# Week #5: Networks

Stochastic Methods Course

Notes by: Francisco Richter

May 15, 2025

## 1 Random Networks

Random network models provide a probabilistic framework for studying the structure and dynamics of complex systems. These models have applications in sociology, epidemiology, computer science, and finance, where phenomena such as information diffusion, epidemic spreading, and social connectivity are analyzed.

### 1.1 Fundamental Concepts

**Definition. Graph**

A graph $G$ is defined as a pair $G = (V, E)$, where:

- $V$ is a set of nodes (or vertices),
- $E$ is a set of edges connecting pairs of nodes.

**Example 1.1.** *Consider a graph with five nodes:*

$$V = \{1, 2, 3, 4, 5\},$$

*and an edge set:*

$$E = \{(1, 2),\ (2, 3),\ (1, 3),\ (2, 4),\ (3, 4),\ (4, 5)\}.$$

*Figure 1 illustrates this graph.*

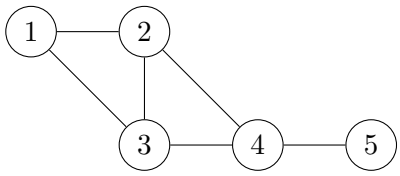


Figure 1: A simple graph with five nodes and six edges.

**Definition. Adjacency Matrix**

For a graph $G = (V, E)$ with $N = |V|$ nodes, the *adjacency matrix* $A$ is an $N \times N$ matrix defined by:

$$A_{uv} = \begin{cases} 1, & \text{if there is an edge between nodes } u \text{ and } v, \\ 0, & \text{otherwise.} \end{cases}$$

**Example 1.2.** *For the graph above, the adjacency matrix is:*

$$A = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

The powers of the adjacency matrix reveal connectivity properties. In particular, $(A^k)_{uv}$ represents the number of distinct walks of length $k$ from node $u$ to node $v$.

**Theorem 1.1.** *Let $A$ be the adjacency matrix of a graph $G$. Then, the entry $(A^k)_{uv}$ equals the number of walks of length $k$ from node $u$ to node $v$.*

*Proof.* We prove by induction on $k$. For $k = 1$, the claim holds by definition. Assume that $(A^k)_{uv}$ counts the number of walks of length $k$ from $u$ to $v$. For $k + 1$, we have:

$$(A^{k+1})_{uv} = \sum_{w \in V} (A^k)_{uw} \, A_{wv}.$$

Each term $(A^k)_{uw}$ counts the number of walks of length $k$ from $u$ to $w$, and $A_{wv}$ indicates the presence of an edge from $w$ to $v$. Summing over all $w$ gives the total number of walks of length $k + 1$, completing the induction. □

**Example 1.3.** *For our graph, computing:*

$$A^2 = \begin{pmatrix} 2 & 1 & 1 & 2 & 0 \\ 1 & 3 & 1 & 1 & 1 \\ 1 & 1 & 3 & 1 & 1 \\ 2 & 1 & 1 & 3 & 0 \\ 0 & 1 & 1 & 0 & 1 \end{pmatrix}.$$

*For instance, $(A^2)_{11} = 2$ indicates that there are two walks of length 2 from node 1 back to itself.*

## 2 Random Network Models

Random network models provide a fundamental framework for generating graphs using simple probabilistic rules. These models help us understand how local random interactions can lead to the emergence of complex global network structures. In this section, we describe three influential models—the Erdős-Rényi model, the Watts-Strogatz model for small-world networks, and the Barabási-Albert model for scale-free networks—with an emphasis on the underlying principles and algorithms.

### Fundamental Definitions

**Definition. Degree Distribution**

For a graph $G = (V, E)$, the *degree* of a node is defined as the number of edges incident to it. The *degree distribution*, denoted by $P(k)$, is the probability that a randomly selected node has degree $k$. Formally,

$$P(k) = \frac{|\{v \in V : \text{degree}(v) = k\}|}{|V|},$$

with the normalization condition $\sum_k P(k) = 1$.

