

Mathematical Models for Infection

A gentle introduction

Chiara Sabatti, during the Covid-19 shelter-in-place, Spring 2020

Goals

- ▶ I am not an epidemiologist
- ▶ The following is only an attempt to gain understanding of terms and data we are presented with daily
- ▶ We start with a simple, in many way unrealistic model
- ▶ There are more sophisticated models that one can use and some might be in use in “professional” projections of the COVID-19 epidemic
- ▶ It might be an interesting research project to try to understand more complex models

A simple model for infection

- ▶ W. O. Kermack and A. G. McKendrick (1927)
- ▶ The goal is to have a sense of how many people are going to be infected over time when a contagious disease enter a population and when the epidemic will end
- ▶ It is a *deterministic* model: it gives a fixed number of infected for each unit of time. Of course, reality is not deterministic, but random, a lot of different outcomes are possible on any given day. But this is a starting point.
- ▶ It was proposed to explain the rapid rise and fall in the number of infected patients observed in epidemics such as the plague (London 1665-1666, Bombay 1906) and cholera (London 1865).

A SIR model

We assume that the population has a fixed size N (i.e., no births, deaths due to disease, or deaths by natural causes)

At any point t in time, the population can be partitioned in three groups:

- ▶ S_t the people that are *susceptible* to be infected, as they have not been infected yet
- ▶ I_t the people that are currently *infected* and capable of transmitting the disease
- ▶ R_t the people that have *recovered* from infection (or have died) and therefore cannot be infected or infect any longer.

$$N = S_t + I_t + R_t$$

Note that we assume that the incubation period of the infectious agent is instantaneous, and duration of infectivity is same as length of the disease.

Model for infections

- ▶ We assume that time is discrete, so that infections happen between one day and the next.
- ▶ We assume that, within a day, everyone has the same chance to interact with anyone else in the population in a way that has the same probability of communicating the disease (this requires a completely homogeneous population with no age, spatial, or social structure). For each affected individual the probability of interacting with another member of the population and infecting him/her is β . So, any day, one contagious individual is expected to infect βN people.
- ▶ Note, however, that only the fraction S_t/N of the population can actually be newly infected.

If today we have I_t infected individuals, tomorrow ($t + 1$)

$$\text{new infections at time } t + 1 = \beta N \times \frac{S_t}{N} \times I_t$$

The interpretation of β

- ▶ Note that β should depend on the population size: if the population is large, the chance of one affected individual to interact with any other specific individual should be smaller.
- ▶ So, often, rather than using β for these models, one uses an alternative parameter B such that $\beta = B/N$ – so that the role of population size is explicitly called out.

We will see more about this later.

Recovery

- ▶ We assume that a portion γ of the infected recovers at each period

$$\text{new recoveries at time } t + 1 = \gamma I_t$$

- ▶ Without going into the mathematics of it, we note that $1/\gamma$ gives you an indication of how long a person is infected: the easiest it is to recover—the larger γ —the shorter the period one is ill.

What happens to the three quantities over time

$$R_t = R_{t-1} + \gamma I_{t-1}$$

$$S_t = S_{t-1} - \beta \times S_{t-1} \times I_{t-1}$$

$$I_t = I_{t-1}(1 - \gamma) + \beta \times S_{t-1} \times I_{t-1}$$

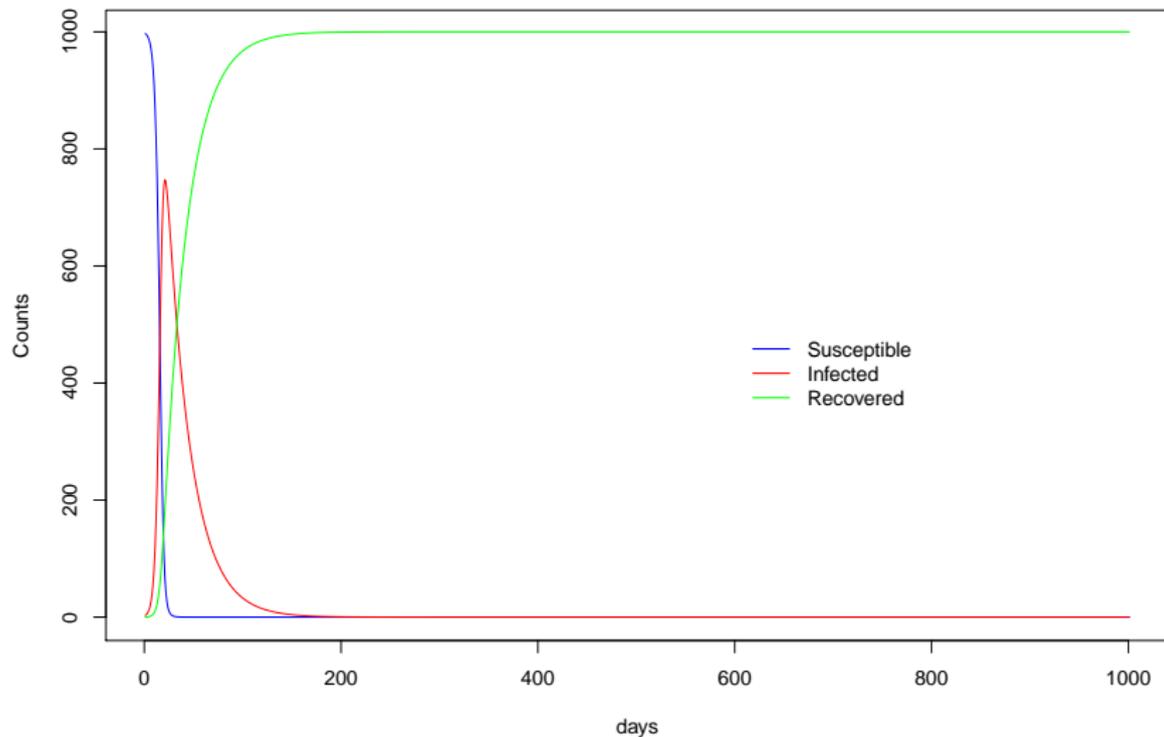
Note: the number of cases etc. should be integers, but here they can be fractional.

