Lecture 1: Basic Epidemic Models, Disease Data, and Challenges in Modeling Epidemics

Michael Li
University of Alberta
www.ualberta.ca/ myli
myli@ualberta.ca

Mathematical Modelling in Biology School
North West University, Potchefstroom, Mar. 20-28, 2023

Outline

- Epidemic Diseases and Endemic Diseases?
- Mathematical Theories of Models of Epidemic Diseases
- Modeling Endemic Diseases: $\mathcal{R}_0$ and the Threshold Theorem
- Parameter Estimation by Fitting Models to Data (Model Training)
- Why Modeling Epidemics Is Difficult: An Classical Example of Modeling Influenza

Materials in this lecture are taken from Michael Li, "An Introduction to Mathematical Modeling of Infectious Diseases", Springer, 2018.

Difference between an Epidemic and an Endemic Disease

Three temporal-spatial forms of infectious diseases:

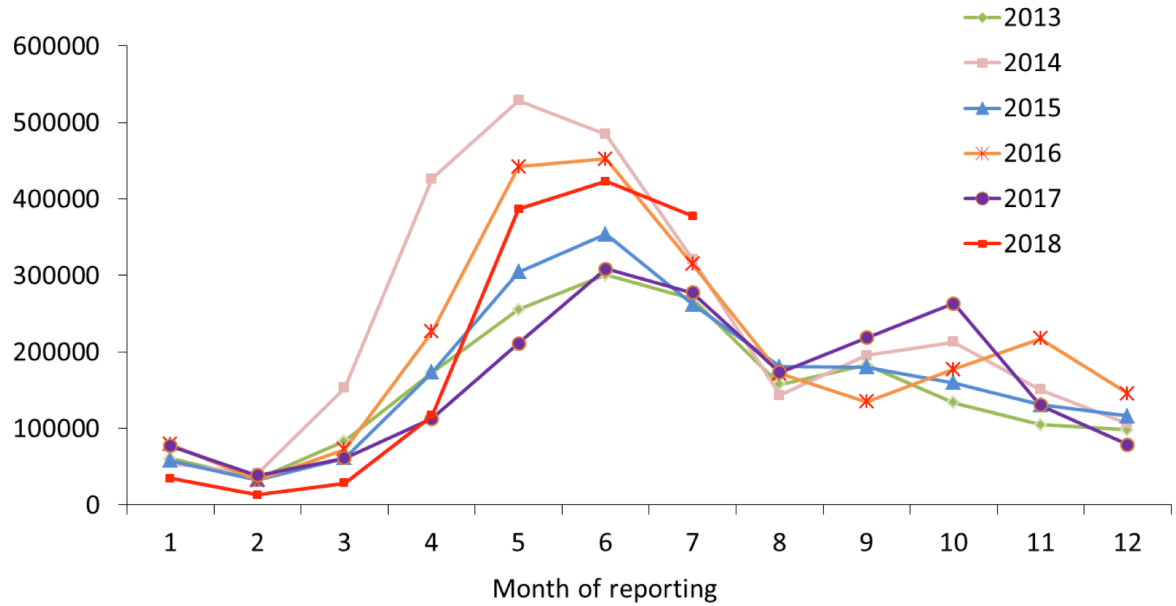
- An outbreak: **sudden increase of number of disease cases** in a small community.
- An epidemic: when an outbreak spreads to a larger area, infects large number of people and **terminates in a short time**, such as SARS-1 in 2003, seasonal influenza, a single wave of COVID-19, etc. When an outbreak spreads to multiple continents, it becomes an pandemic, Spanish flu, H1N1 pandemic in 2009, COVID-19 2020-23.
- Endemic diseases: when an outbreak spread and **persists for a very long time (years, decades)**, such as HIV, TB, Malaria, etc.

Mathematical theory of infectious diseases focuses more on endemic diseases and little on epidemics

The result is modelers faced huge problems in modeling COVID-19 epidemics.

Do We Know an Epidemic When We See One?

Annual data for Foot, Hand and Mouth Disease (FHMD) in a country:



Source: WHO

- Is this an epidemic disease or an endemic disease?
- If it is endemic, what is the best way to model it?
  - For mathematical purposes (mechanisms for long-time periodic patterns)
  - For public health purposes (annual epidemics, how long does it last, when does it peak, and how high is the peak?)

Is It Endemic or Epidemic?

For public health purposes, this disease maybe better modelled as an annual epidemic, to address more relevant public health concerns:

- the severity of the epidemic this year?
- when will the peak number of (case, hospitalizations, etc) occur?
- how high will the peak numbers be?

These questions are critical for health systems planning and resources allocations.

Some of the standard tools for infectious disease models:

- Computing $\mathcal{R}_0$ using next generation matrix
- Equilibria and stability analysis
- Bifurcation analysis
- Simulations

These are designed for asymptotic (long-time) behaviours of endemic diseases, rather than accurate predictions of the time course of a finite-time epidemic.

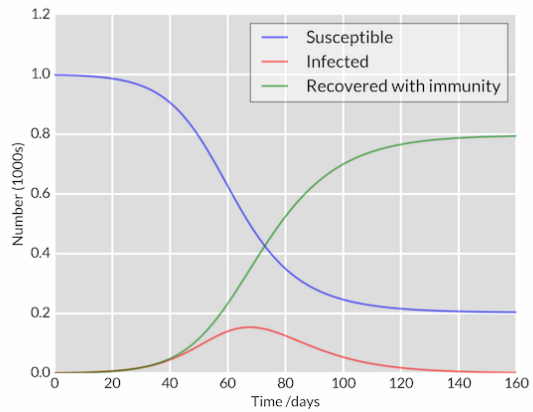
Part I: The Mathematical Theory of Epidemics

We use a simple SIR model to explain the theory.

$$
\begin{aligned}
S(t): &\quad \#\text{ of susceptible people at time } t \\
I(t): &\quad \#\text{ of infected people at time } t \\
R(t): &\quad \#\text{ of recovered people at time } t
\end{aligned}
$$

$$
\begin{aligned}
S' &= -\beta IS \\
I' &= \beta IS - \gamma I - dI \\
R' &= \gamma I
\end{aligned}
$$

initial conditions: $S(0) = S_0 > 0, I(0) = I_0 > 0, R(0) = R_0 \geq 0$.



An important assumption: Natural birth and natural death are negligible. (This captures the characteristics of an epidemic that it does not last forever).

Basic Reproduction Number and Critical Community Size

The basic reproduction number is the average number of secondary infections caused by a single infected during the infectious period within a susceptible population.

$$\mathcal{R}_0 = \frac{\beta}{\gamma + d} S_0 = \frac{S_0}{\rho}$$

where

$$\rho = \frac{\gamma}{\beta + d}$$

is the critical community size to sustain an epidemic:

- if $S_0 > \rho$ (or $\mathcal{R}_0 > 1$), an outbreak can occur
- if $S_0 < \rho$ (or $\mathcal{R}_0 < 1$), an outbreak will not occur

This follows from

$$I'(t) = (\gamma + d)\Big(\frac{S(t)}{\rho} - 1\Big) I(t).$$

Should we Use $\mathcal{R}_0$ or $\mathcal{R}_t$?

