Asymptotics for sparse exponential random graph models

Mei Yin[1]

Department of Mathematics, University of Denver

May 21, 2015

[1]Research supported under NSF grant DMS-1308333. Joint work with Lingjiong Zhu.

In this talk, we study the asymptotics for undirected sparse exponential random graph models where the parameters may depend on the number of vertices of the graph. We obtain exact estimates for the mean and variance of the limiting probability distribution and the limiting normalization constant of the edge-(single)-star model. They are in sharp contrast to the corresponding asymptotics in dense exponential random graph models. Similar analysis is done for directed sparse exponential random graph models parametrized by edges and multiple outward stars.

The hierarchy

grand canonical ensemble	exponential random graph
$\downarrow$	$\downarrow$
canonical ensemble	constrained exponential random graph
$\downarrow$	$\downarrow$
microcanonical ensemble	constrained graph

What is an exponential random graph model?

Probability space: The set $\mathcal{G}_n$ of all simple graphs G_n on n vertices.
Probability mass function:

$$\mathbb{P}_n^{\beta}(G_n) = \exp\left(n^2(\beta_1 t(H_1, G_n) + ... + \beta_k t(H_k, G_n) - \psi_n^{\beta})\right).$$

- $\beta_1, ..., \beta_k$ are real parameters and $H_1, ..., H_k$ are pre-chosen finite simple graphs. Each H_i has vertex set $V(H_i) = \{1, ..., k_i\}$ and edge set $E(H_i)$. By convention, we take H_1 to be a single edge.
- Graph homomorphism $\hom(H_i, G_n)$ is a random vertex map $V(H_i) \to V(G_n)$ that is edge-preserving. Homomorphism density $t(H_i, G_n) = \frac{|\hom(H_i, G_n)|}{|V(G_n)|^{|V(H_i)|}}$.
- Normalization constant:

$$\psi_n^{\beta} = \frac{1}{n^2} \log \sum_{G_n \in \mathcal{G}_n} \exp\left(n^2(\beta_1 t(H_1, G_n) + ... + \beta_k t(H_k, G_n))\right).$$

What is an exponential random graph model?
Probability space: The set $\mathcal{G}_n$ of all simple graphs G_n on n vertices.
Probability mass function:

$$\mathbb{P}_n^{\beta}(G_n) = \exp\left(n^2(\beta_1 t(H_1, G_n) + ... + \beta_k t(H_k, G_n) - \psi_n^{\beta})\right).$$

- $\beta_1, ..., \beta_k$ are real parameters and $H_1, ..., H_k$ are pre-chosen finite simple graphs. Each H_i has vertex set $V(H_i) = \{1, ..., k_i\}$ and edge set $E(H_i)$. By convention, we take H_1 to be a single edge.
- Graph homomorphism $\hom(H_i, G_n)$ is a random vertex map $V(H_i) \to V(G_n)$ that is edge-preserving. Homomorphism density $t(H_i, G_n) = \frac{|\hom(H_i, G_n)|}{|V(G_n)|^{|V(H_i)|}}$.
- Normalization constant:

$$\psi_n^{\beta} = \frac{1}{n^2} \log \sum_{G_n \in \mathcal{G}_n} \exp\left(n^2(\beta_1 t(H_1, G_n) + ... + \beta_k t(H_k, G_n))\right).$$

$\beta_i = 0$ for $i \geq 2$:

$$\begin{aligned}\mathbb{P}_n^\beta(G_n) &= \exp\left(n^2(\beta_1 t(H_1, G_n) - \psi_n^\beta)\right) \\ &= \exp\left(2\beta_1 |E(G_n)| - n^2 \psi_n^\beta\right).\end{aligned}$$

Erdős-Rényi graph $G(n, \rho)$,

$$\mathbb{P}_n^\rho(G_n) = \rho^{|E(G_n)|}(1-\rho)^{\binom{n}{2} - |E(G_n)|}.$$

Include edges independently with parameter $\rho = e^{2\beta_1}/(1 + e^{2\beta_1})$.

