Probabilistic Foundations of Statistical Network Analysis
Chapter 4: Generative models

Harry Crane

Based on Chapter 4 of *Probabilistic Foundations of Statistical Network Analysis*

Table of Contents

Chapter 1  Orientation
2  Binary relational data
3  Network sampling
4  Generative models
5  Statistical modeling paradigm
6  Vertex exchangeability
7  Getting beyond graphons
8  Relative exchangeability
9  Edge exchangeability
10  Relational exchangeability
11  Dynamic network models
Sampling models (Chapter 3) specified by
- candidate distributions describing network variation
- sampling scheme that links the population $Y_N$ to the sample $Y_n = \sum_{n,N} Y_N$

Generative models (Chapter 4) specified by
- candidate distributions
- generative scheme to describe network growth

Describe generative scheme by an *evolution map*. 
Evolution maps (Chapter 4 of FPSNA)

**Definition**

For \( n \leq N \), call \( P : \{0, 1\}^{n \times n} \rightarrow \{0, 1\}^{N \times N} \) an **evolution map** if

\[
P(y)_{|[n]} = y \quad \text{for all } y \in \{0, 1\}^{n \times n}.
\]

An evolution map is an operation by which \( y \in \{0, 1\}^{n \times n} \) ‘evolves’ into \( P(y) \in \{0, 1\}^{N \times N} \) by holding fixed the part of the network that already exists, namely \( y \).

- Let \( \mathcal{P}_{n,N} \) be the set of all evolution maps \( \{0, 1\}^{n \times n} \rightarrow \{0, 1\}^{N \times N} \).
- A **generating scheme** is a random map \( \Pi_{n,N} \) in \( \mathcal{P}_{n,N} \). Distribution can depend on \( Y_n \).
- More precisely, \( \Pi_{n,N} Y_n \) is the network with \( N \) vertices obtained by first generating \( Y_n = y \), putting \( \Pi_{n,N} Y_n = P(y) \), for \( P \in \mathcal{P}_{n,N} \) chosen according to the conditional distribution of \( \Pi_{n,N} \) given \( Y_n = y \).
- The distribution of \( \Pi_{n,N} Y_n \) is computed by

\[
\Pr(\Pi_{n,N} Y_n = y) = \sum_{P \in \mathcal{P}_{n,N}} \Pr(\Pi_{n,N} = P \mid Y_n = y_{|[n]}) \Pr(Y_n = y_{|[n]}) 1(P(y_{|[n]}) = y),
\]

where \( 1(\cdot) \) is the indicator function.
Generative consistency

**Definition (Generative consistency (Definition 4.1 of PFSNA))**

Let $Y_n$ and $Y_N$ be random $\{0, 1\}$-valued arrays and let $\Pi_{n,N}$ be a generating scheme. Then $Y_n$ and $Y_N$ are consistent with respect to $\Pi_{n,N}$ if

$$\Pi_{n,N} Y_n =_{\mathcal{D}} Y_N,$$

for $\Pi_{n,N} Y_n$ defined by the distribution in (1).

**Duality between generative consistency and consistency under selection:**

For any $Y_n$ and generating mechanism $\Pi_{n,N}$, define $Y_N$ by $Y_N = \Pi_{n,N} Y_n$. Then by the defining property of an evolution map, $Y_n$ and $Y_N$ enjoy the relationship

$$S_{n,N} Y_N = S_{n,N} \Pi_{n,N} Y_n = Y_n$$

with probability 1;

that is, $Y_n$ and $\Pi_{n,N} Y_n$ are consistent under selection by default.
Dynamics based on Simon’s preferential attachment scheme for heavy-tailed distributions.

Vertices arrive one at a time and attach preferentially to previous vertices based on their degree.