A simulation for $N = 1000$

```
gamma<-0.04
beta<-0.0005
S<-997
I<-3
R=0
Sc<-S
Ic<-I
Rc<-R
for(i in 1:1000)
{
  Rc<-Rc+gamma*Ic
  NI<-beta*Sc*Ic
  Sc<-Sc-NI
  Ic<-(1-gamma)*Ic+NI
  R<-c(R,Rc)
  S<-c(S,Sc)
  I<-c(I,Ic)
}
```

Simulation results

Epidemic dynamics with $\beta = 5e-04$, $\gamma = 0.04$



What influence the parameters?

- ▶ Suppose that one disease can be communicated only by sharing saliva and another as soon as two people are within 6 feet of each other. Which parameter value should be different in the model for these two diseases?
- ▶ Suppose we find a cure that shortens the period while a person is sick and infective. What parameters would this change?
- ▶ Suppose at some point a vaccine is discovered and administered to a fraction of the population: what would this change?
- ▶ Suppose people practice “social distancing”: what would this change in the model?

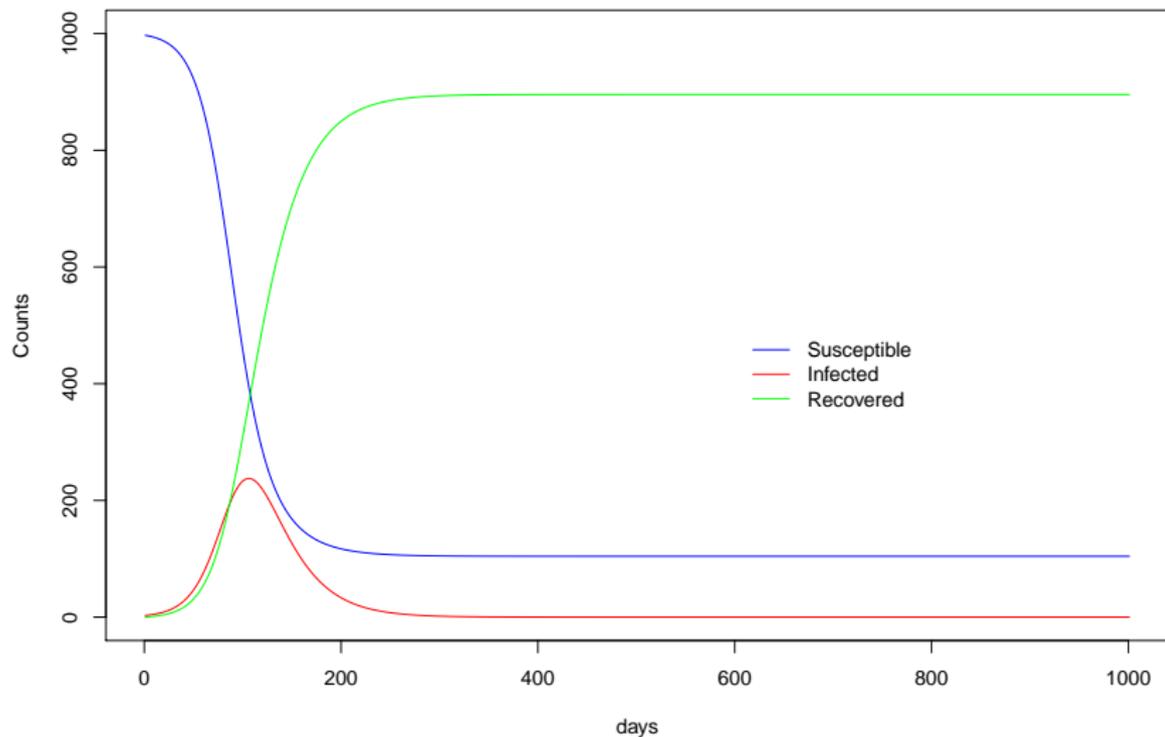
To test out some of our ideas, we are going to change the simulation