## 2.1 Erdős-Rényi Model

**Definition. Erdős-Rényi Model**

The Erdős-Rényi model, denoted by $G(n, p)$, is one of the simplest random graph models. It constructs a graph with $n$ nodes by considering each of the $\binom{n}{2}$ possible edges and including each edge independently with probability $p$. That is, for any pair of distinct nodes $u$ and $v$,

$$P\big((u, v) \in E\big) = p.$$

**Fundamentals of the Model:** The basic idea is that every potential connection between nodes is treated equally, with the same probability $p$ of existing. This simplicity allows for precise mathematical analysis while serving as a baseline for more complex network models.

**Key Properties and Analysis:**

- **Degree Distribution:** In $G(n, p)$, the degree $k$ of any node follows a binomial distribution:

  $$P(k) = \binom{n-1}{k} p^k (1-p)^{n-1-k}.$$

  For large $n$ with small $p$ such that $np$ remains constant, the distribution can be approximated by a Poisson distribution:

  $$P(k) \approx \frac{(np)^k e^{-np}}{k!}.$$

  This reflects the model's inherent randomness in edge formation.
- **Connectivity and Phase Transition:** A critical phenomenon occurs at the threshold

  $$p_c \approx \frac{\ln(n)}{n},$$

  where the graph transitions from having many small disconnected components to containing a single giant component.
- **Clustering Coefficient:** Because edges are formed independently, the clustering coefficient—the probability that two neighbors of a node are connected—is low, roughly equal to $p$.

**Applications and Limitations:** Due to its simplicity, the Erdős-Rényi model is mathematically tractable and provides valuable insights into the fundamental behavior of random graphs. However, its assumptions of independent and uniform edge formation limit its ability to capture clustering and degree heterogeneity observed in many real-world networks.

## 2.2 Small-World Networks: The Watts-Strogatz Model

Small-world networks are notable for their high local clustering and short average path lengths. The Watts-Strogatz model creates such networks by starting with an orderly structure and then introducing a controlled amount of randomness.

**Definition. Regular Ring Lattice**

A *regular ring lattice* with $n$ nodes is constructed by arranging the nodes in a circle. Each node is then connected to its $k$ nearest neighbors: $k/2$ neighbors on its left and $k/2$ neighbors on its right. This results in an undirected graph where every node has exactly $k$ connections.

**Definition.** **Watts-Strogatz Model**

Given a regular ring lattice with $n$ nodes and degree $k$, the Watts-Strogatz model introduces randomness by *rewiring* each edge with a probability $p$. For each edge $(i, j)$ in the lattice, with probability $p$:

1. Remove the edge $(i, j)$.
2. Choose a new node $l$ uniformly at random from all nodes such that $l \neq i$ and there is no existing edge between $i$ and $l$.
3. Add the edge $(i, l)$ to the graph.

If the random number exceeds $p$, the edge remains unchanged. This process avoids self-loops and duplicate edges.

**Simulation Algorithm:** To simulate the Watts-Strogatz model, follow these steps:

1. **Create the Regular Ring Lattice:**
   - Arrange $n$ nodes in a circle.
   - For each node $i$, connect it to its $k/2$ nearest neighbors on either side.
2. **Rewire the Edges:**
   - For each node $i$, iterate over each edge $(i, j)$ where $j$ is one of the $k/2$ neighbors in a single direction (to avoid processing each undirected edge twice).
   - Generate a random number $r \in [0, 1]$.
   - If $r < p$, remove the edge $(i, j)$ and select a new node $l$ uniformly at random from the set of nodes that are neither $i$ nor already connected to $i$. Then add the edge $(i, l)$.

**Key Properties:**

- **Clustering Coefficient:** The initial lattice has high clustering since neighbors of a node are also neighbors of each other. Even for small values of $p$, most local connections remain, preserving a high clustering coefficient.
- **Average Path Length:** The rewiring introduces long-range shortcuts, which greatly reduce the average shortest path length in the network. For $p = 0$, the path length is high (reflecting the regular structure), but for larger $p$, it drops toward values observed in random graphs.
- **Interplay Between Order and Randomness:** By varying $p$, one can smoothly transition from an ordered lattice ($p = 0$) to a random graph ($p = 1$), thereby exploring the balance between local structure and global efficiency.