Let

$$\mathcal{R}_t = \frac{\beta}{\gamma + d} S(t) = \frac{S(t)}{\rho}$$

Then from

$$I'(t) = (\gamma + d)\Big(\frac{S(t)}{\rho} - 1\Big) I(t).$$

we know $\mathcal{R}_t$ satisfies:

- if $S(t) > \rho$ (or $\mathcal{R}_t > 1$), $I(t)$ increases
- if $S(t) < \rho$ (or $\mathcal{R}_t < 1$), $I(t)$ decreases
- when $S(t) = \rho$ (or $\mathcal{R}_t = 1$), $I(t)$ peaks

Differences between $\mathcal{R}_0$ and $\mathcal{R}_t$?

- The value of $\mathcal{R}_0$ impacts the dynamics at the beginning of an epidemic.
- The value of $\mathcal{R}_t$ reflects later phases of an epidemic.

More Properties of Epidemic Models

Property 1. $\lim_{t\to\infty}(S(t), I(t), R(t)) = (S(\infty), I(\infty), R(\infty))$ exists.

From the S equation

$$S'(t) = -\beta I(t)S(t) \leq 0.$$

Therefore, $S(t)$ is decreasing and bounded below by 0, and $S(\infty) \geq 0$ exists (Monotone Convergence Theorem).

Similarly,

$$R'(t) = \gamma I(t) \geq 0$$

and $R(t)$ is increasing and bounded above by N_0. Therefore, $R(\infty) \geq 0$ exists.

From

$$I(t) = N_0 - S(t) - R(t)$$

we know that $I(\infty) = N_0 - S(\infty) - R(\infty) \geq 0$ exists.

More Properties of Epidemic Models

Property 2. $S_0 > 0$ and $I_0 > 0$ imply $0 < S(\infty) < S_0$ and $I(\infty) = 0$.

Biologically, this says that

- An epidemic will terminate ($I(\infty) = 0$).
- At the end of an epidemic, there are individuals who escape the infection ($S(\infty) > 0$), namely, an epidemic will not infect everyone in the population. The limit $S_\infty = S(\infty)$ is called the **final size** of the epidemic.

Assume that $S_0 > 0, I_0 > 0, R_0 \geq 0$. Then $S(t) > 0, I(t) > 0, R(t) > 0$.

Dividing the equations for S and R, we obtain

$$\frac{dS}{dR} = -\frac{\beta}{\gamma} S.$$

Solving this equation for S as a function R, we obtain

$$S(R) = S_0 e^{-\frac{\beta}{\gamma} R} \geq S_0 e^{-\frac{\beta}{\gamma} N_0} > 0. \tag{1}$$

Therefore, $S(t) \geq S_0 e^{-\frac{\beta}{\gamma} N_0}$ for $t > 0$ and thus $S(\infty) \geq S_0 e^{-\frac{\beta}{\gamma} N_0} > 0$.

More Properties of Models of Epidemics

Next, we show that $\lim_{t\to\infty} I(t) = I(\infty) = 0$.

Since $S(\infty)$ and $I(\infty)$ exist, we know from the S equation

$$\lim_{t\to\infty} S'(t) = \lim_{t\to\infty}(-\beta S(t)I(t)) = -\beta I(\infty)\, S(\infty) \le 0 \quad \text{exists.}$$

We claim that $\lim_{t\to\infty} S'(t) = 0$.

Otherwise, if $\lim_{t\to\infty} S'(t) = \alpha < 0$, then $S'(t) < \alpha/2 < 0$ for $t \ge T$, for sufficiently large time T, and thus

$$S(t) < S(T) + \frac{\alpha}{2}(t - T) < 0 \quad \text{for } t > \min\{T, T - \frac{2S(T)}{\alpha}\}.$$

This contradicts $S(t) > 0$ for all $t > 0$. Therefore $\lim_{t\to\infty} S'(t) = 0$.

Then $S(\infty)I(\infty) = 0$. Since $S(\infty) > 0$, we know $I(\infty) = 0$.

First Integral and Global Phase Portrait of SIR Model

Phase Portrait: Sketches of representative trajectories of solutions in the phase space.

First Integral: A scalar-valued function $V(x)$ such that $V(x(t))$ is constant along a solution $x(t)$ of the ODE $x' = f(x)$.

An Example: for a conservative mechanical system, the total energy is a constant along any trajectory. So the total energy is a first integral for the system.

What do we use first integrals for? Dimension reduction. For a first integral $V(x)$ of an n-dimensional ODE $x' = f(x)$, the surface $V(x) = c$ is an invariant surface of dimension $n - 1$, namely, trajectories starting on this surface remain on this surface. We only need to study trajectories on a $n - 1$ dimensional invariant surface.

A First Integral for the SIR Model

Dividing the S, I equations we obtain

$$\frac{dI}{dS} = -1 + \frac{\gamma + d}{\beta}\frac{1}{S} = -1 + \frac{\rho}{S}, \tag{2}$$

where $\rho = (\gamma + d)/\beta$. Integrating (2) we obtain

$$I = -S + \rho \log S + c, \quad \text{for all } t > 0,$$

or

$$\phi(S, I) = I + S - \rho \log S = c, \quad \text{for all } t > 0, \tag{3}$$

where c is an arbitrary integration constant.

The function $\phi(S, I) = I + S - \rho \log S$ is a **first integral** of the SIR model. We can also verify this by

$$\frac{d}{dt}\phi(S(t), I(t)) = I'(t) + S'(t) - \rho\frac{S'(t)}{S(t)} = -(\gamma + d)I(t) + \rho\beta I(t) = 0$$

for all t.

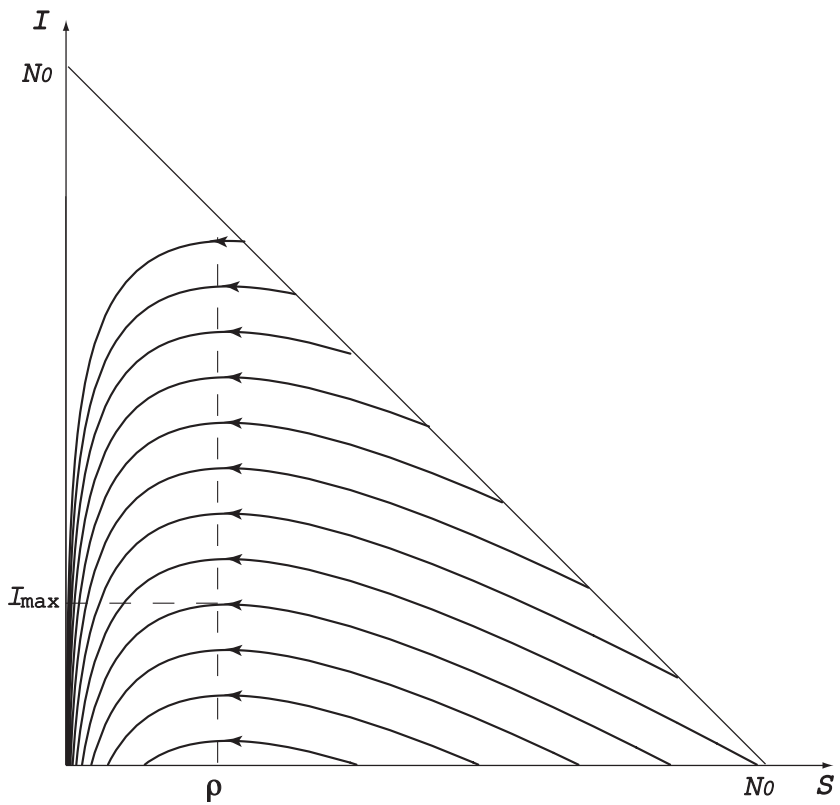
Global Phase Portrait of the SIR Model

Using the first integral $\phi(S, I) = I + S - \rho \log S$, the equation

$$I + S - \rho \log S = c \tag{4}$$

defines an invariant curve in the phase plane (SI-plane) for each c.