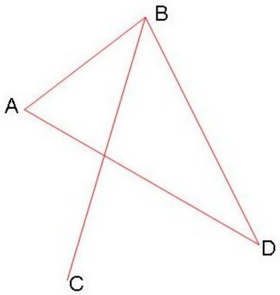
$$\exp(n^2 \psi_n^\beta) = \sum_{G_n \in \mathcal{G}_n} \exp\left(2\beta_1 |E(G_n)|\right) = \left(\frac{1}{1-\rho}\right)^{\binom{n}{2}}.$$

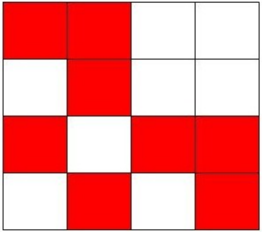
What happens with general β_i?
Problem: Graphs with different numbers of vertices belong to different probability spaces!
Solution: Theory of graph limits (graphons)! (Borgs, Chayes, Kahn, Lovász, Sós, Szegedy, Simonovits, Vesztergombi, Khare, Rajaratnam,...; earlier work of Aldous and Hoover)
Graphon space $\mathcal{W}$ is the space of all symmetric measurable functions $h(x,y)$ from $[0,1]^2$ into $[0,1]$. The interval $[0,1]$ represents a 'continuum' of vertices, and $h(x,y)$ denotes the probability of putting an edge between x and y.

What happens with general β_i?
Problem: Graphs with different numbers of vertices belong to different probability spaces!
Solution: Theory of graph limits (graphons)! (Borgs, Chayes, Kahn, Lovász, Sós, Szegedy, Simonovits, Vesztergombi, Khare, Rajaratnam,...; earlier work of Aldous and Hoover)
Graphon space $\mathcal{W}$ is the space of all symmetric measurable functions $h(x, y)$ from $[0, 1]^2$ into $[0, 1]$. The interval $[0, 1]$ represents a 'continuum' of vertices, and $h(x, y)$ denotes the probability of putting an edge between x and y.

Example: Erdős-Rényi graph $G(n,\rho)$, $h(x,y)=\rho$.
Example: Any $G_n \in \mathcal{G}_n$,

$$h(x,y)=\begin{cases} 1, & \text{if } (\lceil nx \rceil, \lceil ny \rceil) \text{ is an edge in } G_n; \\ 0, & \text{otherwise.} \end{cases}$$





Why are we interested in exponential random graph models?

Dependence between the random edges is defined through certain finite subgraphs H_i, in imitation of the use of potential energy to provide dependence between particle states in a grand canonical ensemble of statistical physics. By varying the activity parameters β_i, one could analyze the extent to which specific values of the subgraph densities interfere with one another.

Why are we interested in exponential random graph models? Dependence between the random edges is defined through certain finite subgraphs H_i, in imitation of the use of potential energy to provide dependence between particle states in a grand canonical ensemble of statistical physics. By varying the activity parameters β_i, one could analyze the extent to which specific values of the subgraph densities interfere with one another.

The hierarchy

$$
\begin{array}{cc}
\text{grand canonical ensemble} & \text{normalization constant} \\
\downarrow & \downarrow \\
\text{canonical ensemble} & \text{conditional normalization constant} \\
\downarrow & \downarrow \\
\text{microcanonical ensemble} & \text{entropy}
\end{array}
$$

The normalization constant encodes useful information about the structure of the measure. By differentiating the normalization constant with respect to appropriate parameters, averages of various quantities of interest may be derived. Computation of the normalization constant is also essential in statistics because it is crucial for carrying out maximum likelihood estimates and Bayesian inference of unknown parameters.

Important contributors:

Pioneers: Holland and Leinhardt (statistics); Frank and Strauss (statistics); Häggström and Jonasson (probability)

Physics: Park and Newman; Zuev, Eisenberg, and Krioukov; Anand and Bianconi

Probability, Statistical Physics: Chatterjee and Diaconis; Bhamidi, Bresler, and Sly; Chatterjee and Dembo; Borgs, Chayes, Cohn, and Zhao; Kenyon, Radin, Ren, and Sadun; Lubetzky and Peres; Lubetzky and Zhao; Aristoff and Zhu

Statistics, Psychology, Sociology: Wasserman and Faust; Snijders, Pattison, Robins, and Handcock; Rinaldo, Fienberg, and Zhou; Hunter, Handcock, Butts, Goodreau, and Morris

...