This detailed description provides the fundamental steps and concepts needed to simulate the Watts-Strogatz model.

## 2.3 Scale-Free Networks: The Barabási-Albert Model

Scale-free networks are distinguished by the presence of a few highly connected nodes, or hubs, alongside many nodes with relatively few connections. The Barabási-Albert model explains this phenomenon through a process of preferential attachment.

**Definition.** **Barabási-Albert Model**

The Barabási-Albert model builds a network using the following iterative algorithm:

1. Start with a small, connected network of $m_0$ nodes.

2. At each time step, add a new node with $m$ ($m \leq m_0$) edges. These new edges connect the new node to $m$ distinct nodes already present in the network.
3. The probability $\Pi(k_i)$ that the new node attaches to an existing node $i$ is proportional to the degree $k_i$ of node $i$:
$$\Pi(k_i) = \frac{k_i}{\sum_j k_j}.$$

**Fundamentals of the Model:** The key principle here is the "rich-get-richer" phenomenon: nodes that already have many connections are more likely to receive new links. This iterative, dynamic process naturally leads to the formation of hubs within the network.

**Key Properties and Analysis:**

- **Degree Distribution:** Due to the preferential attachment mechanism, the degree distribution follows a power-law:
$$P(k) \sim k^{-\gamma},$$
where the exponent $\gamma$ is typically around 3. This heavy-tailed distribution reflects the emergence of a few hubs and many nodes with low connectivity.
- **Robustness and Vulnerability:** Scale-free networks are generally robust against random failures because most nodes have few connections. However, they are particularly sensitive to targeted attacks on the highly connected hubs.
- **Growth Dynamics:** As the network grows, the mechanism of preferential attachment continuously amplifies the connectivity of already well-connected nodes, reinforcing the disparity in node degrees.

**Applications and Limitations:** The Barabási-Albert model has found applications in explaining the structure of the Internet, citation networks, and various social systems. While it successfully models the emergence of hubs, it does not inherently generate the high clustering coefficients observed in some real-world networks, prompting further extensions and refinements of the model.

For a hands-on exploration of these network models, you can directly simulate them using the interactive web application. Visit this interactive app to visualize and experiment with the models in real time.

## 3 The Friendship Paradox

The *friendship paradox* is the counterintuitive phenomenon that, on average, your friends tend to have more friends than you do. Formally, let $G = (V, E)$ be an undirected graph with $N$ nodes, where each node $i$ has degree $d_i$. Define the *average degree* by

$$\mu \;=\; \frac{1}{N}\sum_{i=1}^{N} d_i.$$

When we choose an edge uniformly at random and examine one of its endpoints, the probability that this endpoint has degree $k$ is proportional to $kP(k)$, where $P(k)$ is the probability that a uniformly selected node has degree $k$. Hence, the expected degree of a node chosen in this way (i.e., a "friend") is

$$E[d_{\text{friend}}] \;=\; \frac{\sum_k k^2\,P(k)}{\mu} \;=\; \frac{E[k^2]}{\mu}.$$

Since $E[k^2] \geq \mu^2$ (with strict inequality if the degree distribution is not uniform), it follows that

$$E[d_{\text{friend}}] \;\geq\; \mu.$$

The theorem below captures this more formally.

