Random Graphs:

A general model of a social network.

→ ER (Erdős-Rényi) Random Graphs.

Two Ways of Describing Random Graphs:

Closely related variants of ER Random Graphs.

- G(n, m) Models:
- G(n, p) Models:

A G(n, m) Models:

A graph \( G = (V, E) \) is chosen uniformly at random from the collection of all graphs, which have \( |V| = n \) nodes and \( |E| = m \) edges.

\( G(3, 2) \) - Model:

\[
\begin{align*}
&\quad 1 \\
2 &\quad 3 \\
\end{align*}
\]

\[
\Pr = \{ \frac{1}{3}, \frac{1}{3}, \frac{1}{3}, \}
\]

- Exactly three possible graphs on three vertices and two edges.
- Each is selected with equal probability \( = \frac{1}{3} \).
\( G(n,p) \) Models

A graph \( G = (V,E) \) is constructed by connecting every pair of nodes uniformly randomly with \( Pr[eeE] = p \).

For every pair of vertices \( u, v \in V \) an edge \( e = (u,v) \in E \) is included in the graph with probability \( p \) independent from every other edge.

Equivalently, all graphs with \( |V| = n \) & \( |E| = M \) have equal edge probability of \( p = \frac{M}{\binom{n}{2}} \) = Density of the graph.

\( p = \frac{1}{2} \), all graphs on \( n \) vertices are chosen with equal probability \( = p^k (1-p)^{n-k} = 2^{-\binom{n}{2}} = 2^{-\frac{n(n-1)}{2}} \).

As \( p \) increases from 0 to 1, the model produces denser graphs with higher probability (than sparser graphs).

\( \circ \) Expected Number of Edges:
\[ \langle |E| \rangle = \binom{n}{2} p = \frac{n(n-1)p}{2} \]

\( \circ \) Expected Degree.
\[ \bar{d} = \langle d \rangle = \frac{2 \langle |E| \rangle}{|V|} = \frac{2 \binom{n}{2} p}{n} = \frac{(n-1)p}{2} \]

There are \( (n-1) \) possible other vertices, of which each can be adjacent with probability \( = p \).
\[ d(v) \sim \text{Bin}(n-1, p) \]

\[ \Pr[d(v) = k] = \binom{n-1}{k} p^k (1-p)^{n-1-k}. \]

The degree of a vertex in a graph \( G \in \mathcal{G}(np) \) is distributed as a Binomial.

\[ \mu[d(v)] = (n-1)p \quad \sigma^2[d(v)] = (n-1)p^2. \]

Asymptotic Analysis:

\[ |V| = n \to \infty \]

Random Graphs are often studied in the asymptotic case, as \( |V| \to \infty \) (the number of vertices) tends to infinity.

\( \Rightarrow \) Graphon, Graphlet.

If expected degree \( \bar{d} \) is held constant (independent of \( n \))

\[ \bar{d} = (n-1)p = \text{const} = \lambda \]

\[ p = \frac{\lambda}{n-1} \]

\[ \Pr[d(v) = k] \approx \frac{(n-1)^k}{k!} p^k (1-p)^{n-1-k} \]

\[ \approx \frac{(n-1)^k}{k!} \left( \frac{\lambda}{n-1} \right)^k e^{-(n-1)p} \]

\[ \approx \frac{\lambda^k}{k!} e^{-\lambda} \]

\[ d(v) \sim \text{Poisson}(\lambda) \]

\[ \mu[d(v)] = \lambda \quad \sigma^2[d(v)] = \lambda \]

Poisson Approximation.
Tipping Point

\[ \text{Phase Transition} \]
\[ 0-1 \text{ Law.} \]

Small \( p \): If \( p < \frac{(1-e) \ln n}{n} \), then a graph in \( G(n,p) \) will a.s. contain isolated vertices \( \Rightarrow (\text{DISCONNECTED}) \).

Large \( p \): If \( p > \frac{(1+e) \ln n}{n} \), then a graph in \( G(n,p) \) will a.s. be CONNECTED.

\[ \Pr(\text{connected}) \]

Questions about Random Social Network Models:

1) Does the network have isolated nodes? Cycles? Giant components?

2) What are the probabilities of such events?

3) Asymptotic Analysis:
   Compute Probabilities, as \( n \to \infty \).

Threshold Functions for Connectivity: (Erdős–Rényi 1961).

