Lecture 31: The average degree distribution of preferential attachment graphs

We have shown the first part of Lemma 29.3, that the degree distribution is concentrated around its mean, but we need to show the second part, that the mean converges. Recall the notation and definitions from the last two lectures: Let $(G_n)_{n\geq 1}$ be a sequence of preferential attachment graphs, for $1 \leq i \leq n$, let $D_n(i)$ be the degree of vertex i in G_n, and set

$$P_n(k) := \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}[D_n(i) = k],$$

and

$$N_n(k) := \sum_{i=1}^{n} \mathbb{I}[D_n(i) = k];$$

the proportion and number of vertices with degree k in G_n. The limiting degree distribution is given by $p_k := 4/(k(k+1)(k+2))$.

Proof of Part 2 of Lemma 29.3. We want to show that there is a constant $c > 0$ such that for all $k \geq 1$,

$$\big|\mathbb{E}\big[P_n(k)\big] - p_k\big| \leq \frac{c}{n}.$$

which is the same as

$$\big|\mathbb{E}\big[N_n(k)\big] - np_k\big| \leq c. \tag{31.1}$$

We derive a recursion for $\mathbb{E}\big[N_n(k)\big]$ in n and k by conditioning on G_{n-1}. We initialise the recursion at $n = 2$, where $N_2(1) = 2$ and $N_2(k) = 0$ for $k \neq 2$. For $n \geq 3$, To obtain G_n from G_{n-1}, we add vertex n, which has degree one for its outgoing edge, and then attach an edge from vertex n to a randomly chosen vertex, where the probability the edge is attached to vertex i is $D_{n-1}(i)/(2(n-2))$. Thus we always add a vertex with degree one, and then increase the degree of another vertex by one. If that vertex has degree $k-1$ in G_{n-1}, then we gain a vertex of degree k in G_n, and if it has degree k in G_{n-1}, then we lose a vertex of degree k in G_n. Since the probability of attaching to a vertex with degree j in G_{n-1} is $j/(2(n-2))$, we have for $k \geq 1$,

$$\mathbb{E}\big[N_n(k) - N_{n-1}(k)|G_{n-1}\big] = N_{n-1}(k-1)\frac{k-1}{2(n-2)} - N_{n-1}(k)\frac{k}{2(n-2)} + \mathbb{I}[k=1],$$

and note $N_{n-1}(0) = 0$. Writing $m_n(k) := \mathbb{E}[N_n(k)]$ and taking expectation in the previous equation, we have

$$m_n(k) = m_{n-1}(k)\left(1 - \frac{k}{2(n-2)}\right) + m_{n-1}(k-1)\frac{k-1}{2(n-2)} + \mathbb{I}[k=1]. \tag{31.2}$$

Since we want to show $m_n(k) - np_k$ is bounded, we should try to find a matching recursion for np_k. Some experimenting and easy algebra using the exact formula $p_k = 4/(k(k+1)(k+2))$ leads to (note that this is how you would guess the formula for p_k from (31.2))

$$\begin{aligned} np_k &= (n-1)p_k\left(1 - \frac{k}{2(n-1)}\right) + (n-1)p_{k-1}\frac{k-1}{2(n-1)} + \mathbb{I}[k=1] \\ &= (n-1)p_k\left(1 - \frac{k}{2(n-2)}\right) + (n-1)p_{k-1}\frac{k-1}{2(n-2)} + \mathbb{I}[k=1] + \frac{kp_k - (k-1)p_{k-1}}{2(n-2)}. \end{aligned}$$

Subtracting this equation from (31.2), and setting

$$\begin{aligned} \varepsilon_n(k) &:= m_n(k) - np_k, \\ C(k) &:= kp_k - (k-1)p_{k-1}, \end{aligned}$$

we find

$$\varepsilon_n(k) = \varepsilon_{n-1}(k)\left(1 - \frac{k}{2(n-2)}\right) + \varepsilon_{n-1}(k-1)\frac{k-1}{2(n-2)} + \frac{C(k)}{2(n-2)}, \tag{31.3}$$

and we show that $|\varepsilon_n(k)| \le 2/3$. We perform an induction on n on the statement "$|\varepsilon_n(k)| \le 1$ for all $k \ge 1$". First note that

$$\varepsilon_2(k) = \begin{cases} 2 - 2p_1, & k = 1, \\ 2p_k, & k > 1. \end{cases} \tag{31.4}$$

