The mathematics of diseases

– On Modeling Hong Kong's SARS Outbreak

Dr. Tuen Wai Ng

Department of Mathematics, HKU

Content

- Basic Epidemic Modeling
- SIR Model
- My Recent works on Modeling of the SARS propagation in Hong Kong

Infectious disease modeling

rat flea

Epidemics:

- Black Death

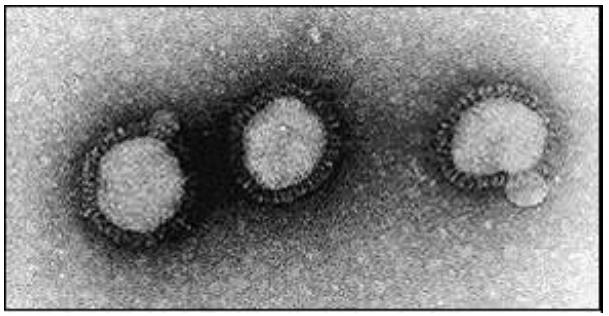
Europe lost 1/3 of population in 1347 - 1350.

Great Plague of London, 1664–66. Killed more than 75,000 of total population of 460,000.

Infectious disease modeling

Epidemics:

- Influenza



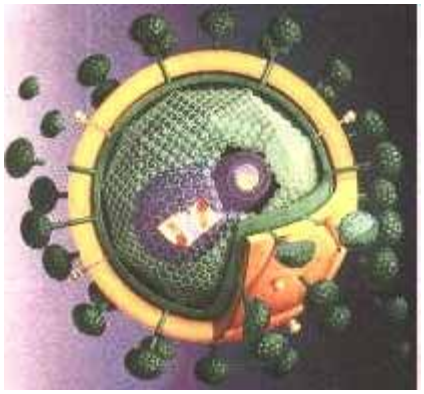
influenza virus

- killed 25 million in 1918-19 in Europe

Infectious disease modeling

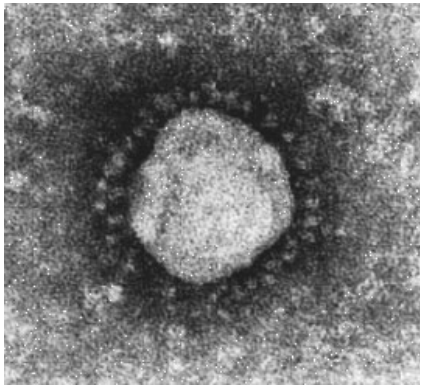
Epidemics:

- AIDS



AIDS virus

- SARS



Coronavirus

Infectious disease modeling

Epidemics:

- bovine spongiform encephalopathy (mad cow disease)

- Chicken flu

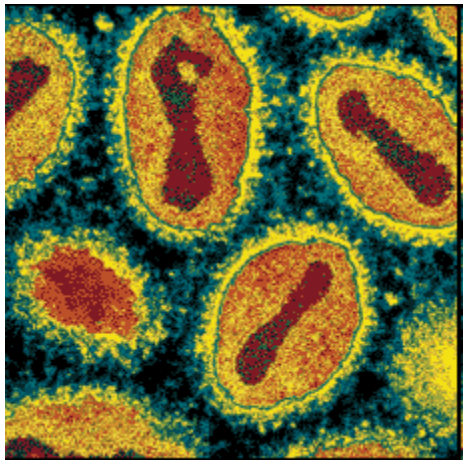
Infectious disease modeling

Mathematical models can:

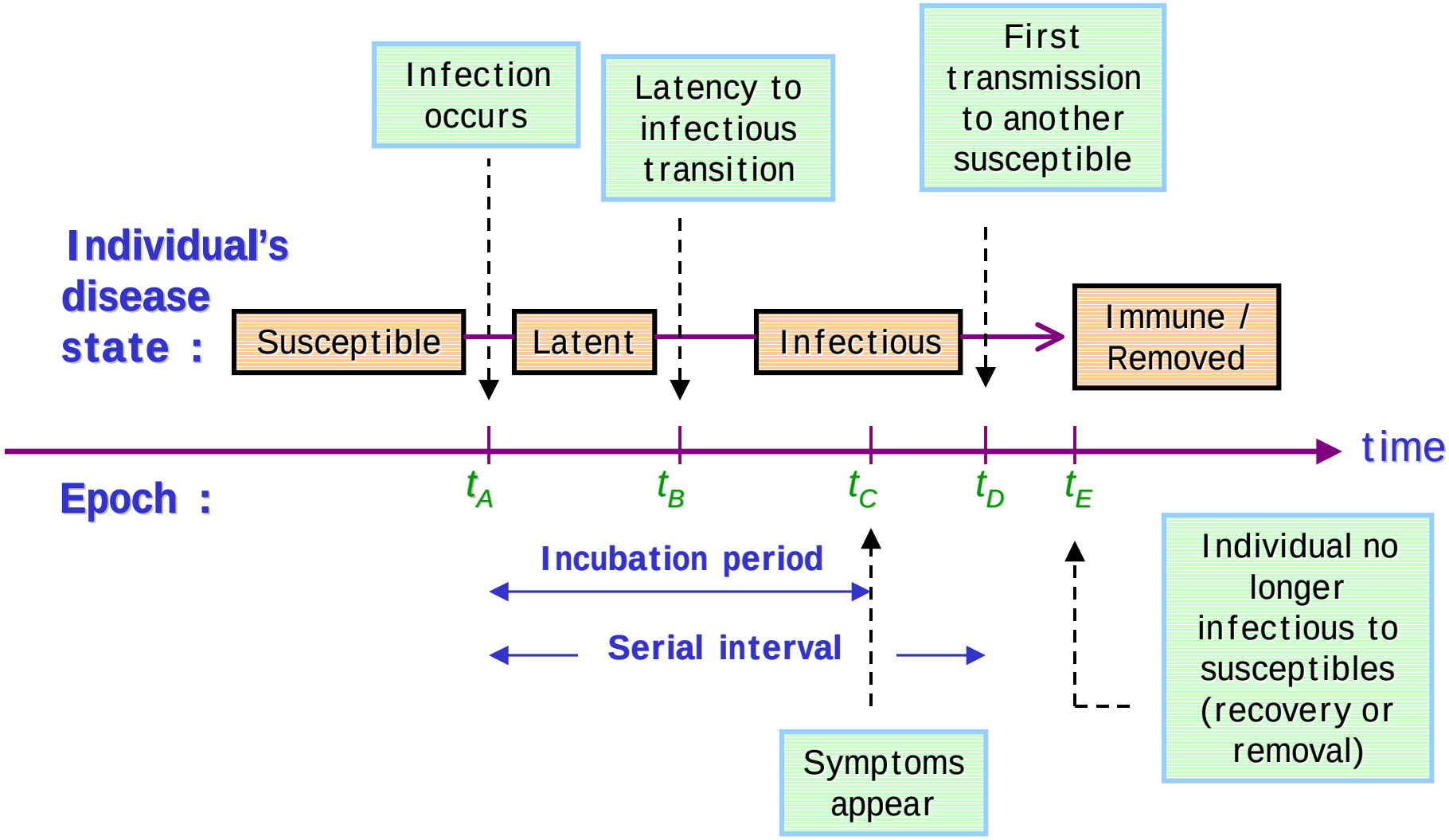
- predict rate of spread, peak, etc., of epidemics
- predict effects of different disease control strategies

The WHO's eradication project reduced smallpox (variola) deaths from two million in 1967 to zero in 1977–80.

