A teachers guide to the modelling of epidemics

Zebedee Mason

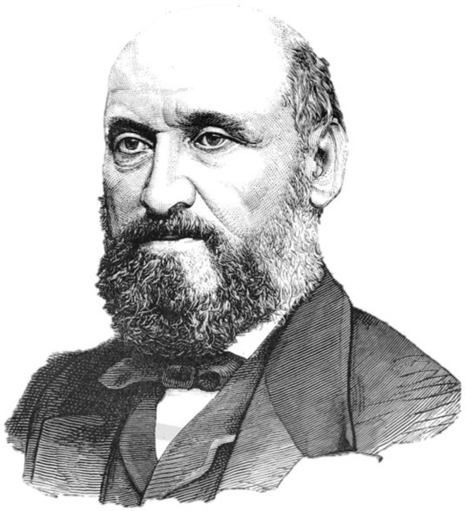
Do you know how an epidemic grows?

Answer at the end

A very brief history

of epidemic modelling

Statistical models



Dr William Farr CB FRS (1807 – 1883)
- English apothecary and statistician
- One of the founders of medical statistics

In 1840 he published work on fitting quadratic and cubic curves to the logarithm of the number of infected.

N.B. The former results in a normal distribution for the number of infections, whereas the latter will be skewed.

Mathematical models

Sir Ronald Ross KCB KCMG FRS FRCS (1857 – 1932)

- 1902 Nobel laureate and Scots Physician

Hilda Phoebe Hudson (1881 - 1965)

- English Mathematician

Developed compartmental models during WWI

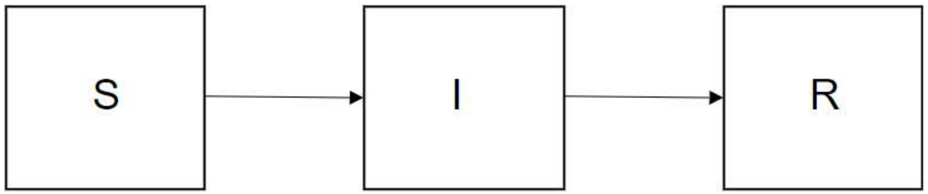
- A compartment represents the state of infection
- There are a set number of individuals in each compartment at any time
- Individuals move between compartments

The SIR model

William Ogilvy Kermack FRS FRSE FRIC (1898 – 1970)
- Blind Scots biochemist

Lt Colonel Anderson Gray McKendrick DSc FRSE (1876 – 1943)
- Mentee of Ross and Scots Physician

Developed the standard SIR model for a well-mixed epidemic
- A constant population N (= S + I + R) of individuals
- S - Susceptible to infection
- I - Infected and infectious
- R - Recovered and immune to re-infection



A Mathematical model isn’t the real world it’s an abstraction

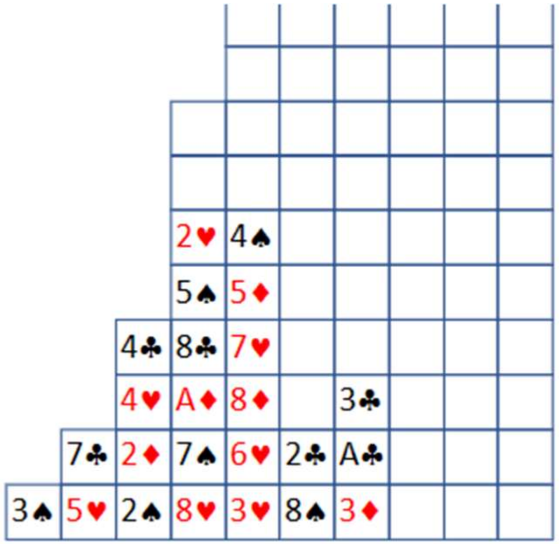
There follow three different implementations of the same model

Epidemiological tag

- When first tagged, a player tags two others
- If tagged again then the player ignores the tag
- Choose one player to be patient zero
- The game ends when the tagging stops

There may be players that have never been tagged
- This is due to an effect known as herd immunity