Large deviation and Concentration of measure:

$$\psi_n^\beta \asymp \max_{h \in \mathcal{W}} \left(\beta_1 t(H_1, h) + \ldots + \beta_k t(H_k, h) - \int_{[0,1]^2} I(h) dx dy \right),$$

where:

$$t(H_i, h) = \int_{[0,1]^{k_i}} \prod_{(i,j) \in E(H_i)} h(x_i, x_j) dx_1 \ldots dx_{k_i},$$

and $I : [0,1] \to \mathbb{R}$ is the function

$$I(u) = \frac{1}{2} u \log u + \frac{1}{2}(1-u)\log(1-u).$$

Let F^* be the set of maximizers. G_n lies close to F^* with high probability for large n. When the parameters β_i are non-negative, the graph behaves like an Erdős-Rényi random graph in the large n limit, where the edge formation probability ρ depends on all parameters. Nevertheless, not much is known when some of the parameters are negative. (Chatterjee and Varadhan; Chatterjee and Diaconis)

The standard model is centered on dense graphs (number of edges comparable to the square of number of vertices), but most networks data are sparse in the real world. What would be a typical random graph drawn from a sparse exponential model?

Let $\beta_i^{(n)} = \beta_i \alpha_n$ where $\alpha_n \to \infty$ as $n \to \infty$. For β_i negative this ensures that $\beta_i^{(n)} \to -\infty$ and translates to sparse graphs.
(sparse) Probability mass function:

$$\mathbb{P}_n^{\beta}(G_n) = \exp\left(n^2(\alpha_n \beta_1 t(H_1, G_n) + ... + \alpha_n \beta_k t(H_k, G_n) - \psi_n^{\beta})\right).$$

(sparse) Normalization constant:

$$\psi_n^{\beta} = \frac{1}{n^2} \log \sum_{G_n \in \mathcal{G}_n} \exp\left(n^2(\alpha_n \beta_1 t(H_1, G_n) + ... + \alpha_n \beta_k t(H_k, G_n))\right).$$

Let $X_{ij} = 1$ when there is an edge between vertex i and vertex j and $X_{ij} = 0$ otherwise. Assume that $\beta_1, ..., \beta_k$ are all negative. If $\lim_{n\to\infty} n^2 e^{2\alpha_n \beta_1} = 0$ and $\lim_{n\to\infty} \frac{\alpha_n}{n} = 0$, then

$$\lim_{n\to\infty} \frac{\mathbb{P}_n^{\beta}(X_{1i} = 1)}{e^{2\alpha_n \beta_1}} = 1.$$

$$\lim_{n\to\infty} \frac{\mathbb{P}_n^{\beta}(X_{1i} = 1, X_{1j} = 1)}{e^{4\alpha_n \beta_1}} = 1, \quad i \neq j.$$

This indicates that when the rate of divergence of α_n is between the order of $\log n$ and n, the graph displays Erdős-Rényi behavior in the large n limit, where the edge formation probability $\rho = e^{2\alpha_n \beta_1}$. It only depends on β_1 and n and decays to 0 as $n \to \infty$. This is in sharp contrast to the standard exponential model where the parameters $\beta_1, \ldots, \beta_k$ are not scaled by α_n and are instead held fixed. (Y and Zhu)

Sketch of Proof.

$$\mathbb{P}_n^\beta(X_{1i}=1)=\frac{2^{\binom{n}{2}}\mathbb{E}[X_{1i}e^{\alpha_n n^2\sum_{p=1}^k \beta_p t(H_p,G_n)}]}{2^{\binom{n}{2}}\mathbb{E}[e^{\alpha_n n^2\sum_{p=1}^k \beta_p t(H_p,G_n)}]}.$$

$$\mathbb{P}_n^\beta(X_{1i}=1,X_{1j}=1)=\frac{2^{\binom{n}{2}}\mathbb{E}\left[X_{1i}X_{1j}e^{\alpha_n n^2\sum_{p=1}^k \beta_p t(H_p,G_n)}\right]}{2^{\binom{n}{2}}\mathbb{E}[e^{\alpha_n n^2\sum_{p=1}^k \beta_p t(H_p,G_n)}]}.$$

Denominator:

$$2^{\binom{n}{2}}\mathbb{E}[e^{\alpha_n n^2\sum_{p=1}^k \beta_p t(H_p,G_n)}]\geq 2^{\binom{n}{2}}\frac{1}{2^{\binom{n}{2}}}=1.$$

$$\begin{aligned}2^{\binom{n}{2}}\mathbb{E}[e^{\alpha_n n^2\sum_{p=1}^k \beta_p t(H_p,G_n)}] &\leq 2^{\binom{n}{2}}\mathbb{E}[e^{\alpha_n n^2\beta_1 t(H_1,G_n)}]\\ &= \left(1+e^{2\alpha_n\beta_1}\right)^{\binom{n}{2}}\to 1.\end{aligned}$$

