

1 Epidemic Modelling Intro

2 SI Model



Mathematics
and Statistics

$$\int_M d\omega = \int_{\partial M} \omega$$

Mathematics 4MB3/6MB3 Mathematical Biology

Instructor: David Earn

Lecture 1
Epidemic Modelling Intro
Monday 7 January 2019

Where to find course information

- The course web page:
<http://www.math.mcmaster.ca/earn/4MB3>
- Click on “Course information sheet”.
- Download pdf or read online.
- Let's [have a look now...](#)

Who is NOT available at these times?

- Monday 9:30-10:30
- Wednesday 9:30-10:30
- Thursday 9:30-10:30
- Thursday 11:30-12:20
- Friday 12:30-1:20
- Friday 2:30-3:20

Group formation

Most work in this course will be done in groups.

- Attempt to form a group of 4 students (you and 3 others) **no later than Wednesday night this week.**
- After you have done your best to form a group of four, **exactly one** member of your group must **e-mail the instructor no later than Wednesday night this week:**
 - Include “Math 4MB3” and your proposed group name in the subject line.
 - **Copy your message to all members of your proposed group so I have everyone’s e-mail in the thread.**
- If you were unable to form a group, then e-mail the instructor explaining what you did to try to form a group, and describe your skills/preferences. (*This is a last resort – please try your best to form a group.*)
- *Instructor may change groups based on survey results.*

Online Surveys

You will be required to fill in online surveys during this course. Doing so in a timely manner contributes to your participation mark.

The first online survey has been posted:

- Go to the [Surveys page](#) on the [course web site](#).
- Follow the link for [Background and Group formation Survey](#).
- Complete the survey **no later than 11:59pm this Wednesday (9 Jan 2019)**.
- It should take only ~ 5 minutes.
- Note that *surveys sometimes fail to save*.
 - Type long answers into a file first and paste them into the survey. Then you won't get as frustrated if it fails to save.

Software

- **ASAP**, install the software discussed on the [Software page](#) on the [course web site](#):

- L^AT_EX



- R



- RStudio



- XPPAUT

- Emacs

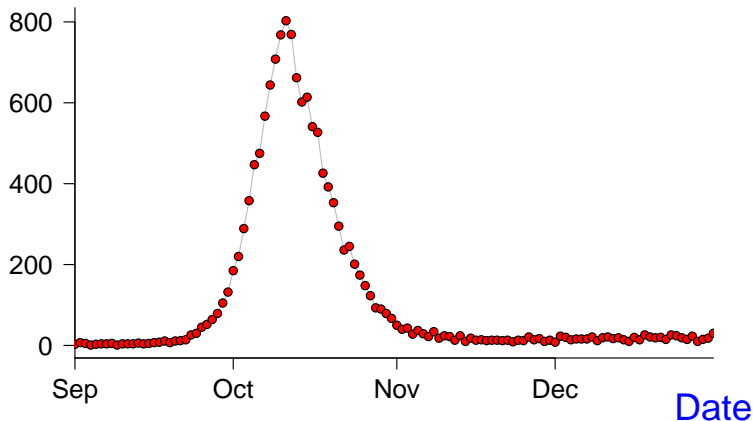


- If you have installation problems, please contact [Ken Moyle](#) <moylek@mcmaster.ca>, who created the [Software page](#).
- **Note:** the [Software page](#) also contains some info about spell-checking and counting words in L^AT_EX documents.

Epidemic Modelling

Pneumonia & Influenza Mortality, Philadelphia, 1918

P&I Deaths



Modelling challenge

Develop a model that helps us understand the [graph on the previous slide](#), based on mechanisms of disease spread.

- Only one variable is observed (P&I deaths per day) so construct a model containing only one variable.
- Think about how disease spreads and express your thoughts with mathematical notation.
- Derive a differential equation that models the process of disease transmission.
- Analyze the model and determine its strengths and weaknesses/limitations.

Make (Biological) Assumptions Clear

- 1 *Assume* the disease is transmitted by contact between an infected individual and a susceptible individual.
- 2 *Assume* the latent period (delay between being infected and becoming infectious) is so short that it can be ignored (technically assume it is zero).
- 3 *Assume* all members of the population are identical and respond identically to the disease. In particular, all susceptible individuals are equally susceptible and all infected individuals are equally infectious.
- 4 *Assume* the population size is fixed during the epidemic, *i.e.*, ignore births, migration, and deaths from causes other than the disease, and count individuals who have died from the disease as part of the population.

