{E}(X_{i,j}). \tag{2}$$

We next analyze $\mathrm{Cov}\left(X_{i,j}, X_{i,j'}\right)$. Assume first that $\{i,j\} \cap \{i',j'\} \ne \emptyset$; without loss of generality we assume that $i = i'$. In this case we will use the following trivial bound

$$\begin{aligned}
\mathrm{Cov}\left(X_{i,j}, X_{i,j'}\right) &= \mathbb{E}\left(X_{i,j} \cdot X_{i,j'}\right) - \mathbb{E}(X_{i,j}) \cdot \mathbb{E}\left(X_{i,j'}\right) \le \mathbb{E}\left(X_{i,j} \cdot X_{i,j'}\right) \\
&= \Pr\left(X_{i,j} \cdot X_{i,j'} = 1\right) \le 1.
\end{aligned} \tag{3}$$

Now, assume that $\{i,j\} \cap \{i',j'\} = \emptyset$. Observe that $X_{i,j} \cdot X_{i',j'} = 1$ if and only if $\mathrm{dist}(i,j) > 2$ and $\mathrm{dist}(i',j') > 2$, that is, if and only if all of the following four events occur.

(a) $ij \notin E(G)$ and $i'j' \notin E(G)$;

(b) $|E(G) \cap \{ii', ij'\}| \le 1$, $|E(G) \cap \{ji', jj'\}| \le 1$, $|E(G) \cap \{ii', ji'\}| \le 1$ and $|E(G) \cap \{jj', ij'\}| \le 1$;

(c) $ik \notin E(G)$ or $jk \notin E(G)$ for every $k \in [n] \setminus \{i, i', j, j'\}$;

(d) $i'k \notin E(G)$ or $j'k \notin E(G)$ for every $k \in [n] \setminus \{i, i', j, j'\}$.

Therefore, for sufficiently large n we have

$$\begin{aligned}\mathbb{E}\left(X_{i,j} \cdot X_{i',j'}\right) &= (1-p)^2 \cdot \left[(1-p)^4 + 4p(1-p)^3 + 2p^2(1-p)^2\right] \cdot \left[(1-p^2)^{n-4}\right]^2 \\ &\le (1-p)^6(1-p^2)^{2(n-4)} + 5p.\end{aligned}$$

It follows that

$$\begin{aligned}\mathrm{Cov}\left(X_{i,j}, X_{i',j'}\right) &= \mathbb{E}\left(X_{i,j} \cdot X_{i',j'}\right) - \mathbb{E}\left(X_{i,j}\right) \cdot \mathbb{E}\left(X_{i',j'}\right) \\ &\le (1-p)^6(1-p^2)^{2(n-4)} + 5p - (1-p)^2(1-p^2)^{2(n-2)} \le 5p. \end{aligned} \tag{4}$$

Combining Equations (1) to (4) shows that

$$\begin{aligned}\mathrm{Var}\left(X\right) &= \sum_{1 \le i < j \le n} \mathrm{Var}\left(X_{i,j}\right) + \sum_{\substack{\{i,j\} \neq \{i',j'\} \\ i<j, i'<j'}} \mathrm{Cov}\left(X_{i,j}, X_{i',j'}\right) \\ &\le \sum_{1 \le i < j \le n} \mathbb{E}\left(X_{i,j}\right) + n^3 \cdot 1 + n^4 \cdot 5p \\ &\le \mathbb{E}\left(X\right) + n^3 + n^{3.6} \\ &\le 2n^{3.6} \\ &= o\left(\left(\mathbb{E}\left(X\right)\right)^2\right).\end{aligned}$$

Using the second moment method we conclude that

$$\lim_{n \to \infty} \Pr\left(\mathrm{diam}\left(G\right) \le 2\right) = \lim_{n \to \infty} \Pr\left(X = 0\right) = 0$$

as claimed.