**Formal definition:**

- Take $m \geq 1$ (integer) and $\delta > -m$ (real number) so that each new vertex attaches randomly to $m$ existing vertices with probability increasing with degree.
- Initiate at a graph $y_0$ with $n_0 \geq 1$ vertices, which then evolves successively into $y_1, y_2, \ldots$ by connecting a new vertex to the existing graph at each step.
- For any $y = (y_{ij})_{1 \leq i, j \leq n}$ and every $i = 1, \ldots, n$, the degree of $i$ in $y$ is the number of edges incident to $i$,
  
  $$\text{deg}_y(i) = \sum_{j \neq i} y_{ij}.$$  

- At step $n \geq 1$, a new vertex $v_n$ attaches to $m \geq 1$ vertices in $y_{n-1}$, with each of the $m$ vertices $v'$ chosen independently without replacement with probability proportional to
  
  $$\text{deg}_{y_{n-1}}(v') + \delta/m.$$
In keeping with the notation of Section 4.1, let \( \Pi_{\delta,m}^{k,n} \), \( k \leq n \), denote the generating mechanism for the process parameterized by \( m \geq 1 \) and \( \delta > -m \).

By letting the parameters \( n_0 \geq 1 \), \( m \geq 1 \), and \( \delta > -m \) vary over all permissible values and treating the initial conditions \( y_0 \) and \( n_0 \) as fixed, the above generating mechanism determines a family of distributions for each finite sample size \( n \geq 1 \), where \( n \) is the number of vertices that have been added to \( y_0 \).

For each \( n \geq 1 \), this process gives a collection of distributions \( M_n \) indexed by \( (m, \delta) \), and each distribution in \( M_k \) indexed by \( (m, \delta) \) is related to a distribution in \( M_n \), \( n \geq k \), with the same parameters through the preferential attachment scheme \( \Pi_{\delta,m}^{k,n} \) associated to the model.

For any choice of parameter \( (\delta, m) \), we express the relationship between \( Y_k \) and \( Y_n \), \( n \geq k \), by

\[
Y_n = D \Pi_{\delta,m}^{k,n} Y_k
\]
Barabási–Albert model (Empirical properties)

**Sparsity:**
- Let $y = (y^{(n)})_{n \geq 1}$ be sequence of graphs ($y^{(n)}$ has $n$ vertices).
- Call $y$ sparse if
  \[
  \lim_{n \to \infty} \frac{1}{n(n-1)} \sum_{1 \leq i \neq j \leq n} y^{(n)}_{ij} = 0.
  \]
- Under BA model, $(Y_n)_{n \geq 1}$ grows by adding one vertex at a time with $m$ new edges, so that
  \[
  \frac{1}{n(n-1)} \sum_{1 \leq i \neq j \leq n} Y_{ij} = \frac{1}{n(n-1)} (mn + n_0) \to 0 \quad \text{as } n \to \infty.
  \]
- Networks under BA model are sparse with probability 1.

**Power law degree distribution:**
- For $k \geq 1$, let
  \[
  p_y(k) = n^{-1} \sum_{i=1}^{n} 1(\deg_y(i) = k).
  \]
- A sequence $y = (y^{(n)})_{n \geq 1}$ exhibits power law degree distribution with exponent $\gamma > 1$ if
  \[
  p_y^{(n)}(k) \sim \gamma^{-k} \quad \text{for all large } k \text{ as } n \to \infty,
  \]
  where $a(k) \sim b(k)$ indicates that $a(k)/b(k) \to 1$ as $k \to \infty$.
- BA model with parameter $(\delta, m)$ has power law degree distribution with exponent $3 + \delta/m$ with probability 1.
Many real-world networks believed to exhibit power law, or nearly power law, degree distribution (Barabási–Albert, ...).

Heuristic check: power law degree distribution implies

$$\log p_Y(k) \sim -\gamma \log(k), \quad \text{large } k \geq 1.$$  \hfill (2)

Yule–Simon distribution (dotted) vs. line $-3 \log(k)$ (solid).

