In the SI model, individuals can be only in two states, Susceptible and Infected. Susceptible ones can become Infected, and Infected individuals rest in this state indefinitely. Parameters are:

\[ \tau \quad \text{Infected fraction} \]
\[ \dot{c} \quad \text{Contact rate: average number of contact per person per time} \]
\[ \beta \quad \text{Effective contact rate}, \beta = \tau \dot{c} \text{, number of newly infected individuals by each infected individual in a population in which everyone else is susceptible.} \]

\[ S \rightarrow I \]

**SI - Application**

An example in which this model can be appropriate is for diffusion of innovation: being infected means buying or using a new service, product or technology whose usage becomes widespread in society, e.g., television, cell-phone, internet, Netflix, etc.

**SI - SIR - SIS**

Three of the most popular models of diffusion in epidemiology are the **SI, SIR** and **SIS** models. Letters correspond to the states in which individuals can be according to the model:

- **Susceptible**: Individual is not Infected
- **Infected**: Individual is Infected
- **Recovered/Removed**: Individual cannot be infected again (Considered cured or dead)

All individuals are in one of the states allowed by the model, and we define:

- \( s(t) \) Fraction of individuals in Susceptible state at time \( t \)
- \( i(t) \) Fraction of individuals in Infected state at time \( t \)
- \( r(t) \) Fraction of individuals in Recovered state at time \( t \)
- \( \tau \) Initial (\( t = 0 \)) fraction of infected individuals

For large times, \( i(t) \to 1 - \frac{1}{\beta} \), i.e., the fraction of infected individuals stabilize around a value which depends only of parameters \( \mu \) and \( \beta \).

\[ S \rightarrow I \]

**SIS - characteristics**

Intuitively, the fraction of infected individuals is now reduced by those switching to the susceptible state, more formally:

\[ \frac{di}{dt} = \frac{\beta(1 - i) - \mu}{1 - \beta \hat{c}(1 - i)} i(t) \]

For large times, \( i(t) \to 1 - \frac{1}{\beta} \), i.e., the fraction of infected individuals stabilize around a value which depends only of parameters \( \mu \) and \( \beta \).

\[ S \rightarrow I \]

**SI - Sketch**

Exponential growth.
In the SIS model, an important notion is the \( R_0 \) ratio, also called \( R_0 \).

\[ R_0 = \frac{\beta}{\mu} \]

\( R_0 \) can be understood as the average number of individuals that will be infected by an infected individual, in a population in which all other nodes are Susceptible. \( R_0 \) is a property of the model and do not change with time. Looking at the \( R_0 \) is important in the early stage of the epidemic:

- If \( R_0 > 1 \), there will be an outbreak
- If \( R_0 < 1 \), the epidemic will disappear naturally.

If \( R_0 \) is just above 1, the outbreak also can stop naturally by chance in the early stage.

Epidemic modeling is a large and rich scientific topic, thus those models are nowadays considered toy models, too simple to model real epidemics. Most used model thus include other factors, such as natural population dynamic (birth, natural death, etc. for long term dynamics), population segmentation (\( k \) might differ among subsets of populations, e.g., elderly, maried couples, etc.), population mixing (\( c \) might vary between members of a subpopulation and another), etc.

A natural way to add more details to a spreading process is, instead of considering an homogeneous population, or a population composed of homogeneous subpopulations, to study the diffusion on a network representing the structure of the population.
Heterogeneous Degrees - SI

For the SI model, we know that all nodes are infected in the end, but what may vary is the speed of the process. The speed of diffusion by degree block can be expressed as:

\[ \frac{d\Theta_k}{dt} = \beta k (1 - i_k) \Theta_k \]

with \( \Theta_k \) being the fraction of infected neighbors of a node with degree \( k \).

Heterogeneous Degrees - SI - time scale

From previous equations, it can be shown that the time scale \( \tau \) of the process, i.e., a measure inversely proportional to its speed, is:

\[ \tau = \frac{1}{\beta k} \Theta_k \]

Thus, for a given average degree \( \langle k \rangle \) and a given \( \beta \), the more heterogeneous the degrees, the faster the diffusion.

