Sampling and Inference for Beta Neutral-to-the-Left Models of Sparse Networks

Benjamin Bloem-Reddy, Adam Foster, Emile Mathieu, Yee Whye Teh

Department of Statistics, University of Oxford

Contents

Temporal networks

- Facebook
- StackOverflow
- etc

Sparsity

- For a complete graph, $K_n = O(n^{1/2})$
- For a sparse graph,

$$K_n = O(n^{1/(1+\sigma)}) \tag{1}$$

for $0 \leq \sigma < 1$

Power law degree distribution

A power law distribution of exponent η on $\{1, 2, ...\}$ has

$$p(d) = C_\eta d^{-\eta} \tag{2}$$

where $\eta > 1$.

Power law degree distribution

A power law distribution of exponent η on $\{1, 2, ...\}$ has

$$p(d) = C_\eta d^{-\eta} \tag{2}$$

where $\eta > 1$.
The asymptotic degree distribution has **power law tail with exponent** $\eta > 1$ if

$$\frac{m_n(d)}{K_n} \xrightarrow[n \to \infty]{p} L(d) d^{-\eta} \, , \tag{3}$$

for slowly varying function $L(d)$.

Power laws and sparsity

We have

$$K_n = \sum_{d=1}^{\infty} m_n(d), \tag{4}$$

$$n = \sum_{d=1}^{\infty} d\, m_n(d). \tag{5}$$

Suppose (somewhat informally) that

$$K_n = C_{n,\eta} \sum_{d=1}^{n} d^{-\eta} \tag{6}$$

Power laws and sparsity

We have

$$K_n = \sum_{d=1}^{\infty} m_n(d), \tag{4}$$

$$n = \sum_{d=1}^{\infty} d\, m_n(d). \tag{5}$$

Suppose (somewhat informally) that

$$K_n = C_{n,\eta} \sum_{d=1}^{n} d^{-\eta} \tag{6}$$

then

$$n = C_{n,\eta} \sum_{d=1}^{n} d^{-\eta+1} = K_n \frac{\sum_{d=1}^{n} d^{-\eta+1}}{\sum_{d=1}^{n} d^{-\eta}} \tag{7}$$

Power laws and sparsity

Letting $n \to \infty$ in

$$\frac{K_n}{n} = \frac{\sum_{d=1}^{n} d^{-\eta}}{\sum_{d=1}^{n} d^{-\eta+1}} \tag{8}$$

we see $K_n = O(n)$ if $\eta > 2$, $K_n = o(n)$ if $\eta \in (1, 2]$.

Power laws and sparsity

Letting $n \to \infty$ in

$$\frac{K_n}{n} = \frac{\sum_{d=1}^{n} d^{-\eta}}{\sum_{d=1}^{n} d^{-\eta+1}} \tag{8}$$

we see $K_n = O(n)$ if $\eta > 2$, $K_n = o(n)$ if $\eta \in (1, 2]$.

Summary

For sparse graphs, $\sigma = 0 \leftrightarrow \eta > 2$ and $\sigma > 0 \leftrightarrow \eta \in (1, 2]$.

Empirical study

SNAP dataset Todo insert lots of pictures

Edge exchangeable models

"The probability of all orderings of history are the same"

Pitman-Yor Process

Parameters $\tau \in (0, 1), \theta > -\tau$.

Arrivals

$$\mathbb{P}(T_{j+1} - T_j > t \mid T_j) = \prod_{i=1}^{t} \frac{T_j + t - j\tau}{T_j + t + \theta} \tag{9}$$

Size-biased reinforcement

$$Z_{n+1}|\mathbf{Z}_n, \mathbf{T} = \begin{cases} K_{n+1} \text{ w.p. } 1 & \text{if } n+1 = T_{K_{n+1}} \\ j \text{ w.p. } \propto (d_{j,n} - \tau) & \text{otherwise} \end{cases} \tag{10}$$

Sparsity and power law properties of edge-exchangeable models

Yule-Simon Process

Parameter $\beta \in (0, 1)$.

Arrivals

$$T_{j+1} - T_j \overset{\text{i.i.d.}}{\sim} \text{Geom}\left(\beta\right) \tag{11}$$

Size-biased reinforcement

$$Z_{n+1} | \mathbf{Z}_n, \mathbf{T} = \begin{cases} K_{n+1} \text{ w.p. } 1 & \text{if } n+1 = T_{K_{n+1}} \\ j \text{ w.p. } \propto d_{j,n} & \text{otherwise} \end{cases} \tag{12}$$