Lecture 9: Concentration of Erdős-Rényi degree distribution

Fix $n \in \mathbb{N}$, $\lambda > 0$, set $p = p_n := \lambda/n$, and let $G = G(n,p) \sim \mathrm{ER}(n,p)$. Let D_i be the degree of vertex i in G, and

$$P_n^{(k)} := \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}[D_i = k]. \tag{9.1}$$

We refer to $\big(P_n^{(k)}\big)_{k=0}^{n-1}$ as the *empirical degree distribution* of G. The following result gives the asymptotic behaviour of this statistic.

Proposition 9.1. *As $n \to \infty$, if $n\,\varepsilon_n^2 \to \infty$, then*

$$\mathbb{P}\left(\max_{k\geq 0}\Big|P_n^{(k)} - e^{-\lambda}\frac{\lambda^k}{k!}\Big| \geq \varepsilon_n\right) \to 0.$$

Remark 9.2. The D_i's are all distributed as $\mathrm{Bi}(n-1,p) \approx \mathrm{Po}(\lambda)$. If they were independent (they are not), then $\big(P_n^{(k)}\big)_{k=0}^{n-1}$ would be the *empirical distribution* of i.i.d. copies of $\mathscr{L}(D_1)$, and the statement of the proposition would continue to hold (and easily follow along the same lines as the proof below). The proposition says that the degrees are *almost* independent, as quantified by the same convergence of the empirical degree distribution.

Proof of Proposition 9.1. There are two steps to the proof.

1. The $P_n^{(k)}$ are quantifiably "concentrated" around there means
$$\mu_n^{(k)} := \mathbb{E}P_n^{(k)} = \binom{n-1}{k}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k}$$
on the scale of $\varepsilon_n \gg n^{-1/2}$; note that the right hand side is the order of the standard deviation of $P_n^{(k)}$.
2. The binomial probabilities $\mu_n^{(k)}$ are quantifiably close to the Poisson probabilities $e^{-\lambda}\frac{\lambda^k}{k!}$.

To leverage these observations, we first note using the triangle inequality and a union bound,

$$\begin{aligned}\mathbb{P}&\left(\max_{k\geq 0}\Big|P_n^{(k)} - e^{-\lambda}\frac{\lambda^k}{k!}\Big| \geq \varepsilon_n\right)\\ &\leq \mathbb{P}\left(\max_{k\geq 0}\Big|P_n^{(k)} - \mu_n^{(k)}\Big| \geq \varepsilon_n/2\right) + \mathbb{I}\left(\max_{k\geq 0}\Big|\mu_n^{(k)} - e^{-\lambda}\frac{\lambda^k}{k!}\Big| \geq \varepsilon_n/2\right).\end{aligned} \tag{9.2}$$

The first term on the right hand side corresponds to Item 1, which we bound as

$$\begin{aligned}\mathbb{P}\left(\max_{k\geq 0}\Big|P_n^{(k)} - \mu_n^{(k)}\Big| \geq \varepsilon_n/2\right) &= \mathbb{P}\left(\bigcup_{k\geq 0}\Big|P_n^{(k)} - \mu_n^{(k)}\Big| \geq \varepsilon_n/2\right)\\ &\leq \sum_{k\geq 0}\mathbb{P}\left(\Big|P_n^{(k)} - \mu_n^{(k)}\Big| \geq \varepsilon_n/2\right)\\ &\leq \frac{4}{\varepsilon_n^2}\sum_{k\geq 0}\mathrm{Var}(P_n^{(k)}),\end{aligned} \tag{9.3}$$

where we have used a union bound and Markov's inequality. To bound the variance, use the expression (9.1), the usual formula for the variance of the sum, and that the (D_i, D_j) has the same distribution for any $i \neq j$ to find

$$\mathrm{Var}(P_n^{(k)}) = \frac{1}{n}\mathbb{P}(D_1 = k)(1 - \mathbb{P}(D_1 = k)) + \frac{n-1}{n}\big(\mathbb{P}(D_1 = k, D_2 = k) - \mathbb{P}(D_1 = k)^2\big)$$

$$\leq \frac{1}{n}\mathbb{P}(D_1 = k) + \big(\mathbb{P}(D_1 = k, D_2 = k) - \mathbb{P}(D_1 = k)^2\big).$$