Smallpox was officially declared eradicated in 1979.



smallpox virus



Note : t_D is constrained to lie in the interval (t_B, t_E), so $t_D > t_C$ (as shown) and $t_D < t_C$ are both possible.

Basic Assumptions of the simplest epidemic model, the SIR model

- Population size is large and constant

- No birth,death,immigration or emigration

- No latency

- Homogeneous mixing

- that is each pair of individuals has equal probability of coming into contact with one another (this is reasonable for a school or households in a building).

Basic Assumptions of the SIR model

- We divide the total population N into three groups:

  I) Susceptible class, S_t=S(t)=no. of susceptibles

  — those who may catch the disease but currently are not infected.

  II) Infective class, I_t=I(t)=no. of infectives

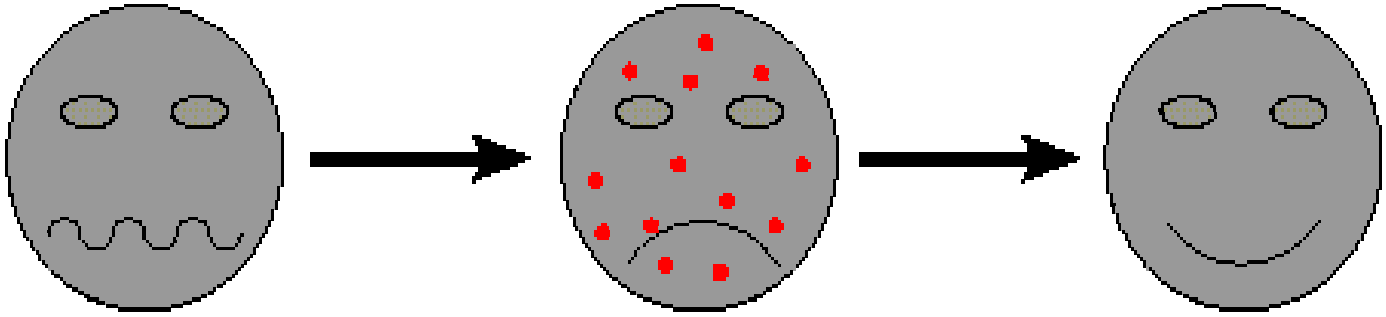
  — those who are infected with the disease and currently contagious.

  III) Removed class, R_t=R(t)=no. of removals

  — those who cannot get the disease, because they either have recovered permanently, are naturally immume, or have died.

Basic Assumptions of the SIR model

- The members of the population progress through the three classes in the following order.



Susceptible S_t → **Infectious** I_t → **Immune/ Removed** R_t

SIR

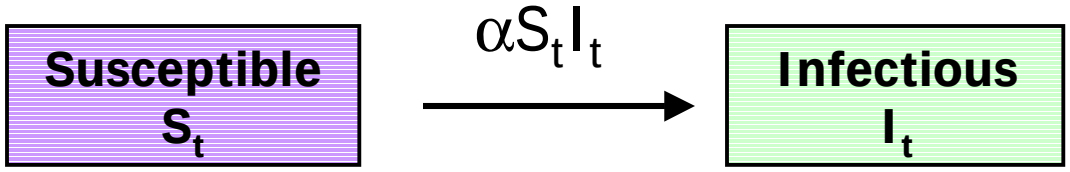
Basic Assumptions of the SIR model

Disease spreads when a susceptible individual comes in contact with an infected individual and subsequently becomes infected.

- Assuming homogeneous mixing, the mass action principle says that the number of encounters between susceptibles and infectives is given by the product S_t I_t.
- However, only a proportion α of the contacts between susceptibles and infectives result in infection. Hence, in the next time interval,

$$S_{t+1} = S_t - \alpha S_t I_t .$$

Basic Assumptions of the SIR model



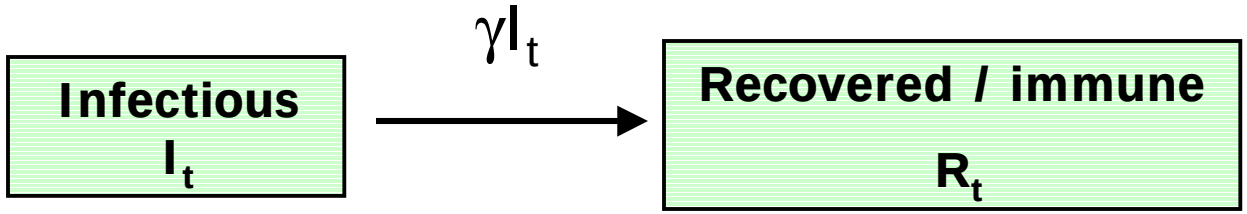
- However, only a proportion α of the contacts between susceptibles and infectives result in infection. Hence, in the next time interval,

$$S_{t+1} = S_t - \alpha S_t I_t .$$

Basic Assumptions of the SIR model

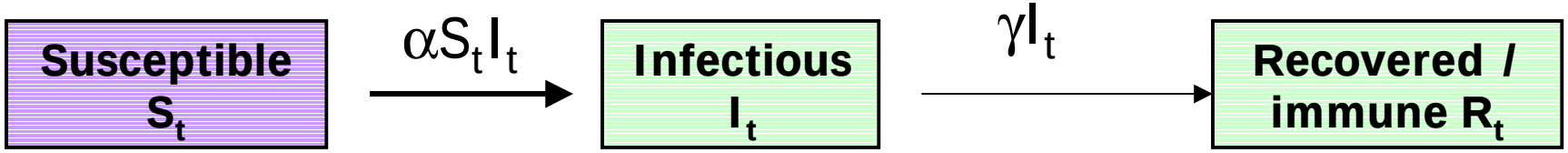
During one time step, the infective class grows by the addition of the newly infected. At the same time, some infectives recover or die, and so progress to the removed stage of the disease.

The removal rate γ measures the proportion of the infective class that ceases to be infective, and thus moves into the removed class, in one step.



Basic Assumptions of the SIR model

Clearly the removed class increases in size by exactly the same amount that infected class decreases.