Plot the curves defined by equation (4) for all real numbers c, we obtain a family of trajectories, which gives us the global phase portrait.




```
%% Plot phase portrait of the Kermack-McKendrick model
%
% S' = -beta IS
% I' = beta IS - gamma I
%
%using the first integral equation
%
%I + S - rho log S = c

%We use contour plot to plot curves defined by the equation for all c.

clear all;
close all;

rho =2;
N0 = 10;

x = linspace(0,N0); % x bounds for the mesh
y = linspace(0,N0);     % y bound for the mesh
[X,Y] = meshgrid(x,y); % define a 2d mesh for the plot
Z = X+ Y - rho*log(X);
contour(X,Y,Z, 15)
xlabel 'S';
ylabel 'I';
```

Epidemic Curves and Their Characteristics

From the phase portrait, we observe the following:

- $S_0 = \rho$ is the critical community size ($\mathcal{R}_0 = 1$ is a threshold value)
- Larger populations (S_0) have a larger epidemic (for similar I_0)
- The larger the population, the smaller the final size S_∞.
- For the same population (S_0), larger I_0 produces a larger epidemic
- All epidemics peak at a time t when $S(t) = \rho$.

The decline of $S(t)$ determines the rise, peaking, and fall of an epidemic.

We will see later that this does not agree well with real-world epidemics.

The Final Size Formula

Number of people infected during the epidemic: $S_0 - S_\infty$

Proportion of infected population: $x = 1 - \frac{S_\infty}{S_0}$.

From the first integral $\phi(S, I) = S - \rho log S + I$, we obtain

$$S(t) - \rho \log S(t) + I(t) = S_0 - \rho \log S_0 + I_0.$$

Assuming $I_0 \approx 0$, and letting $t \to \infty$, we obtain

$$S_\infty - \rho \log S_\infty = S_0 - \rho \log S_0.$$

and thus

$$S_0 - S_\infty = \rho \log \frac{S_0}{S_\infty},$$

or in term of $x = S_\infty / S_0$, using $\mathcal{R}_0 = S_0/\rho$:

$$\mathcal{R}_0(1 - x) + \log x = 0.$$

Threshold Theorem for Endemic Diseases

An endemic disease lasts a very long time (years), and the natural birth and death cannot be ignored.

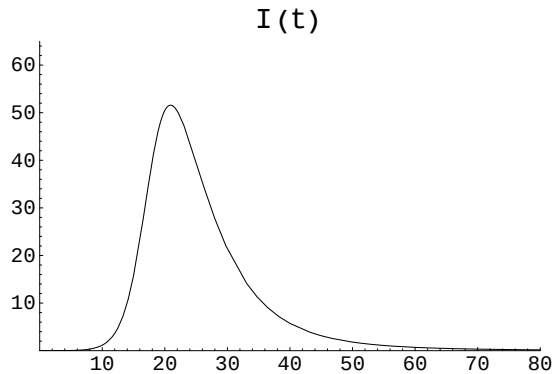
For illustration, we use an SIR model with birth and death to model an endemic disease

$$\begin{aligned} S' &= \Lambda - \beta IS - d_1 S \\ I' &= \beta IS - \gamma I - d_2 I \\ R' &= \gamma I - \ d_3 R \end{aligned}$$

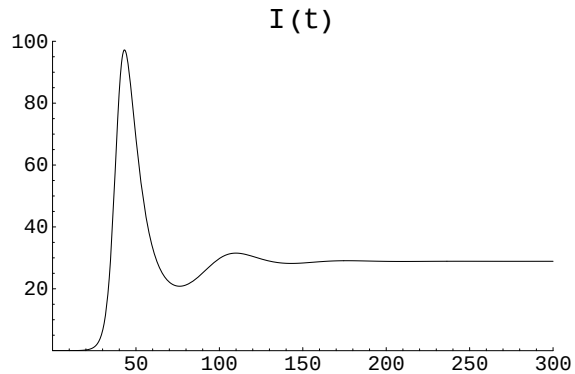
initial conditions: $S(0) = S_0 > 0, I(0) = I_0 > 0,\ R(0) = R_0 = 0.$

$$\mathcal{R}_0 = \frac{\beta}{\gamma + d_2}\frac{\Lambda}{d_1} = \beta \cdot \frac{1}{\gamma + d_2} \cdot \frac{\Lambda}{d_1}.$$

Two possible disease outcomes:



(a) If $\mathcal{R}_0 \leq 1$.



(b) If $\mathcal{R}_0 > 1$.

Threshold Theorem for Endemic Diseases

Threshold Theorem

- If $\mathcal{R}_0 \leq 1$, then the disease-free equilibrium $P_0 = (\bar{S}, 0, 0)$ is globally asymptotically stable, and the disease always dies out irrespective of the initial number I_0.

- If $\mathcal{R}_0 > 1$, then P_0 becomes unstable and a unique endemic (positive) equilibrium $P^* = (S^*, I^*, R^*)$ comes to existence and is always globally asymptotically stable irrespective of the initial number I_0.

Here $\bar{S} = \frac{\Lambda}{d_1}, \quad I^* = 1 - \frac{1}{\mathcal{R}_0}$.

The Threshold Theorem can be illustrated by the bifurcation diagram:

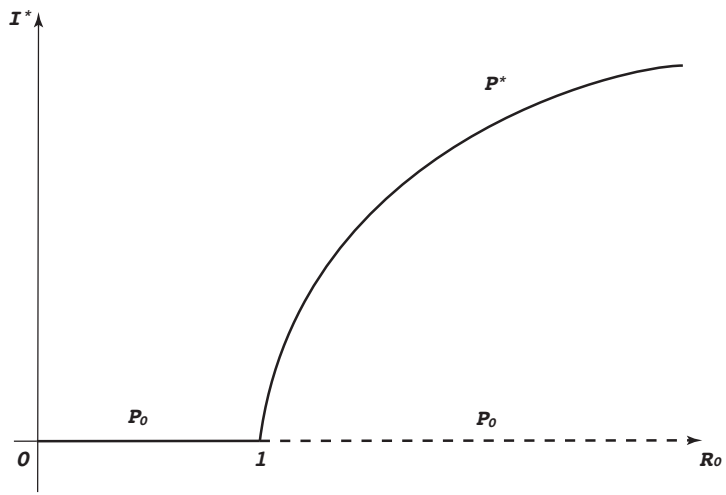


Figure: Bifurcation diagram for a model with demorgraph

Proof I: Using the Poincaré-Bendixson Theorem

Proof 1 Write S and I sub-system

$$\begin{aligned} S' &= \Lambda - \beta IS - d_1 S \\ I' &= \beta IS - \gamma I - d_2 I \end{aligned} \tag{5}$$

as

$$\begin{aligned} S' &= P(S, I) \\ I' &= Q(S, I). \end{aligned} \tag{6}$$

We want to verify the Bendixson-Dulac condition

$$\frac{\partial}{\partial S}(\alpha P) + \frac{\partial}{\partial I}(\alpha Q) < 0 \quad \text{in } \mathbb{R}^2_+.$$