A threshold function for the connectivity of the Erdős–Rényi model \( G(n,p) \) is

\[ t(n) = \frac{\ln n}{n} \]
For a graph $G \in G(n, \frac{\lambda \cdot \ln n}{n})$

\[
\Pr[G = \text{connected}] = \begin{cases} 
0, & \text{if } \lambda < 1; \\
1, & \text{if } \lambda > 1. 
\end{cases}
\]

\[1_i = \begin{cases} 
1, & \text{if node } i \text{ is isolated;} \\
0, & \text{otherwise.}
\end{cases}\]

\[1_i \sim \text{Bernoulli}(\pi) \quad \Rightarrow \quad \pi = \Pr[1_i = 1] = (1-p)^{n-1} = (1-p)^{\frac{\lambda \cdot \ln n}{n}} = e^{-\frac{\lambda \cdot \ln n}{n}} = e^{-\lambda \ln n} = n^{-\lambda}.\]

$X = \sum 1_i = \text{Total } \# \text{ of isolated nodes.}$

\[\mathbb{E}[X], \mathbb{V}[X] = \ln n \ln n\]

\[\mathbb{E}[X] = n \cdot n^{-\lambda} = n^{1-\lambda} \rightarrow \begin{cases} 
\infty, & \text{if } \lambda < 1; \\
0, & \text{if } \lambda > 1. 
\end{cases}\]

Problems

(1) Need to show that $\Pr[x=0] = 0$, if $\lambda < 1$.

(2) Note that $\Pr[x=0] > 0$ does not necessarily imply that the graph is connected.
Assume that \( \lambda < 1 \).

\[
\text{Var}[x] = \sum_i \text{var}(I_i) + \sum_{i \neq j} \text{cov}(I_i, I_j)
\]
\[
= n \text{var}(I_1) + n(n-1) \text{cov}(I_1, I_2)
\]
\[
\begin{align*}
\text{var}(I_1) &= \pi(1-\pi) = \pi - \pi^2 \\
\text{cov}(I_1, I_2) &= E(I_1I_2) - E(I_1)E(I_2) \\
&= (1-\pi)^{2n-3} - (1-\pi)^{n-1}(1-\pi)^{n-1} \\
&= \frac{n^2 \pi^2}{1-\pi} - \pi^2
\end{align*}
\]

\[
\text{Var}[x] = n\pi - n\pi^2 + \frac{n^2 \pi^2}{1-\pi} - n^2 \pi^2
\]
\[
\approx n\pi + n^2 \pi^2 p
\]
\[
= n \cdot n^{-\lambda} + n^2 n^{-2\lambda} p
\]
\[
\approx n^{1-\lambda} \cdot E[x] \quad (\text{if } \lambda < 1)
\]

\[
\text{Var}[x] = \int (x - E[x])^2 p(x) \, dx
\]
\[
\geq (0 - E[x])^2 \text{Pr}[x = 0]
\]
\[
\text{Pr}[x = 0] \leq \frac{\text{Var}[x]}{E[x]} \leq \frac{n^{2\lambda} \text{Var}[x]}{n^{1-\lambda} E[x]^2}
\]
\[
\frac{\text{Var}[x]}{E[x]^2} \approx \frac{1}{E[x]} \rightarrow 0.
\]
"Graph is disconnected"

\[ \Rightarrow \exists V', |V'| = k \quad \text{Connect } \left[ V', V \cup V' \right] \]

\[ k \leq n/2 \]

\[ \Pr \left[ V' \text{ is not connected to } V \cup V', |V'| = k \right] = (1-p)^k \binom{n-k}{k} \]

\[ = (1-p)^k \cdot \frac{k^{n-k}}{k!} \cdot e^{-\lambda k(n-k)/\ln n} \]

\[ \Pr \left[ \exists V', |V'| = k \quad V' \text{ is not connected to } V \cup V' \right] \]

\[ = \left( \frac{n}{k} \right) (1-p)^k \binom{n-k}{k} \]

\[ = \frac{n!}{k! (n-k)!} \cdot e^{-\lambda k(n-k)/\ln n/n} \]

\[ \left\{ \begin{array}{l}
\lambda > \frac{n}{2} \\
\lambda \leq \frac{n}{2}
\end{array} \right\} \]

\[ \approx \sqrt{\frac{2}{\pi n}} \cdot \left( \frac{\eta e}{2} \right)^n \cdot e^{-\frac{\lambda n^2}{4} \cdot \ln n/n} \]

\[ \approx \frac{1}{\sqrt{2\pi n}} \cdot 2^n \cdot n^{-\lambda n/4} \]

\[ \lambda > 1 \quad \left\{ \begin{array}{l}
\approx \sqrt{\frac{2}{\pi n}} \cdot 2^n - n ! \ln n/4
\end{array} \right\} \]

\[ \Pr \left[ \text{Graph is disconnected} \right] \]

\[ \approx \sum_{k=\frac{n}{2}}^{n/2} \binom{n}{k} (1-p)^k \binom{n-k}{k} \]

\[ < \frac{n}{2} \sqrt{\frac{2}{\pi n}} \cdot 2^{[-h \ln n/4 - n]} \]

\[ = \sqrt{\frac{n}{2\pi}} \cdot 2^{[-h \ln n/4 - n]} \]

\[ = \frac{1}{\sqrt{2\pi}} \cdot 2^{[-h \ln n/4 - n - \ln 2]} \]

\[ \rightarrow 0 \quad \text{as } n \rightarrow \infty \]