Epidemic, the card game



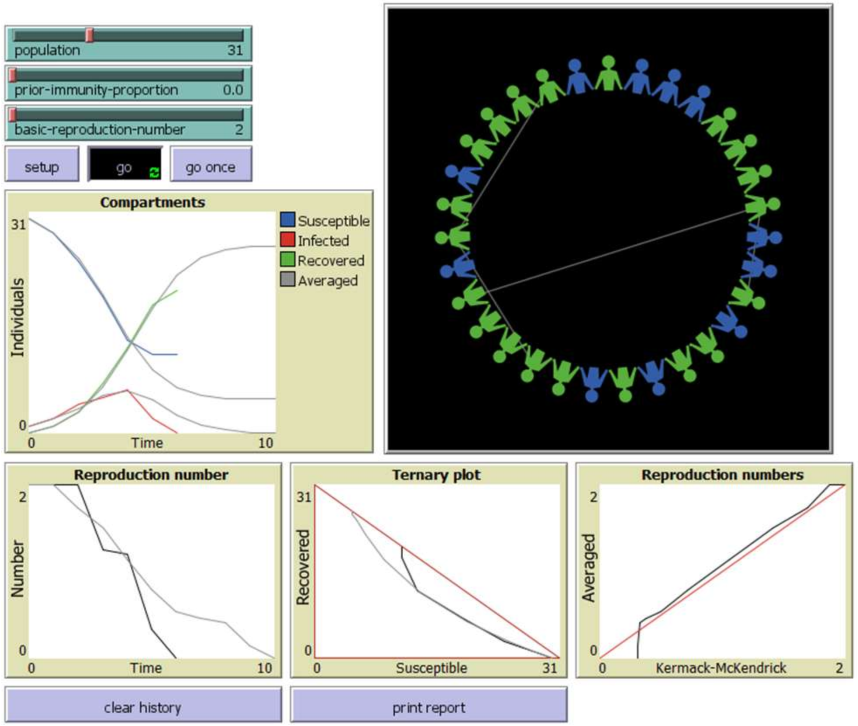
Herd immunity for A♥5♣6♣4♦6♦7♦6♠

- Take two packs of cards
    - One will be used to simulate mixing
    - The other will track the state of individuals
- Cut the packs down to the same 31 cards
    - Chosen for doubling. i.e. $31 = \sum_{i=0}^{4} 2^i$
- Take the mixing pack, shuffle and deal one card - patient zero
- Take the corresponding card from the tracking pack and use it to form the first column

To form a new column, repeat the following for each card in the last dealt column:

- Remove the corresponding card from the mixing pack, shuffle and deal two cards - the two individuals met by the infected
- If they've not been dealt before remove them from the tracking pack and add to the new column
- Return all three cards to the mixing pack

The minimal SIR model in NetLogo



There are sliders for three parameters

- Population N = S + I + R
- The proportion of N at the start in the R compartment
- The number of individuals that one individual will infect in an otherwise susceptible population (S, I, R) = (N-1, 1, 0)

Press setup then go once. Lines will be drawn between the just recovered and the who they try to infect. Press go and the model will run until I = 0. Press setup repeatedly and an average of multiple runs will be computed.

There are plots of

- The split between the SIR compartments versus time
- The ratio of the number of newly infected to old over time
- A ternary plot in which the diagonal line is I = 0
- A comparison of reproduction numbers from the average and from the standard SIR model of Kermack-McKendrick

Probability Theory

Chance of exponential growth

- If the card game resulted in exponential growth of the number of infected then the size of the columns would double - 1, 2, 4, 8, 16
- Create a fraction
  - The numerator is the product of the decreasing number of never infected
  - The denominator is the product of the corresponding size of the shuffle deck

$$\frac{31}{31} \times \frac{30}{30} \times \frac{29}{29} \times \frac{28}{30} \times \frac{27}{29} \times \cdots \times \frac{3}{29} \times \frac{2}{30} \times \frac{1}{29} = \frac{31!}{31 \times 30^{15} \times 29^{15}}$$