Therefore, we have

$$I_{t+1} = I_t + \alpha\, S_t I_t - \gamma\, I_t,$$

$$R_{t+1} = R_t + \gamma\, I_t$$

- **If we let $\Delta S = S_{t+1} - S_t$ and similarly for ΔI and ΔR, then the dynamics of the functions S,I and R are governed by the following equations.**

$$\Delta S = S_{t+1} - S_t = -\alpha S_t I_t,$$

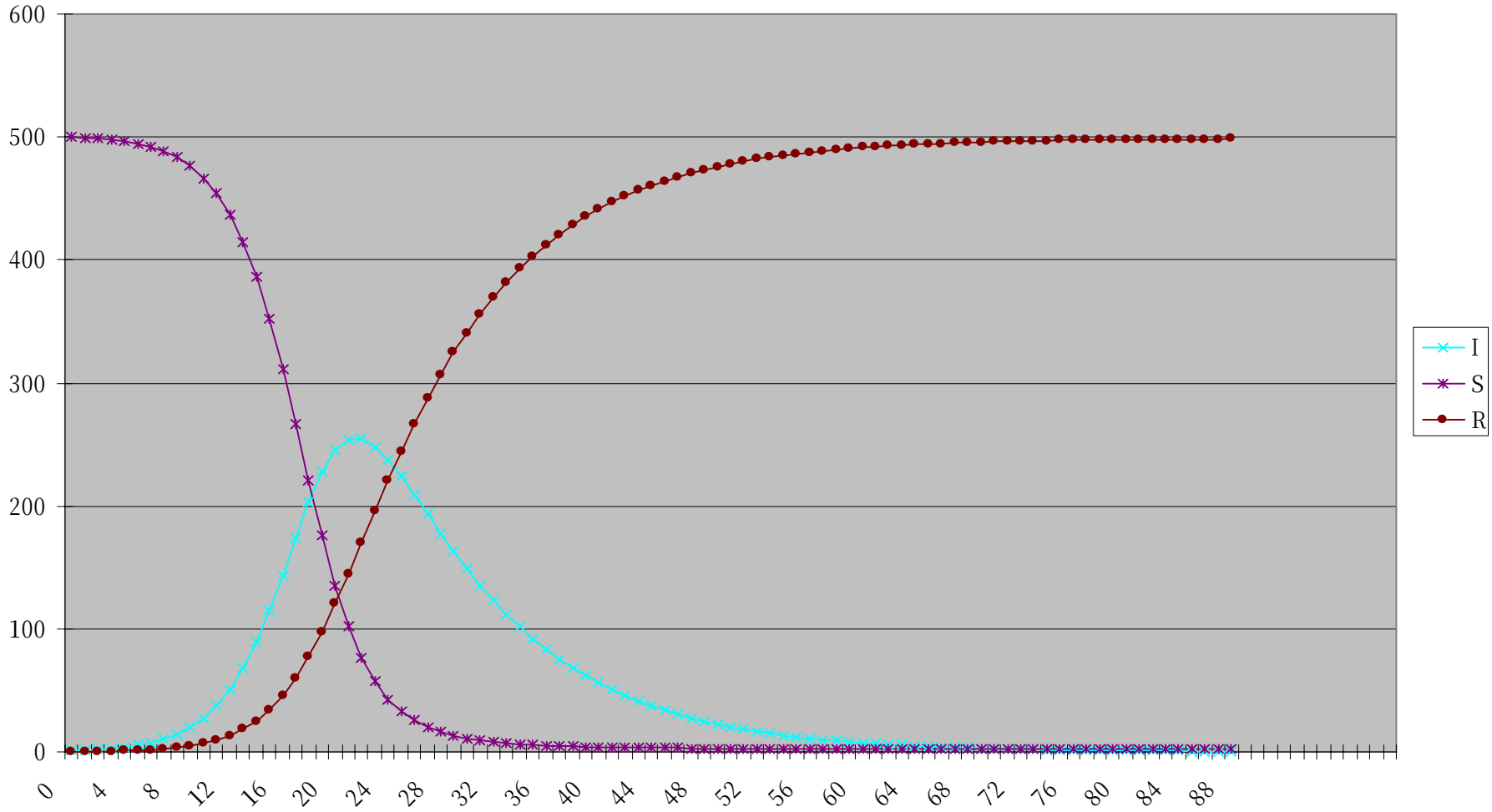
$$\Delta I = I_{t+1} - I_t = \alpha S_t I_t - \gamma I_t,$$

$$\Delta R = R_{t+1} - R_t = \gamma I_t .$$

- **For example, if $I_0 = 1$, $S_0 = 1000$, $R_0 = 0$, $\alpha = 0.001$ and $\gamma = 0.1$, then $S_1 = 1000 - 0.001(1000)(1) = 999$ and $I_1 = 1 + 0.001(1000)(1) - 0.1(1) = 1.9 \sim 2$.**

Exercise 1: Find S_t,I_t and R_t when t =2.

Exercise 2: Use Excel to compute S_t,I_t and R_t for t from 0 to 100 and plot the graph of S_t,I_t and R_t .



Thread Values and Critical Parameters

- We will say an epidemic occurs if $\Delta I > 0$ for some time t (i.e., if at some time t, the number of infectives grows).
- If $\Delta I < 0$ for all times, then the size of the infective class does not increase and no wider outbreak of the disease takes place.
- Therefore, it is important to know when

$$\Delta I = \alpha S_t I_t - \gamma I_t$$
$$= (\alpha S_t - \gamma) I_t$$

is positive, zero, or negative.

- Let's find out when Δ I is zero.

$$\Delta I = \alpha\, S_t I_t - \gamma\, I_t = (\alpha\, S_t - \gamma\,)\, I_t$$

- ΔI =0 once I_t=0 (as the population is disease free).
- Now assume I_t >0, then we have

$$\text{If} \quad S_t > \frac{\gamma}{\alpha}, \ \text{then} \quad \Delta\, I > 0\ .$$

$$\text{If} \quad S_t = \frac{\gamma}{\alpha}, \ \text{then} \quad \Delta\, I = 0\ .$$

$$\text{If} \quad S_t < \frac{\gamma}{\alpha}, \ \text{then} \quad \Delta\, I < 0\ .$$

- Note that S_t is a non-increasing function in t. Therefore, if $S_0 < \frac{\gamma}{\alpha}$, then for all t, $S_t < \frac{\gamma}{\alpha}$.
- Thus, if S_0 is below the value $\frac{\gamma}{\alpha}$, then $\Delta I < 0$ for all times, and the disease decreases in the population.
- However, when $S_0 > \frac{\gamma}{\alpha}$, the number of infective will grow and an epidemic occurs. In other words, we have an outbreak if and only if

$$R_0 = \frac{\alpha}{\gamma} S_0 > 1$$

$$R_0 = \frac{\alpha}{\gamma} S_0$$

- The above expression is called the *basic reproduction number* of the infection.
- Let's consider $R_0 = \frac{\alpha}{\gamma} S_0 = (\alpha S_0)(\frac{1}{\gamma})$ from a more biological viewpoint.
- The term $\alpha S_0 I_0$ measures the number of individuals that become infected at the outset of an epidemic.
- If we divide by I_0 , we obtain a "per-infective" measurement: αS_0 is the number of individuals who become infected by contact with a single ill individual during the initial time step.