Let's change the simulation

```
gamma<-0.04
beta<-0.0005/5
S<-997
I<-3
R=0
Sc<-S
Ic<-I
Rc<-R
for(i in 1:1000)
{
  Rc<-Rc+gamma*Ic
  NI<-beta*Sc*Ic
  Sc<-Sc-NI
  Ic<-(1-gamma)*Ic+NI
  R<-c(R,Rc)
  S<-c(S,Sc)
  I<-c(I,Ic)
}
```

Simulation results

Epidemic dynamics with $\beta = 1e-04$, $\gamma = 0.04$

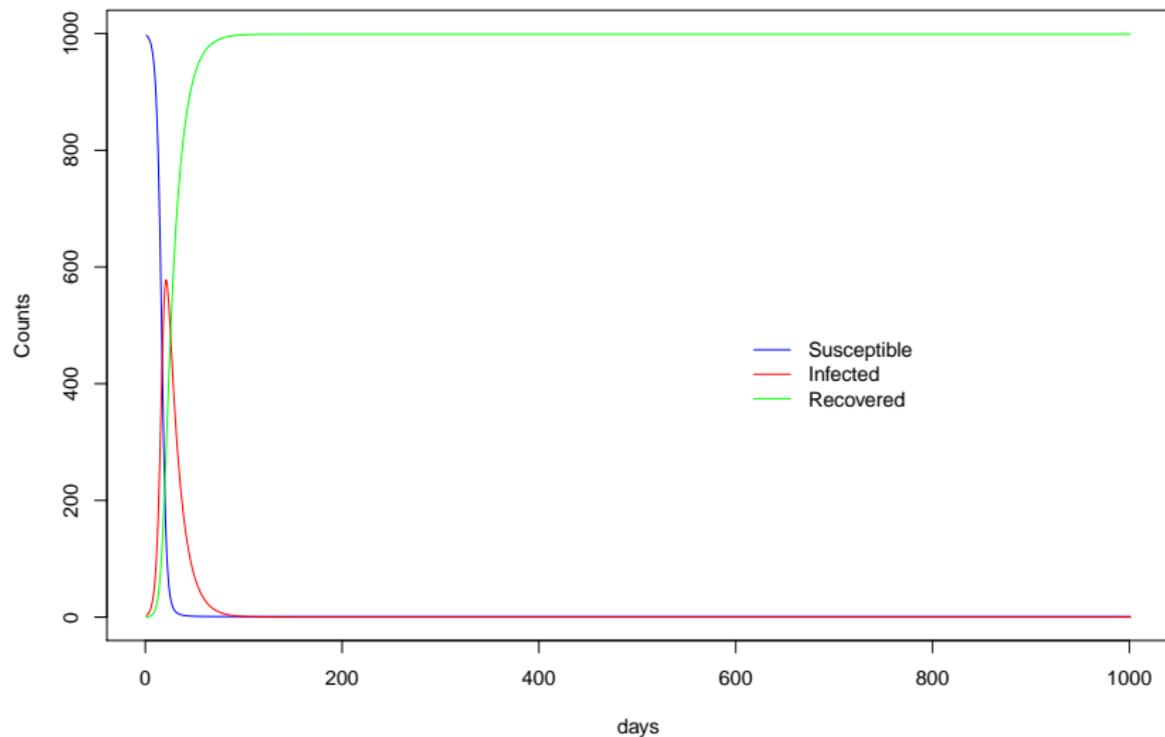


Let's change the simulation again

```
gamma<-0.04*2
beta<-0.0005
S<-997
I<-3
R=0
Sc<-S
Ic<-I
Rc<-R
for(i in 1:1000)
{
  Rc<-Rc+gamma*Ic
  NI<-beta*Sc*Ic
  Sc<-Sc-NI
  Ic<-(1-gamma)*Ic+NI
  R<-c(R,Rc)
  S<-c(S,Sc)
  I<-c(I,Ic)
}
```

Simulation results

Epidemic dynamics with $\beta = 5e-04$, $\gamma = 0.08$



Getting a bit closer to the meaning of parameters

Let's say we are at the beginning of the epidemic, so that the number of susceptible individuals is $\approx N$: there are no recovered, and the infected are a transcurable number (this is particularly true if N is large).

Then, one infectious individual will infect $\frac{\beta}{\gamma}N$ people. Why?

