DH301: Basic Epidemiology

Mathematical Epidemiology

(Lecture 5)

Sandip Mandal

Email: sandipccmb@gmail.com

Indian Institute of Technology - Bombay

30 October 2024

Concepts we'll cover

- **Boxes and arrows**
  - Building a simple compartmental model
  - Relation between a model diagram and its equations
- **Competing hazards**
  - Modelling different possible outcomes
- **Force of infection**
  - What makes an infectious disease model
  - The basic reproduction number (R_0)
- **Interventions**
  - Vaccination and treatment
- **More complex models**
- **Model calibration**
  - **Basic parameter estimation**

Discussion on some problems

Problem 1

In a city of population one lakh an epidemic occurred last year and the number of symptomatic cases were recorded as below.

1 Mar	6 Mar	11 Mar	16 Mar	21 Mar	26 Mar	31 Mar	5 Apr	10 Apr	15 Apr
1	38	3085	27441	15302	5572	1880	623	206	68

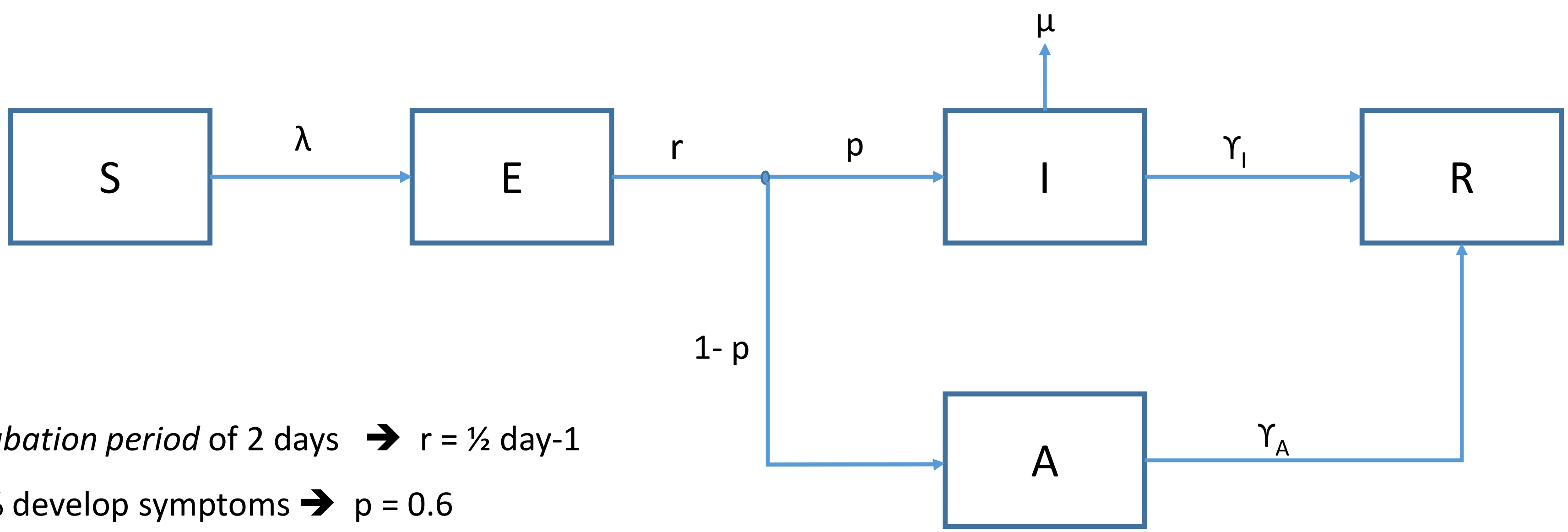
The following factors are known about the disease:

- Upon infection, there is an *incubation period* of 2 days on average (non-infectious)
- After the incubation period, 60% develop symptoms, the rest remain asymptomatic
- Asymptomatic infections are *half* as infectious as symptomatic infection
- 10% of the symptomatic individual die due to the disease
- Average duration of symptomatic infection is 5 days
- Average duration of asymptomatic infection is 3 days

Problem 1

1. Construct a model diagram and indicate all parameter values associated with the model
2. Write down the model equations
3. Write a code to solve the equations and plot all the state variables together.
4. What is the value of beta (transmission rate) that gives the best fit with the observed data?
5. Describe how the epidemic curve varies by varying transmission rate.
6. If asymptomatic infections are not infectious then what would be the change in disease pattern.

