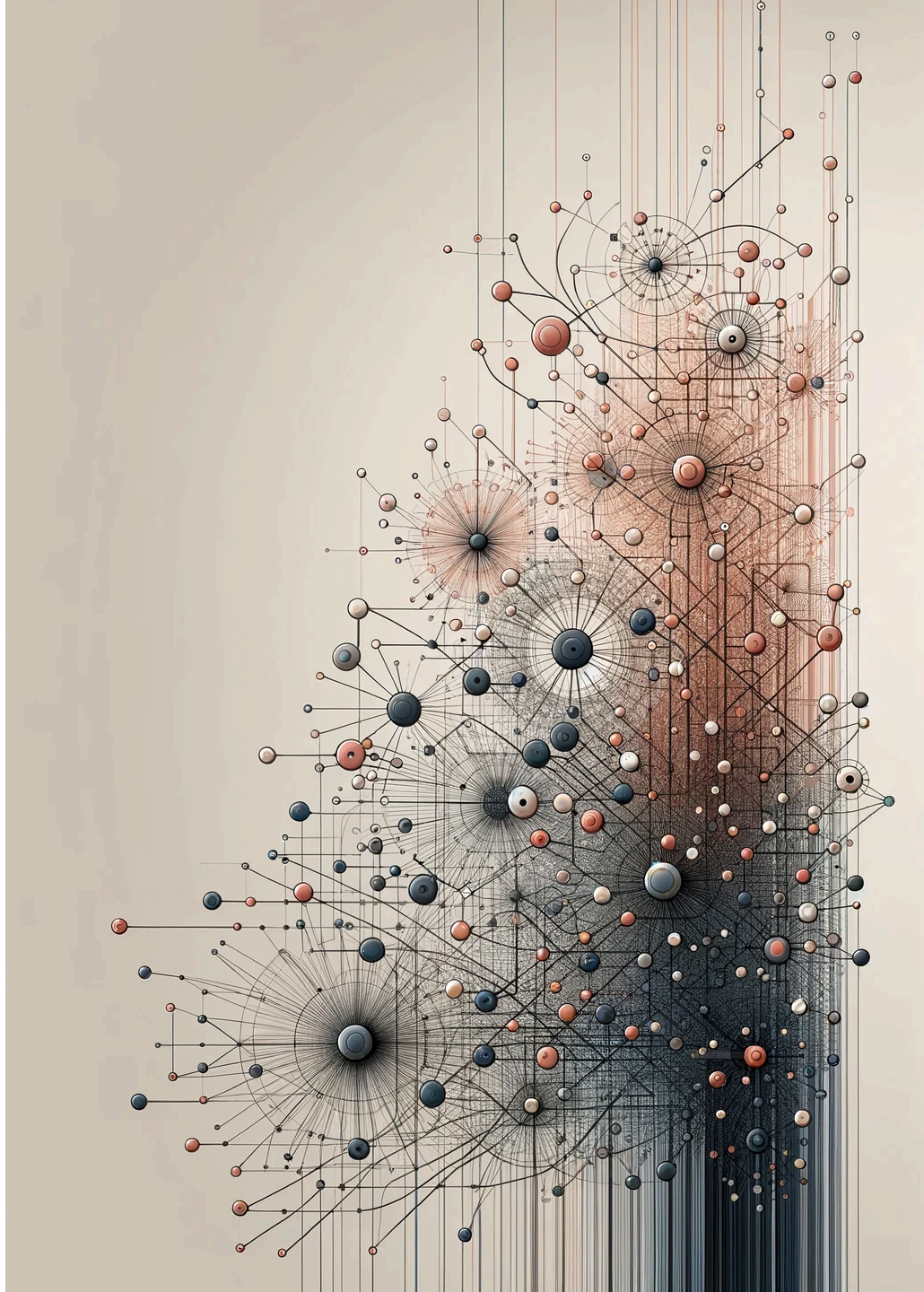
Analisi e Visualizzazione delle Reti Complesse

NS22 - Compartmental Models in Computational Epidemiology

Prof. Rossano Schifanella



Introduction to Computational Epidemiology

What is Computational Epidemiology?

- **Definition**: An interdisciplinary field that combines mathematics, computer science, data science, and epidemiology to model, analyze, and predict health-related dynamics in populations.
- **Core Components**:
  - Mathematical modeling of disease spread dynamics
  - Computer simulations of population-level interactions
  - Data analysis techniques for outbreak detection and tracking
  - Predictive algorithms for forecasting epidemic trajectories
- **Goal**: To develop quantitative tools and simulations that support **evidence-based decision-making** in public health—spanning **infectious diseases**, **chronic conditions**, **environmental exposures**, and **behavioral trends**.