**Figure:** Dotted line shows log-log plot of the Yule–Simon distribution for $\gamma = 3$. Solid line shows the linear approximation in (2) by approximating $\Gamma(\gamma)/\Gamma(k + \gamma) \sim \gamma^{-k}$, which holds asymptotically for large values of $k$. 
Random walk (RW) models

- Add a new edge at each step (instead of new vertex as in BA model).

- Start with initial graph $y_0$ and evolve $y_1, y_2, \ldots$ as follows.
  
  - At step $n \geq 1$, choose vertex $v_n$ in $y_{n-1}$ randomly with distribution $F_n$ (which can depend on $y_{n-1}$). Then draw a random nonnegative integer $L_n$ from distribution also depending on $y_{n-1}$.
  
  - Given $v_n$ and $L_n$, perform a simple random walk on $y_{n-1}$ for $L_n$ steps starting at $v_n$.
  
  - If after $L_n$ steps the random walk is at $v^* \neq v_n$, then add edge between $v^*$ and $v_n$; otherwise, add new vertex $v^{**}$ and put edge between $v^{**}$ and $v_n$.

- Choosing $v_n$ by degree-biased distribution on $y_{n-1}$ and taking $L_n$ to be large simulates BA model.

- For more details on these models see Bloem-Reddy and Orbanz (https://arxiv.org/abs/1612.06404), Bollobas, et al (2003), and related work.
Erdős–Rényi–Gilbert model

- Classical Erdős–Rényi–Gilbert model includes each edge in random graph independently with fixed probability $\theta$.
- Generative description: For any $\theta \in [0, 1]$, define $\Pi_{\theta}^{n,N}$ as the generating scheme which acts on $\{0, 1\}^{n \times n}$ by

$$y \mapsto \Pi_{\theta}^{n,N}(y)$$

which fixes the upper $n \times n$ submatrix to be $y$ and fills in the rest of the off-diagonal entries with i.i.d. Bernoulli random variables $(B_{ij})_{1 \leq i \neq j \leq N}$ with success probability $\theta$. 
Above examples start with a base case \( Y_0 \), from which a family of networks \( Y_1, Y_2, \ldots \) is constructed inductively according to a random scheme.

A generic way to specify a generative network model is to specify a conditional distribution for \( Y_n \) given \( Y_{n-1} \) such that \( Y_n \mid [n-1] = Y_{n-1} \) with probability 1.

Conditional distribution \( \Pr(Y_n = \cdot \mid Y_{n-1}) \) determines the distribution of a random generating mechanism \( \Pi_{n-1,n} \) in \( \mathcal{P}_{n-1,n} \)

\[ \implies Y_n \text{ can be expressed as } Y_n = \Pi_{n-1,n} Y_{n-1} \text{ for every } n \geq 1. \]

Composing these actions for successive values of \( n \) determines the generating mechanism \( \Pi_{n,N}, n \leq N \), by the law of iterated conditioning:

\[ \implies \text{Given } Y_n, \text{ construct } Y_N = \Pi_{n,N} Y_n \text{ by } \]

\[ Y_N = \Pi_{N-1,N}(\Pi_{N-2,N-1}(\cdots (\Pi_{n,n+1} Y_n))). \]

The conditional distribution of \( Y_N \) given \( Y_n \) computed by

\[
\Pr(Y_N = y^* \mid Y_n = y^* \mid [n]) = \\
= \Pr(Y_N = y^* \mid Y_{N-1} = y^* \mid [N-1]) \times \Pr(Y_{N-1} = y^* \mid [N-1] \mid Y_n = y^* \mid [n]) \\
= \prod_{i=1}^{N-n} \Pr(\Pi_{N-i,N-i+1} y^* \mid [N-i]) = y^* \mid [N-i+1] \mid Y_{N-i} = y^* \mid [N-i]).
\]
Network modeling paradigm (Chapter 5) gives framework to handle sampling models (Chapter 3) and generative models (Chapter 4).

See Chapters 3–5 of *Probabilistic Foundations of Statistical Network Analysis*