- The fraction evaluates to roughly two in ten to the power of twelve
- There is a roughly even chance for the first 8 cards
- The definition of the basic reproduction number corresponds to certainties for the first three cards

Hypergeometric distribution

- The shuffling and dealing corresponds to selecting from the Hypergeometric distribution
- For a population size N with K items of interest
- Select n items of which k turn out to be of interest
- The probability that the random variable X equals k is
  - $\Pr(X = k) = \frac{K!(N-K)!n!(N-n)!}{k!(K-k)!(n-k)!(N-K-n+k)!N!}$
- For exponential growth in cards (N, K, n, k) = (30, S, 2, 2)
  - Hence $\Pr(X = 2) = \frac{K!\,(N-2)!}{(K-2)!\,N!} = \frac{S}{30} \times \frac{S-1}{29}$

Exponential growth

Is very poorly understood

What is it?

- Exponential growth arises in the card game when a dealt card is never returned to the pack
  - The height of the next column is double the previous
- It goes to infinity which raises the question
  - What happens and when to cause a finite variable to leave a curve to infinity?
- Which has the usual answer
  - Occam's Razor says it was never growing to infinity in the first place
- Exponential growth either
  - Requires infinite resource
  - Or comes to a hard stop when it uses all the finite resource

Examples

- Compound interest
  - The cynic says that more money can always be printed
  - Dates back to ancient Babylon if not before
- Grains of wheat on a chessboard (according to folklore)
  - Although there isn’t enough wheat in the world to get to the last square
- Pyramid schemes
  - Eventually there is no one left to fool

Common mistaken examples

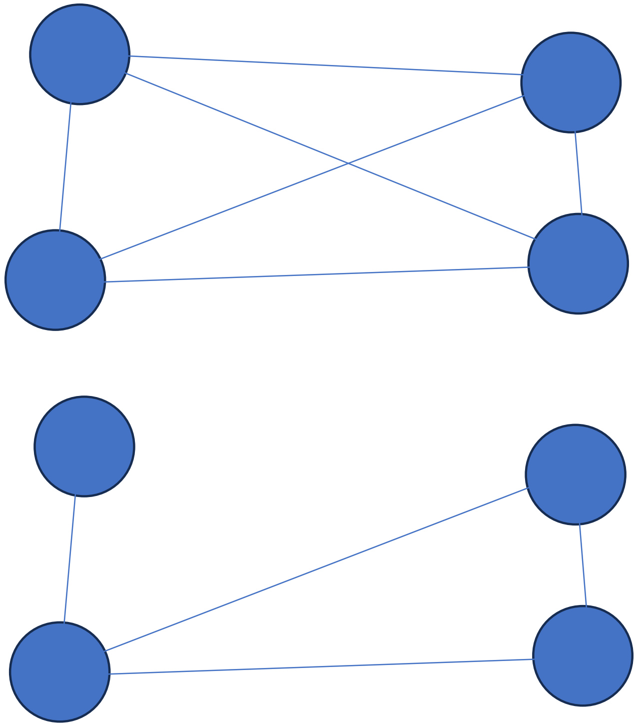
- Nuclear explosions
  - There are unused isotopes in the fallout

Why do people get it wrong?

- Exponential growth requires doubling at a constant rate, without that constant rate it is merely an order of magnitude increase
  - Some people mean an order of magnitude increase
- It is easy to draw a straight line on a log-plot
- It simplifies the Mathematics to make the equations solvable
- They mistake the exponential curve for what is actually a sigmoid (S shaped) curve
- e.g. The logistic function is close to the exponential at its beginning
  - $\frac{1}{1+e^{-t}}$ is approximately e^t for t large and negative