Why Computational Epidemiology Matters

- **Informed Decision Making**: Provides evidence-based guidance for public health officials and policymakers during outbreaks (e.g., COVID-19 pandemic response).
- **Intervention Design**: Enables testing of control measures before implementation:
  - Vaccination strategies
  - Social distancing policies
  - Travel restrictions
  - Resource allocation
- **Preparedness Planning**: Helps healthcare systems anticipate and prepare for future outbreaks through:
  - Scenario modeling
  - Early warning systems
  - Resource requirement forecasts

Applications of Computational Epidemiology Beyond Health

- **Information Diffusion**:
  - Modeling the spread of news, rumors, and misinformation on social media platforms.
  - Understanding the viral transmission of digital content, including memes and trends.
  - Analyzing the dynamics of opinion formation and influence in online communities.
- **Cybersecurity**:
  - Simulating the propagation of malware, ransomware, and computer viruses across networks.
  - Assessing vulnerabilities in network infrastructures to predict and mitigate risks.
  - Designing containment strategies to limit the spread of infections in digital systems.

Modeling Approaches in Computational Epidemiology

Deterministic vs. Stochastic Models

- **Deterministic**: Fixed outcomes from initial conditions.
  - Based on differential equations (e.g., SI, SIS, SIR, SEIR).
  - Good for large populations and analytical insights.
- **Stochastic**: Incorporates randomness in outcomes.
  - Based on probability distributions.
  - Essential for small populations and early outbreaks.
  - Captures uncertainty but is more computationally intensive.

Population Mixing Patterns

How Contacts Shape Disease Spread

- **Homogeneous Mixing** (Mass Action):
  - Every individual has equal probability of contact with every other individual
  - Mathematically tractable but **unrealistic in real populations**
  - Enables analytical solutions but oversimplifies actual contact patterns
  - **Key limitation**: Ignores clustering, community structure, and super-spreaders
- **Heterogeneous Mixing**:
  - Individuals have different contact rates and patterns
  - Much closer to real-world social interaction networks
  - **Critical in real epidemics** but more complex to analyze mathematically
  - Requires **network-based** or agent-based approaches
  - Calls for spatially-aware models (geography-based interactions)

Compartmental Models: The Core Framework

Fundamental Concept

- **Definition**: Mathematical models that divide the population into distinct groups (compartments) based on disease status.
- **Historical Development**:
  - First formalized by Kermack and McKendrick (1927)
  - Revolutionized understanding of epidemic thresholds
  - Formed the foundation for modern mathematical epidemiology
- **Basic Principle**:
  - Individuals move between compartments according to defined rates
  - These transitions are modeled as flows, typically using differential equations
  - Each individual belongs to exactly one compartment at any time

Building Blocks of Compartmental Models

- **States (Compartments)**:
  - Mutually exclusive categories based on disease status
  - Common examples: Susceptible (S), Infectious (I), Recovered (R)
  - Extended versions may include: Exposed (E), Vaccinated (V), etc.
- **Transition Rules**:
  - Mathematical expressions governing flow between compartments
  - Based on biological parameters (infection rate, recovery rate)
  - Often represented as rates per unit time

Building Blocks of Compartmental Models (2)

- **Key Parameters**:
  - **Transmission rate (β)**: Rate of infection per contact
  - **Recovery rate (γ)**: Rate at which individuals recover
  - **Basic reproduction number (R_0)**: Average number of secondary infections caused by a single infectious individual in a completely susceptible population.
- **Mathematical Formulation**:
  - Typically ordinary differential equations (ODEs)
  - Sometimes stochastic versions for small populations

Assumptions in Compartmental Models

- **Homogeneous Mixing**:
  - All individuals interact randomly with equal probability
  - No spatial structure or clustering of contacts
- **Markovian Properties**:
  - Future state depends only on current state, not past history
  - Transition probabilities are constant over time
- **Population-Level Focus**:
  - Large population approximation, individual variations are averaged out
  - Deterministic models work with expected values
- **Well-Mixed Population**:
  - No explicit spatial component, geography and movement patterns generally ignored

The SI Model

- **S**: Susceptible
- **I**: Infectious.
- **Total population**: $N = S + I$.
- **Transmission rate**: β (per contact, per unit time).

Differential Equations

$$\frac{dS}{dt} = -\beta \frac{SI}{N},$$
$$\frac{dI}{dt} = \beta \frac{SI}{N}.$$

Why Differential Equations for Epidemic Models?

Mathematical Justification

- Epidemics are **dynamic processes** that evolve continuously over time.
- Changes in population compartments are **rate-based processes**.
- For large populations, discrete changes appear **continuous**.
- ODEs efficiently capture the **deterministic nature** of epidemics at scale.

Practical Benefits

- Well-established mathematical theory (stability, equilibria analysis).
- Efficient numerical methods for solution.
- Clear relationship to underlying biological parameters.
- Ability to derive analytical insights (thresholds, final size).

Analytical Solution

Exact Solution for $I(t)$

The analytical solution for the infectious population over time is given by:

$$I(t) = \frac{I_0 N e^{\beta t}}{N - I_0 + I_0 e^{\beta t}},$$

where I_0 is the initial number of infectious individuals.