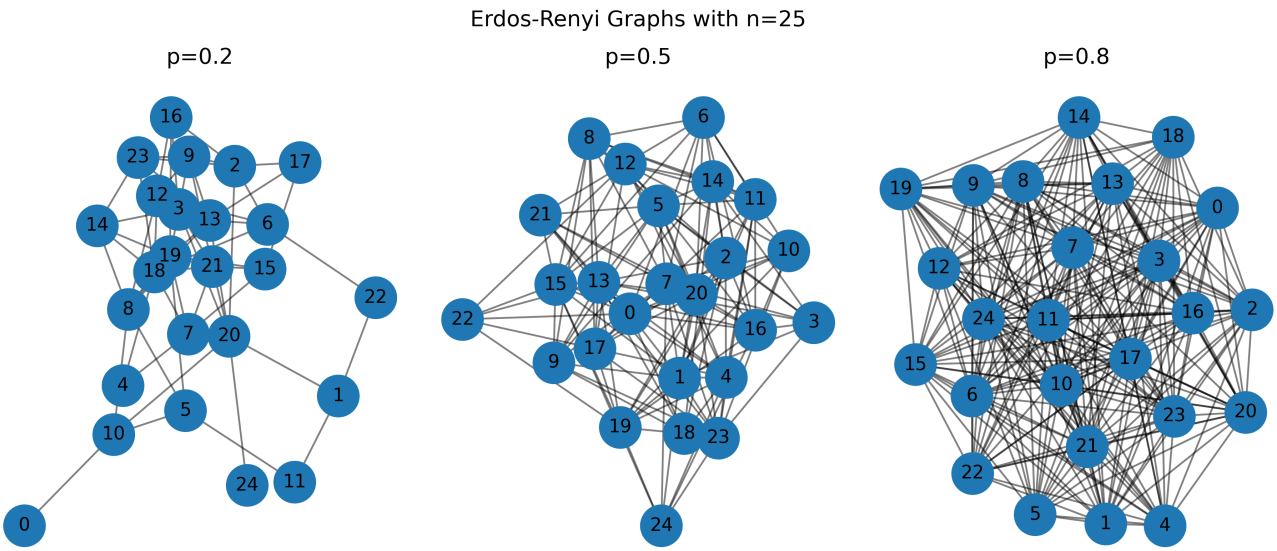


Figure 2: Erdős-Rényi graphs with $N = 25$ nodes for different values of $p$. Lower $p$ produces sparser graphs, whereas higher $p$ yields denser networks.

**Theorem 3.1** (Friendship Paradox)**.** *In any graph whose degree distribution is not uniform, the average degree of a randomly selected neighbor (friend) is strictly larger than the average degree of a randomly selected node:*

$$E[d_{friend}] \;=\; \frac{E[k^2]}{\mu} \;\geq\; \mu,$$

*with equality if and only if all nodes have the same degree.*

*Proof.* Let $P(k)$ be the probability that a uniformly chosen node has degree $k$. Then the average degree is

$$\mu \;=\; \sum_k k\, P(k).$$

Because a node of degree $k$ is $k$ times more likely to appear as an endpoint of a randomly chosen edge, the probability that a randomly selected neighbor has degree $k$ is

$$P_{\text{friend}}(k) \;=\; \frac{k\, P(k)}{\mu}.$$

Hence, the expected degree of a friend is

$$E[d_{\text{friend}}] \;=\; \sum_k k\, P_{\text{friend}}(k) \;=\; \frac{1}{\mu}\sum_k k^2\, P(k) \;=\; \frac{E[k^2]}{\mu}.$$

Writing $E[k^2]$ in terms of $\mu$ and the variance $\sigma^2$, we have $E[k^2] = \mu^2 + \sigma^2$. Thus,

$$E[d_{\text{friend}}] \;=\; \mu + \frac{\sigma^2}{\mu} \;\geq\; \mu,$$

with equality if and only if $\sigma^2 = 0$, which means all nodes have the same degree. □

**Example 3.1.** *For a network with degree distribution* $P(1) = 0.5$ *and* $P(3) = 0.5$*:*

$$\mu = 1 \cdot 0.5 + 3 \cdot 0.5 = 2,$$

$$E[k^2] = 1^2 \cdot 0.5 + 3^2 \cdot 0.5 = 0.5 + 4.5 = 5.$$

*Thus,*

$$E[d_{friend}] = \frac{5}{2} = 2.5,$$

*which exceeds the average degree* $\mu = 2$*.*

# 4 Monte Carlo Integration with Missing Data in Business Networks

In real business networks, some edge information is often missing. Monte Carlo integration allows us to compute the expected value of network metrics by integrating over the uncertainty.

Below, we detail three examples where missing data is modeled by random variables. We describe the mathematical model, the metric of interest, and how to integrate over the uncertainty.