- ▶ At each time periods, the individual infects βN people
- ▶ An individual is contagious for $1/\gamma$ time periods

This quantity is important enough that it gets a name:

$$R_0 = \frac{\beta}{\gamma}S_0 = \frac{\beta}{\gamma}N$$

R_0 and epidemic

R_0 is the average number of people one case infects.

When $R_0 < 1$, each person who contracts the disease will infect fewer than one person before dying or recovering, so the outbreak will peter out.

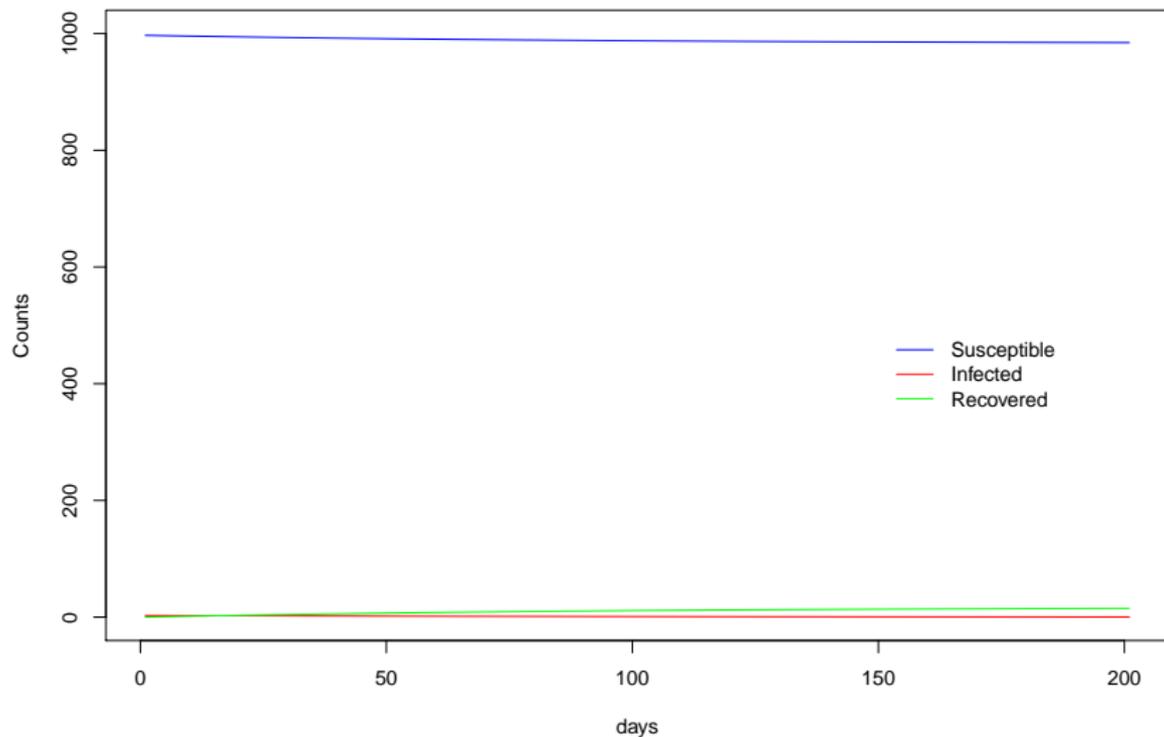
When $R_0 > 1$ each person who gets the disease will infect more than one person, so the epidemic will spread.

To verify this we can look at our simulation and tweak the parameters so that we get one $R_0 < 1$: $R_0 = \frac{0.00005 \times 1000}{0.06} = 0.83$

Note: the results actually hold for a continuous model, so with discrete numbers things might be a bit funny around the edges.

Simulation results: no epidemic

Epidemic dynamics with $\beta = 5e-05$, $\gamma = 0.06$



The early phase of the epidemic

Now, let's look at the beginning of the epidemic, when $R_0 > 1$, and N is large, so that for a while $S_t \approx N$.

Then

$$I_t = I_{t-1} + \beta N I_{t-1} - \gamma I_{t-1} = (1 + \beta N - \gamma) I_{t-1}$$

Which means that

$$I_t = (1 + \beta N - \gamma)^t I_0,$$

that is, there is an **exponential** growth of the number of cases.

Note also that

$$I_t - I_{t-1} = (\beta N - \gamma) I_{t-1},$$

that is the number of new cases in each day is **proportional** to the total number of cases.

Early stage epidemic

Let $\rho = \beta N - \gamma$; note that if $R_0 > 1 \iff \rho > 0$: the role of R_0 in determining if there is an epidemic or not is reflected in ρ .

$$\begin{aligned}I_t &= (1 + \rho)^t I_0 = 1.06^t I_0 \\I_t - I_{t-1} &= (\rho) I_{t-1} \\ \log(I_t - I_{t-1}) &= \log(\rho) + \log(I_{t-1})\end{aligned}$$

Note: we obviously cannot take the $\log()$ when $\rho < 0$.