Conservation Law for $S(t)$

The susceptible population $S(t)$ can be derived using the conservation law:

$$S(t) = N - I(t).$$

Since there is no recovery in this model, the entire population eventually becomes infected as $t \to \infty$.

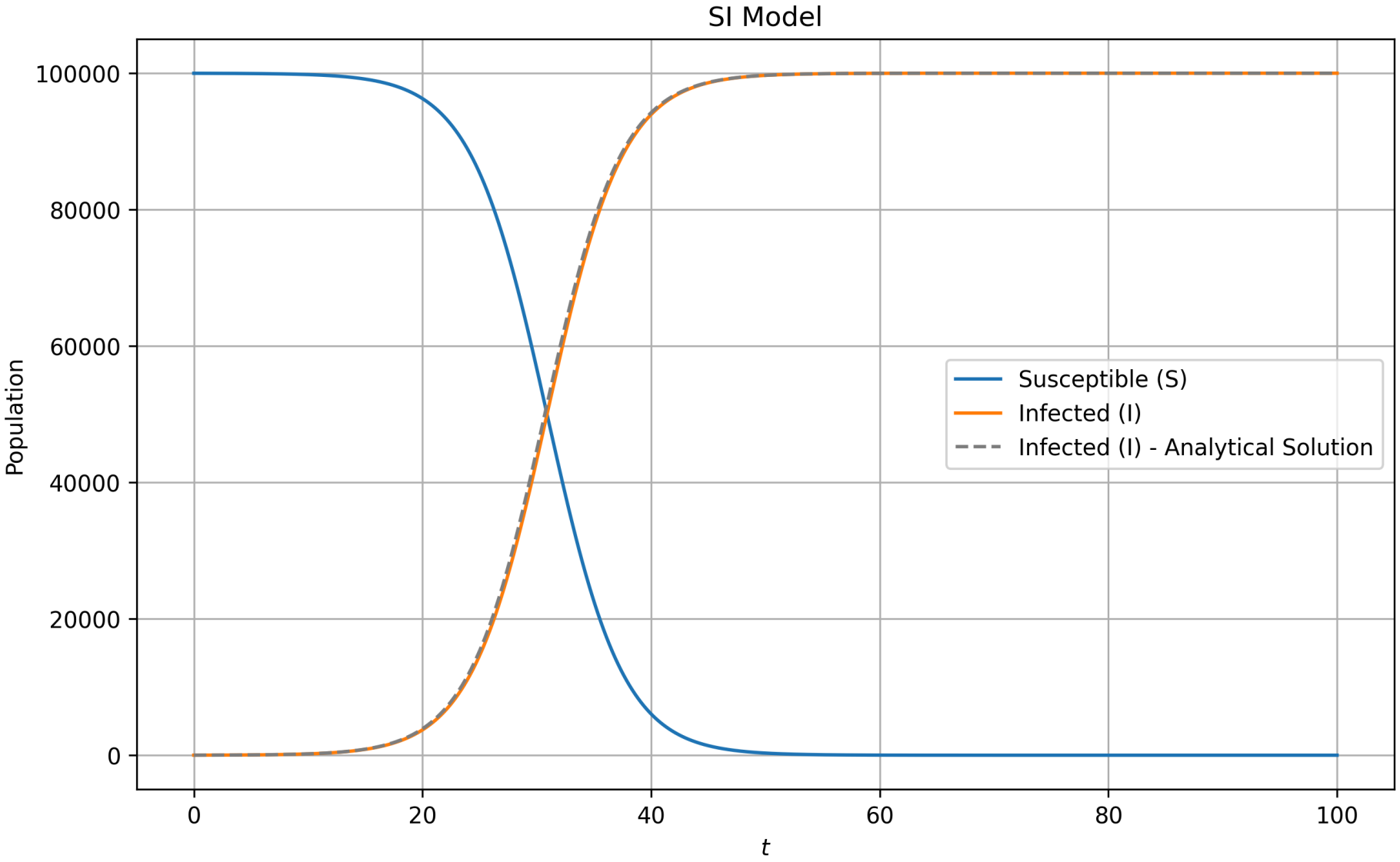
Numerical Approximation Using Euler's Method

For scenarios where an exact solution is not practical, the Euler method can be used to approximate the dynamics:

$$S(t + \Delta t) = S(t) - \Delta t \cdot \beta \frac{S(t)I(t)}{N},$$
$$I(t + \Delta t) = I(t) + \Delta t \cdot \beta \frac{S(t)I(t)}{N}.$$

Steps:

1. Choose a small time step Δt.
2. Initialize $S(0) = N - I_0$ and $I(0) = I_0$.
3. Iteratively compute $S(t)$ and $I(t)$ for each time step.


