

Mathematical Modelling of Infectious Diseases

Barry Rowlingson



Lancaster Medical School

b.rowlingson@lancaster.ac.uk

This lecture is about how and why we construct mathematical models of disease.

I'm one of the research staff in the Medical School, in the CHICAS group, where we do statistical analysis of diseases, including looking at the spread of things like malaria, influenza, dengue fever and so on.



If you've seen The Matrix you'll know its about Mr Anderson, who discovers that the world he thought was real was in fact....



...a computer simulation run by machines for some reason not really adequately explained. He is pulled out of The Matrix and discovers his entire life has been spent in one of these pod things...



CHICAS
centre for
health informatics,
computing, and statistics

...and the “Real World” is an illusion, created by the machines using a vast amount of computing power.

Some people actually think we are living in such a simulation. But let's not entertain such a fantasy and assume this is the real world. Maybe we could build a powerful computer and create a simulation like The Matrix, indistinguishable by its inhabitants from reality. What could we do?

If We Had That Much Computing Power We Could...

- ▶ Run infection experiments
- ▶ Infect “people”
- ▶ Change behaviour
- ▶ Change the environment
- ▶ Observe the impact on the disease process
- ▶ But we don't have that much computing power. Yet.

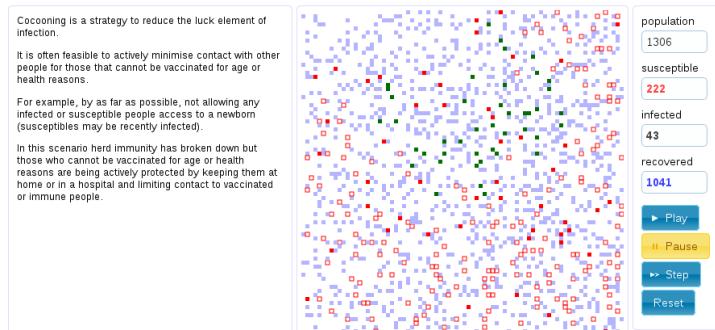


We could do all sorts of megalomaniacal things. We could infect people with diseases and see how it spreads. We could reset The Matrix and try infecting different people, or making the virus more virulent, or seeing what happens under climate change. We could answer questions in the real world, like “If there's 2 degrees of warming, how will that affect malaria transmission in Africa?”

We don't have enough computer power to simulate reality anything like The Matrix. So we have to simplify...

How Simple Can We Make This?

► <http://op12no2.me/toys/herd/>

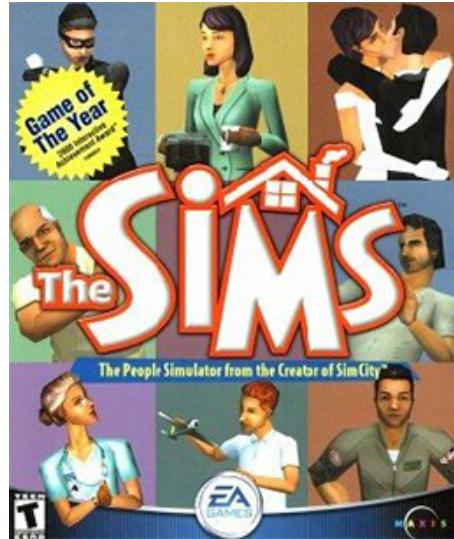


- People are squares on a grid
- Illustrates transmission and immunity



Here's a mini-universe, featuring a bunch of people who look a bit like squares, they run around and bump into other squares. Some of them are infected. If they bump into someone else, they can infect them too.

This web site lets you run some different scenarios with vaccinated people and various different immunity scenarios.



On a slightly more graphically realistic but more frivolous front, there's *The Sims*, where semi-autonomous people wander about a reality you can create.

“Guinea Pig Disease”



And The Sims world has disease. You can get a pet guinea pig for your Sim, and that guinea pig can catch “Guinea Pig Disease”.

"Guinea Pig Disease"



▶ Example of a
"Zoonosis"



GPD is a zoonosis, a disease of animals that can be transmitted to humans. In this case, it used to be fatal. But the creators of The Sims created a lot of bad feeling because people's Sims were dying. So they released a patch to the program so that GPD just made them sick instead. A bit like creating a cure.

Dysentery



...can be passed on to other sims



Someone even created a “mod” - a user-contributed add-on - for The Sims 3, which gives Sims contagious dysentery.

Agent-based Modelling

- ▶ Smallest modelled unit is an individual
- ▶ Mathematical model of agent behaviour
- ▶ Multiple simulations to get probabilistic answers to "if" questions.
- ▶ Found in other disciplines, eg traffic modelling, finance

But what if we have too many agents to model?



What we've got in these examples are "agent based models". Each individual person (or guinea pig) that could get infected is treated separately; mathematics is used to decide how things behave and transmit; you can run multiple simulations to get answers to "if this happens, what's the chance of that happening?" questions.

The agent-based approach is also found in traffic modelling, where vehicles and pedestrians can be the agents, and in finance, where traders are the agents.

But what if you have so many agents you don't have enough computation power? It was never a problem in The Matrix...



For those who don't get the reference, this is Keanu Reeves' character fighting off hundreds of "Agent Smith" duplicates.

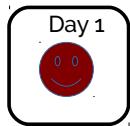
Numerical Models

- ▶ Start with one infectious person
- ▶ Each day, every infectious person infects two more people
- ▶ People are infectious for two days
- ▶ What happens?



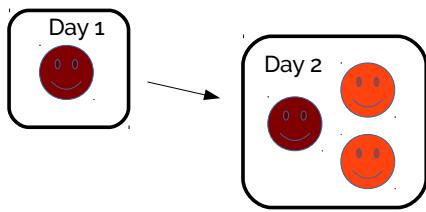
We can instead create numerical models of epidemics. Lets start with one infectious person, who infects two people every day, and people are infectious for two days. What happens?

Infectious Model



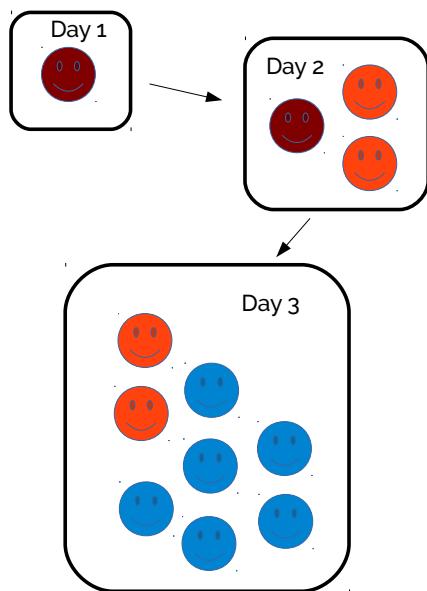
Day one, there's one infectious person, this red face.

