Random graphs and networks

Jan M. Swart

April 17, 2019

The Configuration Model

Def $[n] := \{1, \ldots, n\}$.
In deterministic $\mathrm{CM}_n(d)$, for each n, fix degrees $d_1, \ldots, d_n$. $(d_1^{(n)}, \ldots, d_n^{(n)})$.
Choose U unif. in $[n]$, def $D_n := d_U$.
Assume $\exists D$ s.t.

- $\mathbb{P}[D_n \in \cdot\,] \underset{n\to\infty}{\Longrightarrow} \mathbb{P}[D \in \cdot\,]$,
- $\mathbb{E}[D_n] \underset{n\to\infty}{\longrightarrow} \mathbb{E}[D]$,
- $\mathbb{E}[D_n^2] \underset{n\to\infty}{\longrightarrow} \mathbb{E}[D^2]$.

In $\mathrm{CM}_n(d)$ with i.i.d. degrees, choose $d_1, \ldots, d_n$ i.i.d. $d_i \stackrel{d}{=} D$.
If U indep. of (d_i), then $D_n := d_U \stackrel{d}{=} d_1$.
Def $\mathrm{CM}_n(d)$ *random* graph:

- If $\sum_{i\in[n]} d_i$ not even, add one to d_n.
- $\forall i$, draw d_i half-edges out of i.
- Enumerate the half edges.
- Pair the first half edge to a unif chosen free partner.
- Continue till no half-edges left.

Result: multigraph: may contain multiple edges and loops.
Def $\ell_n := \sum_{i\in[n]} d_i$.
Let $H :=$ set of half-edges. $H = \{h_i : i \in [\ell_n]\}$ enumeration of half-edges.

(i) Each matching of H has probab. $\frac{1}{\ell_n-1}\frac{1}{\ell_n-3}\cdots 1 = \frac{1}{(\ell_n-1)!!}$.

(ii) Uniform matching.

(iii) Law of $\mathrm{CM}_n(d)$ independent of enumeration of H.

(iv) $\mathrm{CM}_n(d)$ not uniform in set of multigraphs with prescribed degrees. Example: $\mathrm{CM}_2(3,3)$.

(v) $\mathrm{CM}_n(d)$ conditioned on being simple is uniform in set of graphs on $[n]$ with prescribed degrees. Proof Number the half-edges. Then for each simple graph with prescribed d_i's there are $\prod_{i=1}^n d_i!$ corresponding matchings.

Loops and multiple edges

Proposition 7.13 Assume $\mathbb{E}[D_n^2] \to \mathbb{E}[D^2] < \infty$.
Let $\nu := \mathbb{E}[\frac{1}{2}D(D-1)]/E[D]$.

$$S_n := \#\text{self-loops} \quad M_n := \#\text{multiple edges.}$$

Then $(S_n, M_n) \underset{n\to\infty}{\Longrightarrow} (S, M)$, where S, M independent Poisson with mean ν resp. ν^2.

Proof idea
Number vertices $1 \le i \le n$.
Number half-edges at given vertex $1 \le s \le d_i$.
$\mathcal{I}_1 := \{(st, i) : 1 \le i \le n,\ 1 \le s < t \le d_i\}$ pairs of half-edges that can form a loop.
$|\mathcal{I}_1| = m_n := \sum_{i\in[n]} \frac{1}{2} d_i(d_i - 1)$.
$I_{st,i}$ indicator loop (st, i) present.
$\mathbb{E}[I_{st,i}] = (\ell_n - 1)^{-1}$ with $\ell_n := \sum_{i\in[n]} \frac{1}{2} d_i$.
"Almost independence" $\Rightarrow$ # loops $\approx$ Poisson with mean

$$\approx \frac{m_n}{\ell_n} = \frac{\sum_{i\in[n]} \frac{1}{2} d_i(d_i-1)}{\sum_{i\in[n]} \frac{1}{2} d_i} = \frac{n^{-1}\sum_{i\in[n]} \frac{1}{2} d_i(d_i-1)}{n^{-1}\sum_{i\in[n]} \frac{1}{2} d_i} \underset{n\to\infty}{\longrightarrow} \nu.$$

$\mathcal{I}_2 := \{(s_1t_1, s_2t_2, i, j) : 1 \le i < j \le n,\ 1 \le s_1 < s_2 \le d_i, 1 \le t_1 \ne t_n \le d_j\}$ possibilities to form a double edge.
$|\mathcal{I}_2| \approx 2 \cdot \frac{1}{2} m_n(m_n - 1) \approx m_n^2$.
$I_{s_1t_1,s_2t_2,i,j}$ indicator multiple edge (s_1t_1, s_2t_2, i, j) present.
$\mathbb{E}[I_{s_1t_1,s_2t_2,i,j}] = \dfrac{1}{(\ell_n-1)(\ell_n-3)}$.
"Almost independence" $\Rightarrow$ # multiple edges $\approx$ Poisson with mean $\approx \dfrac{m_n^2}{\ell_n^2} \underset{n\to\infty}{\longrightarrow} \nu^2$.