SI Model
Population
100000
80000
60000
40000
20000
0
0
20
40
60
80
100
t
Susceptible (S)
Infected (I)
Infected (I) - Analytical Solution

Early Growth Follows an Exponential Trend

At the start of an epidemic, the number of susceptible individuals (S) is approximately equal to the total population (N), as very few individuals are infected. This simplifies the differential equation for $I(t)$:

$$\frac{dI}{dt} = \beta \frac{SI}{N} \approx \beta \frac{NI}{N} = \beta I.$$

This is a **linear differential equation** with respect to I, and its solution is:

$$I(t) = I_0 e^{\beta t},$$

where I_0 is the initial number of infectious individuals.

Key Points

- **Exponential Growth**: The number of infectious individuals grows exponentially because the rate of new infections is proportional to the current number of infectious individuals.
- **Assumption**: This behavior holds only in the early stages when $S \approx N$ and the depletion of susceptibles is negligible.
- **Implication**: The early exponential growth highlights the importance of rapid intervention to prevent uncontrolled spread.

Key Metrics

Understanding the dynamics of an epidemic often requires quantifying key temporal metrics. These metrics provide insights into the speed and progression of disease spread, aiding in decision-making and intervention planning.

- **Doubling Time (t_d)**:
  The time it takes for the number of infectious individuals to double in the early exponential growth phase.

$$t_d = \frac{\ln(2)}{\beta}$$

  - **Interpretation**: A smaller doubling time indicates a faster-spreading epidemic.
  - **Application**: Helps assess the urgency of implementing control measures.

- **Half-Infection Time ($t_{50\%}$)**:
  The time required for half of the total population to become infected.

$$t_{50\%} = \frac{1}{\beta} \ln\left(\frac{N - I_0}{I_0}\right)$$

SI Model: Applications and Practical Interpretation

Applications

- **Permanent Infections**: Models diseases or conditions where recovery is not possible (e.g., chronic infections).
- **Digital Contagion**: Simulates the spread of computer viruses or irreversible information cascades.
- **Baseline Model**: Serves as a foundation for understanding and comparing more complex epidemic dynamics.

Key Takeaway

The SI model provides a simple yet powerful framework for understanding the dynamics of irreversible spreading processes, offering insights into both biological and digital contagion phenomena.

SIS Model

- **S**: Susceptible, **I**: Infectious → **S**usceptible again
- **Total population**: $N = S + I$ (constant)
- **Transmission rate**: β
- **Recovery rate**: γ
- **"Get sick, recover, get sick again"**:
  - Individuals recover but remain susceptible
  - Can become infected multiple times, like catching the common cold repeatedly

Differential Equations:

$$\frac{dS}{dt} = -\beta \frac{SI}{N} + \gamma I,$$
$$\frac{dI}{dt} = \beta \frac{SI}{N} - \gamma I.$$

Analytical Solution

$$I(t) = \frac{K I_0 e^{(\beta-\gamma)t}}{K + I_0(e^{(\beta-\gamma)t} - 1)}$$

where $K = N - \frac{\gamma N}{\beta} = N(1 - \frac{1}{R_0})$

Numerical Approximation Using Euler's Method

For scenarios where an exact solution is not practical, the Euler method can be used to approximate the dynamics:

$$S(t + \Delta t) = S(t) - \Delta t \cdot (\beta \frac{S(t)I(t)}{N} - \gamma I(t)),$$

$$I(t + \Delta t) = I(t) + \Delta t \cdot (\beta \frac{S(t)I(t)}{N} - \gamma I(t)).$$

Basic Reproduction Number (R_0)

- **Definition**: The basic reproduction number, denoted as R_0, is a key epidemiological metric that represents the average number of secondary infections caused by a single infectious individual in a completely susceptible population. Mathematically, it is expressed as:

$$R_0 = \frac{\beta}{\gamma}$$

  where:
  - β: Transmission rate (rate of infection per contact per unit time),
  - γ: Recovery rate (rate at which individuals recover or are removed from the infectious state).