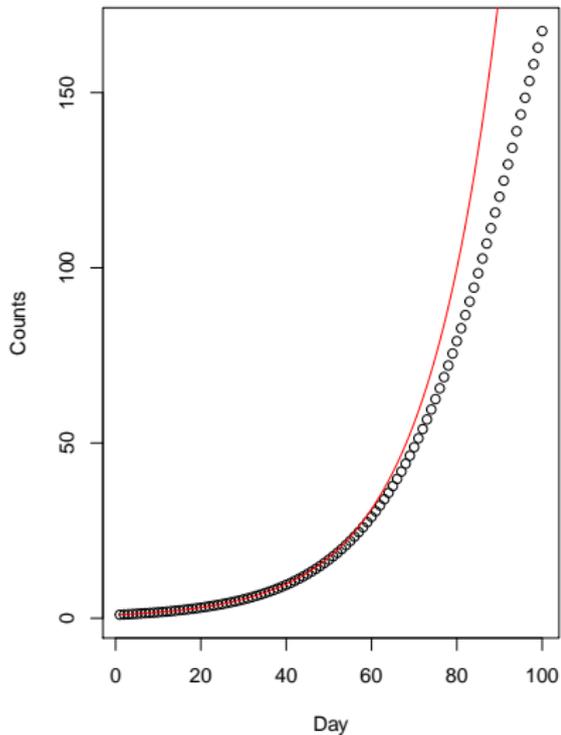
Let's choose some parameter values that give a "slow" growth, so that we can have a "long" early stage in our small population $N = 1000$: $\beta = 0.0001$, $\alpha = 0.04$.

Note: In our case $\rho = 0.0001 \times 1000 - 0.04 = 0.06$ and the number of cases here will be very small and it will look unrealistic, but this is what we need with a small population to have a longish first fase.

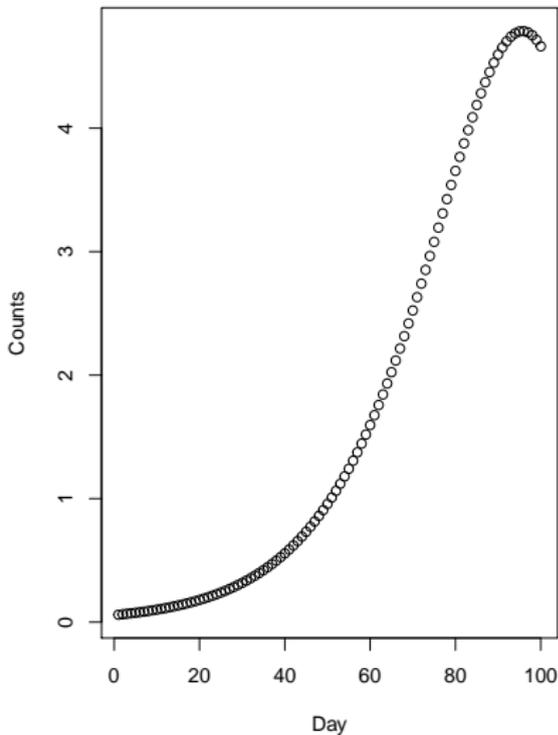
Exponential growth of cases in early stages: simulation