### Example 1: Business Partnership Network with Missing Partnership Strengths

**Model Setup:** Consider a network $G = (V, E)$ where each node represents a business and each edge $(u, v)$ represents a partnership with an associated strength $w_{uv}$. For some partnerships, the weight is missing. We model a missing weight as a continuous random variable $W$ with a probability density function $f_W(w)$. For instance, if weights are normalized between 0 and 1, $W$ might follow a Beta distribution.

**Metric: Weighted Degree.** For a node $v$, the weighted degree is defined by:

$$D(v) = \sum_{u:(u,v)\in E} w_{uv}.$$

If the weight on an edge $(u, v)$ is missing and denoted by $W$, then the conditional weighted degree is:

$$D(v \mid W) = \sum_{(x,v)\in E\setminus\{(u,v)\}} w_{xv} + W.$$

The expected weighted degree of $v$ is given by:

$$E[D(v)] = \sum_{(x,v)\in E\setminus\{(u,v)\}} w_{xv} + E[W],$$

where

$$E[W] = \int_{w_{\min}}^{w_{\max}} w\, f_W(w)\, dw.$$

For several missing edges with independent missing values $\mathbf{W} = (W_1, W_2, \ldots, W_m)$, the expectation becomes:

$$E[D(v)] = \int_{\mathbb{R}^m} D(v \mid \mathbf{w}) \prod_{i=1}^{m} f_{W_i}(w_i)\, d\mathbf{w}.$$

This integral is typically approximated using Monte Carlo methods by drawing samples $\mathbf{w}^{(j)}$ and computing the average of $D(v \mid \mathbf{w}^{(j)})$.

### Example 2: Supply Chain Network with Uncertain Supplier-Customer Relationships

**Model Setup:** In a supply chain network, nodes represent companies and an edge $(u, v)$ indicates a supplier–customer relationship. For some potential edges, the existence of the relationship is uncertain. We model the existence of a missing edge $e = (u, v)$ as a Bernoulli random variable $X_e$ with parameter $p$:

$$P(X_e = 1) = p, \quad P(X_e = 0) = 1 - p.$$

**Metric: Shortest Path Length.** Suppose we are interested in the shortest path length $d(i, j)$ between two companies $i$ and $j$. If an uncertain edge is involved, its existence alters the path length. For a single uncertain edge:

$$E[d(i, j)] = p\, d(i, j \mid X_e = 1) + (1 - p)\, d(i, j \mid X_e = 0).$$

For multiple uncertain edges, denote $\mathbf{X} = (X_1, X_2, \ldots, X_m)$ and the expected shortest path is:

$$E[d(i, j)] = \sum_{\mathbf{x} \in \{0,1\}^m} d(i, j \mid \mathbf{X} = \mathbf{x})\, P(\mathbf{X} = \mathbf{x}),$$

with

$$P(\mathbf{X} = \mathbf{x}) = \prod_{e \in M} p_e^{x_e} (1 - p_e)^{1 - x_e}.$$

Monte Carlo integration approximates this expectation by sampling $\mathbf{x}$ from the Bernoulli distributions and averaging the corresponding $d(i, j \mid \mathbf{x})$.

### Example 3: Financial Risk Network with Missing Transaction Data

**Model Setup:** Consider a financial network where nodes represent institutions and an edge $(u, v)$ represents a financial transaction or exposure $w_{uv}$. If some transactions are missing, we model them as continuous random variables. For example, suppose a missing transaction follows a log-normal distribution:

$$W \sim \text{LogNormal}(\mu, \sigma^2),$$

with density

$$f_W(w) = \frac{1}{w\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln w - \mu)^2}{2\sigma^2}\right), \quad w > 0.$$

**Metric: Total Financial Exposure.** Let the total exposure in the network be:

$$R(G) = \sum_{(u,v) \in E} w_{uv}.$$

If the set $M \subset E$ represents edges with missing data, then conditioned on the missing data $\mathbf{W} = (W_e)_{e \in M}$,

$$R(G \mid \mathbf{W}) = \sum_{(u,v) \in E \setminus M} w_{uv} + \sum_{e \in M} W_e.$$