1. Model diagram and parameter values



Incubation period of 2 days ➔ r = ½ day-1

60% develop symptoms ➔ p = 0.6

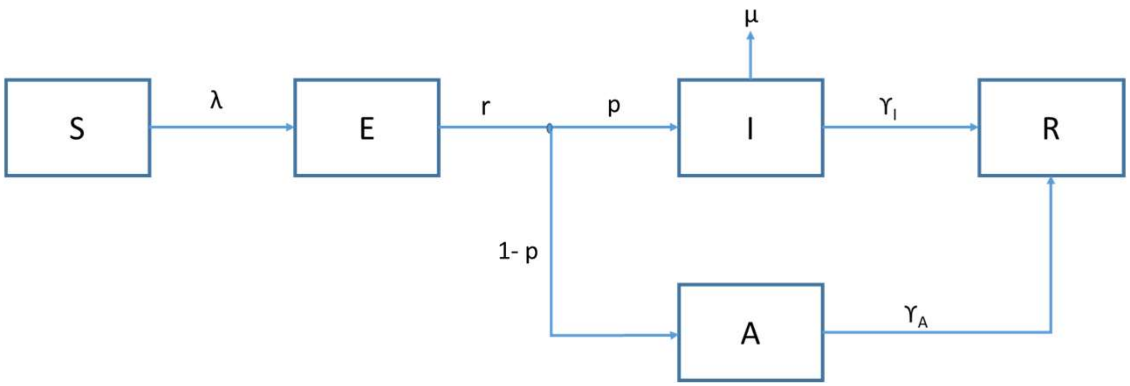
Average duration of symptomatic infection is 5 days ➔ Υ_I = 1/5 day-1

Average duration of asymptomatic infection is 3 days ➔ Υ_A = 1/3 day-1

10% of the symptomatic individual die due to the disease ➔ CFR = 0.1; $\frac{\mu}{\mu+\Upsilon_I} = 0.1; \Rightarrow$ μ = 0.0022