$$I_t = (1 + \rho)^t I_0 = 1.06^t I_0$$

Total cases

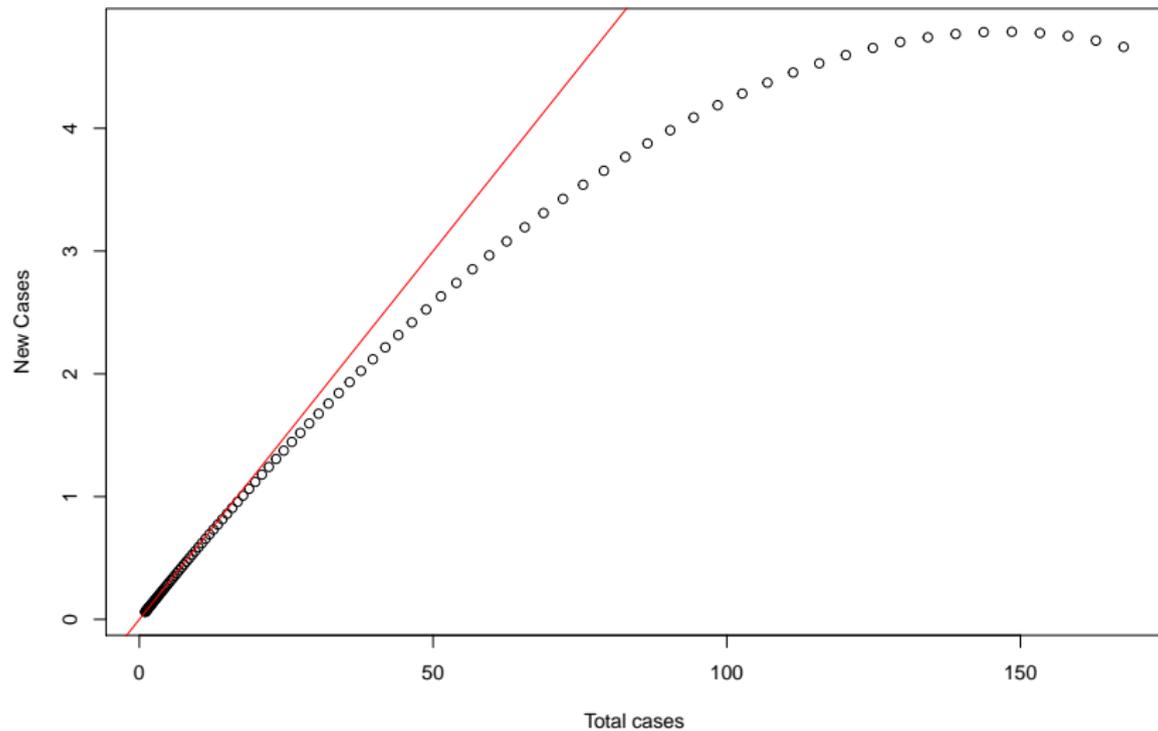


New cases



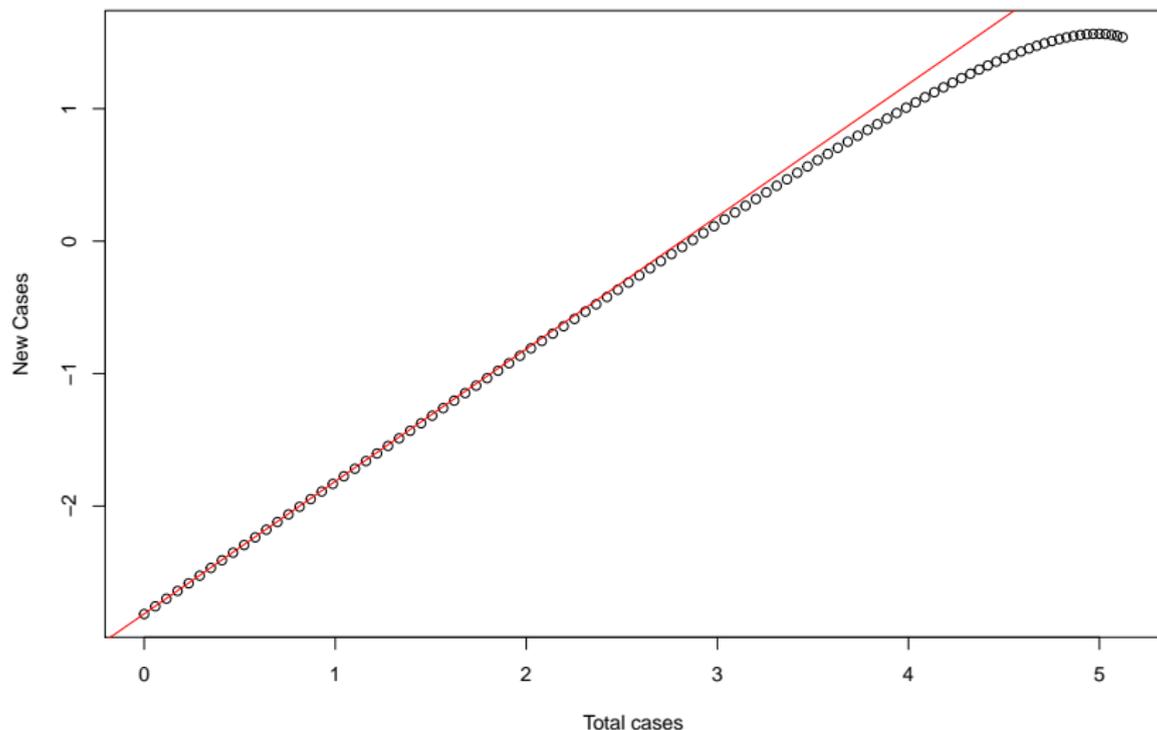
New cases and existing cases in early stages

$$I_t - I_{t-1} = \rho I_{t-1} = 0.06 I_{t-1}$$



New cases and existing cases in early stages

$$\log(I_t - I_{t-1}) = \log(\rho) + \log(I_{t-1}) = \log(0.06) + \log(I_{t-1})$$



New cases and existing cases in early stages - log plot

- ▶ The role of β , N and γ here is simply in defining the intercept $\log(\rho)$ of the line with slope 1.
- ▶ It is easier to resolve on the same graph the very first stages when the number of cases is small and the later stages when the number of cases is high
- ▶ It is easier to detect when we depart from the slope of line 1