If the degree distribution follows a power law of exponent \( \alpha < 3 \), we have seen that \( k^2 \) diverge towards infinity, thus \( \tau \) tends toward 0, thus the diffusion is nearly instantaneous. This can be understood as follows: if a node is connected to nearly every other node, then it has an extremely high probability to become infected immediately, and can then infect the rest of the network extremely quickly.

\[ ^{\alpha}\text{Barrat, Barthelemy, and Vespignani 2008} \]

Heterogeneous Degrees - \( \lambda \)

For ISIs and SIR models, it can also be shown that the epidemic threshold \( \lambda \) (or \( R_0 \)) is not reached when \( \lambda = \frac{\langle k \rangle}{\beta} > 1 \) as in homogeneous networks, but when \( \lambda > \langle k^2 \rangle / 2 \). This means that in a very heterogeneous network, an outbreak can start even if \( \lambda \) is very small, and below 1. Intuitively, even if people recover faster than they spread the virus in average, some nodes (hubs) will nevertheless become infected, and since they can infect many others, the contagion will spread.

\[ ^{\alpha}\text{Barrat, Barthelemy, and Vespignani 2008} \]

Heterogeneous Degrees - \( \gamma \)

\[ \Theta_k \] represents the fraction of infected neighbors of a node with degree \( k \). If the network is homogeneous, \( \Theta_k(t) = i(t) \). If the graph is heterogeneous, nodes with different degrees have different probabilities of being infected: higher degree nodes are, by definition, more exposed, since they have more chances of being infected at each step. This is important for two reasons:

- Due to the friendship paradox, nodes are more likely to be connected to large nodes than to small ones.
- In real networks, we have seen that there is often a degree assortativity, thus nodes of a given degree have different degrees in their neighborhood.

We can thus define \( \Theta_k \) as the probability to be connected to nodes of a given degree and or their respective probability of being infected \( \Theta_k = \sum_{k'} P(k'|k) i_{k'} \), with \( P(k'|k) \) the probability that a node with degree \( k \) connects to a node with degree \( k' \). For simplicity, we assume no degree-degree correlation. \( P(k'|k) \) can then be expressed as the fraction of all edge stubs attached to nodes of degree \( k' \), independent of the \( k \) under study:

\[ P(k'|k) = \frac{k' P(k')}{\sum_{k''} k'' P(k'')} \]

And:

\[ \Theta_k = \sum_{k'} \frac{k' P(k') i_{k'}}{\langle k \rangle} \]

SIR - Scale Free

In this experiment, we compare an ER network to Configuration Models with power law degree distributions.

**Network parameters** \( n = 1000, \langle k \rangle = 5 \). We vary \( \langle k \rangle \) from \( 1 \) to \( 5 \), keeping the exponent of the distribution. We observe that the more marked the communities, the less efficient the spreading process.

**SIR parameters**: \( \theta = 0.2, \gamma = 0.5 \). The initial number of infected nodes is 5, all of them in the same community structure.

SIR - Experimental

When analytical solutions cannot be simply derived, empirical simulations can be used to observe the effect of network properties on diffusion processes. In particular, these properties can be used to assess the effect of typical heterogeneity: degree-heterogeneity, belonging to blocks, spatial heterogeneity, etc.
In this experiment, we compare an ER network to Watts Strogatz random graphs, varying the probability of rewiring edges. It can be understood as a model of spatial proximity: with $p = 0$, each node is connected only to its direct neighbors in the 1 dimensional space. If $p = 1$, each node is connected to exactly $k$ random nodes.

**Network parameters** \( n = 1000, \langle k \rangle = 5 \)

**SIR parameters** \( \theta = 0.2, \gamma = 0.5 \). The initial number of infected nodes is 5, being 5 direct neighbors.

---

**Application of diffusion models**

Diffusion models can be used for several applications:

- Model fitting: better understand an actual epidemic by fitting parameters on real observations
- Predicting trends of evolution
- Control of epidemics: Given an epidemic model and a supporting network, find an optimal solution to control (accelerate or slow-down) the epidemic.

---

**Optimal node/edge removal**

One way to slow-down an epidemic consist in removing nodes (e.g., vaccination). The problem can be formulated as a budget constrained removal, i.e., if we can remove only \( x \) nodes/edges, which one should we choose? Based on theoretical and experimental results, heuristic solutions consist in removing: highest degree nodes, highest betweenness nodes/edges (isolating communities), long-distance edges (shortcuts) in spatial networks.

---

**Friendship paradox and node removal**

It has been proposed\(^\circ\) that the friendship paradox could be used to apply budget-constrained high degree nodes preferential vaccination in real networks where finding such nodes is not possible because the whole network is unknown: instead of targeting random individuals, one could vaccinate random contacts of random individuals, thus greatly increasing the average degree of vaccinated persons.

\(^\circ\)Cohen, Havlin, and Ben-Avraham 2003

---

**Going further**

- Book *Dynamic processes on Networks*: Barrat, Barthelemy, and Vespignani 2008
- Surveys: 
  - Analysis and Control of Epidemics: Nowzari, Preciado, and Pappas 2016
  - Diffusion in networks: Lamberson 2016
  - Impact of community structure: Stegehuis, Van Der Hofstad, and Van Leeuwaarden 2016

---

**References**