Precise proof

$$S_n := \sum_{m \in \mathcal{I}_1} I_m \quad M_n := \sum_{m \in \mathcal{I}_2} I_m \quad \text{number of loops and multiple edges.}$$

$$(X)_r := X(X-1)\cdots(X-r+1) \quad \text{factorial moment.}$$

$$\mathbb{E}[(S_n)_r] = \sum_{m_1 \in \mathcal{I}_1} \sum_{m_2 \in \mathcal{I}_1 \setminus \{m_1\}} \cdots \sum_{m_r \in \mathcal{I}_1 \setminus \{m_1, \ldots, m_{r-1}\}} \mathbb{P}[I_{m_1} = \cdots = I_{m_r} = 1]$$
$$=: \sum_{m_1, \ldots, m_r \in \mathcal{I}_1}^{*} \mathbb{P}[I_{m_1} = \cdots = I_{m_r} = 1].$$

<u>Theorem 2.6</u> $\mathbb{E}[(X_n)_r] \to \lambda^r$ $(r \geq 1)$ implies $X_n \underset{n\to\infty}{\Longrightarrow} \mathrm{Pois}(\lambda)$.
Similarly, $\mathbb{E}[(X_n)_s (Y_n)_r] \to \lambda^2 \mu^r$ implies $(X_n, Y_n) \underset{n\to\infty}{\Longrightarrow} (\mathrm{Pois}(\lambda), (\mathrm{Pois}(\mu))$.
Need to control, for $m_{1,1}, \ldots, m_{1,s} \in \mathcal{I}_1$ and $m_{2,1}, \ldots, m_{2,s} \in \mathcal{I}_1$

$$\mathbb{P}[I_{m_{1,1}} = \cdots = I_{m_{1,s}} = I_{m_{2,1}} = \cdots = I_{m_{2,r}} = 1]$$
$$= \frac{1}{(\ell_n - 1)(\ell_n - 3)\cdots(\ell_n - 1 - 2s - 4r)},$$

or $= 0$ if events incompatible, i.e., want same half-edge to connect to two different half-edges. Diligent counting completes the proof. $\square$

The Erased Configuration Model

Def $D_n^{\mathrm{er}} :=$ degree of $U_n \in [n]$ after erasing loops and multiple edges.

<u>Theorem 7.10</u> Assume $\mathbb{E}[D_n] \underset{n\to\infty}{\longrightarrow} \mathbb{E}[D] < \infty$. Then $D_n^{\mathrm{er}} \underset{n\to\infty}{\Longrightarrow} D$.

<u>Proof</u> Need to show no loops and multiple edges at U_n as $n \to \infty$.
$\mathbb{E}[\#\text{ loops at } U_n \mid D_n = k] = \frac{\frac{1}{2}k(k-1)}{\ell_n - 1} \underset{n\to\infty}{\longrightarrow} 0.$

$$\limsup_{n\to\infty} \mathbb{P}[\exists \text{ loop at } U_n] \leq \limsup_{n\to\infty} \left\{ \frac{\frac{1}{2}k(k-1)}{\ell_n - 1} \mathbb{P}[D_n \leq k] + \mathbb{P}[D_n > k] \right\}$$
$$\leq \mathbb{P}[D > k] \underset{k\to\infty}{\longrightarrow} 0.$$

If $d_i = k$, then
$$\mathbb{E}[\#\text{ multiple edges at } i] = \frac{1}{(\ell_n - 1)(\ell_n - 3)} \tfrac{1}{2}k(k-1) \sum_{j \in [n] \setminus \{i\}} d_j(d_j - 1).$$

$$\limsup_{n\to\infty} \mathbb{P}[\exists \text{ multiple edge at } U_n]$$
$$\leq \limsup_{n\to\infty} \left\{ \tfrac{1}{2}k(k-1) \underbrace{\left(\frac{1}{\ell_n^2} \sum_{i \in [n]} d_i^2 \right)}_{=:\, P_n} \mathbb{P}[D_n \leq k] + \mathbb{P}[D_n > k] \right\}.$$

$P_n =$ probab. two unif. chosen half-edges are in same vertex.
Let $p_n(k) := \mathbb{P}[D_n = k]$ and let $\mathbb{P}[\hat{D}_n = k] := \hat{p}_n(k) := \frac{1}{\mathbb{E}[D_n]} k p_n(k)$ size-biased law. Then

$$P_n = \sum_k \hat{p}_n(m) \frac{m}{\ell_n} \leq \{ \frac{m}{\ell_n} \mathbb{P}[\hat{D}_n \leq m] + \mathbb{P}[\hat{D}_n > m] \}$$

and

$$\limsup_{n\to\infty} \mathbb{P}[\exists \text{ multiple edge at } U_n] \leq \tfrac{1}{2} k(k-1) \mathbb{P}[\hat{D} > m] + \mathbb{P}[D > k].$$

First $m \to \infty$, then $k \to \infty$ gives ≤ 0. □

Conditioning i.i.d. (d_i) on $d_i \leq a_n$ with $a_n \to \infty$ has no influence on the limit law of D_n.
Theorem 7.22 Assume $\mathbb{P}[D_n \leq a_n] = 1$ with $a_n = o(n)$. Then $D_n^{\text{er}} \underset{n\to\infty}{\Longrightarrow} D$.

