GIANT COMPONENT.

(A) When $p(n) \ll \frac{\ln n}{n}$ — edge probability

the Erdős-Rényi Graphs $G(n, p(n))$ is d.s. disconnected.

(B) In this regime [i.e., $p(n) \ll \frac{\ln n}{n}$]

$$X = \sum_{i \neq j} I\{i \not\sim j\}, \quad E[X] = n(n-1)(1-p(n))^{n-2}$$

$$E[X] = n \left(1 - p(n)\right)^{n-1} p(n) = n e^{-p(n)n} = n e^{-(\ln 2)(1-e^{-\ln 2})} = n e^{-p(n)n} \to n e^{-\mu}$$

as $n \to \infty$

(C) Thus, when $p(n) \ll \frac{\ln n}{n}$

the graph has an arbitrarily large number of connected components.

We will study another interesting regime that gives rise to GIANT COMPONENTS.
Two Regimes:

\[ p(n) = \begin{cases} \frac{\lambda}{n} & \text{if } \lambda < 1 \\ \lambda & \text{if } \lambda > 1 \end{cases} \]

For \( \lambda < 1 \), all components of the graph are "small".
For \( \lambda > 1 \), one component of the graph is a "giant." (UNIQUE)

A unique giant component
A component that remains a constant
growth of individuals in the social network.

Best way to think about this structure:

**BREADTH-FIRST SEARCH**
Think of two processes (related)
Assume \( \lambda < 1 \).
Graph Process: \( Z_k^G \) = # individuals at stage \( k \) of the graph.

Branching Process: \( Z_k^B \) = # individuals in a pure branching process.

\( Z_k^G < Z_k^B \) (e.g., triadic closure)
Note

Expected # of children for a node
\[ = n p(n) = n \cdot \frac{\lambda}{m} = \lambda \]

\[ E[\mathbb{Z}_k^B] = \lambda^k \]

\[ S_i = \# \text{ nodes in the Erdős-Rényi graph connected to individual } 1 \]

\[ E[S_i] = \sum_k E[\mathbb{Z}_k^B] < \sum_k E[\mathbb{Z}_k^B] = \sum_k \lambda^k = \frac{1}{1 - \lambda} \]

Theorem

Let \( p(n) = \frac{\lambda}{n} \quad (\lambda < 1) \)

For all sufficiently large \( \alpha > 0 \), we have

\[ P(\max_{1 \leq i \leq n} |S_i| \geq \alpha \ln n) \to 0 \text{ as } n \to \infty. \]

where \( |S_i| = \text{ size of the component containing the individual i} \)

proof: Omitted. \( \square \)
Assume \( \lambda > 1 \).

Giant component with \( p(n) = \frac{\lambda}{n} \) (\( \lambda > 1 \)).

**Claim**
\[ Z_k^6 \geq Z_k^5 \quad \text{when} \quad \lambda^k \leq o(\sqrt{n}). \quad \square \]

Estimate the expected number of conflicts at stage \( k+1 \).

Two of the “friends” at stage \( k \) have a common friend at stage \( k+1 \). (Triadic Closure).

\[
E \left[ \text{Number of conflicts at stage } k+1 \right] = E \left[ \binom{Z_k}{2} \left( \frac{Z_k}{n} \right) p^2 \right] \leq np^2 E[Z_k^2] \\
= np^2 \left\{ E[Z_k]^2 + \text{Var}[Z_k] \right\} \\
\text{where} \quad Z_k \sim \text{Poisson} \left( \lambda^k \right), \quad E[Z_k] = \text{Var}[Z_k] = \lambda^k \\
= np^2 \left( \frac{\lambda^k + \lambda^k}{2} \right) \\
\leq \frac{\lambda^k}{n} \cdot \frac{n^2 \lambda^k}{2} = \frac{\lambda^{2(k+1)}}{2n}
\]

Thus conflicts become non-negligible when \( \lambda^k \sim \sqrt{n} \).
THEOREM

Let $p(n) = \frac{c}{n}$ ($c > 1$) in an Erdős–Rényi random graph $G(n, p(n))$. Then, there exists some $c > 0$ such that

$$\Pr\left[ \exists \text{ a component of size } \geq c\sqrt{n} \text{ nodes} \right] \rightarrow 1 \text{ as } n \to \infty$$

Thus, between any two components of size $\sqrt{n}$, the probability of having at least one link

$$\Pr\left[ \text{ there exists a link between } C_1 \text{ and } C_2 \right]$$

$$\geq 1 - (1 - p(n))^{\left| C_1 \right| \cdot \left| C_2 \right|}$$

$$= 1 - (1 - \frac{c}{n})^{\frac{\sqrt{n}}{c}} = 1 - \left(1 - \frac{c}{n}\right)^{\frac{n}{c}} \cdot e^2$$

$$= 1 - e^{-e^2} > 0$$

A positive constant independent of $n$.

Thus, components of size $\leq \sqrt{n}$ connect to each other forming a connected component of size $9n$ for some $0 < q < 1$.

GIANT COMPONENT.
CONTAGION AND DIFFUSION

Assume:

- A social network of $n$ individuals.
- In which a randomly chosen individual is infected with a contagion virus.
- The social network is described by an Erdős–Rényi random graph with link probability $p(n)$.
- Any individual is immune to the virus with probability $\pi$.

Generate Erdős–Rényi random graph $G(n, p)$.
Delete $\pi n$ of the nodes at random.
Identify the component that the initially infected individual lies in...

$$G \left( (1-\pi)n, \frac{\binom{n}{2} p(n)}{(1-\pi)n} \right)$$
Three regimes:

(I) \( p(1-\pi)n < 1 \)

\[ E \left[ \text{Size (as a fraction)} \right] < \frac{\ln n}{n} \to 0 \]

(II) \( 1 < p(1-\pi)n < \log [(1-\pi)n] \)

\[ E \left[ \text{Size (as a fraction)} \right] = \frac{q \cdot q(1-\pi)n + (1-q) \cdot \ln n}{n} = q^2(1-\pi) \]

(III) \( p(1-\pi)n > \log [(1-\pi)n] \)

\[ E \left[ \text{Size (as a fraction)} \right] = (1-\pi) \]

\( q \): Fraction of nodes in the giant component of the random graph with \((1-\pi)n\) nodes:

\[ q = 1 - e^{-q(1-\pi)n}p \]
RANDOM SURFER MODEL

Imagine a web surfer bouncing along randomly following the hyper-link graph of the web.

When the surfer arrives at a node he chooses at random, the hyper-links (directed edge) to a new node.

Asymptotically, the proportion of time the random surfer spends on a given node/page is a measure of relevance (relative importance) of that node.

Dangling Nodes $\rightarrow$ Sinks
or Periodicity in the Graph may get the surfer trapped in a limit state or a limit cycle.

$\rightarrow$ Stochastic Teleportation
PAGE RANK.

Each node is important if it is cited by other important nodes.

Node $j \Rightarrow W(j) =$ Page Rank Value.

$A =$ Adjacency Matrix

d$(i)$ = Out-degree of node $i$

$W(j) = \sum \frac{w(i)}{dout(i)} A_{ij}$

Let $P_{ij} = \frac{A_{ij}}{\text{dout}(i)} \Rightarrow P =$ Stochastic Matrix

$\omega^T = \omega^T P$

WITH TELEPORTATION:
Jump to a random node with prob $(1-s)$, $0.8 \approx 0.9$

$\omega^T = s \omega^T P + \frac{(1-s)}{n} \text{e}^T$

Iterative Algorithm:

$W_{k+1}^T = s W_k^T P + \frac{(1-s)}{n} \text{e}^T$