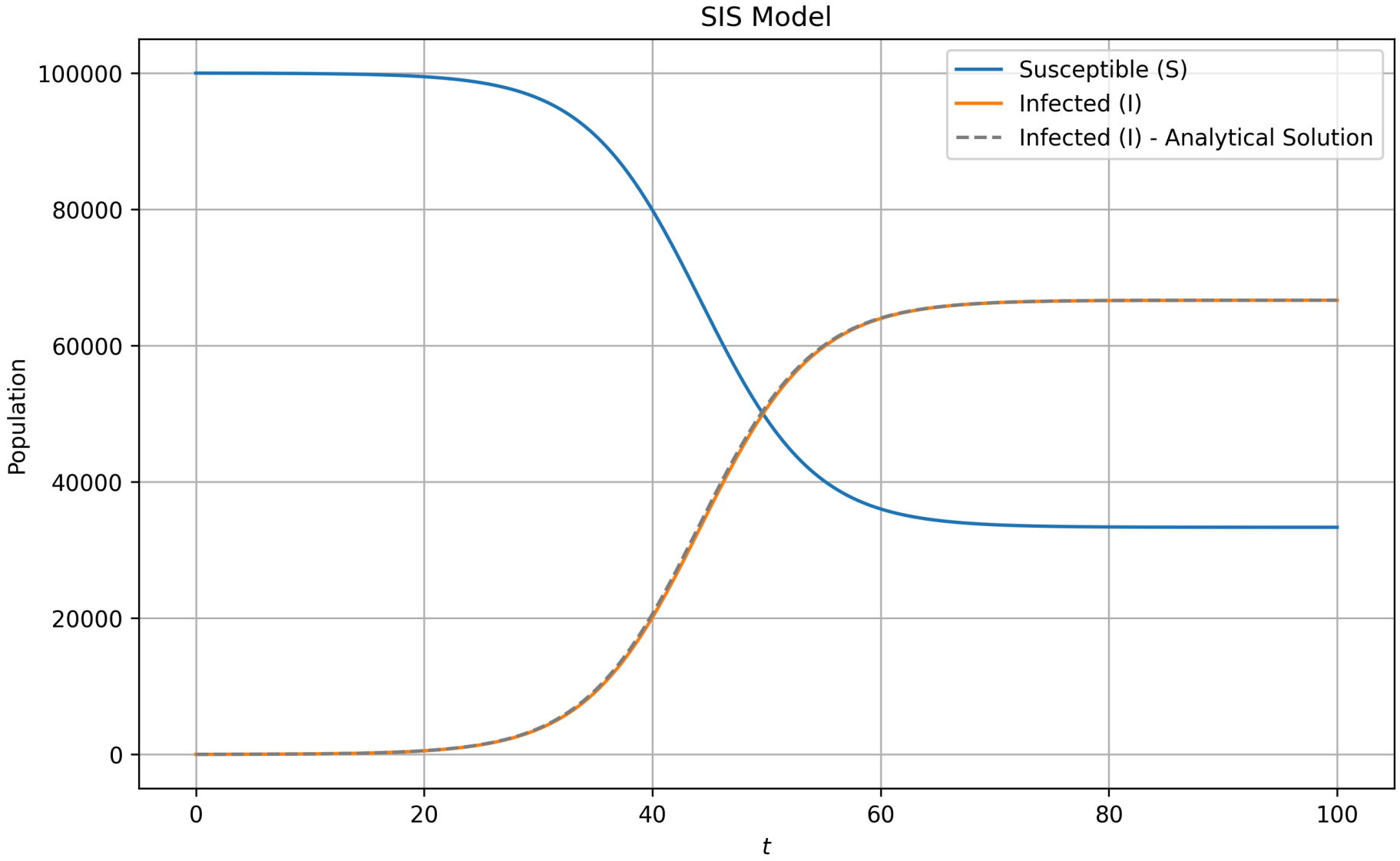
Context and Importance:

- R_0 serves as a threshold parameter that determines whether an infectious disease will spread or die out in a population:
  - If $R_0 < 1$: Each infected individual causes less than one new infection on average, leading to the eventual extinction of the disease.
  - If $R_0 > 1$: Each infected individual causes more than one new infection on average, allowing the disease to spread and potentially persist in the population.
- The critical value of $R_0 = 1$ marks the **epidemic threshold**, separating the two regimes of disease dynamics:
  - Below the threshold ($R_0 < 1$), the disease cannot sustain itself.
  - Above the threshold ($R_0 > 1$), the disease can grow exponentially, leading to an outbreak or epidemic.

Applications:

- R_0 is used to estimate the intensity of interventions required to control an epidemic, such as vaccination coverage or social distancing measures.
- It provides insights into the transmissibility of a disease and helps compare the spread potential of different pathogens (e.g., measles has a high R_0 of 12–18, while seasonal influenza has a lower R_0 of 1.2–1.8).

Key Takeaway: Understanding and estimating R_0 is crucial for designing effective public health strategies to prevent and control infectious disease outbreaks.


SIS Model
Susceptible (S)
Infected (I)
Infected (I) - Analytical Solution
Population
100000
80000
60000
40000
20000
0
0
20
40
60
80
100
t

Threshold Behavior in the SIS Model

The **basic reproduction number (R_0)** determines whether the disease will die out or persist.

Threshold Interpretation:

1. **If $R_0 < 1$:**
   - The disease cannot sustain itself.
   - Infected individuals recover faster than they can infect others.
   - The population converges to the **disease-free equilibrium** ($I^* = 0$).
2. **If $R_0 > 1$:**
   - The disease spreads and persists in the population.
   - The population converges to the **endemic equilibrium** ($I^* > 0$).