Infectious Model



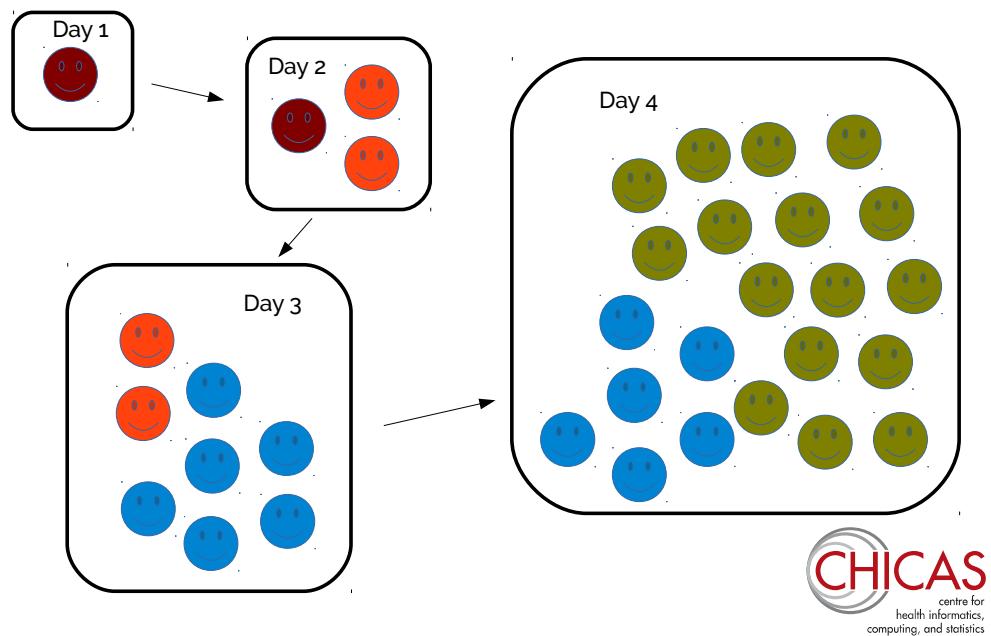
Day 2, red face has infected the two orange faces,
and all three of them are infectious.

Infectious Model



The three infectious faces from day 2 generate six infectious cases, the blue faces. Orange faces are still infectious, but red face is now either dead or recovered.

Infectious Model

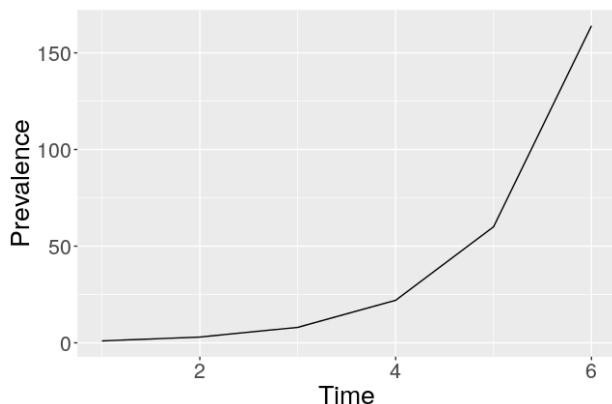


CHICAS
centre for
health informatics,
computing, and statistics

Day 4 we get 16 new cases plus the 6 blue infectious faces from the previous day. And we can carry on this process, but drawing faces isn't a good way to visualise this...

Rapid Growth

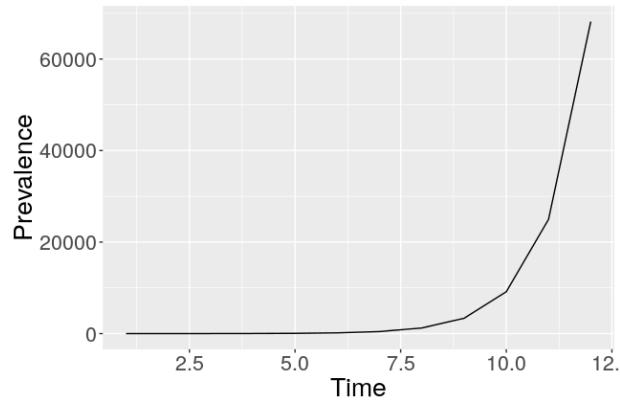
Time	Prevalence
1	1
2	3
3	8
4	22
5	60
6	164



Plotting number of infectious cases against time, you can see there is increasingly rapid growth

Rapid Growth.....

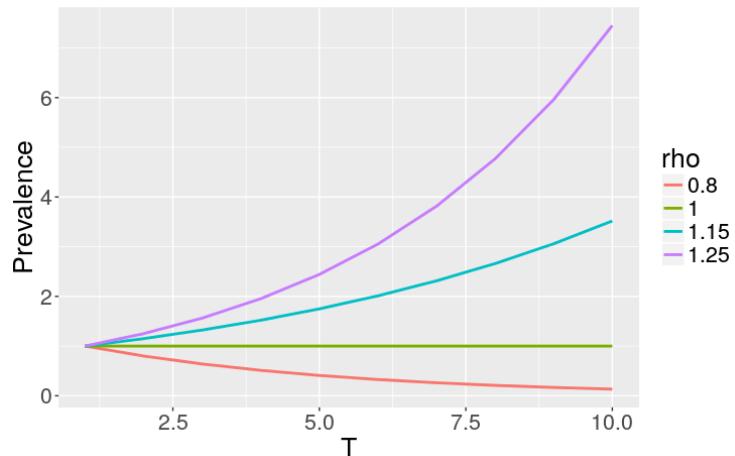
Time	Prevalence
1	1
2	3
3	8
4	22
5	60
6	164
...	...
12	68192
...	...
24	11 billion



Carrying on, after 12 steps its up to 68,000 infectious cases, and 12 steps after that there's 11 billion.

This kind of growth is referred to as exponential growth.

Reproduction number



The rate of exponential growth is controlled by the reproduction number, which is usually given the greek letter rho. If rho is less than one you get exponential decay, rather than growth.

Is Exponential Growth Valid?

- ▶ Yes, when:
 - ▶ Initial Stages of an Infection
 - ▶ Large Susceptible Population
 - ▶ Thorough Mixing of Infectious Cases
- ▶ No, when:
 - ▶ Finite Population
 - ▶ Few Susceptible Population
 - ▶ Poor Mixing of Infectious Cases and Population

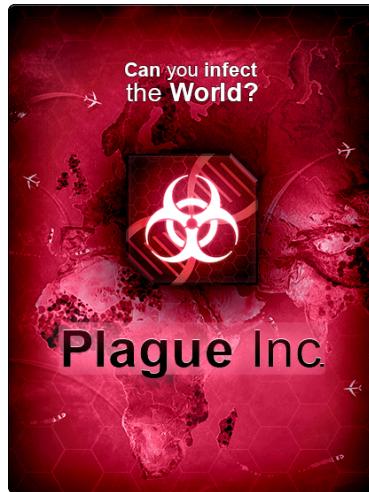


Clearly exponential growth like that can't be sustainable for very long in a disease outbreak.
Where is it valid? How does it break down?

Its a good model in the early stages of an infection, where the number of infectious cases is small compared to the population, or where there's a large susceptible population with thorough mixing of cases and population.

But if you have a finite population, or there's few susceptible people in the population, or the infectious people don't have access to the susceptibles, then the exponential growth model becomes invalid.