- Actually, if we introduce one infective into a wholly susceptible population S_0, this ill individual may eventually infect many more than αS_0 others, because an infective may remain contagious for many time steps.
- For example, suppose a young child remains contagious with chickenpox for about 7 days. Then, using a time step of 1 day, this child would infect about (αS_0) (7) susceptibles over the course of a week.
- Moreover, if the period of contagion lasts 7 days, then each day we expect roughly or approximately 14% of the total number of infectives to move from the infective class I_t into the removed class R_t.

- Because the removal rate γ measures the fraction of the infective class “cured” during a single time step, we have found a good estimate for γ; we take

$$\gamma = \frac{1}{7} \approx 0.142$$

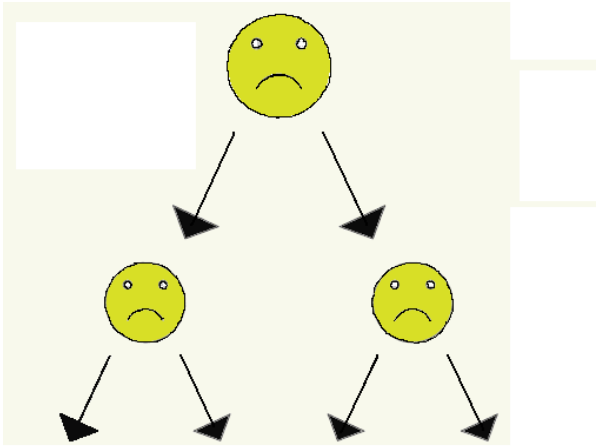
- At the same time, we have found a good interpretation for $1/\gamma$: it is the average duration of infective class I_t into the removed class R_t .
- In fact, we can estimate γ for real diseases by observing infected individuals and determining the mean infectious period $1/\gamma$ first.

- In summary, we have

$$R_0 = (\alpha S_0)\left(\frac{1}{\gamma}\right)$$

$$= \begin{pmatrix} \text{no. of new cases arising from one} \\ \text{infective per unit time} \end{pmatrix} \begin{pmatrix} \text{average duration} \\ \text{of infection} \end{pmatrix}.$$

- Thus, R_0 is interpreted as the average number of secondary infections that would be produced by an infective in a wholly susceptible population of size S_0 .



- Note that, from this point of view, the critical value of $R_0 = 1$ makes good biological sense.
- If $R_0 > 1$, then a primary case of disease induces more than one secondary case of the illness, the size of the infective class increases, and an epidemic results.
- If $R_0 = 1$, then a diseased individual produces only one new case of the disease, and no epidemic can occur; there can be no growth in the number of infectives.
- When $R_0 < 1$, the disease dies out.

Basic Reproductive Number (R_0)

- In a population

  if $R_0 > 1$: epidemic

  if $R_0 = 1$: endemic stage

  if $R_0 < 1$: sucessful control of infection

- If population is completely susceptible

  measles : $R_0 = 15\text{-}20$

  smallpox : $R_0 = 3 – 5$

  SARS: ???

Continuous model

- Note that so far we are using discrete time intervals (e.g. one day). Now if we let the time interval to be very small, say one second. Then ΔI is almost equal to the instinct change of I.
- Therefore, one may replace ΔI by dI/dt, which is the rate of change of I (also called the derivative of I). Similarly, we may replace ΔS and ΔR by dS/dt and dR/dt respectively.
- With these notations, our system of equations can be replaced by the following system of ordinary differential equations.

A system of three ordinary differential equations describes the SIR model:

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

$$\frac{dI}{dt} = \alpha I(t)S(t) - \gamma I(t)$$

$$\frac{dR}{dt} = \gamma I(t)$$

where is α the infection rate and γ the removal rate of infectives.

- Given such a system of differential equations, one would like to solve it (i.e. find functions S,I and R which satisfy the equations).
- In general, this system of differential equations does not have any closed form solution. However, for given α and γ, we can solve the systems of differential equations by using some mathematical software like Mathlab.
- In general, the graphs of the S,I,R have the following shapes.

Graphs of S,I,R functions

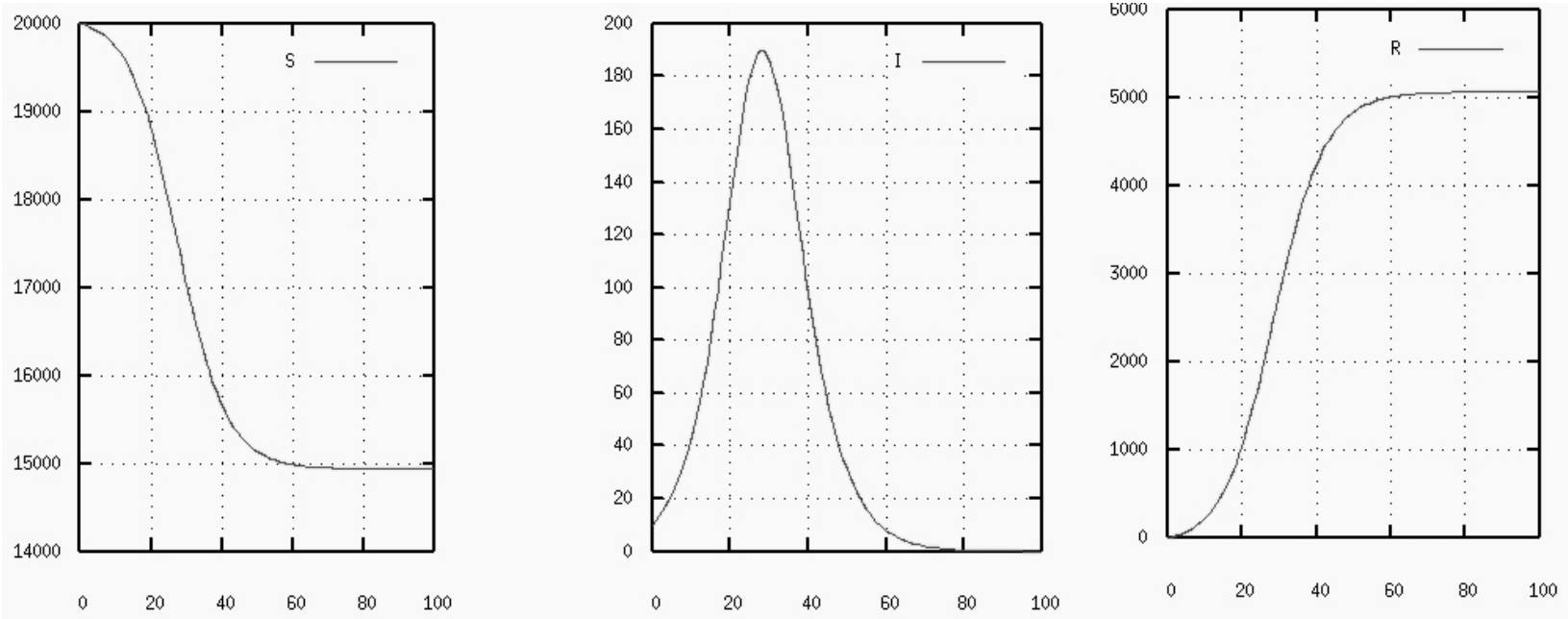


Figure 1: Typical dynamics for the SIR model.