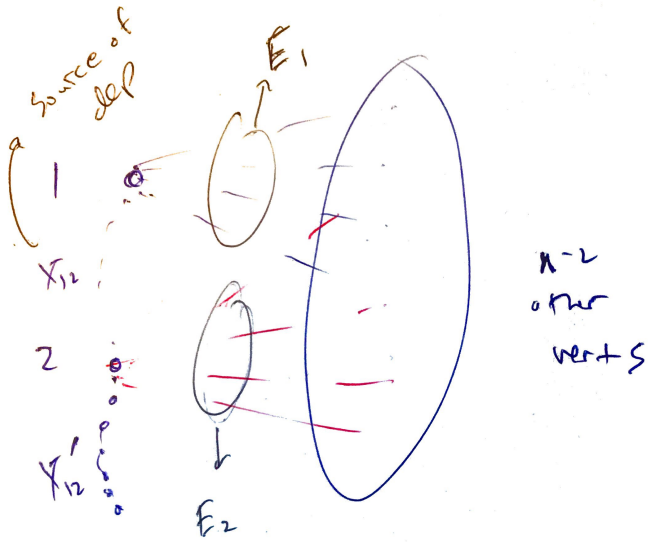
We need both these terms to be order $1/n$ after summing in k; the first obviously satisfies this. For the second, we use a coupling. The idea is that D_1 and D_2 only depend on each other through the edge connecting them, so they are *almost* independent. Denoting the indicator that there is an edge between vertices i and j by X_{ij}, we have

$$D_1 = X_{12} + \sum_{j=3}^{n} X_{1j} =: X_{12} + E_1, \quad D_2 = X_{12} + \sum_{j=3}^{n} X_{2j} =: X_{12} + E_2,$$

and, letting $X'_{12} \sim \mathrm{Be}(p)$ be independent of all variables already defined, define

$$D'_2 := X'_{12} + \sum_{j=3}^{n} X_{2j} = X'_{12} + E_2.$$

The key properties of this construction is that $E_1, E_2, X_{12}, X'_{12}$ are independent, and thus so are D_1 and D'_2. The picture may help:



Now, using that D'_2 is independent of D_1 in the first equality below, that $X_{12}, X'_{12}, E_1, E_2$ are independent in the third and fourth equality, and that that $E_1 \stackrel{d}{=} E_2$ in the fourth, we have

$$\begin{aligned}
&\mathbb{P}(D_1 = k, D_2 = k) - \mathbb{P}(D_1 = k)^2 \\
&\quad = \mathbb{E}\big\{\mathbb{I}[D_1 = k, D_2 = k] - \mathbb{I}[D_1 = k, D'_2 = k]\big\} \\
&\quad = \mathbb{E}\big\{\mathbb{I}[E_1 = k - X_{12}, E_2 = k - X_{12}] - \mathbb{I}[E_1 = k - X_{12}, E_2 = k - X'_{12}]\big\} \\
&\quad = p(1-p)\mathbb{E}\big\{\mathbb{I}[E_1 = k, E_2 = k] - \mathbb{I}[E_1 = k, E_2 = k-1]\big\} \\
&\qquad + p(1-p)\mathbb{E}\big\{\mathbb{I}[E_1 = k-1, E_2 = k-1] - \mathbb{I}[E_1 = k-1, E_2 = k\big\} \\
&\quad = p(1-p)\big(\mathbb{P}(E_1 = k)^2 - \mathbb{P}(E_1 = k)\mathbb{P}(E_1 = k-1)\big) \\
&\qquad p(1-p)\big(\mathbb{P}(E_1 = k-1)^2 - \mathbb{P}(E_1 = k)\mathbb{P}(E_1 = k-1)\big) \\
&\quad \leq p\big(\mathbb{P}(E_1 = k) + \mathbb{P}(E_1 = k-1)\big).
\end{aligned}$$