Finite Population



<http://www.ndemiccreations.com/en/>



Let's look at what happens with a finite population in another game.



Plague Inc is a game where you try to infect the planet with a plague. Nice. The game keeps track of how many people are infected in each country.

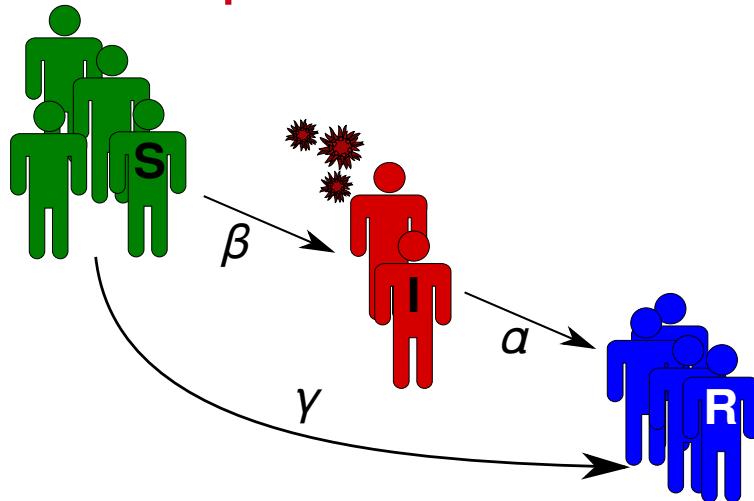
Aggregate Infection Tables

Country	Uninfected	Infectious	Dead
China	903,452,543	1,063,345	8,365
USA	325,753,342	0	1,20546
Brazil	180,537,623	20,663,634	7,653,634
Nigeria	191,836,451	0	0
...			
Pitcairn Islands	0	0	57



So it has a table for the world that tracks how many people in each country are uninfected, infectious, or dead. It also has a model for how much international trade goes on between countries, spreading the disease globally.

Compartment Model



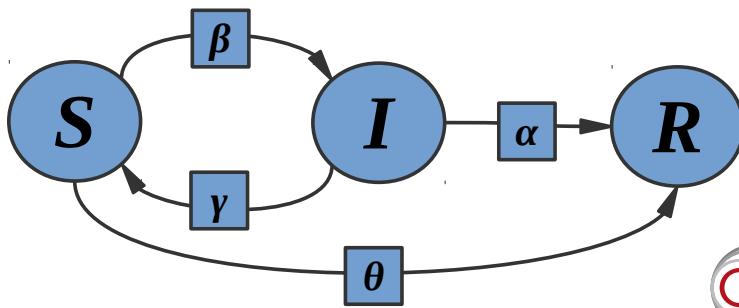
This separation of a finite population into categories is known as a compartment model. The compartments are usually labelled S for susceptible, I for infectious, and R for “Removed”, which in this case could be dead or recovered if there's immunity to further infections.

The possible transitions between compartments are shown as arrows, and the rate of the transition by a greek letter.

Here we have a green susceptible population being infected by the red infectious population. The infectious people die or recover and go into the R compartment. Susceptible people can also die of other causes.

Compartment Model Diagram

- ▶ S = #Susceptible
- ▶ I = #Infectious
- ▶ R = #Removed or Recovered
- ▶ Greek letters are parameters
- ▶ Control rates of change

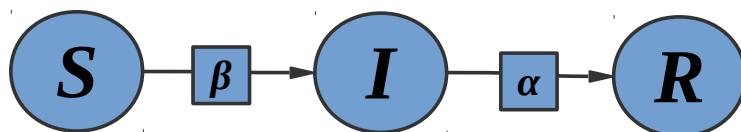


In papers you'll often find simplified diagrams a bit like this. This is similar to the last diagram except there's an extra arrow where infectious people can recover into the S compartment and possibly get infected again.

How complex you make these models depends on what simplifications you can get away with. For example, you might not need the $S \rightarrow R$ transition if few people would die of other causes during your model.

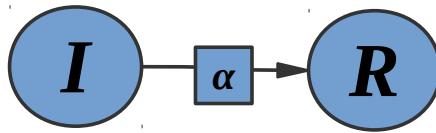
SIR Compartment Model

- ▶ Three states
- ▶ Two parameters
- ▶ No re-infection (no arrows back to S)
- ▶ Ignore other causes of death (no $S \rightarrow R$)



The classic simplest model is the SIR model with two parameters. Let's break this down and see how it works...

Removal/Death Rate



- ▶ Suppose today:
 - ▶ $I = \underline{1000}, R = \underline{0}$
 - ▶ $\alpha = 0.01$
- ▶ Then tomorrow
 - ▶ I changes to $1000 - 1000 \times \alpha = \underline{990}$
 - ▶ R changes to $0 + 1000 \times \alpha = \underline{10}$

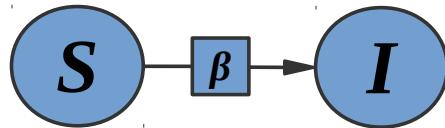


First let's look at the right hand side of the diagram.

Suppose we have 1000 infectious cases in I and none currently in R, and our alpha parameter is 0.01. This means 1% of our infectious cases will transition to R each step, in this case, 10. That will leave us with 990 in I and 10 in R.

Simultaneously though, something is going on with the left hand side of the diagram...

Infection Rate



- ▶ Depends on number of infectious, I
- ▶ Suppose today:
 - ▶ $S = \underline{1000}$, $I = \underline{10}$
 - ▶ $\beta = 0.002$
- ▶ Then tomorrow:
 - ▶ S becomes $1000 - 10 \times 0.002 \times 1000 = \underline{980}$
 - ▶ I becomes $10 + 10 \times 0.002 \times 1000 = \underline{30}$

The change here, unlike before, depends on the number of people in the S and the I compartment, because each infectious person in I can potentially infect people in S . So to compute the number of people that transition from S to I , we multiply the populations in each with the beta parameter.

So supposing we have 1000 susceptibles and 10 infectious and beta is 0.002, that gives us a change of 20. Now we have 980 susceptible and 30 infectious.

To simulate the epidemic we have to do both steps, and repeat several time, but first there's a bit of notation.

Time Notation

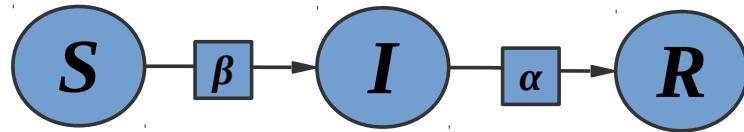
- ▶ S_1 is the number of susceptibles on day 1
 - ▶ Could be week 1, month 1, year 1...
- ▶ S_2 is the number of susceptibles on day 2
- ▶ S_t is the number on some day t
- ▶ S_{t+1} is number on the day after t
- ▶ Time can start at $t=0$



A letter with a subscript refers to a time point. So S with a subscript 1 is the number of susceptibles on day 1. This might not be days, it could be weeks or months – its whatever time step the model is using.

That is then generalised to S with a subscript t , meaning any time point t , and S with a subscript $t+1$, meaning the time point after t .