Sketch of Proof Continued.
Numerator of the mean:

$$
\begin{aligned}
& 2^{\binom{n}{2}}\mathbb{E}\left[X_{1i}e^{\alpha_n n^2 \sum_{p=1}^{k}\beta_p t(H_p,G_n)}\right] \\
\geq\ & 2^{\binom{n}{2}}\mathbb{E}\left[e^{\alpha_n n^2 \sum_{p=1}^{k}\beta_p t(H_p,G_n)}\middle| X_{1i}=1, X_{i'j'}=0, (i',j')\neq(1,i)\right] \\
& \cdot\mathbb{P}\left(X_{1i}=1, X_{i'j'}=0, (i',j')\neq(1,i)\right) \\
=\ & e^{2\alpha_n\beta_1+\alpha_n n^2\sum_{p=2}^{k}\beta_p c_p n^{-|V(H_p)|}} \asymp e^{2\alpha_n\beta_1}.
\end{aligned}
$$

$$
\begin{aligned}
2^{\binom{n}{2}}\mathbb{E}[X_{1i}e^{\alpha_n n^2\sum_{p=1}^{k}\beta_p t(H_p,G_n)}] &\leq 2^{\binom{n}{2}}\mathbb{E}[X_{1i}e^{\alpha_n n^2\beta_1 t(H_1,G_n)}] \\
&= e^{2\alpha_n\beta_1}\left(1+e^{2\alpha_n\beta_1}\right)^{\binom{n}{2}-1} \\
&\asymp e^{2\alpha_n\beta_1}.
\end{aligned}
$$

Sketch of Proof Continued.
Numerator of the variance:

$$2^{\binom{n}{2}}\mathbb{E}\left[X_{1i}X_{1j}e^{\alpha_n n^2\sum_{p=1}^{k}\beta_p t(H_p,G_n)}\right]$$
$$\geq 2^{\binom{n}{2}}\mathbb{E}\left[e^{\alpha_n n^2\sum_{p=1}^{k}\beta_p t(H_p,G_n)}\middle| X_{1i}=X_{1j}=1, X_{i'j'}=0, (i',j')\neq(1,i) \text{ and } (i',j')\neq(1,j)\right]$$
$$\cdot\mathbb{P}\left(X_{1i}=X_{1j}=1, X_{i'j'}=0, (i',j')\neq(1,i) \text{ and } (i',j')\neq(1,j)\right)$$
$$= e^{4\alpha_n\beta_1+\alpha_n n^2\sum_{p=2}^{k}\beta_p c_p n^{-|V(H_p)|}} \asymp e^{4\alpha_n\beta_1}.$$

$$\begin{aligned}
2^{\binom{n}{2}}\mathbb{E}[X_{1i}X_{1j}e^{\alpha_n n^2\sum_{p=1}^{k}\beta_p t(H_p,G_n)}] &\leq 2^{\binom{n}{2}}\mathbb{E}[X_{1i}X_{1j}e^{\alpha_n n^2\beta_1 t(H_1,G_n)}] \\
&= e^{4\alpha_n\beta_1}\left(1+e^{2\alpha_n\beta_1}\right)^{\binom{n}{2}-2} \\
&\asymp e^{4\alpha_n\beta_1}.
\end{aligned}$$

Region of sparsity in the edge-p-star model:
Under suitable assumptions,

$$\psi_n^{\beta} \asymp \frac{1}{2} e^{2\alpha_n \beta_1}, \quad \beta_1 < 0 \text{ and } \beta_1 + \beta_2 \leq 0,$$

coincides with the limiting normalization constant of an Erdős-Rényi random graph with edge formation probability $\rho = e^{2\alpha_n \beta_1}$.

$$\psi_n^{\beta} \asymp \frac{1-p}{2p} \gamma_n e^{\gamma_n}, \quad \beta_1 = 0 \text{ and } \beta_2 < 0,$$

where $2\alpha_n \beta_2 e^{(p-1)\gamma_n} p = \gamma_n$. (Y and Zhu)

Similar analysis may be done for directed sparse exponential random graph models parametrized by edges and multiple outward stars.