Advertisement

- To learn more about differential calculus, you may take the following courses.

  MATH0801-Basic Mathematics I

  MATH0802-Basic Mathematics II

  MATH0803-Basic Mathematics III

- To learn more about differential equations, you may take MATH2405-Differential Equations
- To know more about how mathematics can be applied in biology, you may take MATH0011-Numbers and Patterns in Nature and Life.

Case study: The Hong Kong's SARS Outbreak in 2003

- Since November 2002 (and perhaps earlier) an outbreak of a very contagious atypical pneumonia (now named Severe Acute Respiratory Syndrome,SARS) initiated in the Guangdong Province of China.
- This outbreak started a world-wide epidemic after a medical doctor from Guangzhou infected several persons at a hotel in Hong Kong around February 21st, 2003.
- At the very beginning of SARS outbreak, SARS was believed to be a disease transmitted by respiratory droplets through close person-to-person contact.

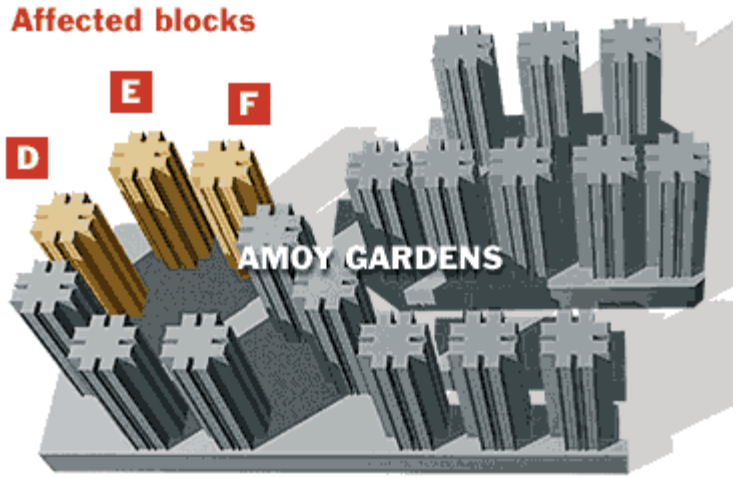
- Respiratory droplets are relatively large-sized particles and thus cannot travel long distances through air and therefore this mode of transmission cannot account for the rapid and wide spread of disease at Amoy Gardens, a housing estate in the Kowloon district of Hong Kong.

- At that time, one may therefore ask if the mode of transmission is airborne.

- It is well-known that influenza is an airborne disease.
- In the 4th March 1978 issue of the British medical Journal, 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. The SIR model was applied by J.D. Murray, to study the spread the flu epidemic in this school and he found that $\alpha = 0.00218$ and $\gamma = 0.440$ and hence $R_0 = 3.78$.

Application of SIR model to the Amoy Garden outbreak

- At that time, in order to see whether SARS is likely to be an airborne disease, together my colleagues, Dr. W.K. Ching and Dr. S.K.Chung, we studied the spread of SARS epidemic at a high rising residential building (Block E of the Amoy Gardens) by applying the SIR model.



- Since from 26th March 2003 to 30th March 2003, most of these confirmed cases are from Block E of the Amoy Gardens, we shall assume all of them are actually from Block E.

  - There are 33 floors, 8 flats each floor in Block E. Hence there are about 792 residents living in the building (we assume here that there are 3 people living in each flat and there are totally $8 \times 33 = 264$ flats in Block E). Therefore, we set $S(0) = 792 - I(0)$. $I(0)$ is unknown at that time.

- It is believed that one infected resident (a super spreader) initiated the epidemic when he visited and stayed with his brother's family, so we assume that $0 < I(0) \leq 4$.
- Note that the infected number of residents I(t) at time t is unobservable (why?) and we assume R(t) is equal to the cumulative number of residents in Block E confirmed with SARS symptoms at time t.

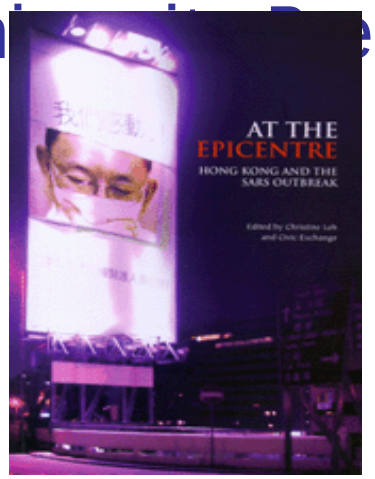
- These numbers of confirmed cases and predicted cases (by the SIR model) are then summarized in the table below.

Day	26/3/03	27/3/03	28/3/03	29/3/03	30/3/03
Confirmed Cases	7	22	56	78	112
Predicted Cases	13	27	47	77	117

- The key idea of estimating the model parameters is to choose the parameters α, γ and I(0) such that they minimize the square of errors between the observed data and the model predictions.
- Recall that we can solve the system of differential equations by standard mathematical software once we have fixed α, γ,S(0) and I(0). In particular, we can find the R(t) curve and compare it with the observed R values.
- In this way, we find that when α= 0.001875 and γ= 0.975 and I(0)=4 (and hence R_0=1.62), the error between the predicted and the observed values of R is minimum.

- For SARS, in the Block E of Amoy Gardens scenario, we found $\alpha = 0.001875$ and $\gamma = 0.975$.
- For influenza, in the broadening school scenario, Murray found $\alpha = 0.00218$ and $\gamma = 0.44$.
- We note that the value of the infection parameter α in this case is very closed to that of SARS at the Amoy Gardens.
- This suggests that both epidemics may have similar ways of transmission.
- At that time, a report of this finding was sent to the WHO.

- The mystery of Amoy Gadrens remains unsolved even though there are several completing theories.
- Therefore, the mode of transmission is still not very clear. SARS appears to be transmitted mainly by person-to-person contact. However, it could also be transmitted by contaminated objects, air, or by other unknown ways.
- See the recent book “At the Epicentre”, published by Hong Kong Un[illegible]ess.



Another puzzle

- The pattern of the SARS outbreak in Hong Kong was puzzling after a residential estate (Amoy Gardens) in Hong Kong was affected, with a huge number of patients infected by the virus causing SARS.
- In particular it appeared that underlying this highly focused outbreak there remained a more or less constant background infection level. This pattern is difficult to explained by the standard SIR epidemic model.