Let $\alpha(S, I) = \frac{1}{I}$ be a Dulac multiplier. Then

$$\begin{aligned} \frac{\partial}{\partial S}(\alpha P) + \frac{\partial}{\partial I}(\alpha Q) &= \frac{\partial}{\partial S}\left(\frac{b}{I} - \beta S - \frac{bS}{I}\right) + \frac{\partial}{\partial I}(\beta S - b - \gamma) \\ &= -\beta - \frac{b}{I} < 0, \quad \text{in } \mathbb{R}^2_+. \end{aligned}$$

Proof II: Using the Method of Lyapunov Functions

For proof of GAS of the disease-free equilibrium P_0, we will see that our first integral $\phi(S, I)$ for models of epidemics becomes an Lyapunov function for models of endemic diseases:

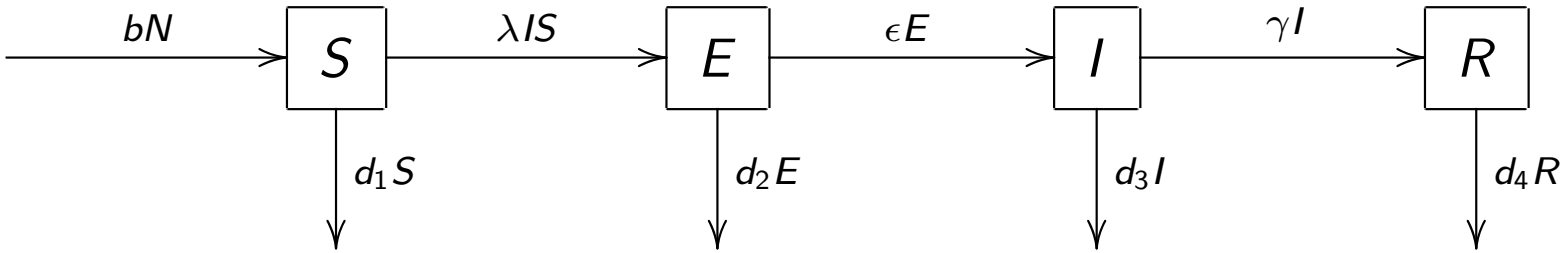
$$V(S, I) = S - \rho \log S + I, \quad \rho = \frac{\gamma + d_2}{\beta}.$$

$$\begin{aligned}
\frac{d}{dt} V(S(t), I(t)) &= S'(t) - \rho \frac{S'(t)}{S(t)} + I'(t) \\
&= \Lambda - \beta S(t) I(t) - d_1 S(t) - \frac{\rho}{S(t)} (\Lambda - \beta S(t) I(t) - d_1 S(t)) \\
&\quad + \beta S(t) I(t) - (\gamma + d_2) I \\
&= \Lambda - 2 d_1 \rho - \frac{\Lambda \rho}{S} + I(\beta \rho - (\gamma + d_2)) \\
&= 2 d_1 \rho - d_1 \rho - \frac{d_1 \rho}{S} = d_1 \rho (2 - \frac{S}{\rho} - \frac{\rho}{S}) \quad \leq \ 0.
\end{aligned}$$

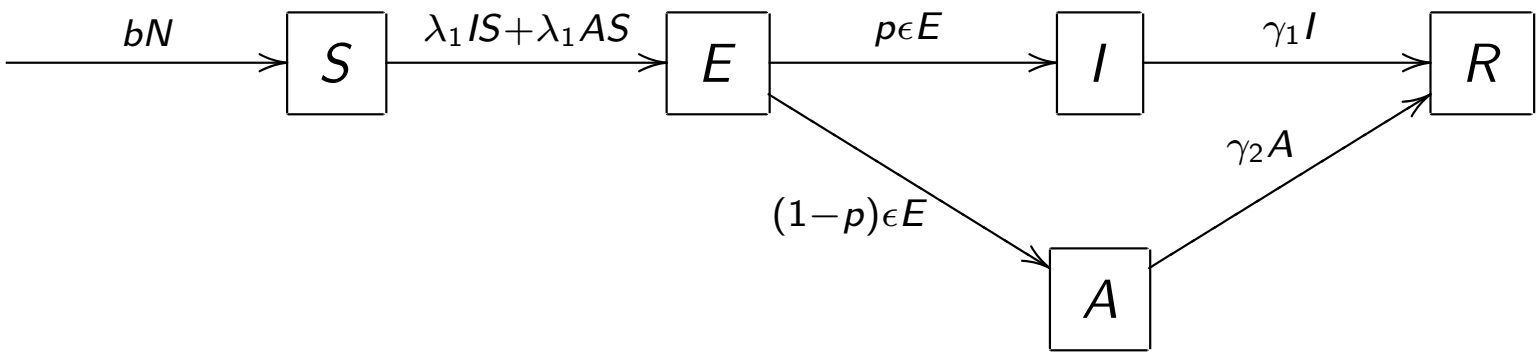
Therefore, $\frac{d}{dt} V(S(t), I(t)) \leq 0$, and we proved global stability of P_0.