In particular, for all $k \ge 1$, we have

$$|\varepsilon_2(k)| \le 2/3, \tag{31.5}$$

which is our base case. To perform the induction step on n, assume that $|\varepsilon_{n-1}(k)| \le 2/3$ for all $k \ge 1$. Then (31.3), the triangle inequality, and the induction hypothesis together imply

$$\begin{aligned} |\varepsilon_n(k)| &\le |\varepsilon_{n-1}(k)|\left(1 - \frac{k}{2(n-2)}\right) + |\varepsilon_{n-1}(k-1)|\frac{k-1}{2(n-2)} + \frac{|C(k)|}{2(n-2)} \\ &\le \frac{2}{3}\left(1 - \frac{k}{2(n-2)} + \frac{k-1}{2(n-2)}\right) + \frac{|kp_k - (k-1)p_{k-1}|}{2(n-2)} \\ &\le \frac{2}{3}\left(1 - \frac{1}{2(n-2)}\right) + \frac{1}{6(n-2)} \\ &= \frac{2}{3} - \frac{1}{6(n-2)} \le \frac{2}{3}, \end{aligned} \tag{31.6}$$

where we have used that

$$|kp_k - (k-1)p_{k-1}| = \frac{8}{k(k+1)(k+2)} \le \frac{1}{3}. \qquad \square$$

We have proved the theorem, but in preparation for our next result, note that we can actually sharpen this estimate to the following lemma.

Lemma 31.1. *Recalling the notation above, for any $\alpha \in (0,1)$, $n \ge 2$, and $k \ge 1$,*

$$\left|\mathbb{E}\big[N_n(k)\big] - np_k\right| \le \left(\frac{2}{1-\alpha}\right)k^{-\alpha}.$$

Proof. Write $C_\alpha := 2/(1-\alpha) > 2$. The case where $k = 1$ is already covered by the previous proof (the constant $c = 2/3 \le 2 < C_\alpha$). For $n = 2$ and $k \ge 2$, (31.4) implies, using the notation there, that

$$\varepsilon_2(k) = 2p_k = 8/(k(k+1)(k+2)) \le 2k^{-\alpha} \le C_\alpha k^{-\alpha}.$$

Following the previous proof, for the general induction step, just as in (31.6), we have

$$\begin{aligned} |\varepsilon_n(k)| &\le |\varepsilon_{n-1}(k)|\left(1 - \frac{k}{2(n-2)}\right) + |\varepsilon_{n-1}(k-1)|\frac{k-1}{2(n-2)} + \frac{|kp_k - (k-1)p_{k-1}|}{2(n-2)} \\ &\le C_\alpha k^{-\alpha}\left(1 - \frac{k}{2(n-2)}\right) + C_\alpha\frac{(k-1)^{1-\alpha}}{2(n-2)} + \frac{|kp_k - (k-1)p_{k-1}|}{2(n-2)} \\ &= C_\alpha k^{-\alpha} - \frac{C_\alpha\big(k^{1-\alpha} - (k-1)^{1-\alpha}\big) - |kp_k - (k-1)p_{k-1}|}{2(n-2)}. \end{aligned}$$

Since $|kp_k - (k-1)p_{k-1}| = \frac{8}{k(k+1)(k+2)}$, we need to show that for $k \ge 2$,

$$C_\alpha\big(k^{1-\alpha} - (k-1)^{1-\alpha}\big) \ge \frac{8}{k(k+1)(k+2)}. \tag{31.7}$$

By the concavity of $x \mapsto x^{1-\alpha}$, we have

$$k^{1-\alpha} - (k-1)^{1-\alpha} \geq (1-\alpha)k^{-\alpha}$$

and so (31.7) is implied by

$$2k^{-\alpha} \geq 8k^{-3},$$

which holds since $k^{3-\alpha} \geq 2^{3-\alpha} > 4$. □