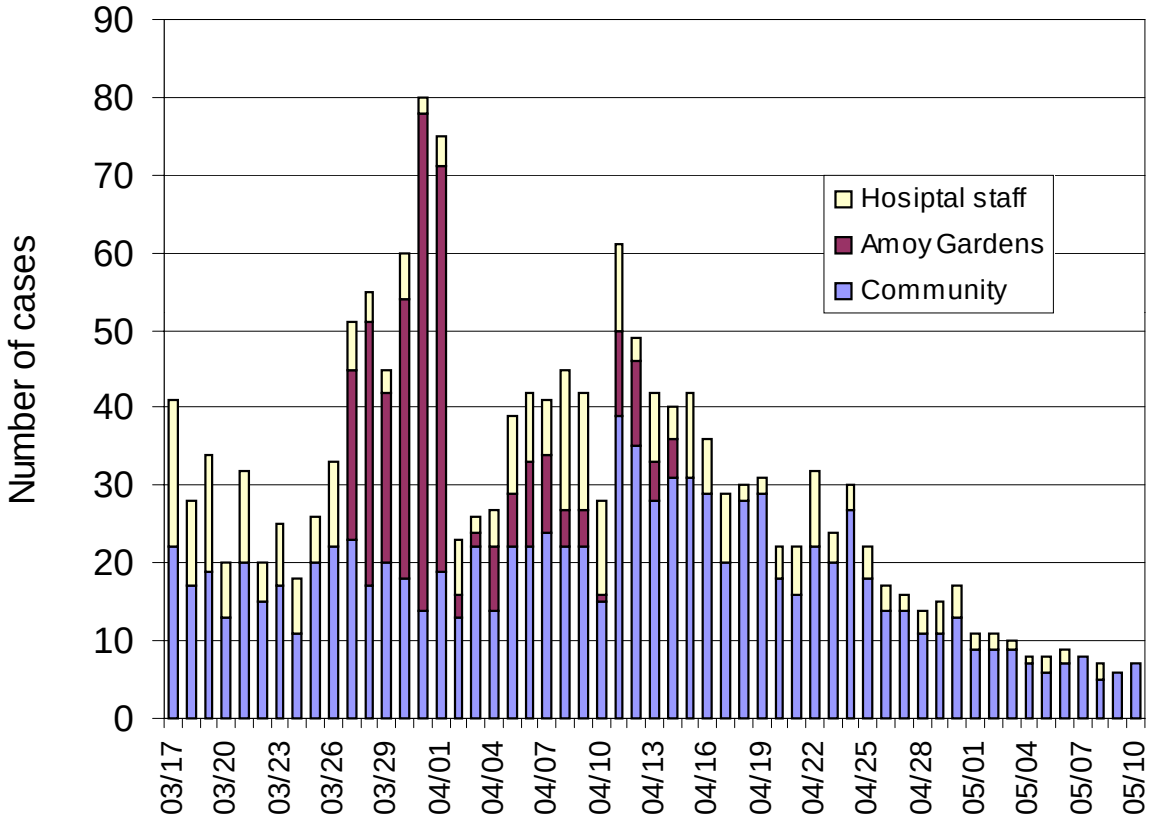


Figure 2: Daily new number of confirmed SARS cases from Hong Kong: hospital, community and the Amoy Gardens.

- This pattern is difficult to explain with the standard SIR model. Try to build another model.

Joint Work With

- Gabriel Turinici

  INRIA, Domaine de Voluceau, Rocquencourt, France

- Antoine Danchin

  Génétique des Génomes Bactériens,

  Institut Pasteur, Paris, France

  (Former director of HKU-Pasteur Institute)

A Double Epidemic Model for the SARS Propagation

- Published in *BMC Infectious Diseases* 2003, **3**:19 (10 September 2003)

- Can be found online at: http://www.biomedcentral.com/1471-2334/3/19

Hypothesis of the Double Epidemic Model for the SARS Propagation

- There are two epidemics, one is SARS caused by a coronavirus virus, call it virus A.
- Another epidemic, which may have appeared before SARS, is assumed to be extremely contagious because of the nature of the virus and of its relative innocuousness, could be propagated by contaminated food and soiled surfaces. It could be caused by some coronavirus, call it virus B. The most likely is that it would cause gastro-enteritis.

Hypothesis of the Double Epidemic Model for the SARS Propagation

- The most likely origin of virus A is a more or less complicated mutation or recombination event from virus B.
- Both epidemics would spread in parallel, and it can be expected that the epidemic caused by virus B which is rather innocuous, protects against SARS (so that naïve regions, not protected by the epidemic B can get SARS large outbreaks).

Motivation of a Double Epidemic Model for the SARS Propagation

- Learning from a set of coronavirus mediated epidemics happened in Europe that affected pigs in the 1983-1985.
- It is known at that time that a virus and its variant caused a double epidemic when the virus changed its tropism from the small intestine to lung.

Motivation of a Double Epidemic Model for the SARS Propagation

- This in a way allowing the first one to provide some protection to part of the exposed population.
- D Rasschaert, M Duarte, H Laude: **Porcine respiratory coronavirus differs from transmissible gastroenteritis virus by a few genomic deletions**. *J Gen Virol* 1990, **71 (Pt 11)**:2599-607.

A Double Epidemic Model for the SARS Propagation

- The hypothesis is based on:

A) the **high mutation** and **recombination rate** of coronaviruses.

SR Compton, SW Barthold, AL Smith: **The cellular and molecular pathogenesis of coronaviruses**. *Lab Anim Sci* 1993, **43**:15-28.

A Double Epidemic Model for the SARS Propagation

B) the observation that **tissue tropism** can be changed by **simple mutations** .

BJ Haijema, H Volders, PJ Rottier: **Switching species tropism: an effective way to manipulate the feline coronavirus genome**. *J Virol* 2003, **77**:4528-38.

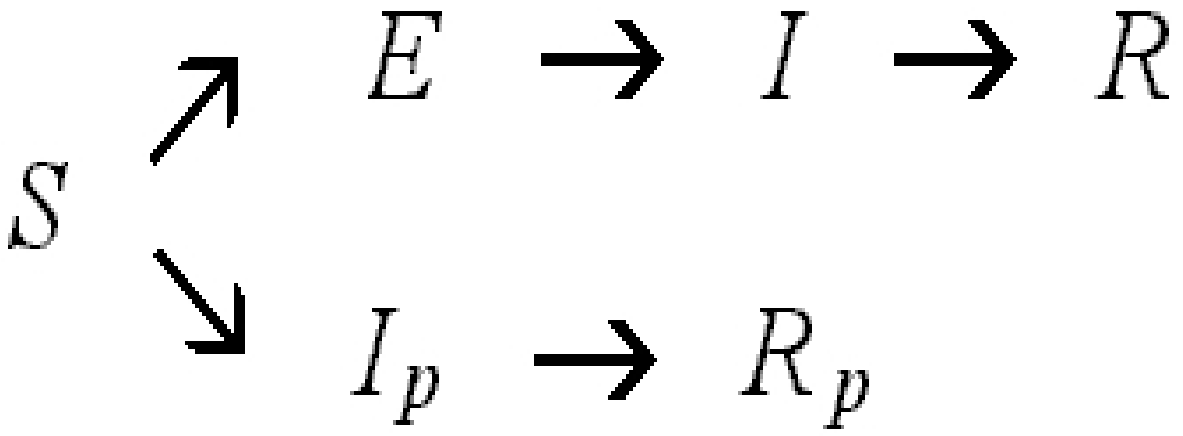
A Double Epidemic SEIRP Model

- Assume that two groups of infected individuals are introduced into a large population.
- One group is infected by virus A.
- The other group is infected by virus B.
- Assume both diseases which, after recovery, confers immunity (which includes deaths: dead individuals are still counted).
- Assumed that catching **disease B** first will protect the individual from **disease *A***.

