Power-Law Distribution

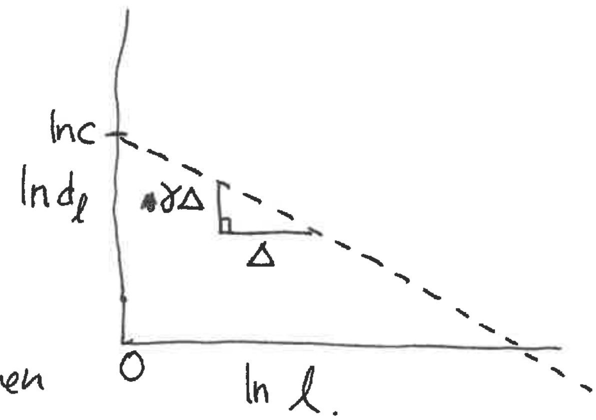
In many real-world networks, the degree distribution obeys

$$d_\ell \approx c \cdot \ell^{-\gamma}$$

for some $\gamma > 0$ and some normalizing constant* $c > 0$.

Taking logs on both sides gives

$$\ln d_\ell \approx -\gamma \ln \ell + \ln c$$



So a power-law degree distribution looks like a line with slope $-\gamma$ when plotted on a log-log plot.

- Called a <u>power-law distribution</u>
  - γ is the <u>degree exponent</u>.
- Observed by Vilfredo Pareto, a 19th century economist who noticed that a small number of wealthy individuals posess most of the world's wealth.

  "20% of people have 80% of the money"

* remember $d_\ell = \dfrac{\#\text{nodes of degree } \ell}{\#\text{nodes}}$, so $\sum_\ell d_\ell = \frac{n}{n} = 1$, which is why we need to normalize by $c = \left(\sum_{\ell=\delta}^{\Delta} \ell^{-\gamma} \right)^{-1}$

Max-degree. min-degree

Role of γ in $d_\ell = c \cdot \ell^{-\gamma}$

d_ℓ decreasy polynomially in ℓ, where γ controls how fast this decrease is.

$0 < \gamma \leq 1$ is unusual, since $\sum_{\ell=1}^{\infty} \ell^{-\gamma} = \infty$ in this case, so there is no possible normalizing constant c.

For $\gamma > 1$, $\sum_{\ell=1}^{\infty} \ell^{-\gamma}$ is finite and is $\approx \int_1^{\infty} x^{-\gamma} dx = \frac{1}{\gamma-1}$, so $c \approx \gamma - 1$

For $\gamma > 2$, we can even compute the (average / expected) degree

$$\sum_{\ell=1}^{\infty} \ell \cdot c \ell^{-\gamma} = c \sum_{\ell=1}^{\infty} \ell^{1-\gamma} \approx c \frac{1}{1-(\gamma-1)} = \frac{c}{\gamma-2} = \frac{\gamma-1}{\gamma-2}.$$

Important: c is only defined for $\gamma > 1$.
average degree only defined for $\gamma > 2$.

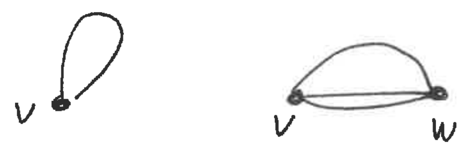
Note: Lots of handwaving here
- integrals only approximate sums.
- $c \cdot \ell^{-\gamma}$ is not necessarily integer, so it can't count anything. Instead we have $\approx n \cdot c \cdot \ell^{-\delta}$ nodes of degree ℓ [for large n].
- It may not be possible to construct a graph that (even approximately) achieves the degree distribut d_ℓ, for ℓ close to 1.

Chung-Lu Model

Take any ~~non-negati~~ positive weights $w_1, \dots, w_n$, generates a random graph with $V(G) = [n]$ such that

$$\mathbb{E}(\deg(i)) = w_i \text{ for each } i \in [n].$$

Caveat: Graph may contain self-loops and parallel edges.



For each $\cancel{\{i,j\} \in \binom{[n]}{2}}$ $i,j \in [n]$ define $p_{ij} = \begin{cases} \dfrac{w_i \cdot w_j}{W} & \text{if } i \neq j \\ \dfrac{w_i^2}{2W} & \text{if } i = j \end{cases}$

where $W = \sum_{i=1}^{n} w_i$

If $p_{ij} < 1$, treat p_{ij} as the probability that $ij \in E(G)$.

In general #copies of edge ij in G. $= \begin{cases} \lceil p_{ij} \rceil & \text{with probability } p_{ij} - \lfloor p_{ij} \rfloor \\ \lfloor p_{ij} \rfloor & \text{with probability } 1 - (p_{ij} - \lfloor p_{ij} \rfloor) \end{cases}$

$$\mathbb{E}(\deg(i)) = \sum_{j\in[n]\setminus i} p_{ij} + 2p_{ii} = \sum_{j\in[n]\setminus i} \frac{w_i w_j}{W} + 2\cdot\frac{w_i^2}{2W}$$

$$= \sum_{j\in[n]} \frac{w_i w_j}{W} = \frac{w_i}{W}\cdot\sum_{j\in[n]} w_j = \frac{w_i}{W}\cdot W = w_i .$$

In real applications we can expect $w_i \geq 1$ for all $i\in[n]$ and $\sum_{i\in[n]} w_i << \binom{n}{2}$ and $w_i \leq n$.

So $\quad n \leq W \leq \binom{n}{2}$ and $\frac{w_i}{W} \leq 1$, so $p_{ij} \leq 1$.

We can approximate any degree-distribution $d_1, d_2, \dots d_n$ we want including power-law-like distubutions:

$$\text{set } w_i = \ell \quad \text{for } \approx d_\ell \cdot n \text{ values of } i\in\{1,\dots,n\}$$

If we take $w_i = pn$ for each $i\in[n]$ and ignore self-loops then we get $G_{n,p}$.

Chung-Lu: Efficient Implementation.

- Obvious implementation requires (at least) $\Omega(n^2)$ time since we have to consider $i,j \in [n]$.
- First generate a binomial$\left(\frac{W}{\binom{n}{2}}, \binom{n}{2}\right)$ random variable M.
- Randomly choose M edges where $Pr(ij) = p_{ij} = \frac{w_i \cdot w_j}{W}$ $\left(\text{or } p_{ii} = \frac{w_i^2}{2W}\right)$
  - to do this, choose i with probability $\frac{w_i}{W}$
  - choose j with probability $\frac{w_j}{W}$ $\left(\text{or } i \text{ with probability } \frac{w_i}{2W}\right)$

Requires a fast method to choose from a non-uniform distribution

$$\frac{w_1}{W}, \frac{w_2}{W}, \ldots, \frac{w_n}{W}$$

For this we can preprocess using Walker's Method (a.k.a the Alias method) described on page 107 of L. Devroye, Non-uniform Random Variate Generation, Springer-Verlag, 1986. See the link on the course webpage.