Proof W.l.o.g. $d_i \geq 1$ for all i. Then $\ell_n \geq n$ and hence

$$P_n \leq \{ \frac{a_n}{\ell_n} \mathbb{P}[\hat{D}_n \leq a_n] + \mathbb{P}[\hat{D}_n > a_n] \} \underset{n\to\infty}{\longrightarrow} 0.$$

□

Consequence We can construct erased configuration models with arbitrary degree distribution.

Heavy tails

Theorem 7.24 Assume $(d_i)_{i\in[n]}$ i.i.d. with

$$\mathbb{P}[D \geq k] = k^{1-\tau} L(k),$$

where $\tau \in (1, 2)$ and L *slowly varying*, i.e., $L(ck)/L(k) \to 1$ for all $c > 0$. Then

$$\mathbb{P}[D_n^{\text{er}} = k] \underset{n\to\infty}{\longrightarrow} \mathbb{P}[D^{\text{er}} = k] \quad \text{with} \quad \mathbb{P}[D^{\text{er}} \leq k] \leq c k^{-1}$$

for some $c < \infty$.
Note This says $\approx$ the limit law has $\tau \geq 2$.

Conjecture D^{er} has $\tau = 2$.

Proof Order the degrees as $d_{(1)} \geq d_{(2)} \geq d_{(3)} \geq \cdots$.
Theorem 2.33 says that there exists a u_n of the form $u_n = n^{1/(\tau-1)} l_n$ with l_n slowly varying, s.t.

$$\frac{1}{u_n} \big(\ell_n, d_{(1)}, d_{(2)}, d_{(3)}, \ldots\big) \underset{n\to\infty}{\Longrightarrow} (\eta, \xi_1, \xi_2, \xi_3, \ldots),$$

where $\{\xi_1 > \xi_2 > \cdots\}$ is Poisson point set on $[0,\infty)$ with intensity measure $\mu\big([x,\infty)\big) = x^{1-\tau}$ and $\eta := \sum_{j=1}^{\infty} \xi_j$.
Let Q be the random probab. law defined $Q_j := \xi_j/\eta$.

Conditional on Q, let $I_1, I_2, \ldots$ be i.i.d. with law Q and let

$$K(m,k) := \mathbb{P}\big[\#\Delta_m = k\big] \quad \text{with} \quad \Delta_m : \{i : \exists 1 \le l \le m \text{ s.t. } I_l = i\}$$

Thm 7.23 says that

$$\mathbb{P}[D_n^{\mathrm{er}} = k] \underset{n\to\infty}{\longrightarrow} \mathbb{P}[D^{\mathrm{er}} = k] := \sum_{m=0}^{\infty} p_m K(m,k).$$

"Proof" All half edges at a typical vertex connect to vertices of high degree. Now $K(m,k) = \lim_{n\to\infty} \mathbb{P}[D_n^{\mathrm{er}} = k \,|\, D_n = m]$.

Missing lemma $\#\Delta_m \sim c m^{\tau-1}$ with high probability.

Consequence $D^{\mathrm{er}} \approx D^{\tau-1}$ when both are large, so

$$\begin{aligned} \mathbb{P}[D^{\mathrm{er}} \ge k] &\approx \mathbb{P}[D^{\tau-1} \ge k] = \mathbb{P}[D \ge k^{1/(\tau-1)}] \\ &= (k^{1/(\tau-1)})^{1-\tau} L(k^{1/(\tau-1)}) = k^{-1} L'(k) \end{aligned}$$

with L, L' slowly varying.

Proof of Lemma? Divide the interval $[0,\eta]$ in pieces of length $\xi_1, \xi_j, \ldots$. Choose m points uniformly on $[0,\eta]$. Then $\#\Delta_m$ is the number of intervals that contains at least one point. For large m, the m points look like a Poisson points set with intensity m/η, so

$$\mathbb{E}[\#\Delta_m] \approx \mathbb{E}\Big[\sum_{j=1}^{\infty} \big(1 - e^{-(m\xi_j/\eta)}\big)\Big] \approx \int_0^{\infty} \big(1 - e^{-\frac{m}{\eta}x}\big)\mu(\mathrm{d}x).$$

Forgetting about multiplicative constants,

$$\approx \int_0^{1/m} x\mu(\mathrm{d}x) + \int_{1/m}^{\infty} \mu(\mathrm{d}x) \approx \int_0^{1/m} x \cdot x^{-\tau}\mathrm{d}x + (1/m)^{1-\tau} \approx m^{\tau-2} + m^{\tau-1}.$$

If we believe the law of $\#\Delta_m$ to be concentrated near its mean, then this "proves" the lemma. □