A Double Epidemic SEIRP Model

- We divide the population into six groups:
  - Susceptible individuals, S(t)
  - Exposed individuals for virus A, E(t)
  - Infective individuals for virus A, I(t)
  - Recovered individuals for virus A, R(t)
  - Infective individuals for virus B, Ip(t)
  - Recovered individuals for virus B, Rp(t)

The progress of individuals is schematically described by the following diagram.



The system of ordinary differential equations describes the SEIRP model:

$$\frac{dS}{dt} = -rS(t)I(t) - r_p S(t)I_P(t) \quad (1)$$

$$\frac{dE}{dt} = rS(t)I(t) - bE(t) \quad (2)$$

$$\frac{dI}{dt} = bE(t) - aI(t) \quad (3)$$

$$\frac{dR}{dt} = aI(t) \quad (4)$$

$$\frac{dI_P}{dt} = r_P S(t)I_P(t) - a_P I_P(t) \quad (5)$$

$$\frac{dR_P}{dt} = a_P I_P(t) \quad (6)$$

Meaning of some parameters

- It can be shown that the fraction of people remaining in the exposed class E s time unit after entering class E is e^{-bs}, so the length of the latent period is distributed exponentially with mean equals to

$$\int_0^\infty e^{-bs}\,ds = 1/b$$

Meaning of some parameters

- It can be shown that the fraction of people remaining in the infective class I s time unit after entering class I is e^{-as}, so the length of the infectious period is distributed exponentially with mean equals to

$$\int_0^\infty e^{-as} ds = 1/a$$

Meaning of some parameters

The incubation period (the time from first infection to the appearances of symptoms) plus the onset to admission interval is equal to the sum of the latent period and the infectious period and is therefore equal to $1/b + 1/a$.

Empirical Statistics

- CA Donnelly, et al., **Epidemiological determinants of spread of causal agent of severe acute respiratory syndrome in Hong Kong**, *The Lancet*, 2003.
- The observed mean of the incubation period for SARS is 6.37.
- The observed mean of the time from onset to admission is about 3.75.
- Therefore, the estimated $1/a + 1/b$ has to be close to $6.37+3.75=10.12$.

Parameter Estimations

- Since we do not know how many Hong Kong people are infected by virus B, we shall consider the following two scenarios.
- Case a: Assume Ip(0)=0.5 million, S(0)=6.8-0.5=6.3 million,E(0)=100,I(0)=50.
- Case b: Assume Ip(0)=10, S(0)=6.8 million,E(0)=100,I(0)=50.

Parameter Estimations

- We fit the model with the total number of confirmed cases from 17 March, 2003 to 10 May, 2003 (totally 55 days).
- The parameters are obtained by the gradient-based optimization algorithm.
- The resulting curve for *R* fits very well with the observed total number of confirmed cases of SARS from the community.

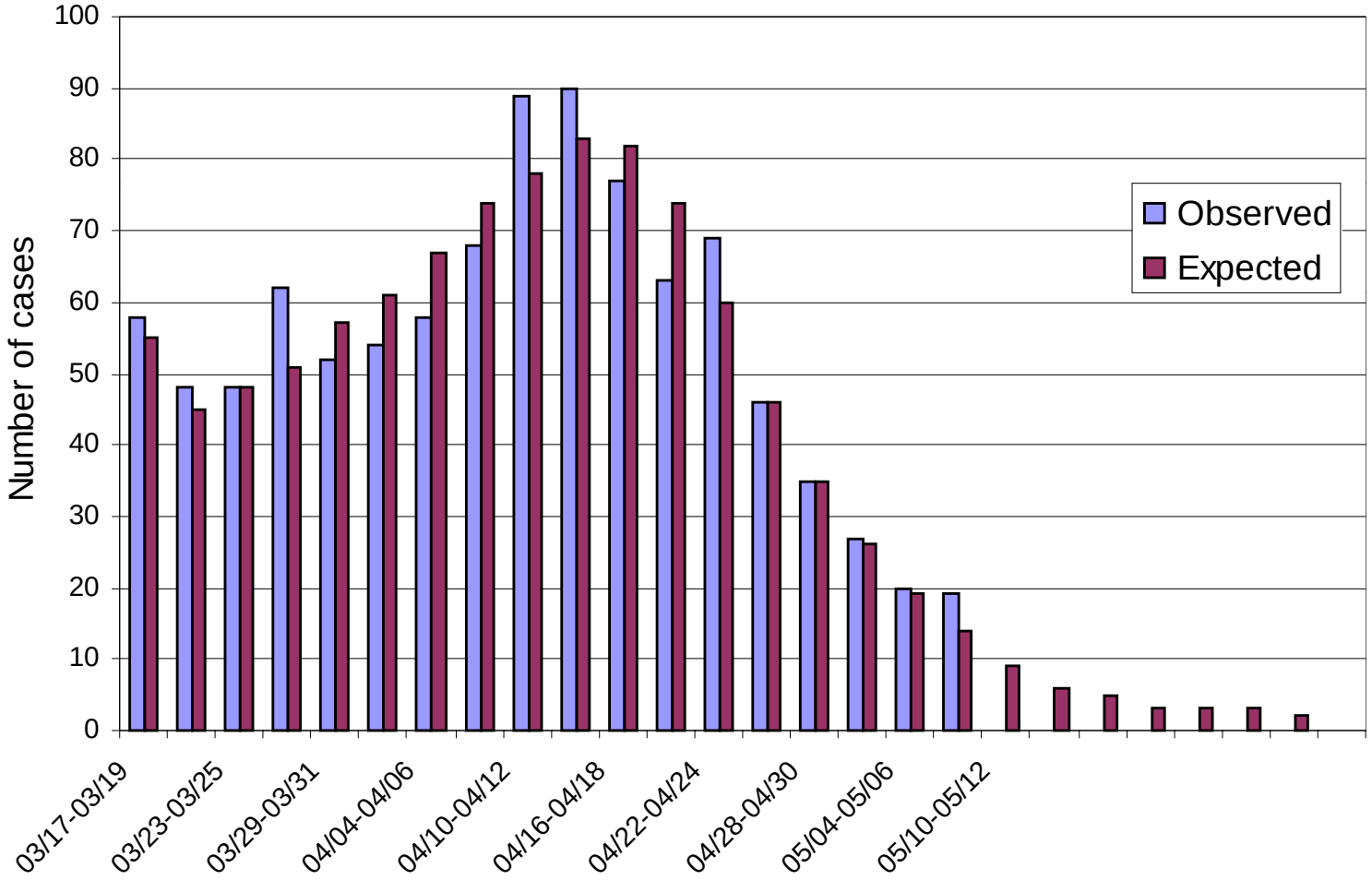


Figure 3: Number of SARS cases in Hong-Kong community (and the simulated case "a") per three days.