Updating Steps



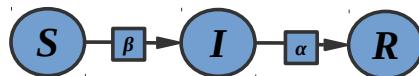
- ▶ $S_t - (\beta \times I_t \times S_t) \rightarrow S_{t+1}$
- ▶ $R_t + (\alpha \times I_t) \rightarrow R_{t+1}$
- ▶ $I_t - (\alpha \times I_t) + (\beta \times I_t \times S_t) \rightarrow I_{t+1}$
- ▶ Increase t , and repeat...



Now we can write some updating formulae for our SIR model. These expressions are how we compute the next value of S, I, and R from the current time step.

Example 1

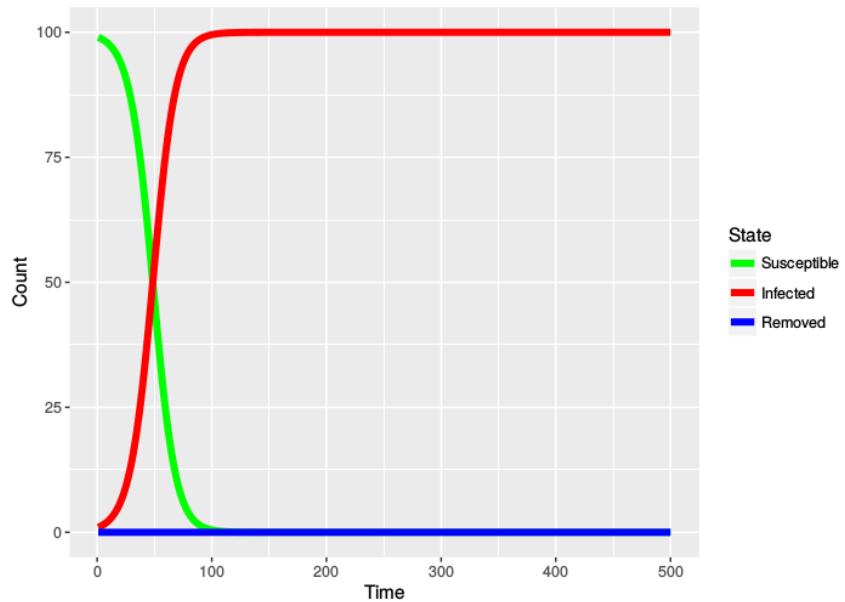
- ▶ 100 people in total
- ▶ Start with one infection
- ▶ 99 susceptibles
- ▶ Set $\alpha=0$
 - ▶ Nobody recovers
 - ▶ Effectively an **SI** model
- ▶ Set $\beta=0.001$



So let's set one up and see how it progresses. We'll take 99 susceptibles and one infectious person. Let's set alpha to zero – then nobody recovers, and we get an “SI” model.

We'll set the infection rate to 0.001.

Epidemic Chart



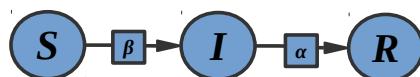
When we run that model you'll see this happen. The number of infectious cases rises, exponentially at first, but then levels out as there's fewer susceptible population to infect. After about 100 days everyone is infected.

The R line stays at zero.

Since we have a finite population, and everyone is either in S or I, the sum of these two lines must be a constant, so as I goes up, S goes down equally.

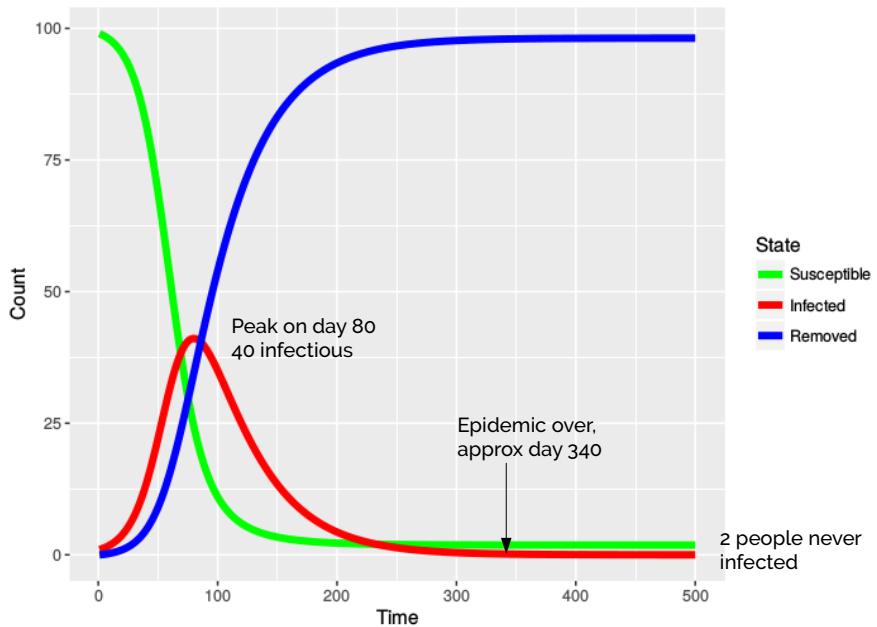
Example 2 - recovery

- ▶ 100 people in total
- ▶ Start with one infection
- ▶ 99 susceptibles
- ▶ Set $\alpha=0.025$
 - ▶ Some recovery
- ▶ Set $\beta=0.001$



Now we'll add some recovery. Same as before, but now alpha is 0.025.

Epidemic Chart

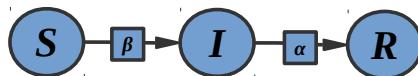


What happens then is that we see an epidemic peak, here it is on about day 80 with 40 people infectious.

By day 340 there's no infectious people, and the epidemic is over. At that point there are 2 people in the susceptibles group, the lucky ones who didn't get the infection.

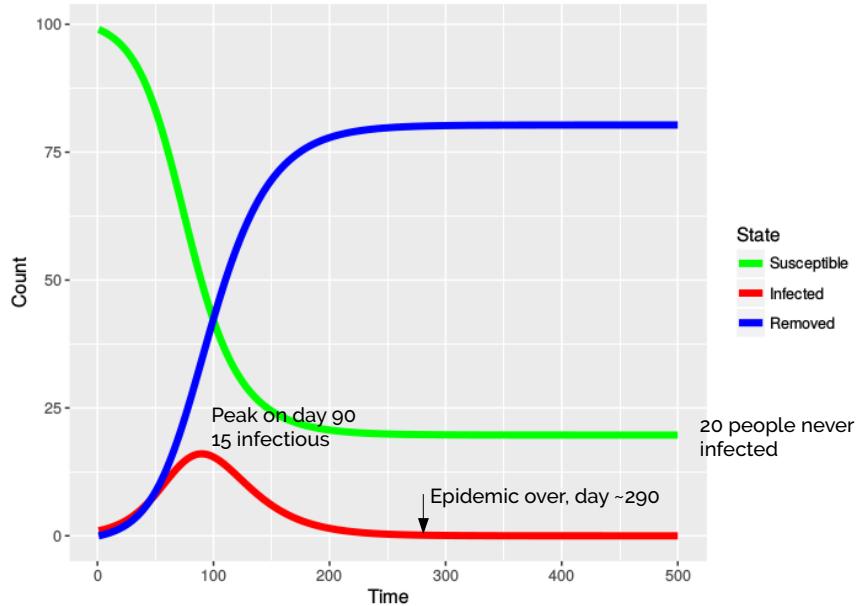
Example 3 – faster recovery

- ▶ 100 people in total
- ▶ Start with one infection
- ▶ 99 susceptibles
- ▶ Set $\alpha=0.050$
 - ▶ Double previous value
- ▶ Set $\beta=0.001$



What happens if we double the recovery parameter?
Do we get twice as many recovered? Does the
epidemic peak earlier or later?