Some Examples of More Complex Epidemic Models

I. SEIR Models for diseases with latency:

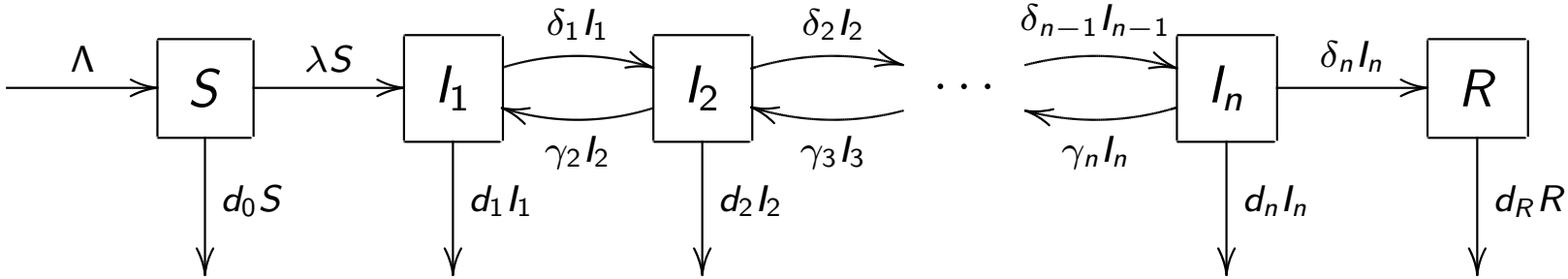


II. SEIAR Models for diseases with latency and asymptomatic state:

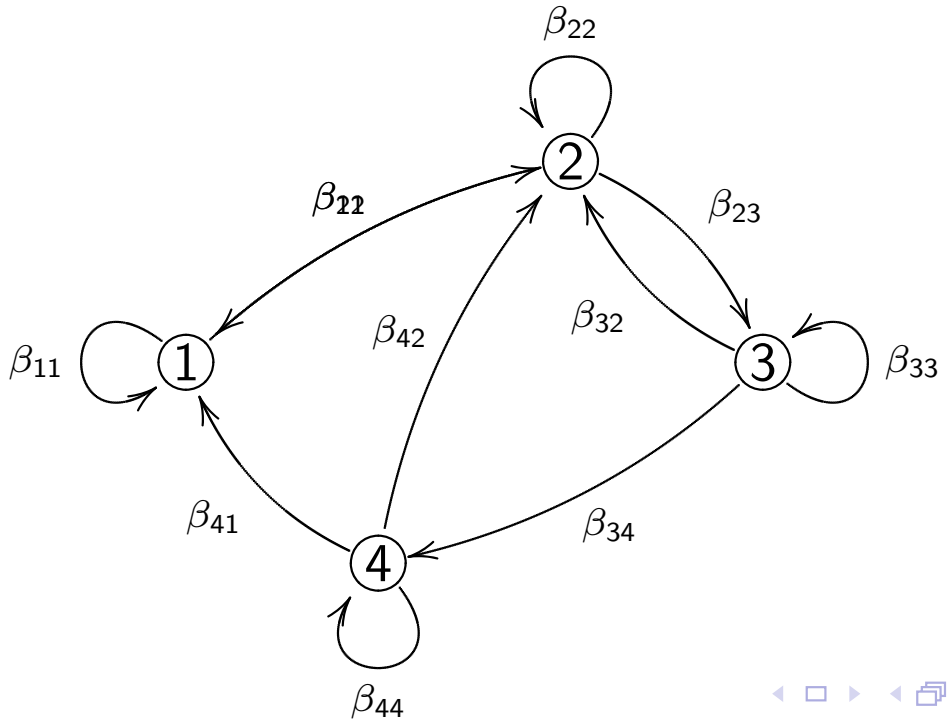


Some Examples of More Complex Epidemic Models

III. Stated Progression Models for diseases with long infectious period:

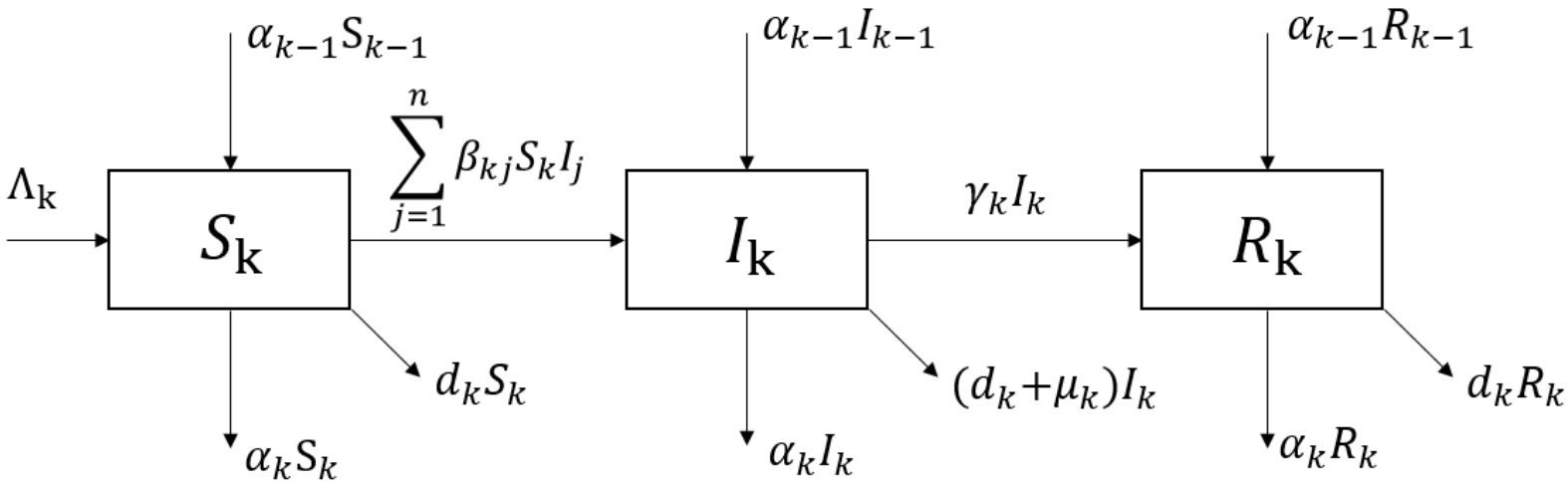


IV. Multi-group Models for diseases in heterogeneous population:



Some Examples of More Complex Epidemic Models

V. Age-Group Models Incorporating age-structure in diseases transmission: Divide the population into n discrete age groups: for each $1 \leq k \leq n$:



Some Examples of More Complex Epidemic Models

VI. Multi-City Models incorporating spatial movement in disease transmission. An example of such a model is:

$$S_i' = \Lambda_i - \beta_i S_i I_i - d_i^S S_i + \sum_{j=1}^{n} a_{ij} S_j - \sum_{j=1}^{n} a_{ji} S_i,$$

$$I_i' = \beta_i S_i I_i - (d_i^I + \gamma_i) I_i + \sum_{j=1}^{n} b_{ij} I_j - \sum_{j=1}^{n} b_{ji} I_i,$$

$$R_i' = \gamma_i I_i - d_i^R R_i + \sum_{j=1}^{n} c_{ij} R_j - \sum_{j=1}^{n} c_{ji} R_i,$$

$$i = 1, 2, \ldots, n.$$

Parameter Estimation

- Model projections critically depend on the values of parameters
- Parameters for epidemic models include both the rate constants and initial conditions (which are typically unknown).
- Values of some parameters can be **estimated directly from data**:
  - Natural death rate can be estimated as $1/L$, where L is the life expectancy
  - Recovery rate can be estimated as $1/D$, where D is the mean infectious period
- Values of some key parameters such as transmission coefficient (β) and initial conditions are difficult to estimate directly from data.
- The values of these parameters need to be **estimated by fitting model output to known data**
- Several model fitting approaches exist, e.g. least squares method, maximum likelihood method, Bayesian inference-based Monte Carlo Markov chain (MCMC) method, and others
- Model fitting methods are similar to methods of regression for Statistical models, and model training for Machine Learning models
- To demonstrate principles of model fitting, we will demonstrate the Least Squares method.

Parameter Estimation for Epidemic Models

Suppose that our epidemic model is described by the initial value problem of a system of differential equations:

$$\begin{aligned} x' &= f(x, \theta), \quad x \in \mathbb{R}^d, \ t \in [0, t_{\max}], \\ x(0) &= x_0. \end{aligned} \tag{7}$$

- $\theta \in \mathbb{R}^m$ is an m-dimensional parameter
- $[0, t_{\max}]$ is the time interval in which we investigate the epidemic.
- x_0 is the initial conditions (some components are often unknown)

The disease data is often given at discrete observation time points $t_1, \ t_2, \ \cdots, \ t_p \in [0, t_{\max}]$ in the form

$$(t_1, g(x^{(1)})), \ (t_2, g(x^{(2)})), \ \cdots, \ (t_p, g(x^{(p)})). \tag{8}$$

Not all states in the model are directly measurable. The function $g(x) : \mathbb{R}^d \to \mathbb{R}^n$ represents measurable quantities of the state variable x, also called the **output function**.