About Assumptions. . .

- Note that the first assumption on the [previous slide](#) is actually correct.
- The other assumptions are wrong, but are reasonable approximations.
- It is best to *start as simple as possible and add complexity later*, in order to:
 - obtain a model that actually succeeds in explaining [the data](#) with as few assumptions as possible;
 - check that inferences we draw from our model(s) are robust to the inclusion of more biological details/realism.

What variables should we include in our model?

- Independent variable: time (t)
- Dependent variable: Many options, e.g.,
 - Incidence (number of new infections per unit time)
 - Prevalence (total current number of infected individuals)
 - Death rate (number of deaths per unit time)
 - Death toll (number of deaths so far)
- So, what would be best?
- Not deaths, because whether or not you die may be unrelated to how much you transmit.
- But deaths are what **we observe!** What to do?!?
- Make another assumption. . .

Additional assumption(s)

- We actually want to know incidence or prevalence, but we **observe deaths**.
- Under what circumstances would daily deaths be a good estimate of incidence? (*i.e.*, What must we assume in addition to the **assumptions we have already made**.)
 - 5 **Assume** that the time from infection to death is exactly the same (a certain number of days) for every individual who dies.
 - 6 **Assume** that the probability of dying from the disease is the same for every individual who is infected.
- Then daily death counts are proportional to daily incidence a certain number of days in the past, *i.e.*, the “mortality curve” that **we observe** is a translated and scaled version of the “epidemic curve” (new cases per day).

So... what variables should we include in our model?

- Independent variable: time (t)
- Dependent variable: one of:
 - Incidence (number of new infections per unit time)
 - Prevalence (total current number of infected individuals)
- Which one?
- Choose prevalence (I) because anybody who is currently infectious can infect others, so it will probably be easier to formulate a transmission model based on prevalence. (Try not to lose sight of underlying biological mechanisms.)
- But our **mortality curve** is related to incidence, not prevalence!?! Argh. What to do?!?
- Let's work with prevalence and see how it works out. Maybe we'll be able to derive the incidence curve from a model based on prevalence.

Notational note

- We use I for prevalence because prevalence is the number of infected individuals.

- So, let's try to write down a model...



Mathematics
and Statistics

$$\int_M d\omega = \int_{\partial M} \omega$$

Mathematics 4MB3/6MB3 Mathematical Biology

Instructor: David Earn

Lecture 2

SI Model

Wednesday 9 January 2019

Announcements

- We will meet tomorrow (Thursday @ 10:30am) as scheduled, but:
 - Going forward we will have a two-hour class on Mondays, 9:30–11:20am, in HH-410.
 - We will not have a class on Thursdays, but:
 - That will be a great time to meet with your group since you are all definitely available then.
- [Assignment 1](#) is due when class starts on Monday 21 Jan 2019.
- Links to GitHub and Dropbox are posted on the [Software page](#). There are many other tools for online collaboration, some specific to \LaTeX .
- Have you successfully installed the required software?

A first (naïve) attempt at an epidemic model

- Variables: time t , prevalence $I(t)$
- How does I increase?
- Start with I_0 infected individuals at time $t = 0$. What happens for $t > 0$.
- Let $B =$ average number of contacts with susceptible individuals that lead to a new infective per unit time per infective in the population (and suppose B is constant). Then