Epidemic Chart

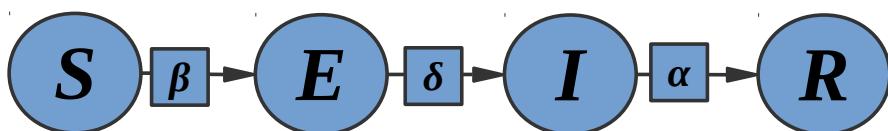


In this case the epidemic peaks a bit later, with fewer cases, and is over quicker.

Also instead of 2, we find 20 people have never been infected.

More Complex Models

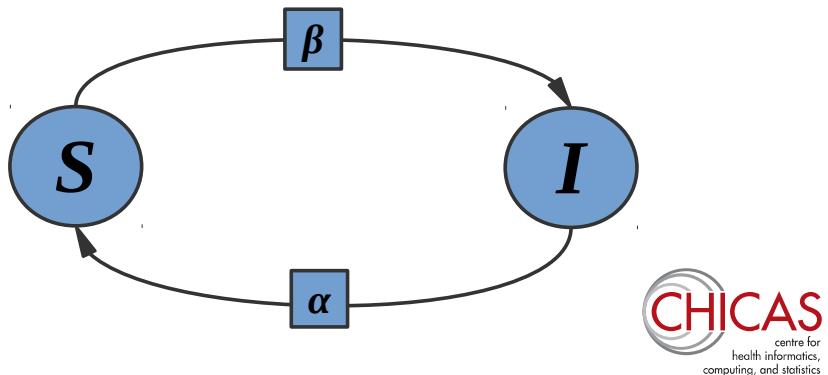
- ▶ Add a “Latent” infected state
- ▶ Infected, and will progress to infectious at some rate δ
- ▶ Usually “E”



Compartment models can have more than three compartments. A common extension is the SEIR model, where the E compartment contains people who are infected but not yet infectious. There's another parameter that controls the rate of change to the infectious state.

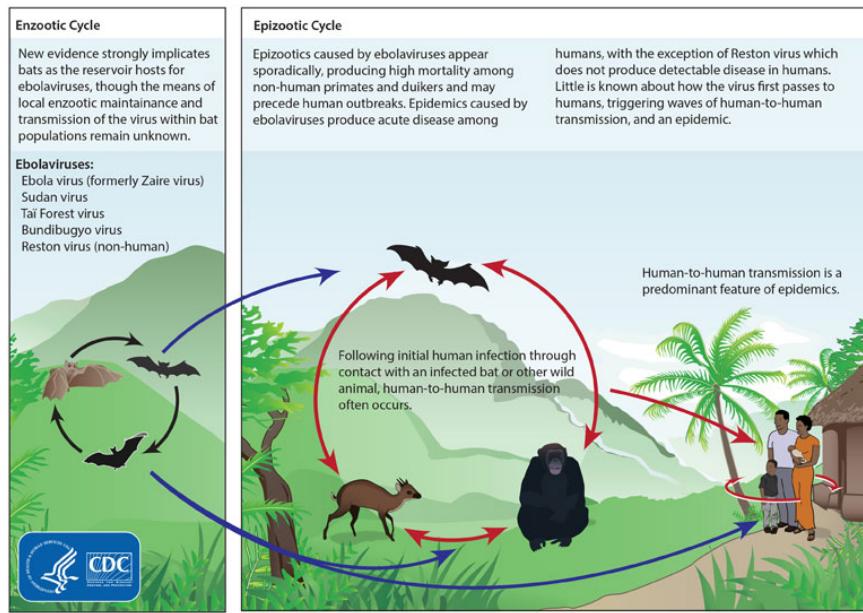
SIS Model

- ▶ Susceptible people become infectious
- ▶ Infected people recover
- ▶ How does the system end up?



Another model is the SIS model, where infected people can recover and be reinfected.

More Species



Epidemic models can be extended to multiple species. This illustration shows how the Ebola virus can be transferred from a reservoir in bats, into other mammals, and then into humans.

This results in several coupled SIR-type models where the infection rates in one species depends on the number infected in another species.

Human Ectoparasites

Human ectoparasites and the spread of plague in Europe during the Second Pandemic



Katharine R. Dean, Fabienne Krauer, Lars Walløe, Ole Christian Lingjærde, Barbara Bramanti, Nils Chr. Stenseth, and Boris V. Schmid

PNAS 2018; published ahead of print January 16, 2018, <https://doi.org/10.1073/pnas.1715640115>

Contributed by Nils Chr. Stenseth, December 4, 2017 (sent for review September 4, 2017; reviewed by Xavier Didelot and Kenneth L. Gage)

► Can we blame the rats?



This recent paper aimed to investigate if the epidemic pattern of plague could be explained by the accepted model of rats carrying infected fleas.