Key Metrics in the SIS Model

- **Growth Rate**:
  - The rate of change in infectious individuals is:

$$\frac{dI}{dt} = \beta \frac{SI}{N} - \gamma I$$

  - During the early epidemic phase when $S \approx N$:

$$\frac{dI}{dt} \approx \beta \frac{NI}{N} - \gamma I = (\beta - \gamma) I$$

  - The coefficient $r = \beta - \gamma$ is the growth rate
  - Using $R_0 = \frac{\beta}{\gamma}$, we can rewrite:

$$r = \gamma(R_0 - 1)$$

- Determines how quickly the number of infections grows or declines.

- **Time to Equilibrium**:
  - Approximate time to reach equilibrium:

$$t_{\text{eq}} \approx \frac{5}{|r|} = \frac{5}{|\beta - \gamma|}.$$

- **Proportion Infected at Equilibrium**:
  - For $R_0 > 1$, the proportion of the population infected at equilibrium is:

$$\frac{I^*}{N} = 1 - \frac{1}{R_0}.$$

SIR Model

- **S**: Susceptible
- **I**: Infectious
- **R**: Recovered/Removed
- **Total population**: $N = S + I + R$
- **Transmission rate**: β
- **Recovery rate**: γ

Differential Equations:

$$\begin{aligned}
\frac{dS}{dt} &= -\beta \frac{SI}{N}, \\
\frac{dI}{dt} &= \beta \frac{SI}{N} - \gamma I, \\
\frac{dR}{dt} &= \gamma I.
\end{aligned}$$

SIR Model: Analytical Insights

Unlike the SI and SIS models, the SIR model has **no closed-form solution** due to the nonlinear nature of the equations.

Numerical Approximation Using Euler's Method

Since no closed-form solution exists, the SIR model is typically solved using numerical methods:

$$S(t+\Delta t) = S(t) - \Delta t \cdot \beta \frac{S(t)I(t)}{N}$$

$$I(t+\Delta t) = I(t) + \Delta t \cdot (\beta \frac{S(t)I(t)}{N} - \gamma I(t))$$

$$R(t+\Delta t) = R(t) + \Delta t \cdot \gamma I(t)$$

SIR Model: Basic Reproduction Number (R_0)