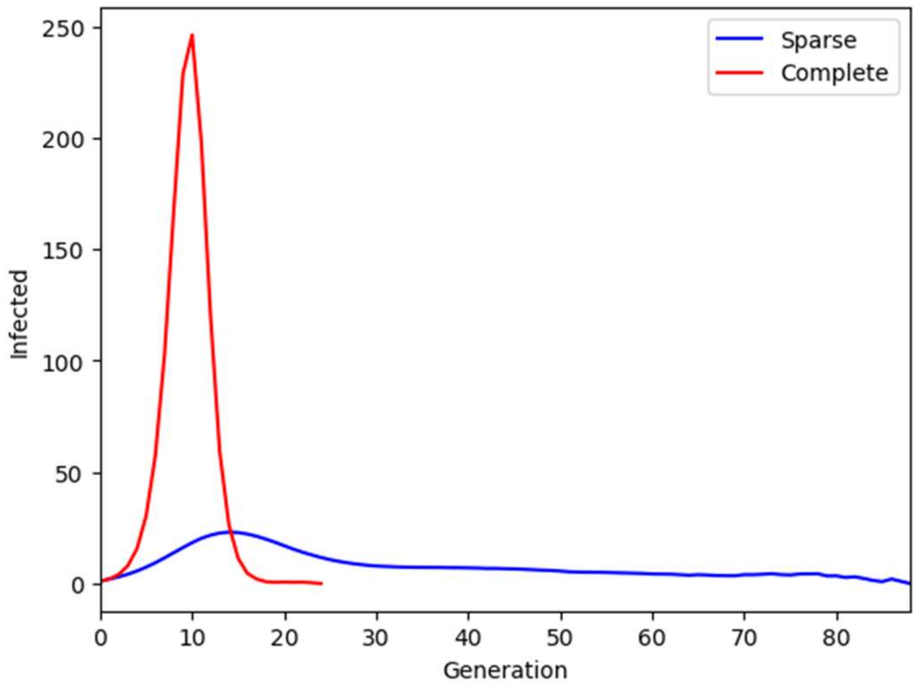
Graph Theory

Well mixed



- In the model the individuals are well mixed
  - Each has equal chance of bumping into another
- Hence a graph (network) can be drawn in which each node (individual) is connected to every other node.
  - This is known as a complete graph.
- Can drop connections to form a sparse graph
  - Adapt the Minimal SIR model for sparse graphs
  - i.e. Deal the cards from a reduced pack

Social networks



- Voles in Northumbria aren’t very social
  - They don’t have Facebook friends
- The plot compares the average epidemic on the voles network to that of a complete graph with the same number of individuals
- Apart from taking longer, fewer individuals get infected by the end
- What does this different mixing mean for lockdowns? For seasonality?

Ordinary Differential Equations

Continuity

- Given a large number of individuals the variables S, I and R can be treated as real numbers rather than integers
- This allows the Mathematician to chop an individual up into pieces smaller than that allowed by the laws of Physics never mind the laws of man
- i.e. Infinitesimals, which allows for differential equations

Derive the model

- Differentiate the constant $N = S + I + R$ with respect to time t
- Re-arrange $0 = \frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt}$
- To $\frac{dI}{dt} = \left(-\frac{dS}{dR} - 1\right)\frac{dR}{dt}$ where $-\frac{dS}{dR} \geq 0$ is the reproduction number
- Or, without loss of generality, $\frac{dI}{dt} = (R_0 X - 1)\frac{dR}{dt}$
  - Where R_0 is, confusingly not $R(t = 0)$, but the basic reproduction number
  - And X is a number between 0 and 1, the proportion
- Guess X and $\frac{dR}{dt}$ to 'close' the model

The standard SIR model

- Kermack-McKendrick (1927)
- $\frac{dS}{dt} = -R_0 \frac{S}{N} I\gamma$
  - Substitute $N - S$ for I and integrate to get the logistic function
- $\frac{dI}{dt} = \left(R_0 \frac{S}{N} - 1\right) I\gamma$ hence $X = \frac{S}{N}$
  - An imbalance between logistic growth and exponential decay
- $\frac{dR}{dt} = I\gamma$
- Integrate $-\frac{dS}{dR} = R_0 \frac{S}{N}$ to get an exponential curve for the ternary plot