Plugging this into (9.3), noting $p = \lambda/n$ and $n\varepsilon_n^2 \to \infty$ as $n \to \infty$, we have that

$$\mathbb{P}\left(\max_{k\geq 0}\left|P_n^{(k)} - \mu_n^{(k)}\right| \geq \varepsilon_n/2\right) \leq \frac{4(1+2\lambda)}{n\varepsilon_n^2} \xrightarrow[n\to\infty]{} 0.$$

The second term of (9.2) corresponds to Item 2. Since $\mu_n^{(k)} = \mathbb{P}(D_1 = k)$, and $D_1 \sim \mathrm{Bi}(n-1, \lambda/n)$, we bound the maximum by the sum and use the definition of total variation distance to find

$$\begin{aligned}\mathbb{I}\Big(\max_{k\geq 0}\Big|\mu_n^{(k)} - e^{-\lambda}\frac{\lambda^k}{k!}\Big| \geq \varepsilon_n/2\Big) &\leq \mathbb{I}\Big(\sum_{k\geq 0}\Big|\mu_n^{(k)} - e^{-\lambda}\frac{\lambda^k}{k!}\Big| \geq \varepsilon_n/2\Big) \\ &\leq \mathbb{I}\Big(d_{\mathrm{TV}}\big(\mathrm{Bi}(n-1,\lambda/n), \mathrm{Po}(\lambda)\big) \geq \varepsilon_n/4\Big) \\ &\leq \frac{4}{\varepsilon_n}\, d_{\mathrm{TV}}\big(\mathrm{Bi}(n-1,\lambda/n), \mathrm{Po}(\lambda)\big).\end{aligned} \tag{9.4}$$

We first bound the total variation distance using the triangle inequality

$$\begin{aligned}d_{\mathrm{TV}}\big(\mathrm{Bi}(n-1,\lambda/n), \mathrm{Po}(\lambda)\big) &\leq d_{\mathrm{TV}}\big(\mathrm{Bi}(n-1,\lambda/n), \mathrm{Bi}(n,\lambda/n)\big) + d_{\mathrm{TV}}\big(\mathrm{Bi}(n,\lambda/n), \mathrm{Po}(\lambda)\big) \\ &\leq d_{\mathrm{TV}}\big(\mathrm{Bi}(n-1,\lambda/n), \mathrm{Bi}(n,\lambda/n)\big) + \frac{\lambda^2}{n},\end{aligned}$$

where the second inequality was derived in Lecture 8. To bound the first term of the right hand side above, we make a coupling and use the coupling definition of total variation distance. Let $X_1, \ldots, X_n$ be i.i.d. distributed as $\mathrm{Be}(\lambda/n)$. Then for $m = n, n-1$, $\sum_{i=1}^m X_i \sim \mathrm{Bi}(m, \lambda/n)$, and so

$$d_{\mathrm{TV}}\big(\mathrm{Bi}(n-1,\lambda/n), \mathrm{Bi}(n,\lambda/n)\big) \leq \mathbb{P}\Big(\sum_{i=1}^{n-1} X_i \neq \sum_{i=1}^{n} X_i\Big) = \mathbb{P}(X_n \neq 0) = \frac{\lambda}{n}.$$

Plugging the last two displays into (9.4) implies

$$\mathbb{I}\Big(\max_{k\geq 0}\Big|\mu_n^{(k)} - e^{-\lambda}\frac{\lambda^k}{k!}\Big| \geq \varepsilon_n/2\Big) \leq \frac{4\lambda(\lambda+1)}{n\varepsilon_n} \xrightarrow[n\to\infty]{} 0,$$

since $n\varepsilon_n \to \infty$. We have showed that both terms in (9.2) tend to zero, and so the proof is complete. □