$$I(t + \Delta t) \simeq I(t) + B I(t)\Delta t$$

- In the limit $\Delta t \rightarrow 0$, we have

$$\frac{dI}{dt} = BI \quad \Longrightarrow \quad I(t) = I_0 e^{Bt}$$

Beware: implicit assumptions that should be explicit

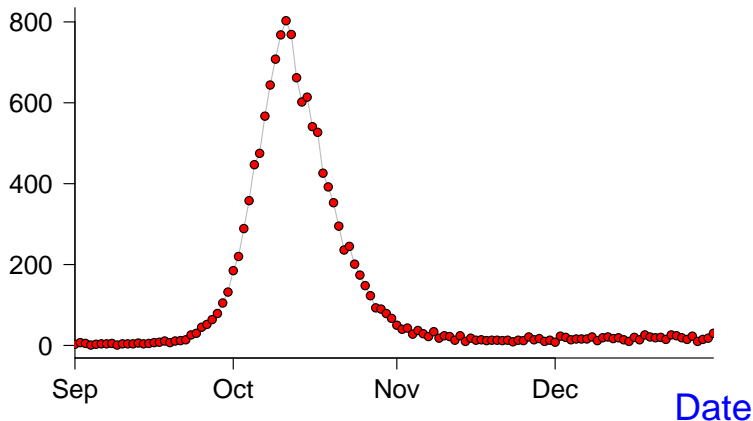
- Ignored discrete nature of individuals when taking limit.
- *Ignored finite infectious periods!*
 - Sometimes it isn't obvious that we've made some assumptions until after we see what the model predicts.

How can we tell if our model is good?

- Compare model predictions with **data**.
- What is the best way to do that?
- Depends on what predictions we're trying to test.
- Model predicts exponential growth.
How should we test that prediction?
- Transforming **the data** might help.

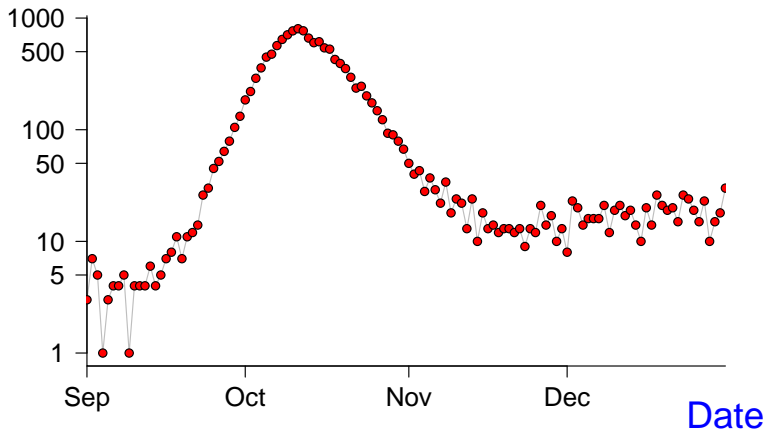
Original data: P&I Mortality, Philadelphia, 1918

P&I Deaths



Logarithmic scale: P&I Mortality, Philadelphia, 1918

P&I Deaths



Parameter estimation

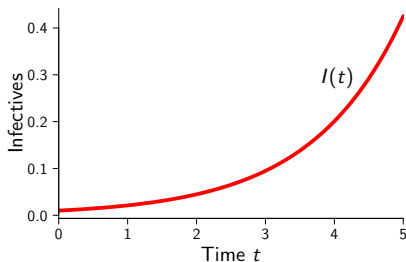
How can we estimate the model parameters, I_0 and B , from the P&I data?

- Fit a straight line through the part of the **logarithmic mortality curve** that looks straight.
- The slope of the line is B .
- The “intercept” is $\log I_0$.
 - “Intercept” in quotes because we need to define $t = 0$ as the time when exponential growth begins.
- **Note:** Parameter estimation is, in general, a very tricky business and deserves a great deal of attention (beyond the scope of this course).

Naïve epidemic model

- Variables: time t , prevalence $I(t)$
- Parameter B = average number of contacts with susceptible individuals that lead to a new infective per unit time per infective in the population

$$\frac{dI}{dt} = BI \quad \Longrightarrow \quad I(t) = I_0 e^{Bt}$$



Naïve model: the good and the bad

- Good:
 - Makes clear predictions
 - Predictions can be tested
 - Estimation of parameter (B) is easy
 - B is the slope of the straight portion of the epidemic curve on the log scale. (Why?)
 - Remember we are imagining that the mortality curve is equivalent to the epidemic curve after translation and scaling.
 - Why do translation and scaling not affect the estimate of B ? Assignment 2. . .
- Bad:
 - Model is consistent only with exponential growth phase.
 - Absurd long-term prediction: unbounded growth in $I(t)$
 - Implicitly assumed that population size $N = \infty$.