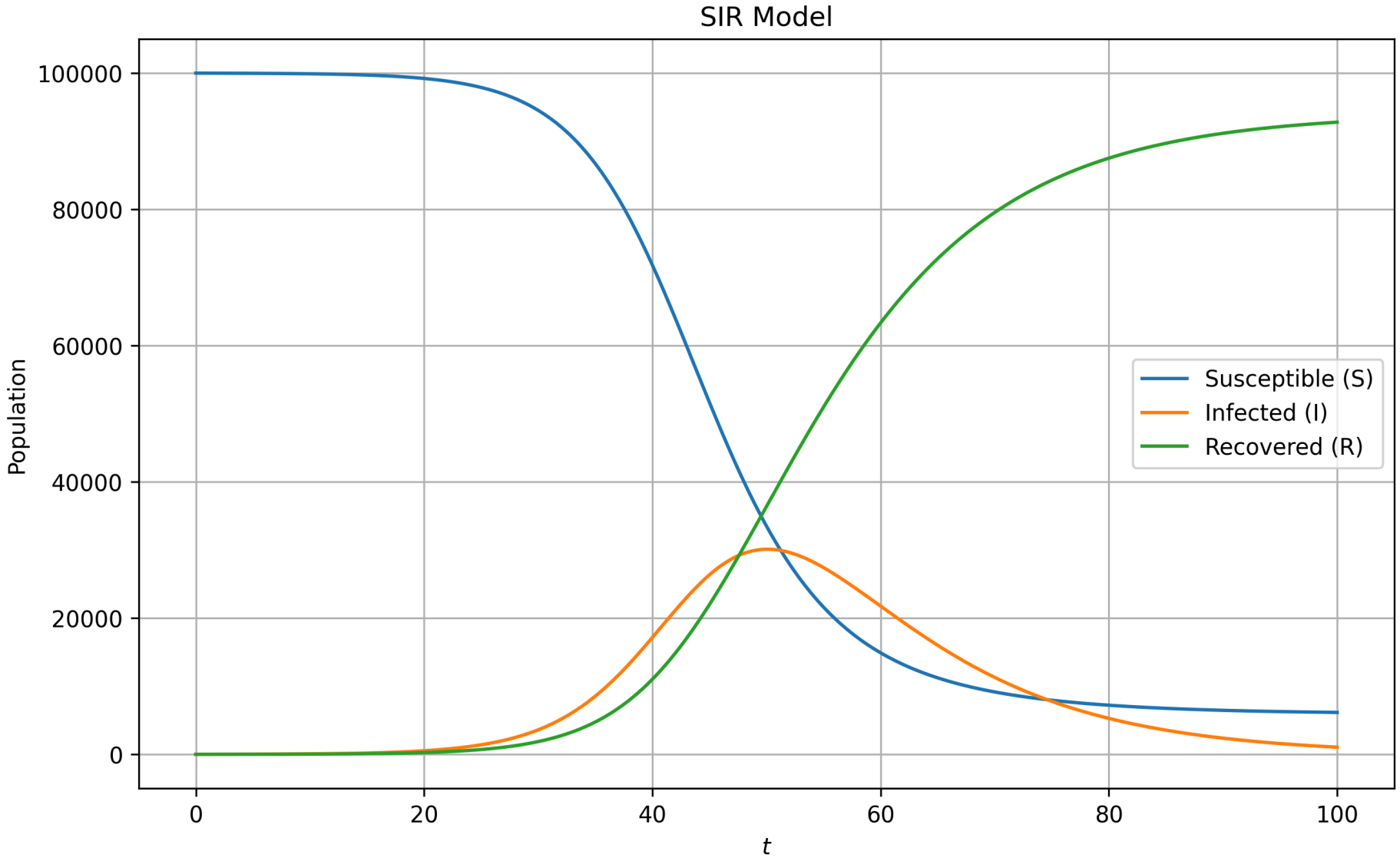
$$R_0 = \frac{\beta}{\gamma}$$

Interpretation in SIR vs SIS Models

- **SIS Model**: R_0 determines disease persistence
  - $R_0 < 1$: Disease dies out
  - $R_0 > 1$: Disease becomes endemic
- **SIR Model**: R_0 determines epidemic occurrence
  - $R_0 < 1$: No epidemic occurs
  - $R_0 > 1$: Epidemic occurs but eventually ends

Key Distinction

- **SIS**: Recovered individuals become susceptible again, allowing persistence
- **SIR**: Permanent immunity depletes susceptible pool, ensuring epidemic end


SIR Model
Population
100000
80000
60000
40000
20000
0
0
20
40
60
80
100
t
Susceptible (S)
Infected (I)
Recovered (R)

SIR Model: Three Phases of the Epidemic

Growth Phase ($S > \frac{\gamma N}{\beta}$)

- $\frac{dI}{dt} > 0$: Infections exceed recoveries
- **Initial exponential growth in infectious population**

Peak Phase ($S = \frac{\gamma N}{\beta}$)

- $\frac{dI}{dt} = 0$: Infections equal recoveries
- **Maximum number of simultaneous infections reached**

Decline Phase ($S < \frac{\gamma N}{\beta}$)

- $\frac{dI}{dt} < 0$: Recoveries exceed infections
- **Epidemic naturally subsides**

SIR Model: Critical Thresholds

Epidemic Turning Point

- **Critical threshold**: $S_{threshold} = \frac{\gamma N}{\beta} = \frac{N}{R_0}$
  - Interpretation: This is the point where the number of new infections starts to decline because the susceptible population has dropped below the level needed to sustain the epidemic.
  - Implication: For diseases with $R_0 > 2$, the peak of the epidemic occurs before most of the population is infected, highlighting the importance of early interventions to reduce R_0.

Key Epidemic Measures: Attack Rate

- **Attack Rate**: Final proportion of the population that gets infected during the epidemic.
  - Formula:

$$1 - \frac{S_\infty}{N}$$

  - S_∞ represents the number of susceptible individuals remaining at the end of the epidemic. These are the people who were never infected throughout the course of the epidemic, either because they were naturally resistant, avoided exposure, or benefited from interventions like vaccination.
  - Interpretation: This metric quantifies the total impact of the epidemic by measuring how many individuals were ultimately infected.
  - Implication: A higher R_0 typically leads to a higher attack rate, emphasizing the need for interventions to reduce transmission.

Key Epidemic Measures: Peak Prevalence

- **Peak Prevalence**: Maximum percentage of the population infected simultaneously.
  - Approximation:

$$1 - \frac{1 + \ln R_0}{R_0}$$

  - Interpretation: This represents the worst-case burden on healthcare systems, as it indicates the highest number of active cases at any given time.
  - Implication: Flattening the curve reduces peak prevalence, preventing healthcare system overload.

Key Epidemic Measures: Doubling Time

- **Doubling Time**: Time it takes for the number of cases to double during the early exponential growth phase.
  - Formula:

$$t_d \approx \frac{\ln(2)}{\gamma(R_0 - 1)}$$

  - Interpretation: A shorter doubling time indicates a faster-spreading epidemic.
  - Implication: This metric is crucial for assessing the urgency of implementing control measures.

Final Size Relation

- **Final Size Relation**: Links the initial conditions of the epidemic to the total number of infections without solving the full dynamics.
  - Formula:

$$\ln\left(\frac{S_0}{S_\infty}\right) = R_0\left(1 - \frac{S_\infty}{N}\right)$$

  - Interpretation: This equation allows us to estimate the final size of the epidemic (total infections) based on R_0 and the initial susceptible population.
  - Implication: Useful for long-term planning and evaluating the effectiveness of interventions.

