Sampling power law degree distributions

June 4, 2021

1 Unclustered scale-free networks

The degree distribution of scale-free networks is given by a power law with exponent $\gamma > 2$,

$$p(x) = (\gamma - 1)x_{\min}^{\gamma-1}x^{-\gamma} \ , \tag{1}$$

for $x \geq x_{\min}$. In finite systems, the maximum degree is limited by the fact that a node cannot have more neighbors than available nodes. This natural cut-off is given by

$$\int_{x_{\max}}^{\infty} p(x)\mathrm{d}x < \frac{1}{N} \ , \tag{2}$$

with N the system size. In short, Eq. (2) states that the number of nodes with degree $x_{\max}$ or higher must be smaller than N. Solving Eq. (2) yields

$$x_{\max} > x_{\min}N^{\frac{1}{\gamma-1}} \ . \tag{3}$$

In practice, we sample the degree sequence from a discrete probability function

$$P(k) = Ak^{-\gamma} \ , \tag{4}$$

for $k_{\min} \leq k \leq k_{\max}$. In order to guarantee a connected network we choose $k_{\min} = 2$, and from Eq. (3) we find

$$k_{\max} = \left\lfloor k_{\min}N^{\frac{1}{\gamma-1}} \right\rfloor \ . \tag{5}$$

Finally, the normalization constant yields $A = \zeta(\gamma, k_{\min}) - \zeta(\gamma, k_{\max}+1)$, with $\zeta(x, \ell)$ the Hurwitz zeta function

$$\zeta(x, \ell) = \sum_{k=0}^{\infty}(k+\ell)^{-x} = \sum_{k=\ell}^{\infty} k^{-x} \ . \tag{6}$$

2 Rejection method

It is very simple to sample the continuous distribution given in (1). Sampling a uniform random number u_1 we solve for the complementary cumulative function

$$u_1 = \int_x^{\infty} p(x)\mathrm{d}x \rightarrow x = x_{\min} u_1^{\frac{1}{1-\gamma}} \ . \tag{7}$$

The discrete distribution is a bit trickier, but feasible using an adequate rejection method. First we make a continuous version of the discrete distribution $P(k)$ such that $\hat{P}(x) = P(k)\ \forall x \in [k, k+1)$. Next we realize that $p(x)$ is generally larger than $\hat{P}(x)$ except for an area around $x = x_0 + 1$ (see Figure 1).[1] If we construct the function $q(x) = Cp(x)$, with $C = P(k_0)/p(x_0+1)$, we see that $q(x)$ is always larger than $\hat{P}(x)$.

So we follow this algorithm:

1. Sample $u_1 \in U(0,1)$ and sample x.
2. Evaluate the integer part $k = \lfloor x \rfloor$.
3. If $k > k_{\max}$, reject the candidate k.
4. Else, sample $u_2 \in U(0,1)$ and evaluate $y = u_2 C p(x)$.
   - If $y \leq P(k)$, accept the sampled degree k.
   - Else, reject the candidate k.

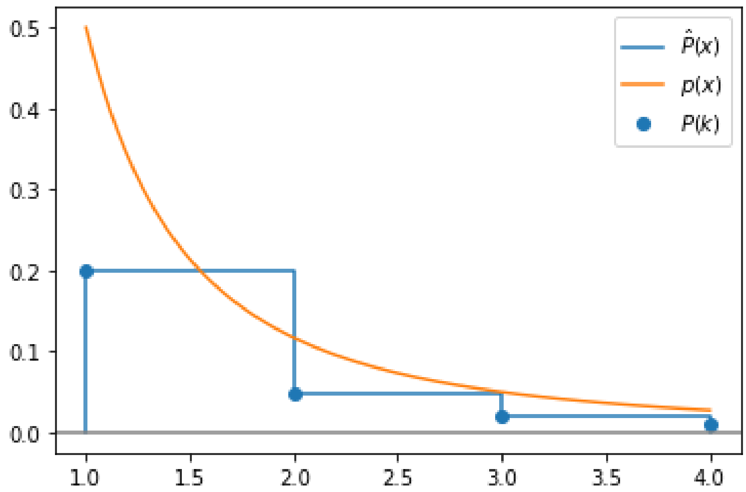

Figure 1: Schematic representation of $p(x)$, $P(k)$ and $\hat{P}(x)$ with $k_0 = x_0 1$.

[1]This is technically not true. Other points can also show this behavior, depending on the value of γ. However, the largest difference occurs always around $x = x_0 + 1$, so the factor C that we have computed guarantees that all the differences are solved.