How can we improve our model?

- Insist that population size is finite ($N < \infty$).
- Keep track of both **infectives** $I(t)$ and **susceptibles** $S(t)$.
- Assume individuals who are *not infected* are **susceptible**:

$$I(t) + S(t) = N = \text{constant.}$$

New model parameter(s)?

- B = average number of contacts with **susceptible** individuals that lead to a new **infective** *per unit time per infective*
- In the naïve model, we assumed $B = \text{constant}$.
Is B really constant?
- B depends on how many **susceptibles** there are.
- $B = \beta S(t)$
- β = average number of contacts between **susceptibles** and **infectives** that lead to a new **infective**
per unit time
per infective
per susceptible
- β is called the **transmission rate**.

Revised epidemic model: "SI model"

$$\frac{dI}{dt} = \beta S(t)I(t)$$

- Two state variables. One equation. Problem? No:

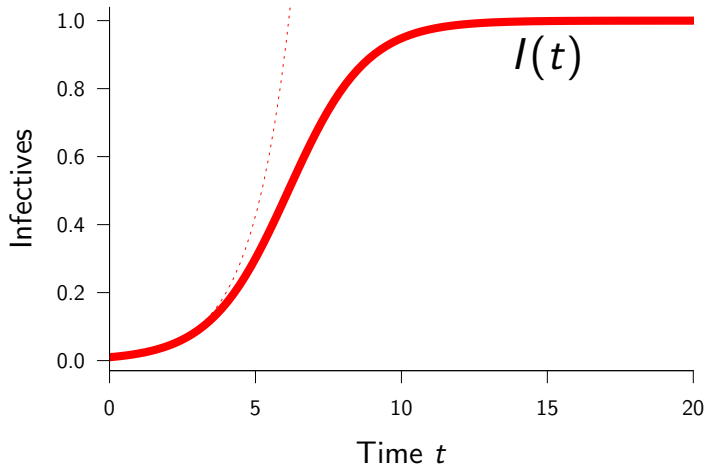
$$\frac{dS}{dt} = -\beta S(t)I(t)$$

- But $S(t) = N - I(t) \implies I(t)$ is still the only variable:

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

- Is this a better model?
- What does it predict?

SI model: Example solution



SI model: Analysis

- We can find the exact solution. How?

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

- But exact solution is not particularly enlightening.
- Qualitative analysis:
 - Initially $I \ll N$. What does the model predict then?
Exponential growth. Great!
 - As I grows, growth rate slows. Why?
Fewer and fewer **susceptibles** to infect.
 - Asymptotic behaviour? Equilibria? Periodic orbits?
(periodic orbit = recurrent epidemics)
 - (Non-trivial) periodic orbits impossible in one dimension
(existence-uniqueness theorem).
 - Consider equilibria. . .

SI model: Equilibrium Analysis

$$\frac{dI}{dt} = \beta I(N - I), \quad I \in [0, N]$$

- Two equilibria:
 - $I = 0$ Disease Free Equilibrium (**DFE**)
 - $I = N$ Endemic Equilibrium (**EE**)
- Stability:
 - DFE is unstable ($0 < I < N \implies dI/dt > 0$)
 - EE is locally asymptotically stable (**LAS**)
 - EE is globally asymptotically stable (**GAS**)
 (stability of EE follows from $0 < I < N \implies dI/dt > 0$)
 (GAS requires a little more analysis... Assignment 1)
 - *Note:* In one dimension, global analysis always easy.
 In higher dimensions, often try to find Lyapunov function.
 (Lyapunov function for EE of SI model?... Assignment 1)
- Conclusions identical for any $\beta > 0$.

SI model: Biological Inferences

- For *any* transmission rate β :
 - Initially, exponential growth of cases.
 - Eventually, convergence to equilibrium (EE) at which *everyone* in the population is **infective**. hmmm...
- Is this model better than our first naïve model?
YES.
 - Still correctly predict initial exponential growth.
 - Can match epidemic curve for longer (up to the peak).
 - Does not predict absurd unbounded growth in **infective** population.
 - But this model cannot explain the decline of the epidemic.
- What should we do? Two obvious options:
 - 1 Get depressed, drop the course.
 - 2 Try to improve the model.