Early Growth Approximation

- **Early Growth Approximation**: Describes the exponential growth of infections when $S \approx S_0$ and I is small.
  - Formula:

$$I(t) \approx I_0 e^{(\beta \frac{S_0}{N} - \gamma)t} = I_0 e^{\gamma(R_0 - 1)t}$$

  - Interpretation: This shows how quickly the number of infections grows in the early stages of the epidemic when $R_0 > 1$.
  - Implication: Highlights the critical need for early interventions to slow the exponential growth phase.

Peak Infection Time

- **Peak Infection Time**: The time at which the infection rate reaches its maximum.
  - Condition:

$$\frac{dI}{dt} = 0 \implies S = \frac{\gamma N}{\beta} = \frac{N}{R_0}$$

  - Interpretation: The infection rate peaks when the susceptible population drops to the critical threshold $S_{threshold}$.
  - Implication: Understanding this timing helps in planning healthcare resource allocation and intervention strategies.

SIR Model: Control Parameters

Herd Immunity

- Occurs when a sufficient proportion of a population is immune to a disease, making its spread to non-immune individuals unlikely.
- **Herd immunity threshold**: $p_c = 1 - \frac{1}{R_0}$
  - Minimum percentage of immune population needed to prevent epidemic spread
  - Higher R_0 requires higher immunity level
- **Vaccination coverage needed**: $v_c = \frac{R_0 - 1}{R_0}$
  - Proportion that must be vaccinated to prevent outbreaks
  - Example: If $R_0 = 4$, need 75% vaccination coverage

Key Insights

- **Epidemic Threshold**
  - Disease spreads only when $R_0 > 1$
  - Minor changes in transmission near threshold have significant impact
- **Self-Limiting Nature**
  - Unlike SIS model, epidemics naturally end as susceptible pool depletes
  - Final epidemic size determined by R_0 and initial conditions
- **Intervention Timing**
  - Early interventions yield maximum effectiveness
  - Most infections have already occurred once peak is reached
- **Flattening the Curve**
  - Reducing β extends epidemic duration while lowering peak prevalence
  - May result in similar cumulative cases, but distributed over longer timeframe

Model Extensions

SEIR Model

- **Adds an Exposed (E) compartment**:
  - Represents individuals who are infected but not yet infectious (latent period).
- **Key Use Case**: Diseases with a significant incubation period (e.g., COVID-19, measles).
- **Dynamics**:
  - $S \rightarrow E$: Susceptible individuals become exposed upon infection.
  - $E \rightarrow I$: Exposed individuals transition to infectious after the incubation period.
  - $I \rightarrow R$: Infectious individuals recover or are removed.
- **Practical Insights**:
  - Captures delays in disease spread due to incubation.
  - Useful for modeling interventions like quarantine during the latent phase.

SIRS Model

- **Adds waning immunity**:
  - Recovered individuals lose immunity over time and return to the susceptible state.
- **Key Use Case**: Diseases where immunity is temporary (e.g., seasonal flu, reinfections).
- **Dynamics**:
  - $R \rightarrow S$: Immunity wanes, and individuals become susceptible again.
- **Practical Insights**:
  - Models recurring outbreaks or seasonal epidemics.
  - Highlights the importance of booster vaccinations.

SEIRS Model

- **Combines SEIR and SIRS**:
  - Includes both a latent period and waning immunity.
- **Key Use Case**: Diseases with incubation periods and temporary immunity (e.g., COVID-19 with waning immunity).
- **Practical Insights**:
  - Captures complex dynamics of reinfections and delayed spread.
  - Useful for long-term epidemic forecasting.

SIR with Demography

- **Adds births and deaths**:
  - Accounts for population turnover (e.g., births replenish susceptibles, deaths remove individuals from all compartments).
- **Key Use Case**: Long-term modeling of endemic diseases (e.g., measles, malaria).
- **Dynamics**:
  - Births add to S, and natural deaths reduce S, I, and R.
- **Practical Insights**:
  - Models the steady-state behavior of diseases in populations with constant turnover.
  - Highlights the role of vaccination in maintaining herd immunity.

Age-Structured SIR

- **Adds age-specific compartments**:
  - Divides S, I, and R into age groups with different contact rates and disease dynamics.
- **Key Use Case**: Diseases with age-dependent transmission or severity (e.g., COVID-19, influenza).
- **Practical Insights**:
  - Captures heterogeneity in disease spread and outcomes across age groups.
  - Useful for designing targeted interventions (e.g., vaccinating high-risk age groups).

Network-Based Models

- **Replaces homogeneous mixing assumption**:
  - Models disease spread on contact networks (e.g., social networks, transportation networks).
- **Key Use Case**: Diseases with structured transmission patterns (e.g., sexually transmitted infections, localized outbreaks).
- **Practical Insights**:
  - Captures the role of network structure in disease dynamics.
  - Useful for identifying super-spreaders and optimizing targeted interventions.
  - **Recognizes that real disease spread follows existing social and physical contact patterns**

Q&A