Probability space: The set $\mathcal{X}_n$ of all simple digraphs X_n on n vertices. Let $X_{ij} = 1$ when there is a directed edge from vertex i to vertex j and $X_{ij} = 0$ otherwise.

Outward directed p-star homomorphism density of X_n:

$$s_p(X_n) = n^{-p-1} \sum_{1 \le i, j_1, \dots, j_p \le n} X_{ij_1} \cdots X_{ij_p} = n^{-p-1} \sum_{i=1}^{n} \left(\sum_{j=1}^{n} X_{ij} \right)^p .$$

Similar analysis may be done for directed sparse exponential random graph models parametrized by edges and multiple outward stars.

Probability space: The set $\mathcal{X}_n$ of all simple digraphs X_n on n vertices. Let $X_{ij} = 1$ when there is a directed edge from vertex i to vertex j and $X_{ij} = 0$ otherwise.

Outward directed p-star homomorphism density of X_n:

$$s_p(X_n) = n^{-p-1} \sum_{1 \le i, j_1, \dots, j_p \le n} X_{ij_1} \cdots X_{ij_p} = n^{-p-1} \sum_{i=1}^{n} \left(\sum_{j=1}^{n} X_{ij} \right)^p .$$

Let $\beta_i^{(n)} = \beta_i \alpha_n$ where $\alpha_n \to \infty$ as $n \to \infty$. For β_i negative this ensures that $\beta_i^{(n)} \to -\infty$ and translates to sparse graphs.
(sparse) Probability mass function:

$$\mathbb{P}_n^{\beta}(X_n) = \exp\left(n^2(\alpha_n \beta_1 s_1(X_n) + ... + \alpha_n \beta_k s_k(X_n) - \psi_n^{\beta})\right).$$

(sparse) Normalization constant:

$$\psi_n^{\beta} = \frac{1}{n^2} \sum_{X_n \in \mathcal{X}_n} \exp\left(n^2(\alpha_n \beta_1 s_1(X_n) + ... + \alpha_n \beta_k s_k(X_n))\right).$$

Assume that $\beta_1, ..., \beta_k$ are all negative. If $\lim_{n\to\infty} ne^{\alpha_n\beta_1} = 0$ and $\lim_{n\to\infty} \frac{\alpha_n}{n} = 0$, then

$$\lim_{n\to\infty} \frac{\mathbb{P}_n^{\beta}(X_{1i} = 1)}{e^{\alpha_n\beta_1}} = 1.$$

$$\lim_{n\to\infty} \frac{\mathbb{P}_n^{\beta}(X_{1i} = 1, X_{1j} = 1)}{e^{2\alpha_n\beta_1}} = 1, \quad i \neq j.$$

This indicates that when the rate of divergence of α_n is between the order of $\log n$ and n, the graph displays Erdős-Rényi behavior in the large n limit, where the edge formation probability $\rho = e^{\alpha_n\beta_1}$. It only depends on β_1 and n and decays to 0 as $n \to \infty$. (Y and Zhu)

Consider an Erdős-Rényi random graph on n vertices with edge formation probability ρ. The distribution of the degree of any vertex i is Binomial with parameters n and ρ. A known fact is that for n large, ρ small and $n\rho$ a constant, Binomial distribution with these parameters tends to a Poisson distribution with parameter $n\rho$. Correspondingly, if $ne^{\alpha_n \beta_1}$ approaches a constant $\lambda \in (0, \infty)$ as $n \to \infty$, i.e., when the divergence rate of α_n is of the order of $\log n$, will the graph display Poisson behavior?

Assume that $\beta_1, ..., \beta_k$ are all negative. If $\lim_{n\to\infty} ne^{\alpha_n \beta_1} = \lambda \in (0, \infty)$, then

$$\lim_{n\to\infty} \frac{\mathbb{P}_n^{\beta}(X_{1i} = 1)}{\lambda n^{-1}} = 1.$$

$$\lim_{n\to\infty} \frac{\mathbb{P}_n^{\beta}(X_{1i} = 1, X_{1j} = 1)}{\lambda^2 n^{-2}} = 1, \quad i \neq j.$$

Moreover, the degree of any vertex is asymptotically Poisson with parameter λ,

$$\sum_{i=1}^{n} X_{1i} \to \text{Poisson}(\lambda).$$

(Y and Zhu)

Thank You!:)