Matching Model Outputs with the Corresponding Data

To fit with data, it is natural that we only consider values of the solution $x(t, \theta)$ at the observational time points:

$$x(t, \theta) \approx (x(t_1, \theta), x(t_2, \theta), \cdots, x(t_p, \theta))^T.$$

and match the observable model outputs:

$$g(x(t, \theta)) \approx g((x(t_1, \theta)), g(x(t_2, \theta)), \cdots, g(x(t_p, \theta))^T.$$

with the observed data points:

$$y = (g(x^{(1)}), g(x^{(2)}), \cdots, g(x^{(p)}))^T,$$

We can form the squared sum of errors (SSE) between the observable parts of the solution and data:

$$\mathrm{SSE}(\theta) = ||g(x(t, \theta)) - y||^2 = \sum_{i=1}^{p} ||g(x(t_i, \theta)) - g(x^{(i)})||^2. \qquad (9)$$

where $||\cdot||$ is any vector norm of $\mathbb{R}^n$.

Initial Conditions as Parameters

- In our discussion, we consider that the initial conditions are known and fixed.
- In reality, some components of the initial conditions are unknown (e.g. $I(0)$)
- In practice, unknown initial conditions are also considered as parameters and will be estimated from fitting model to data.

Compare Apples to Apples, Oranges to Oranges: When forming the SSE, it is crucial to measure differences in quantities of the same type. Mismatching model output with data is a common mistake people makes.

The Nonlinear Least Squares Fitting: We look for a value $\hat{\theta}$ of model parameter θ such that $\mathrm{SSE}(\theta)$ is the minimum.

$$\mathrm{SSE}(\hat{\theta}) = \min\{\mathrm{SSE}(\theta) \ : \ \theta \in G \subset \mathbb{R}^p\}.$$

Such a problem is a nonlinear least-squares problem, since the dependence of a solution $x(t, \theta)$ on the parameter θ is through a highly nonlinear system of differential equations.

Matlab Functions for Nonlinear Least Squares

Several Matlab functions can be used for parameter estimation.

- **lsqcurvefit** requires the following inputs: the model equation, an initial guess for the parameters to be fitted, and the time points and data points. It then solves the nonlinear least-squares problem directly.
- **nlinfit** is another nonlinear regression routine that uses an iterative least-squares estimation with an initial value for the parameters.
- **fminsearch** is used to minimizes $\mathrm{SSE}(\theta)$. The *fminsearch* takes $\mathrm{SSE}(\theta)$ and an initial guess θ_0 of the parameter value, and uses a direct search routine to find a minimum value of SSE. Each step involves numerically solving the ODE model with a new set of parameter values. To ensure the minimum value returned by *fminsearch* is not just a local minimum, the process can be repeated with several choices of initial guess.

An Epidemic Model Example

We use the Kermack-McKendrick SIR model for demonstration:

$$\begin{aligned} S' &= -\lambda IS \\ I' &= \lambda IS - \gamma I \\ R' &= \gamma I. \end{aligned} \tag{10}$$

Our objective is to estimate the parameters λ and γ by fitting the model to disease data.

Our data is on the number of infected $I(t)$ and total population $N(t)$.

Step I: Codes for Solving ODE Models

1. Define a function f that takes three inputs.

```
function f=SIRmodel(t,
  y,par)
%  Label the
  parameters and state
   variables
   lambda = par(1);
   gamma = par(2);
   S=y(1);
   I=y(2);
   R=y(3);
   N=S+I+R;
% Input the
  differential
  equations
   Sdot=-lambda*I*S;
   Idot=lambda*I*S-
   gamma*I;
   Rdot=gamma*I;
   f=[Sdot Idot Rdot]';

end
```

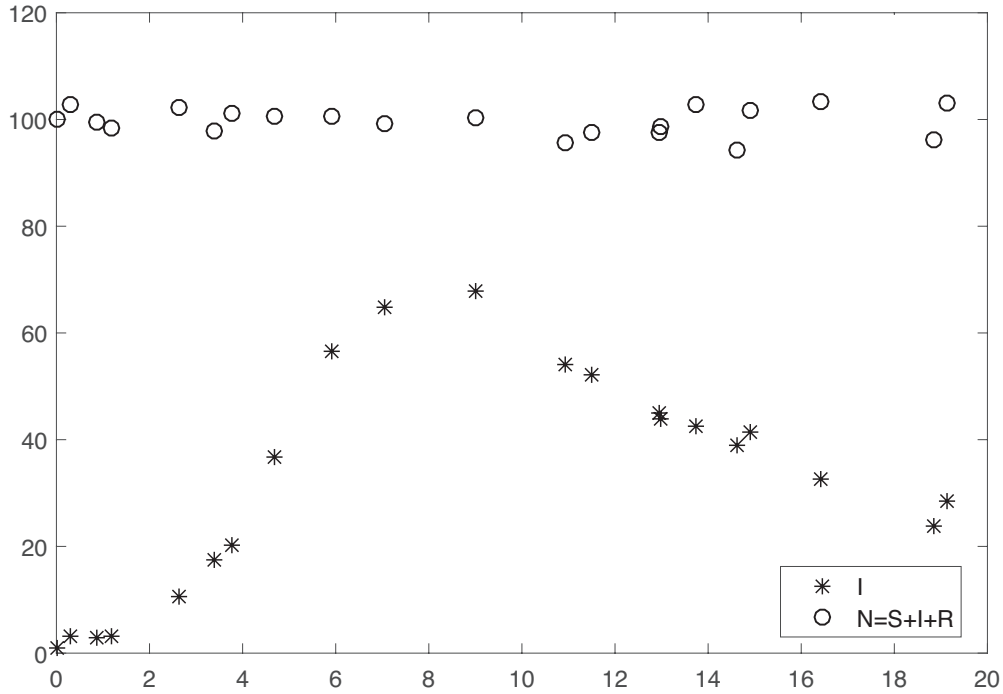
2. Use the ODE solver ODE45.

```
function sol=SIRSol(
  par,IC,t)
% disp(num2str(par))
   DeHandle=@(t,y)
   SIRModel(t,y,par);
   [~, Y]=ode45(
   DeHandle,t,IC);
   sol=Y';
end
```

Step II: Format Data

- For demonstration purpose, we will artificially generate data at 20 time points
- Data is generated from the model with prescribed values for λ, γ.
- Random noise is added to the generated data.
- We will be able to compare the best-fit values $\hat{\lambda}, \hat{\gamma}$ to the prescribed values.