The expected total exposure is

$$E[R(G)] = \sum_{(u,v) \in E \setminus M} w_{uv} + \sum_{e \in M} E[W_e],$$

where for each missing edge,

$$E[W_e] = \exp\left(\mu + \frac{\sigma^2}{2}\right).$$

If multiple missing values are present, we have:

$$E[R(G)] = \int_{\mathbb{R}_+^{|M|}} \left( \sum_{(u,v)\in E\setminus M} w_{uv} + \sum_{e\in M} w_e \right) \prod_{e\in M} f_{W_e}(w_e)\, d\mathbf{w}.$$

This integral is often estimated by Monte Carlo methods, sampling many realizations of $\mathbf{W}$ and averaging the corresponding total exposure.

# 5 Exercises

**Exercise 1:** Consider a business partnership network $G = (V, E)$ where each edge $(u, v)$ is assigned a strength $w_{uv}$ representing the quality of the partnership. Some partnership strengths may be missing. Each missing weight $W$ is modeled by a continuous random variable with density

$$f_W(w) = 12\, w^2(1 - w), \quad 0 \le w \le 1.$$

For a node $v$, the weighted degree is defined as

$$D(v) = \sum_{(u,v)\in E} w_{uv}.$$

(a) Suppose node $v$ has three partnerships. Two edges have observed strengths $w_1 = 0.7$ and $w_2 = 0.5$, and the third edge's strength is missing. Calculate $E[D(v)]$.

(b) Now consider a more general scenario: node $v$ has $k$ observed partnerships with weights $w_1, w_2, \ldots, w_k$ and $m$ missing partnerships, where the missing weights $\{W_1, W_2, \ldots, W_m\}$ are independent random variables with density $f_W(w) = 12\, w^2(1 - w)$. Derive an expression for $E[D(v)]$ in terms of the observed weights and $m$. Then, outline an algorithm to approximate $E[D(v)]$ using Monte Carlo simulation when $m > 1$.

**Exercise 2:** In a supply chain network, nodes represent companies and an edge $(u, v)$ indicates a supplier–customer relationship. Some potential edges are uncertain. An edge $e = (u, v)$ exists with probability $p$ and does not exist with probability $1 - p$. The metric of interest is the shortest path length between two companies. Suppose that when the edge exists the shortest path length is $d(A, B \mid X = 1)$ and when it does not exist it is $d(A, B \mid X = 0)$.

(a) Given that for companies $A$ and $B$ the edge exists with probability 0.3, $d(A, B \mid X = 1) = 2$ and $d(A, B \mid X = 0) = 5$, compute $E[d(A, B)]$.

(b) Now assume there are two independent uncertain edges with existence probabilities $p_1$ and $p_2$. Let $d_{ij}$ be the shortest path length when the first edge is present if $i = 1$ (or absent if $i = 0$) and the second edge is present if $j = 1$ (or absent if $j = 0$). Derive an expression for $E[d]$ in terms of $p_1$, $p_2$, and $d_{ij}$.

**Exercise 3:** In a financial network, nodes represent financial institutions, and each edge $(u, v)$ represents a transaction with a reported value $w_{uv}$. However, the reported transaction values are measured with error. Assume that the true transaction value is given by

$$W_{uv} = w_{uv} \times \epsilon_{uv},$$

where the multiplicative error factor $\epsilon_{uv}$ is an independent random variable with density

$$f_\epsilon(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right), \quad x > 0.$$

The total exposure of the network is defined as

$$R(G) = \sum_{(u,v)\in E} W_{uv}.$$

Suppose there are two transactions with reported values $w_1 = 100$ and $w_2 = 150$, and the error factors $\epsilon_1$ and $\epsilon_2$ are independent with parameters $\mu = 0$ and $\sigma^2 = 0.25$.

(a) Compute the expected true transaction value $E[W_{uv}]$ for each transaction, and then calculate $E[R(G)]$.

(b) Outline an algorithm to estimate the distribution of $R(G)$ using Monte Carlo simulation when multiple transactions are present. Explain how you would use this simulation to approximate the risk measure $P(R(G) > T)$ for a given threshold $T$.