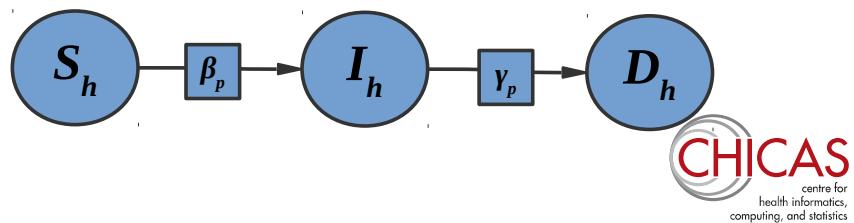
Pneumonic Model

► Human-Human Transmission

$$\frac{dS_h}{dt} = -\beta_p \frac{S_h I_h}{N_h},$$

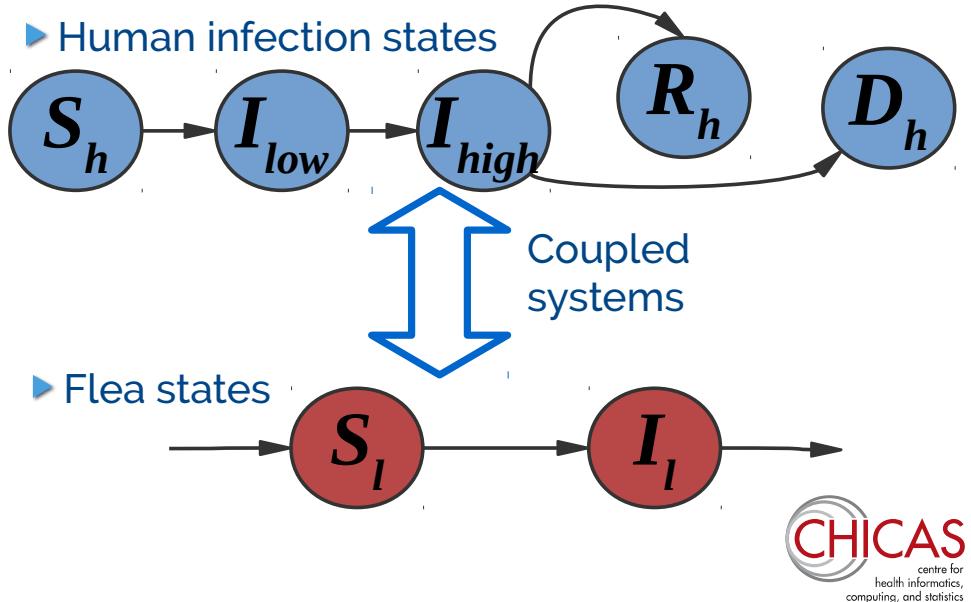
$$\frac{dI_h}{dt} = \beta_p \frac{S_h I_h}{N_h} - \gamma_p I_h,$$

$$\frac{dD_h}{dt} = \gamma_p I_h.$$



They compared three models – a simple human-human model based on SIR, and controlled by these three equations which are similar to those updating formulae for the SIR model I gave earlier.

Human Ectoparasites



The second model is where the infection is transmitted by human fleas or lice – no rats are involved.

The model has two compartment models for humans and fleas, and these are coupled so that more infected fleas mean more infected humans.

Human Ectoparasites

$$\begin{aligned}\frac{dS_h}{dt} &= -\beta_l \frac{S_h I_l}{N_h}, \\ \frac{dI_{\text{low}}}{dt} &= \beta_l \frac{S_h I_l}{N_h} - \sigma_b I_{\text{low}}, \\ \frac{dI_{\text{high}}}{dt} &= (1 - g_h) \sigma_b I_{\text{low}} - \gamma_b I_{\text{high}}, \\ \frac{dR_h}{dt} &= g_h \sigma_b I_{\text{low}}, \\ \frac{dD_h}{dt} &= \gamma_b I_{\text{high}}, \\ \frac{dS_l}{dt} &= r_l S_l \left(1 - \frac{N_l}{K_l}\right) - \left[\left(\beta_{\text{low}} I_{\text{low}} + \beta_{\text{high}} I_{\text{high}}\right) \frac{S_l}{N_h}\right], \\ \frac{dI_l}{dt} &= \left[\left(\beta_{\text{low}} I_{\text{low}} + \beta_{\text{high}} I_{\text{high}}\right) \frac{S_l}{N_h}\right] - \gamma_l I_l.\end{aligned}$$



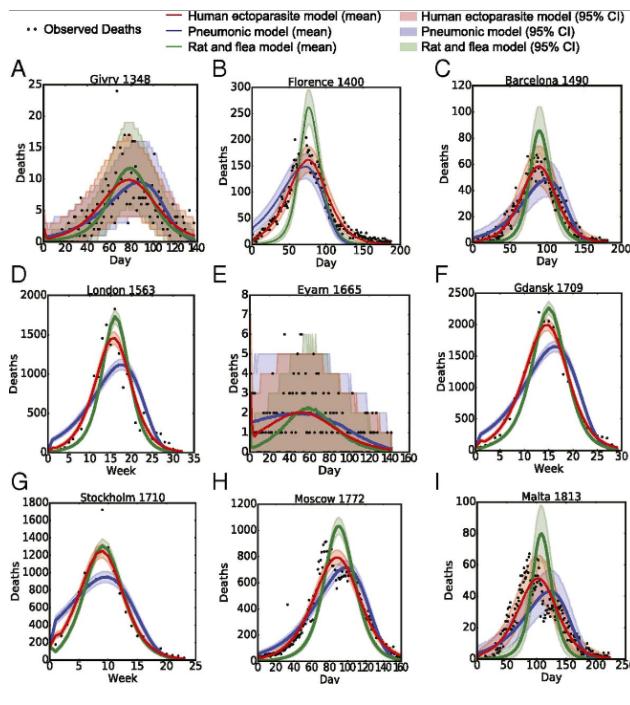
That translates to seven equations...

Human/Rat/Flea Model

$$\begin{aligned}\frac{dS_r}{dt} &= -\beta_r \frac{S_r F}{N_r} [1 - e^{-aN_r}], & \frac{dF}{dt} &= (1 - g_r) \gamma_r I_r H - d_f F, \\ \frac{dI_r}{dt} &= \beta_r \frac{S_r F}{N_r} [1 - e^{-aN_r}] - \gamma_r I_r, & \frac{dS_h}{dt} &= -\beta_h \frac{S_h F}{N_h} [e^{-aN_r}], \\ \frac{dR_r}{dt} &= g_r \gamma_r I_r, & \frac{dI_h}{dt} &= \beta_h \frac{S_h F}{N_h} [e^{-aN_r}] - \gamma_h I_h, \\ \frac{dD_r}{dt} &= (1 - g_r) \gamma_r I_r, & \frac{dR_h}{dt} &= g_h \gamma_h I_h, \\ \frac{dH}{dt} &= r_f H \left(1 - \frac{H}{K_f}\right), & \frac{dD_h}{dt} &= (1 - g_h) \gamma_h I_h.\end{aligned}$$



The accepted human/rat/flea model is expressed by ten equations, and three separate compartment models (for humans, rats, and fleas).

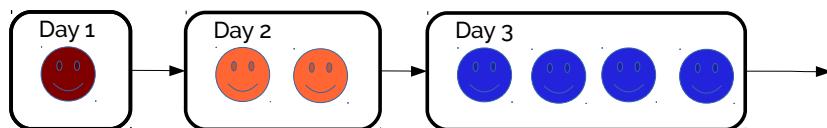


CHICAS
centre for
health informatics,
computing, and statistics

When the authors fit death data from several cities to the three models they claim that the human flea/louse model (in red) fits the data better than the others. Perhaps the rats weren't to blame?

Deterministic

- ▶ Each infectious person infects 2 new people
- ▶ Infectious period is one day



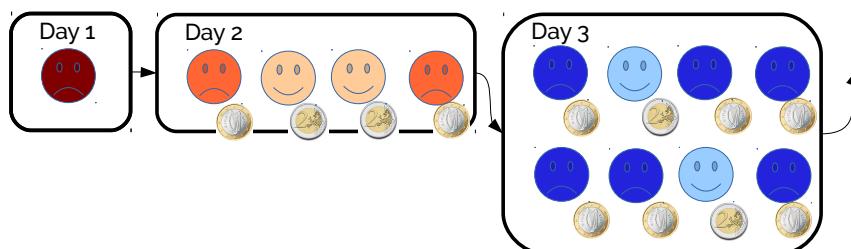
Time	1	2	3	4	5	6	7	8
Count	1	2	4	8	16	32	64	128



Currently our simulations have been “deterministic”, meaning if I run the simulation again I get exactly the same result. So suppose I have a scenario where the infectious population doubles every day like this. That seems a little unlikely.