This the data we generated



Step II: Generating Data

1. First take 20 randomly chosen time points:

```
numpts=20;
tdata=[0 sort(20*rand
  (1,numpts))];
```

2. Then we generate normally distributed noise:

```
width=0.1;
ndataSIR=20*[ 0
  normrnd(0,width,[1,
  numpts]);
      0 normrnd(0,
  width,[1,numpts]);
      0 normrnd(0,
  width,[1,numpts])];
```

3. Add the noise term to the model outputs with $\lambda = 0.01$ and $\gamma = 0.1$ and initial conditions $S(0) = 50$, $I(0) = 1$, and $R(0) = 0$ to produce a set of artificial data:

```
numpts=20;
tdata=[0 sort(20*rand
  (1,numpts))];
```

4. Generate data for $I(t)$ and $N(t)$:

```
SIRData=[0 1 0 ; 1 1
  1]*SIRData;
```

Step III: Define SSE and Call fminsearch

1. Define SSE:

```
SIRparSol= @(par,t) [0 1 0 ; 1 1 1]*SIRSol([par(1) par(2)
  ], IC, t);
SumSquaresSIR = @(par) sum(sum((SIRparSol(par,tdata)-
  SIRData).^2));
```

2. Call *fminsearch* with an initial guess for the parameter values $\lambda = 8, \gamma = 0.02$:

```
[SIRtheta,fval,exitflag] = fminsearch(SumSquaresSIR,[8
  0.02]);
SIRsol=SIRparSol(SIRtheta,tsol);
```

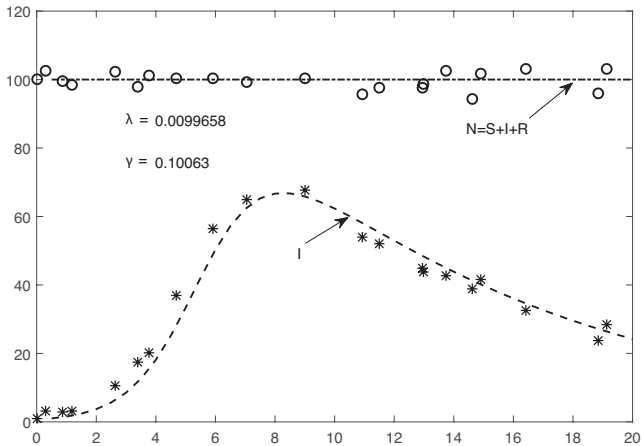
3. Plot the solutions in comparison to data

```
figure;
plot(tdata,SIRData,'.');
hold on;
plot(tsol,SIRsol,'--');
```

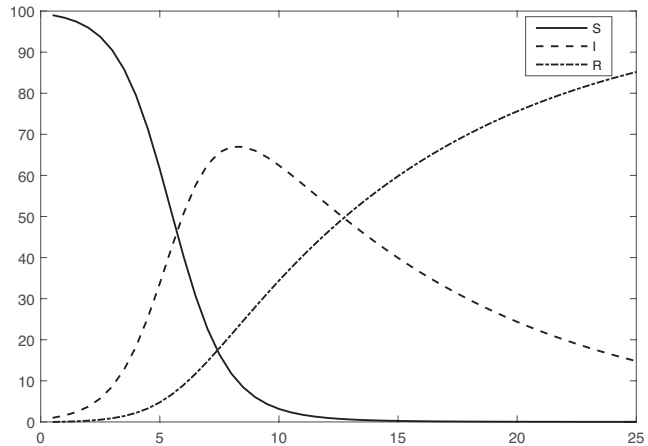
Fitting Result

The best-fit values are $\lambda = 0.0099658$ and $\gamma = 0.10063$, which are very close to the pre-assigned values $\lambda = 0.01$ and $\gamma = 0.1$.

Visualization of the fitting results:



(a) Fitting of model outputs with data



(b) Solutions with Best-fit parameter values

Matlab Codes for the fitting is available.

Challenges in Parameter Estimation of Epidemic Models

- In the the standard text of epidemic modeling and research literature, it is has been a standard practice to assume that we can directly observe the infected $I(t)$.
- In Public Health practices, disease data are reported cases.
- An infectious disease case is an individual who is **infected**, has **taken** a diagnostic test, and has a **positive** test result, and the result has be **recorded**, and **entered** into the public health data system.
- Any of the above step break down, the infected individual is not part of the case data.
- A case has a very strict case definition. When cases definition changes, the recorded data also changes.
- Disease cases is a (often small) fraction of the infected population.

A Classic Example of Mis-Fitting

Modeling a flu epidemic in a boy's boarding school in the UK in 1978.

Influenza Epidemic in an English Boarding School 1978

In 1978 in the British medical journal, *The Lancet*, there was a report with detailed statistics of a flu epidemic in a boys' boarding school with a total of 763 boys. Of these 512 were confined to bed during the epidemic, which lasted from 22nd January to 4th February 1978. It seems that one infected boy initiated the epidemic. This situation has many of the requirements assumed in the above model derivation. Here, however, the epidemic was severe and the full system has to be used. Also, when a boy was infected he was put to bed and so we have $I(t)$ directly from the data. Since in this case we have no analytical solution for comparison with the data, a best fit numerical technique was used directly on the equations (10.1)–(10.3) for comparison of the data. Figure 10.3 illustrates the resulting time evolution for the infectives, $I(t)$, together with the epidemic statistics. The R-equation (10.3) is uncoupled; the solution for $R(t)$ is simply proportional to the area under the $I(t)$ curve.

A Classic Example of Mis-Fitting

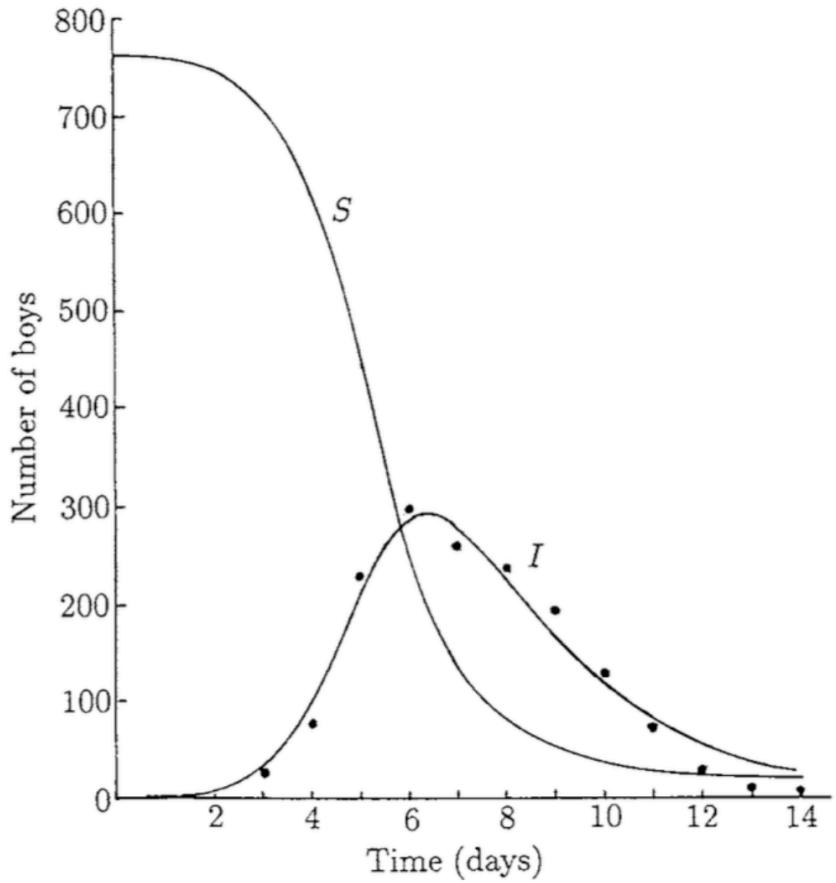


Figure 10.3. Influenza epidemic data (•) for a boys' boarding school as reported in the British medical journal, *The Lancet*, 4th March 1978. The continuous curves for the infectives (I) and susceptibles (S) were obtained from a best fit numerical solution of the *SIR* system (10.1)–(10.3): parameter values $N = 763$, $S_0 = 762$, $I = 1$, $\rho = 202$, $r = 2.18 \times 10^{-3}$/day. The conditions for an epidemic to occur, namely, $S_0 > \rho$, are clearly satisfied and the epidemic is severe since R/ρ is not small.