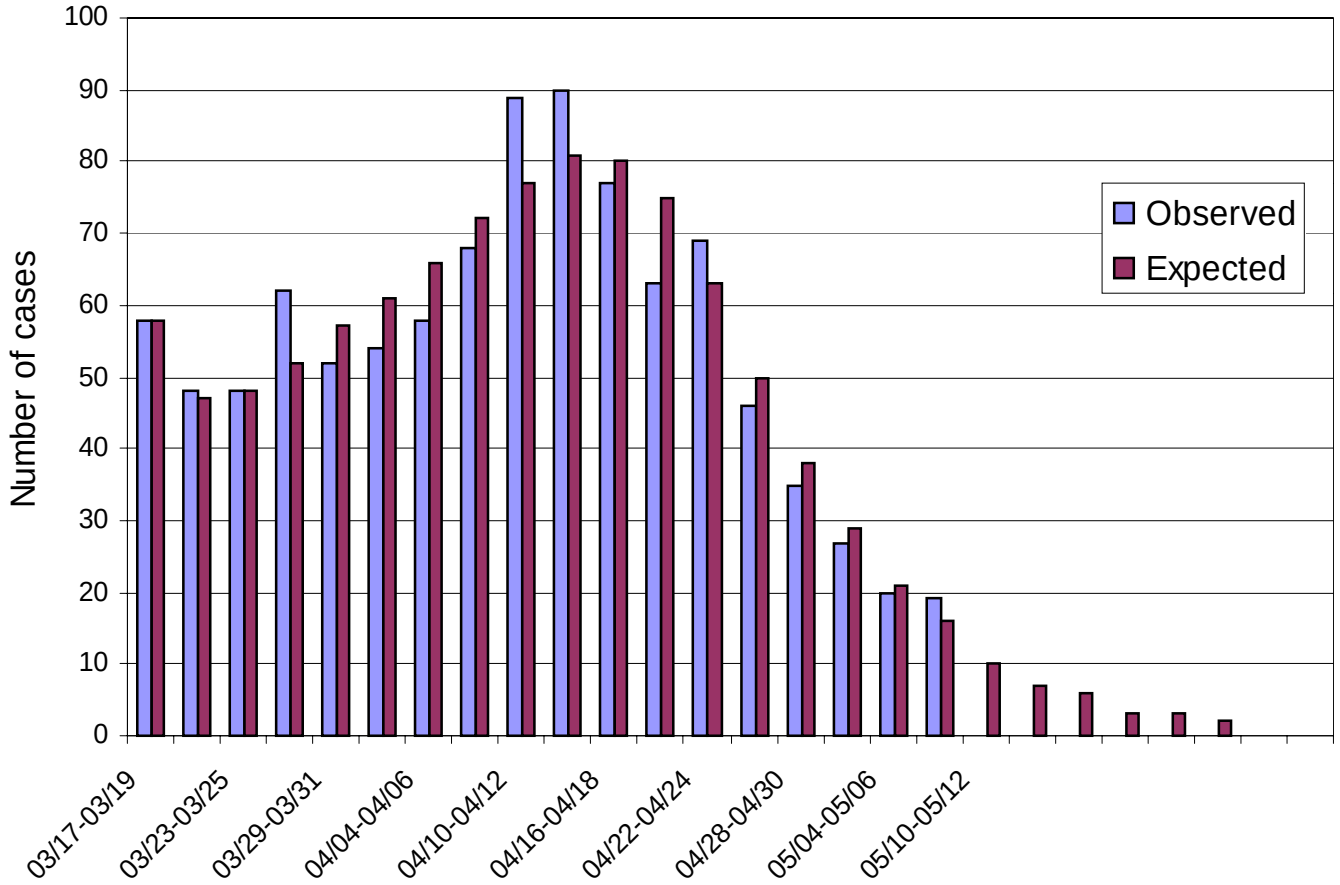


Figure 4: Number of SARS cases in Hong-Kong community (and the simulated case "b") per three days.

Parameter Estimations

- Case a: Assume Ip(0)=0.5 million, S(0)=6.3 million,E(0)=100,I(0)=50.

- $r=10.19\times10^{-8}$, $r\_p=7.079\times10^{-8}$.
- a=0.47,a_p=0.461,b=0.103.
- Estimated 1/a + 1/b = 11.83 (quite close to the observed 1/a+1/b= 10.12).

Parameter Estimations

- Case b: Assume Ip(0)=10,
- S(0)=6.8 million,E(0)=100,I(0)=50.

- $r=10.08\times10^{-8}$, $r\_p=7.94\times10^{-8}$.
- a=0.52,a_p=0.12,b=0.105.
- Estimated 1/a + 1/b = 11.44 (quite close to the observed 1/a+1/b= 10.12).

Basic reproductive factor

- We define the basic reproductive factor R_0 as

$$R_0 = rS(0)/a.$$

- R_0 is the number of secondary infections produced by one primary infection in a whole susceptible population.
- Case a: $R_0 = 1.37$.
- Case b: $R_0 = 1.32$.

Conclusion

- We did not explore the intricacies of the mathematical solutions of this new epidemiological model, but, rather, tried to test with very crude hypotheses whether a new mode of transmission might account for surprising aspects of some epidemics.
- Unlike the SIR model, for the SEIRP model we cannot say that the epidemic is under control when the number of admission per day decreases.
- Indeed in the SEIRP models, it may happen that momentarily the number of people in the Infective class is low while the Exposed class is still high (they have not yet been infectious);

- Thus the epidemic may seem stopped but will then be out of control again when in people in the Exposed class migrate to the Infected class and will start contaminating other people (especially if sanitary security policy has been relaxed). Thus an effective policy necessarily takes into account the time required for the Exposed (E) class to become infectious and will require zero new cases during all the period.
- The double epidemic can have a flat, extended peak and short tail compared to a single epidemic, and it may have more than one peak because of the latency so that claims of success may be premature.

- This model assumes that a mild epidemic protects against SARS would predict that a vaccine is possible, and may soon be created.
- It also suggests that there might exist a SARS precursor in a large reservoir, prompting for implementation of precautionary measures when the weather cools down.