Infectious disease modeling

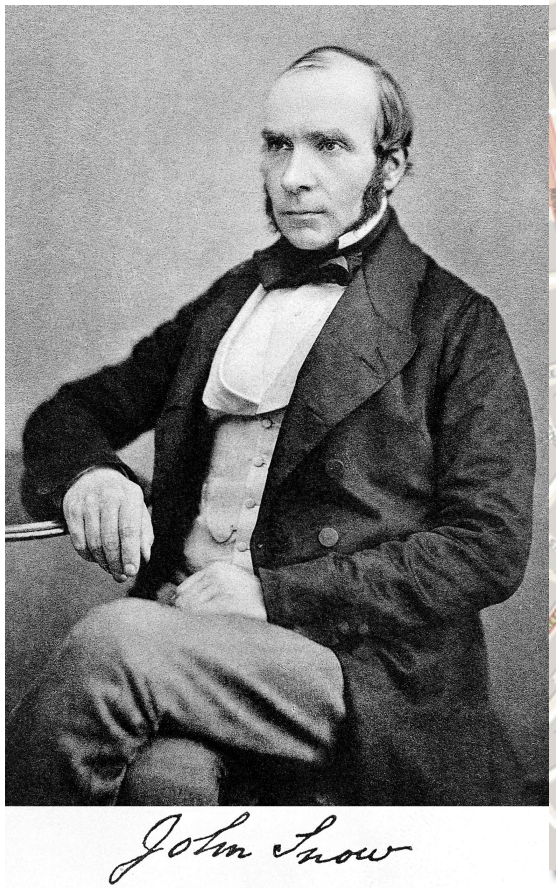
Matthew Macauley

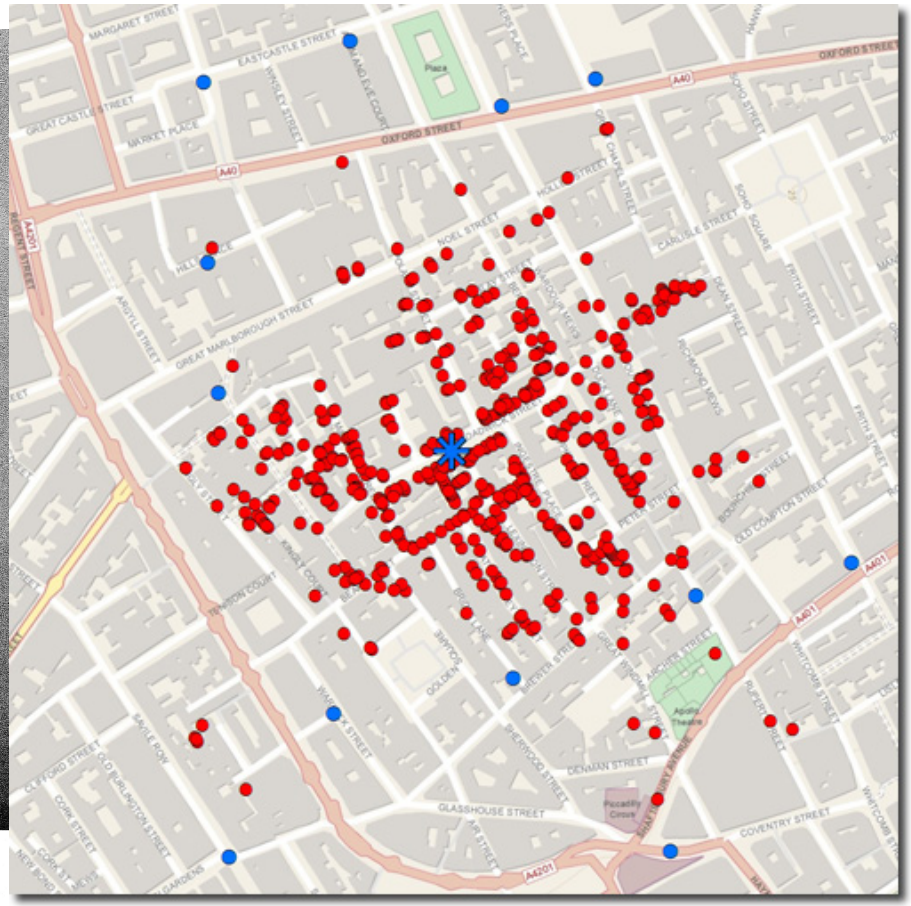
Department of Mathematical Sciences
Clemson University
http://www.math.clemson.edu/~macaule/

Math 4500, Spring 2017

Some history

John Snow (1813–1858) is widely considered to be the "father of epidemiology." In 1854, an outbreak of cholera struck the Soho neighborhood of London. Snow identified the source of the outbreak to be the Broad Street water pump.