Stochastic

- ▶ Each infectious person → 4 new contacts
- ▶ Chance of new infection is $\frac{1}{2}$



Time	1	2	3	4	5	6	7	8
Count	1	2	6	15	27	54	111	207

CHICAS
centre for
health informatics,
computing, and statistics

Instead we can run a stochastic, or random, model.

In this case we have each infectious person having four new contacts, and each of those has a 50% chance of catching the infection. So on day one those four faces are contacted by red face, but only two get infected. From day 2 those two infected faces contact eight blue faces but only six get infected based on a random coin toss.

If I keep doing this I get a table of infectious cases over time.

Repeated Randomness

Time	1	2	3	4	5	6	7	8
1	2	6	15	27	54	111	207	
1	3	9	16	31	68	120	250	
1	2	4	6	17	23	50	115	
1	1	2	1	0	0	0	0	
1	1	3	7	17	32	69	118	
1	3	6	12	26	43	83	156	

► Deterministic:

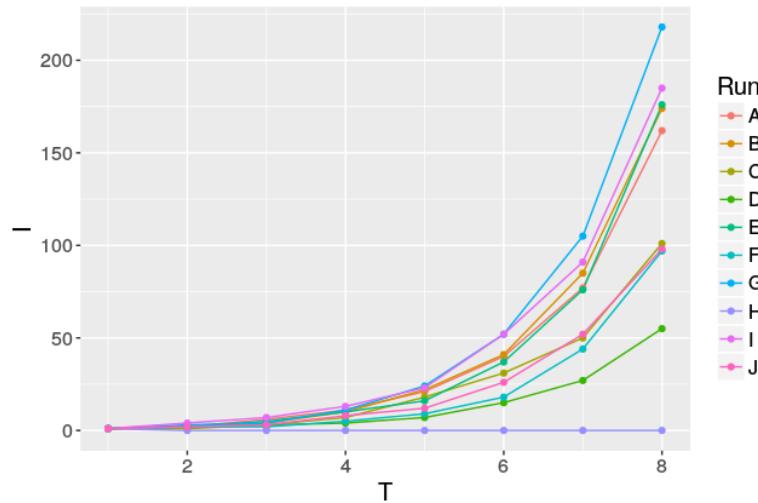
Time	1	2	3	4	5	6	7	8
Count	1	2	4	8	16	32	64	128



But now if I do it again my coin tosses land differently, and I get a different epidemic. And if I happen to toss all tails at any point, the epidemic dies out, and once its died out it can't come back.

Compare with the deterministic model, that ends up with 128 infectious after 8 steps – the random model ranges from 0 to 250 in six simulations

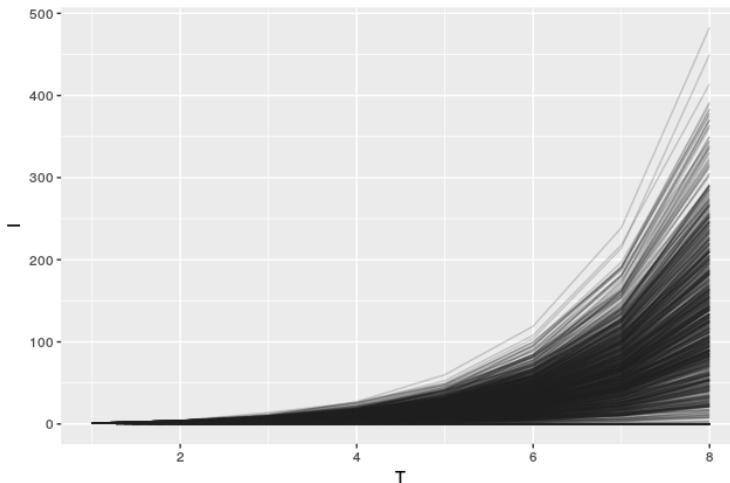
Results



If we plot ten of these runs on a graph you can see the possible range of variation from the random model.

Note the deterministic model would be a curve going up to 128 at $T=8$.

More Results



Here's a few hundred runs – this time the largest produced nearly 500 infectious cases, but you can see that the average is going to be around 128.

Stochastic models like this, by incorporating our understanding of variability, let us answer questions like "What's the chances of this affecting X% of the population?" - the sort of question that's important for policy.

Randomness

- ▶ Previously...
 - ▶ Each infectious person → 4 new contacts
 - ▶ Who meets the same number of people every day?
 - ▶ The number of people you meet is the result of a *probability distribution*



Lets add another random step to our model.

Previously we had each infectious person meeting four new contacts with a 50% chance of infection.
Who meets the same number of people every day?
The number of people you meet is the result of a probability distribution.

Probability Distributions

► Uniform

$$P = 1/2$$



$$P = 1/6$$



$$P = 1/10$$



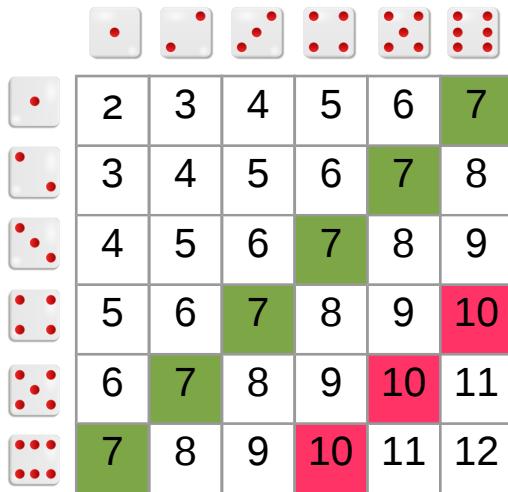
Some probability distributions we know already. The Uniform distribution is where each outcome has the same chance – like heads or tails, or any of the numbers 1 to 6 on a single die, or 1 to 10 on a D10

Non-Uniform



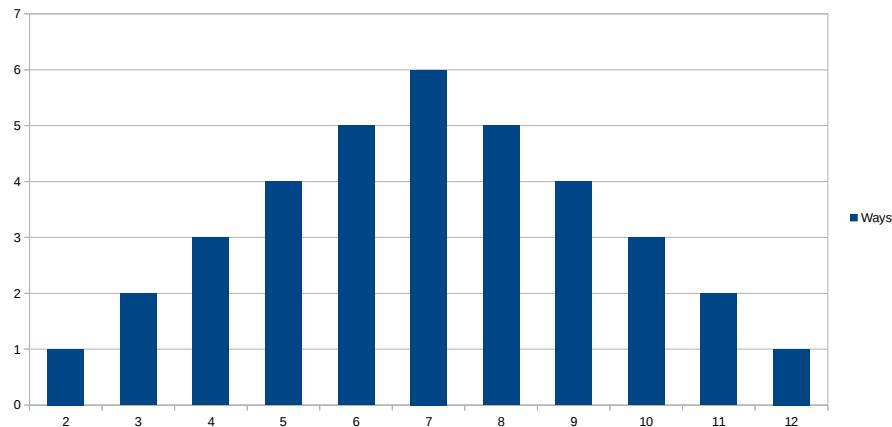
But what about the sum of two dice? That's not uniform.