The Original Source - Paper in the BMJ, 1978

EPIDEMIOLOGY

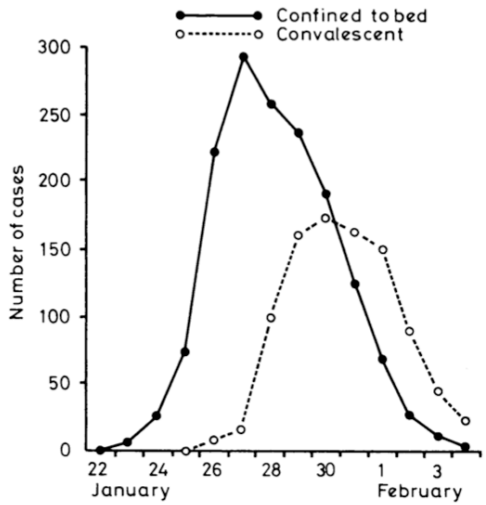
Influenza in a boarding school

The following notes are compiled by the Communicable Disease Surveillance Centre (Public Health Laboratory Service) and the Communicable Diseases (Scotland) Unit from reports submitted by microbiological laboratories, community physicians, and environmental health officers.

During January an epidemic of influenza occurred in a boarding school in the north of England. A total of 763 boys between the ages of 10 and 18 were at risk, all except 30 being full boarders; the staff were from the surrounding villages. There were 113 boys between the ages of 10 and 13 in the junior house, while the rest were divided into 10 houses of about 60 boys each.

The Easter term began on 10 January, with boys returning from all over Britain and some from Europe and the Far East. One boy from Hong Kong had a transient febrile illness from 15 to 18 January. On Sunday 22 January three boys were in the college infirmary. The graph shows the daily total number confined to bed or convalescent during the epidemic: 512 boys (67%) spent between three and seven days away from class, and 83% of the boys in the junior house were affected. Of about 130 adults who had some contact with the boys, only one, a house matron, developed similar symptoms.

Most of the boys who became ill first complained of feeling very tired, with headache as fever developed, and sore throat and tracheitis being the rule. The temperature was usually 100°-102°F (38°-39°C) and often higher in the morning. Three boys with no other abnormal



signs had temperatures of 105°-106°F (40°-41°C). Many had mild reddening of the anterior pillars of the fauces, but the throat never looked as inflamed as symptoms suggested. In only five boys were there abnormal signs on chest examination. Symptoms subsided quickly once the boys were confined to bed. They were allowed up 36 hours after their temperatures had returned to normal and back to classes two to four days later, depending on the severity of the attack. The average time off sick was five to six days.

One boy of 13 was readmitted after two days with probable bacterial pneumonia, with a temperature of 104°F (40°C), pulse rate of 110/min, respiration rate of 22/min, and moist sounds in his right lung. He was given ampicillin and by next morning his temperature was 99°F (37°C) and his chest clear. Five days later he went home to convalesce. Four boys developed wheezy bronchitis. Two received ampicillin and two tetracycline. All recovered quickly and were back at work in seven to eight days. Four boys with otitis media, with bulging red ear drums, responded to ampicillin within 48 hours and none had any aural discharge. One boy had sinusitis, which again responded to ampicillin. He was in bed for seven days and off work for ten days. In all, only 10 of the 512 boys who became ill received antibiotics.

Throat swabs were taken from eight boys, and influenza A viruses similar to A/USSR/90/77 (H1N1) were isolated from six. The spread of this virus through the school was much more rapid than in the outbreaks due to influenza B in November 1954 and to influenza A (Asian flu) H2N2 in October 1957. These two epidemics reached their peak in two weeks and lasted four weeks. This year's epidemic reached a peak in seven days and was over in 13 days. Influenza vaccine (Fluvirin) had been given to 630 boys in October 1977—as had been the practice for some years. The incidence of influenza among the boys had been low except in those years in which a definite antigenic shift occurred. The fact that this is the first major outbreak of influenza at the school since the Asian flu suggests that influenza vaccination has a useful role in a boarding school. Had it been possible to include the H1N1 strain in the vaccine a major outbreak might well have been avoided.

What Did the Original Source Say About the Epidemic?

During January an epidemic of influenza occurred in a boarding school in the north of England. A total of 763 boys between the ages of 10 and 18 were at risk, all except 30 being full boarders; the staff were from the surrounding villages. There were 113 boys between the ages of 10 and 13 in the junior house, while the rest were divided into 10 houses of about 60 boys each.

The Easter term began on 10 January, with boys returning from all over Britain and some from Europe and the Far East. One boy from Hong Kong had a transient febrile illness from 15 to 18 January. On Sunday 22 January three boys were in the college infirmary. The graph shows the daily total number confined to bed or convalescent during the epidemic: 512 boys (67 %) spent between three and seven days away from class, and 83 % of the boys in the junior house were affected. Of about 130 adults who had some contact with the boys, only one, a house matron, developed similar symptoms.

Most of the boys who became ill first complained of feeling very tired, with headache as fever developed, and sore throat and tracheitis

How to Correctly Interpret the Data?

From the source article we know the following

- The curve show the daily number of boys who were sick and were confined to bed in the infirmary.
- During any particular day, there are infected boys in the infirmary, but also infected boys who are going about in the school with not apparent symptoms, but may be infectious.
- the curve is a standard epidemic curve: curve of **daily new identified cases**.
- In our lectures, we have been calling $I(t)$ the epidemic curves, but in reality, it is a curve of new cases.
- We should not fit $I(t)$ to the epidemic curve.
- Which term in the SIR model should we fit to the epidemic curve?
- Some people fit the βIS term to the epidemic curve, since they both represent "incidence".

Our SIR Model Does Not Predict Cases!

Our SIR model:

$$\begin{aligned} S' &= -\beta IS \\ I' &= \beta IS - \gamma I \\ R' &= \gamma I. \end{aligned} \tag{11}$$

Terms in the model:

- βIS: Incidence of infection. Different from incidence of cases (daily new case reports) in epidemiology
- $I(t)$: Prevalence of infection (# of people living with the infection). Different from prevalence of cases (cumulative number of cases) in epidemiology

Our SIR model predict infections, not identified cases!

What is Actually Wrong in the Example? Model Validation

Question: If we ignore the mis-matching in the example (we shouldn't), and assume that the fitting allowed an accurate description of $I(t)$, would that be enough?
The answer: No.

Important: Fitting between model output and data is only model training. Another important part of modeling is model validation.

Methods of Model Validation

- Method of Cross-Validation: split the data into two parts: training data and validation data. Use training data for model training, and validate the model using validation data.
- Using data that is independent of training data for validation

For the example, we know the final size of the epidemic: 512 (67%) of the students are infected. From the plot, we see that total number of infected students predicted by the model is 730 (96%)

While the model "accurately" describe the $I(t)$, but it seriously over-predicted the final size.