The SIR model

Consider a disease spreading through a population, with the following assumptions:

- N people.
- 3 states: **S**usceptible, **I**nfected, **R**ecovered.
- Transition: $S \xrightarrow{\alpha} I \xrightarrow{\gamma} R$.
- $S_t + I_t + R_t = N$.
- No births or deaths.
- Population is homogeneously mixed.

This is just one example. These assumptions can be changed in other similar models.

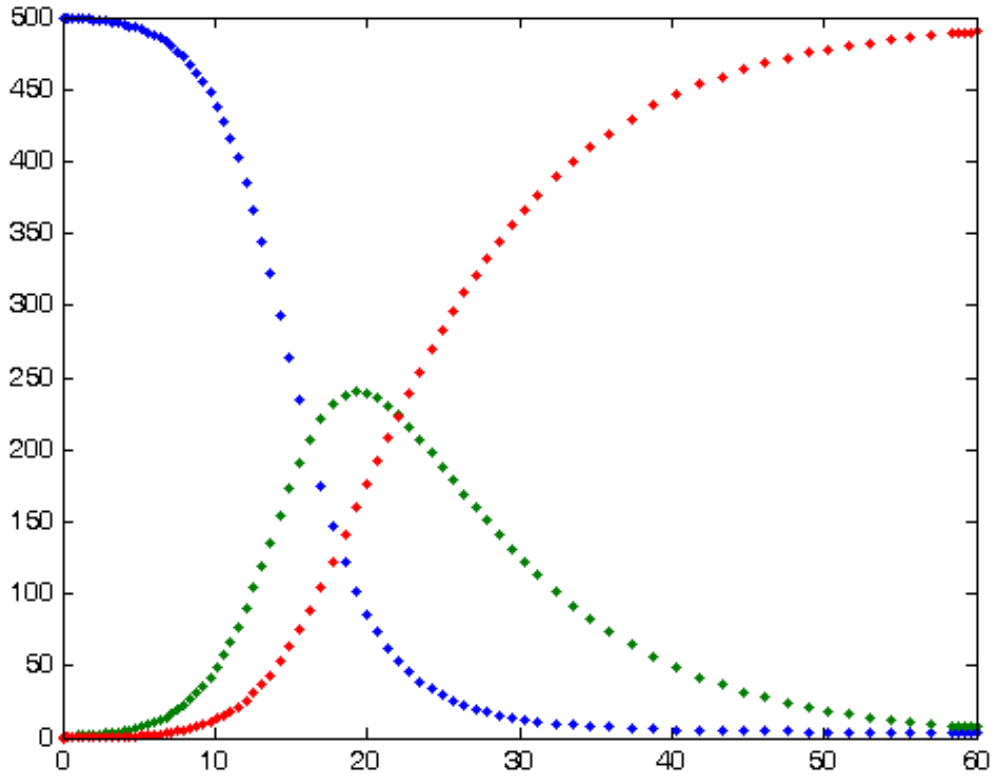
SIR model

The following is the "standard" SIR model, using ODEs (left) and difference equations (right):

$$\begin{cases} S' = -\alpha SI \\ I' = \alpha SI - \gamma I \\ R' = \gamma I \end{cases} \qquad \begin{cases} \Delta S = -\alpha SI \\ \Delta I = \alpha SI - \gamma I \\ \Delta R = \gamma I \end{cases}$$

SIR model: $S' = -\alpha SI;\ \ I' = \alpha SI - \gamma I;\ \ R' = \gamma I$

The following is what we should expect $S(t)$, $I(t)$, and $R(t)$ to look like, qualitatively.



Key aspect: $I' = (\alpha S - \gamma)I = \gamma\big(\frac{\alpha}{\gamma}S - 1\big)I$.