Data

Kermack and McKendrick (1927) A contribution to the mathematical theory of epidemics, *Proceedings of the Royal Society*, A

MATHEMATICAL THEORY OF EPIDEMICS—I 49

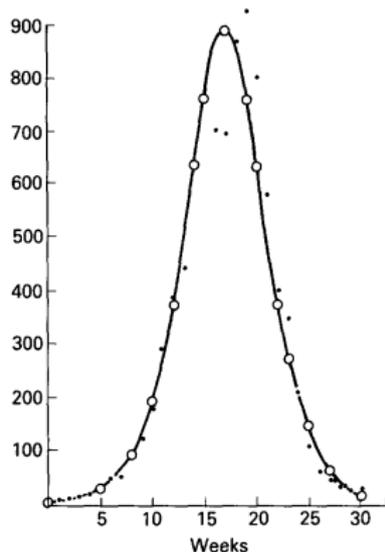
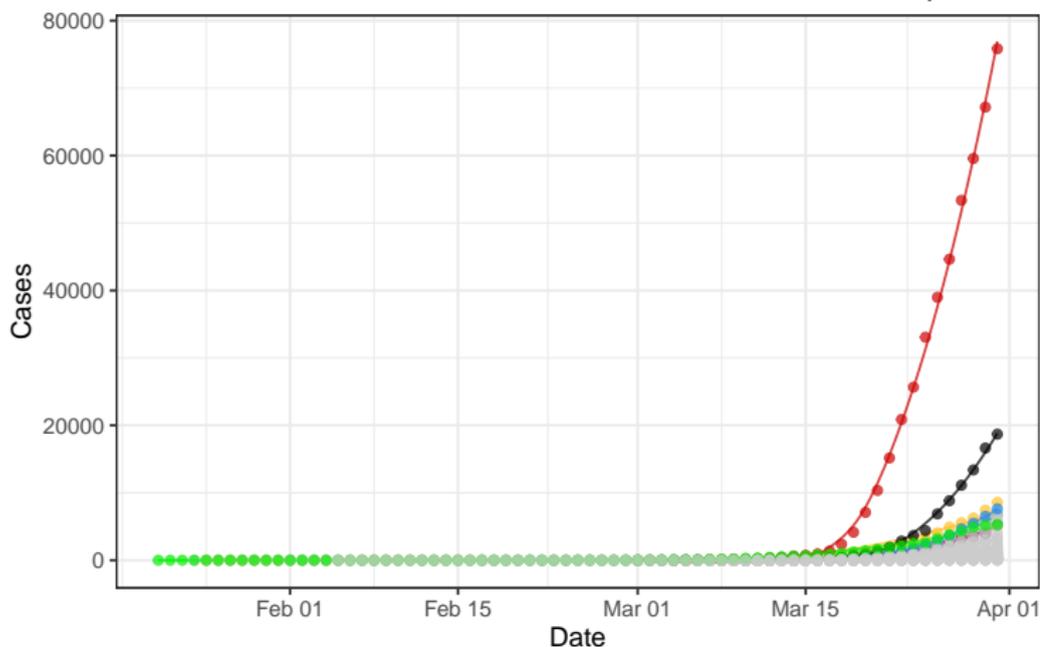


Figure 1. Deaths from plague in the island of Bombay over the period 17 December 1905 to 21 July 1906. The ordinate represents the number of deaths per week, and the abscissa denotes the time in weeks. As at least 80–90% of the cases reported terminate fatally, the ordinate may be taken as approximately representing dz/dt as a function of t . The calculated curve is drawn from the formula:

Data from: <https://github.com/nytimes/covid-19-data>

Each US state corresponds to a series of connected points

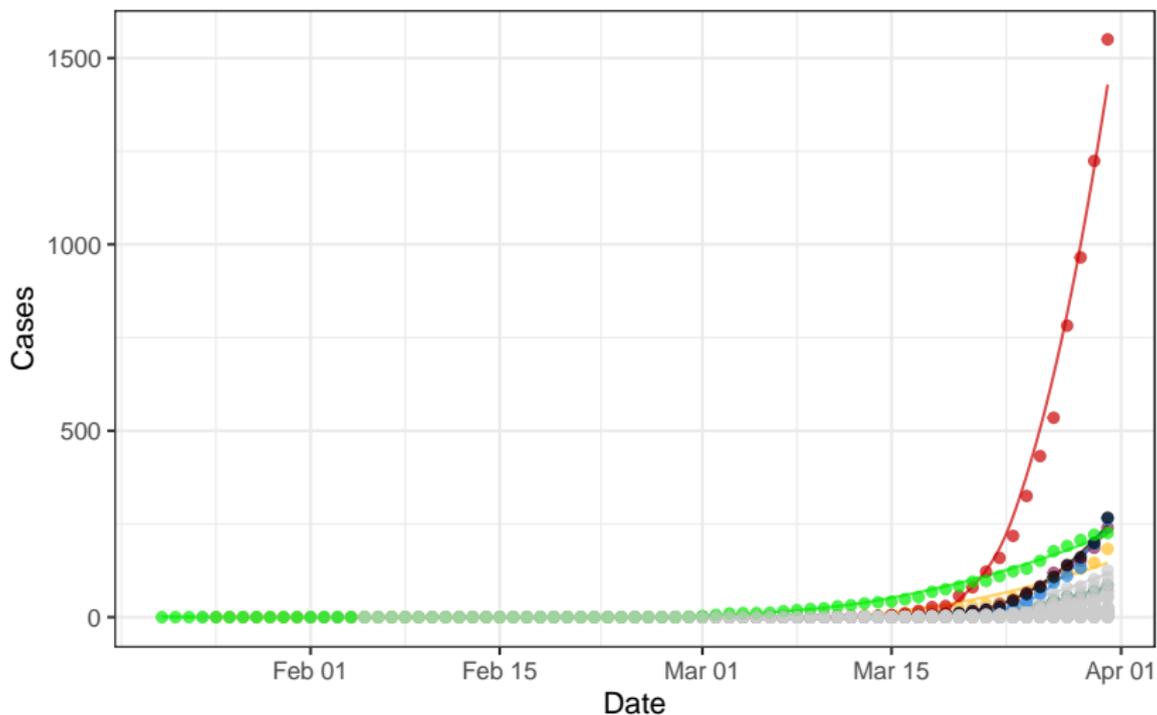
Total COVID-19 confirmed cases over the course of the epidemic



Red is NY, black NJ, golden CA, green WA, blue MI

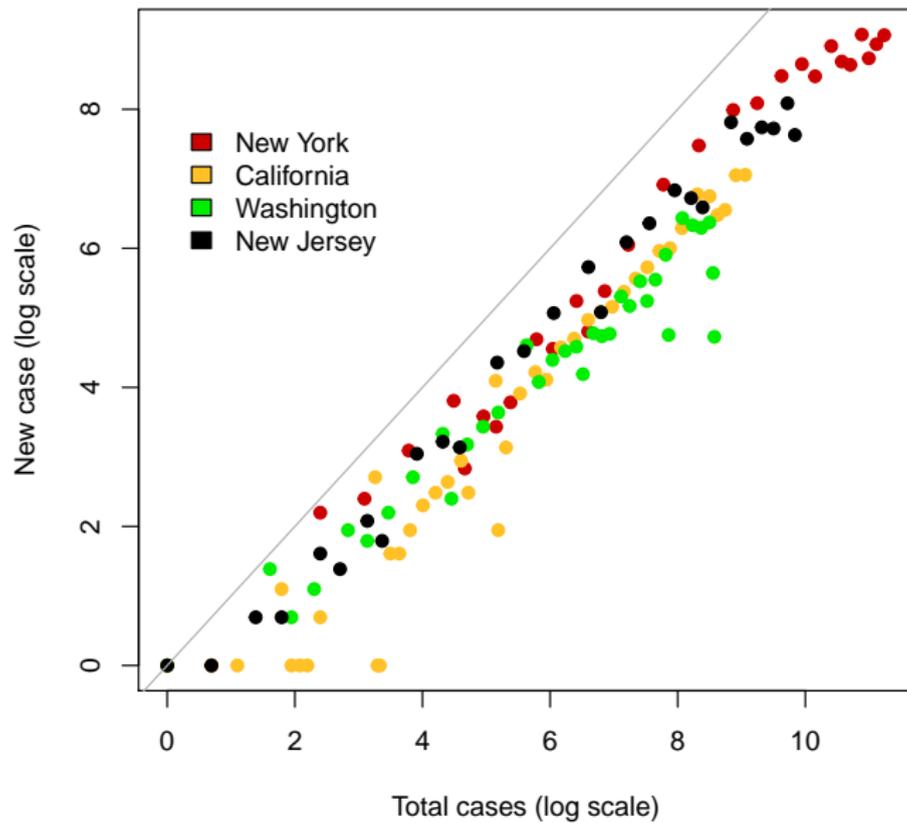
USA data

Total COVID-19 deaths over the course of the epidemic



Red is NY, black NJ, golden CA, green WA, blue MI

Few USA states



Looking at data

- ▶ In the early stages, the role of γ and β is indistinguishable: we only get to see the effect of $\rho = \beta N - \gamma$.
- ▶ Is COVID-19 spreading less in California than in New York ($\beta_{CA} < \beta_{NY}$?)
- ▶ It is difficult to interpret the data, as we do not have a complete count of the number of cases; this strongly depends on the number of individuals who have been tested.

Few USA states –looking at deaths