Non-Uniform



If you look at the possible outcomes of the sum of two dice you see there's more ways of scoring 7 than 10, and only one way of scoring 2 or 12.

Rolling Two Dice



That produces a distribution that looks like this. Could we use something like this to get a distribution of the number of contacts?

Simple Count Distribution

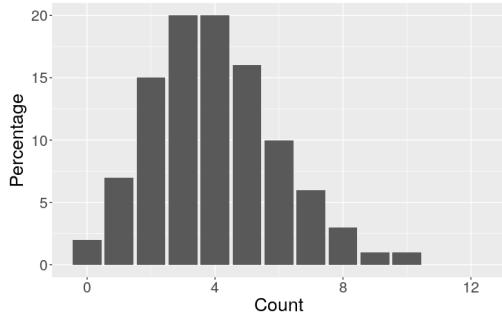
- ▶ Need a distribution that:
 - ▶ Is always greater than or equal to zero
 - ▶ Has no upper bound
 - ▶ As few parameters as possible



What we want is a distribution that has these properties. Statisticians have worked out a nice probability distribution with these properties.

Poisson Distribution, mean=4

$$\frac{\lambda^k e^{-\lambda}}{k!}$$



Count	Percentage
0	2
1	7
2	15
3	20
4	20
5	16
6	10
7	6
8	3
9	1
10	1
11	0
12	0
...	0

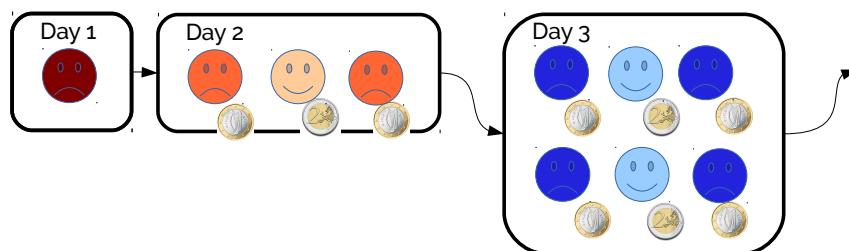


The Poisson distribution has a single parameter that defines the mean of the numbers from the distribution, and it comes from a fairly simple formula. The plot and the table show the distribution.

So if the mean number of contacts is four, there's a 2 in 100 chance of contacting 0 people, and a 1 in 100 chance of contacting 10 people.

Poisson Contacts

- ▶ Each infectious person → P(4) new contacts
- ▶ Chance of new infection is $\frac{1}{2}$

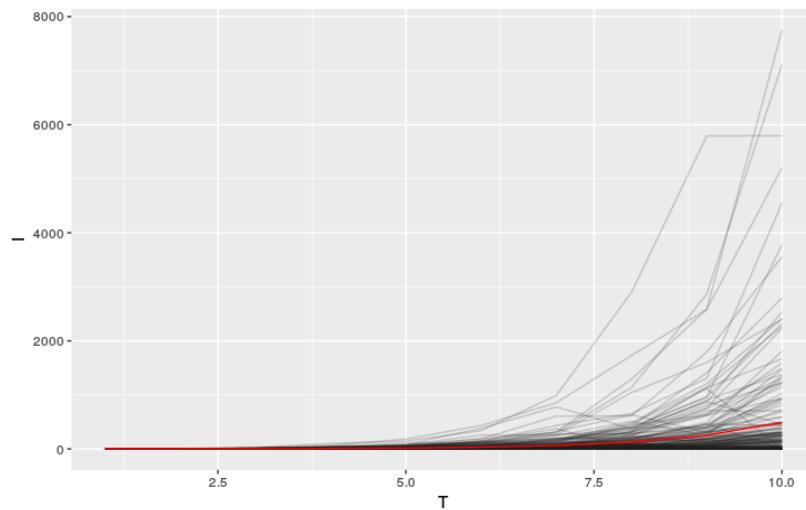


Time	1	2	3	...
Count	1	2	4	...

CHICAS
centre for
health informatics,
computing, and statistics

So we can run our simulation but instead of fixing it at four contacts, we take a number from the Poisson distribution with mean 4, and then the contacts are infected by a coin toss.

t up to 10

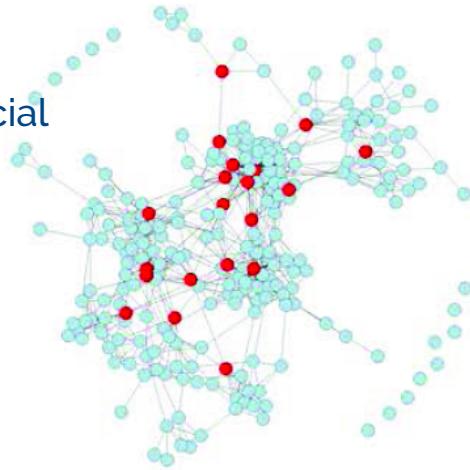


This results in epidemics that still have a similar average value as a deterministic process, as seen by the red line, but can produce even larger rare events.

The epidemic can also produce some very large downward drops too.

Network Model Analysis

- ▶ People interact with neighbours
- ▶ People group in cliques
- ▶ Some people highly social



Of course in reality we don't interact with a random number of random people – we have contact groups, like colleagues, friends etc, and children in school classes. A lot of research goes into seeing how these groups affect disease spread.

BBC [Sign in](#) [News](#) [Sport](#) [Weather](#) [iPlayer](#) [TV](#) [Radio](#) [More...](#) [Search](#)

NEWS HEALTH

Home World UK England N.Ireland Scotland Wales Business Politics Health **Education** Sci/Environment Technology Entertainment

18 January 2015 Last updated at 00:03 [Share](#) [Facebook](#) [Twitter](#) [Email](#)

Popular medical students 'should get flu jab first'

[COMMENTS \(168\)](#)



The government wants three-quarters of healthcare workers to be vaccinated

Prioritising medical students with lots of friends for flu jabs could help increase the number of healthcare workers protected against the virus, say Lancaster University researchers.

In a study in *The Lancet*, they calculated that vaccination rates would rise if people with large social networks influenced their peers.

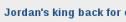
Related Stories

[Flu cases 'highest for three years'](#)
['Too few' toddlers](#)

Top Stories

 [NZ judge to lead child abuse inquiry](#)

 [Jordan's king back for crisis trip](#)

 [Taiwan plane in deadly river](#)

 [Chilcot rejects Iraq timetable](#)

Features

 [Dark horse](#)
Can you still offend them?

 [Rail wave](#)
The lines run to the sea

 [Husband](#)
Man battles pregnant wife

 [Rodney](#)
Tactics for someone's

 [Been and gone](#)
A pioneering photographer

CHICAS
centre for
health informatics,
computing, and statistics

So this was the BBC article on some research I was involved in a few years ago about how popular students should be vaccinated first!

Conclusions

- ▶ Disease modelling helps understand the epidemic process
- ▶ Simple models can be useful
- ▶ Complex models give more insight, but take longer to compute



So a quick summing up.

Disease modelling helps us understand the epidemic process

Simple models can be useful.

Complex models can give more insight, but take longer to compute.