- If $S > \gamma/\alpha$, then $I' > 0$ (epidemic is *growing*).
- If $S < \gamma/\alpha$, then $I' < 0$ (epidemic is *shrinking*).

The value $\rho := \gamma/\alpha$ is a *threshold*, called the "*relative removal rate*".

SIR model: $S' = -\alpha SI;\ \ I' = \alpha SI - \gamma I;\ \ R' = \gamma I$

Definition

Initially, $I'(0) = (\alpha S(0) - \gamma) I(0)$. Define $\mathcal{R}_0 := \frac{\alpha}{\gamma} S(0)$, the "*basic reproductive number*".

- If $\mathcal{R}_0 > 1$, then $I'(0) > 0 \Longrightarrow$ epidemic occurs.
- If $\mathcal{R}_0 < 1$, then $I'(0) < 0 \Longrightarrow$ no epidemic occurs.

Key point

$\mathcal{R}_0$ represents the expected number of people an initially infected person will infect.

We'll see *why* this is true shortly.

First, here are estimates of $\mathcal{R}_0$ for some well-known diseases:

- 12–18 for measles
- 5–7 for smallpox and polio
- 2–5 for HIV and SARS
- 2–3 for the 1918 Spanish flu
- 1.5–2.5 for Ebola

SIR model: $\Delta S = -\alpha SI;\ \Delta I = \alpha SI - \gamma I;\ \Delta R = \gamma I$

Let's analyze $\mathcal{R}_0 := \frac{\alpha}{\gamma} S_0$ in the (discrete) setting of difference equations.

$\alpha S_0 I_0 = \#$ people infected in the first time-step.

$\alpha S_0 = \#$ people infected per-capita in the first time-step.

$\frac{1}{\gamma} \approx$ average duration of infection.

Putting this together, we get

$$\begin{aligned} \mathcal{R}_0 = \frac{\alpha}{\gamma} S_0 &= (\alpha S_0)\left(\frac{1}{\gamma}\right) \\ &= (\#\text{new infections per person per day})(\text{ave. duration}) \\ &= = \#\text{ secondary infections from one sick person} \end{aligned}$$

Policy goal

Inact steps to reduce $\mathcal{R}_0$ to be below 1. Note that:

- γ is more a property of the actual disease and hard to change;
- α depends both on the disease and on social factors (vaccines, hand-washing, isolation, etc.)

An example: $\Delta S = -\alpha SI;\ \Delta I = \alpha SI - \gamma I;\ \Delta R = \gamma I$

Consider an epidemic with the following properties, with a timestep of 1-day:

- Initially there are $S_0 = 500$ susceptibles.
- There is a 0.1% chance of transmission ($\alpha = .001$).
- 10-day illness ($\gamma = .1$).

The basic reproductive number is easily computed:

$$\mathcal{R}_0 = \frac{\alpha}{\gamma} S_0 = \frac{.001}{.1}(500) = 5\,.$$

On day 1, the expected number of infections will be

$$I_1 = \alpha S_0 I_0 + (1-\gamma) I_0 = .001(500)(1) + .9(1) = 1.4\,.$$

If measures (e.g., vaccinations) can be taken to reduce $S_0 = 90$, then $\mathcal{R}_0 = .9$ and the epidemic will be averted.

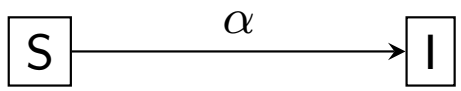
The epidemic subsides when $S_t < \rho = \frac{\gamma}{\alpha} = 100$ (i.e., when four-fifths of the population has contracted the disease).

Computationally, $\lim_{t\to\infty} S_t = 2.15$, which means that not everyone gets sick.