All models are wrong …

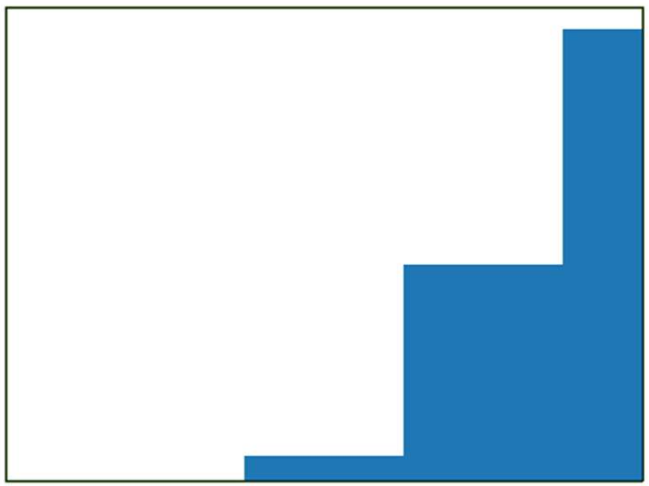
- but some are useful. George Box, Statistician
- The reproduction rate is always less than the basic reproduction rate contrary to its definition
  - $R_0 \frac{S}{N} < R_0$ since $S < N$ for all time
- The number of infected decays to zero but is never zero
  - There’s always a fraction of an individual left
  - Hence the epidemic never ends
- Some aspects of reality are lost but the power of differential calculus is gained

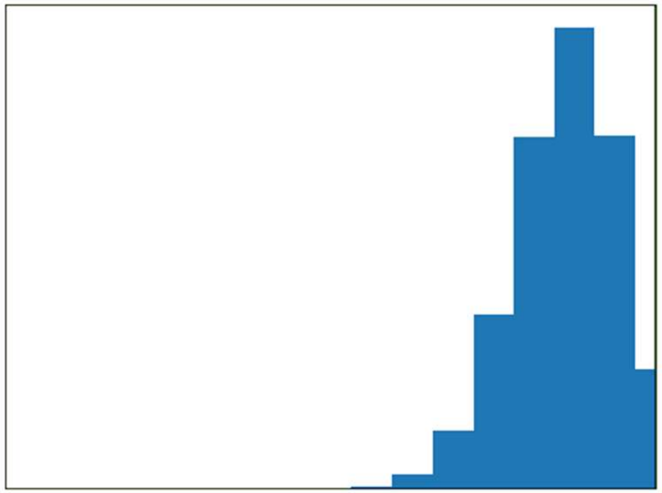
Stochastic Differential Equations

A taste for questioning minds

Discrete distribution

- Each column in the card game is a sample from a distribution
- That is composed from multiple Hypergeometric distributions
- By increasing the population the discrete distribution goes continuous
  - As long as S, I and R are kept in proportion





Conjecture

- It is guessed that X is logit-normally distributed
  - With mean corresponding to logistic growth
- $logit(X) \sim \boldsymbol{N}(\mu, \sigma^2), E(X) = \frac{S}{N}$
- μ, σ^2 are functions of $\frac{S}{N}, R_0 \frac{I}{N}$
- $logit(x) = log\left(\frac{x}{1-x}\right)$, the inverse of the logistic function

Simulation

- Assuming μ, σ^2 are known
- To create X
  - Create a uniformly distributed number between 0 and 1
  - Put that value through the inverse CDF of the normal distribution
  - Put that value through the logistic function
- Use the discretised equations
  - $S(t + 1/\gamma) = S(t) - I(t + 1/\gamma)$
  - I$(t + 1/\gamma) = R_0 X I(t)$
  - $R(t + 1/\gamma) = R(t) + I(t)$
- Note that N is conserved

Resources

- The NetLogo model is available from
  - https://github.com/stochanswers/netlogo
- Where you will find a link to run it online in NetLogo Web
  - https://www.netlogoweb.org/launch
- Alternatively
  - https://www.modelingcommons.org/browse/one_model/7697

Do you know how a well-mixed epidemic grows?

One infects two, two infect four or three or two or one or none. All with different probabilities and a weighted average of logistic growth.

A less mixed epidemic will grow more slowly.