Other epidemic models

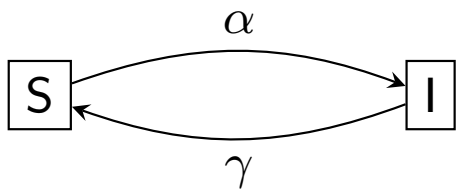
- SI model (e.g., herpes, HIV).

$$\begin{cases} S' = -\alpha SI \\ I' = \alpha SI \end{cases}$$



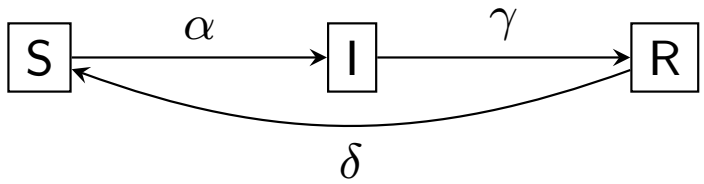
- SIS model. Disease w/o immunity (e.g., chlamydia).

$$\begin{cases} S' = -\alpha SI + \gamma I \\ I' = \alpha SI - \gamma I \end{cases}$$



- SIRS model. Finite-time immunity (e.g., common cold).

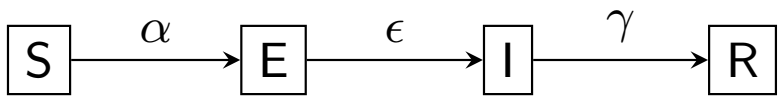
$$\begin{cases} S' = -\alpha SI + \delta R \\ I' = \alpha SI - \gamma I \\ R' = \gamma I + \delta R \end{cases}$$



Other epidemic models

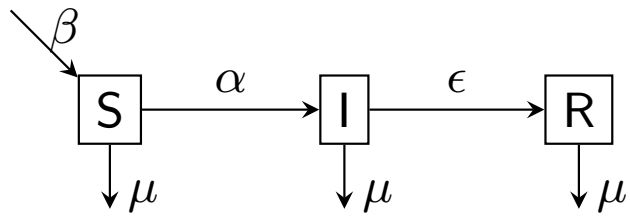
- SEIR model. E = exposed (incubation period, no symptoms).

$$\begin{cases} S' = -\alpha SI \\ E' = \alpha SI - \epsilon E \\ I' = \epsilon E - \gamma I \\ R' = \gamma I \end{cases}$$



- SIR model with birth and death rate.

$$\begin{cases} S' = -\alpha SI + \beta N - \mu S \\ I' = \beta SI - \gamma I - \mu I \\ R' = \gamma I - \mu R \end{cases}$$



Other approaches

Differential equations are continuous time, continuous space.

Difference equations are discrete time, continuous space.

Consider an SIR model that is discrete time and discrete space.

Let X be a social network:

- vertices: people
- edges: contacts
- edge weights: contact hours

Every person (node) has a state: S, I, or R (e.g., $x_j = 0, 1, 2$).

How to model agent-based disease transmission?

1. Bernoulli trials (weighted "coin flips")

Suppose i is infectious, and j is susceptible ($x_i = I$, $x_j = S$)

Let p be the probability that i infects j after 1 hour together:

$$\Pr(i \text{ infects } j) = p \qquad \Longrightarrow \qquad \Pr(i \text{ doesn't infect } j) = 1 - p\,.$$

If i and j spend $t > 1$ hours together, then

$$\Pr(i \text{ infects } j) = 1 - \Pr(i \text{ doesn't infect } j) = 1 - (1-p)^t$$

(assuming each hour is an independent event).

Now, suppose j comes in contact with individuals $i_1, \ldots i_k$, for a duration of $t_1, \ldots, t_k$, respectively.

$$\Pr(j \text{ gets infected}) = 1 - \Pr(\text{nobody infects } j) = 1 - \prod_{i=1}^{k} (1-p)^{t_i}\,.$$

How to model agent-based disease transmission?

2. Exponential distribution.

Suppose i is in contact with j for $t_{ij} \in [0, 24]$ hours during a given day, and say

$$\Pr(i \text{ infects } j) = 1 - e^{-rt_{ij}}.$$

Assume that the probability of getting infected by two different people are independent events, thus

$$\Pr(i \text{ not infected}) = \prod_{\text{edges } \{i,j\}} \left(1 - (1 - e^{-rt_{ij}})\right)$$

$$\Pr(i \text{ infected}) = 1 - \prod_{\text{edges } \{i,j\}} \left(1 - (1 - e^{-rt_{ij}})\right)$$

Regardless of which approch is taken, the computations are usually done with the aid of a computer, and rely more on simulation and statistical analysis.

Question

What are the pros and cons of using an agent-based model versus an *ODE-based model*?