$\lambda = \beta * (I + 0.5 * A)/N$

2. Model equations



$$\lambda = \beta * (I + 0.5 * A)/N$$

$$\frac{dS}{dt} = -\lambda S$$

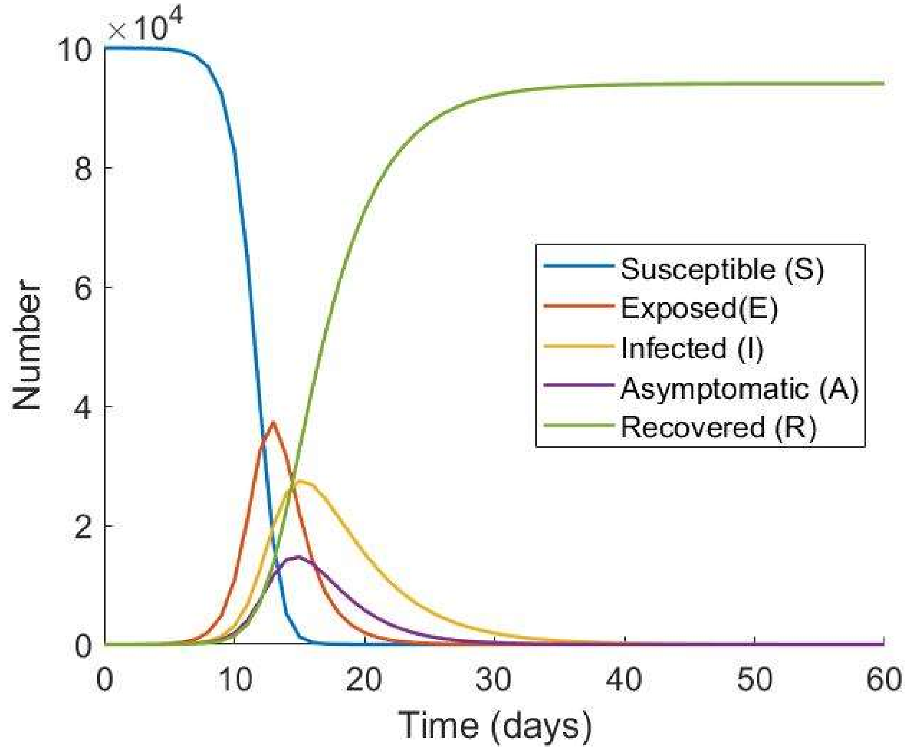
$$\frac{dE}{dt} = \lambda S - rE$$

$$\frac{dI}{dt} = p\, r\, E - \gamma_I I \;\; - \mu\, I$$

$$\frac{dA}{dt} = (1 - p)\, r\, E - \gamma_A A$$

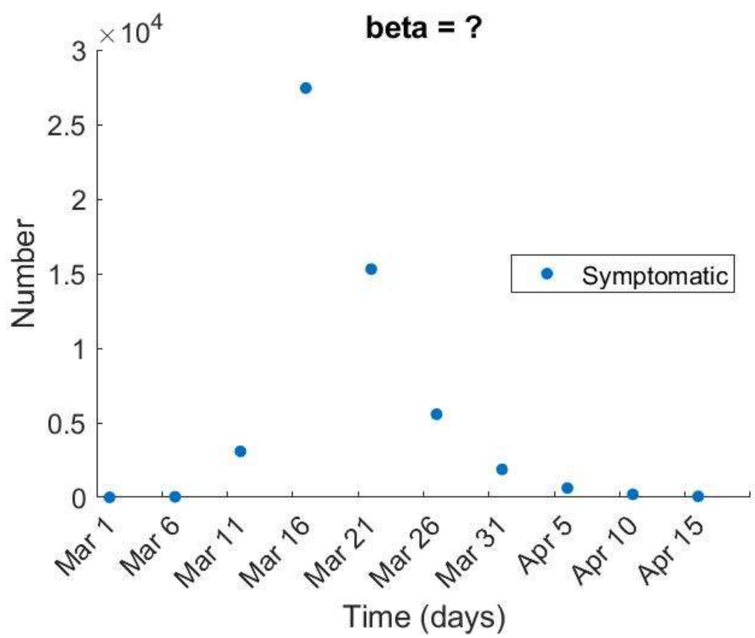
$$\frac{dR}{dt} = \gamma_I I + \gamma_A A$$

3.Write a code to solve the equations and plot all the state variables together.

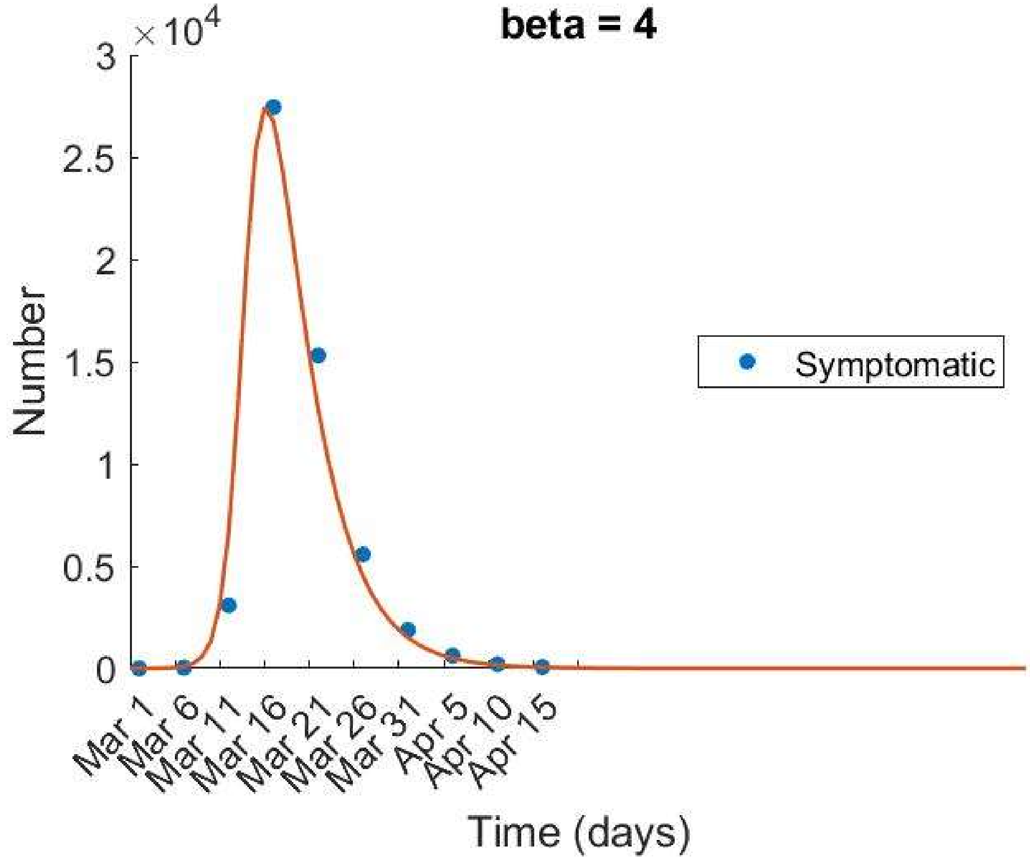


4. What is the value of beta (transmission rate) that gives the best fit with the observed data?

1 Mar	6 Mar	11 Mar	16 Mar	21 Mar	26 Mar	31 Mar	5 Apr	10 Apr	15 Apr
1	38	3085	27441	15302	5572	1880	623	206	68

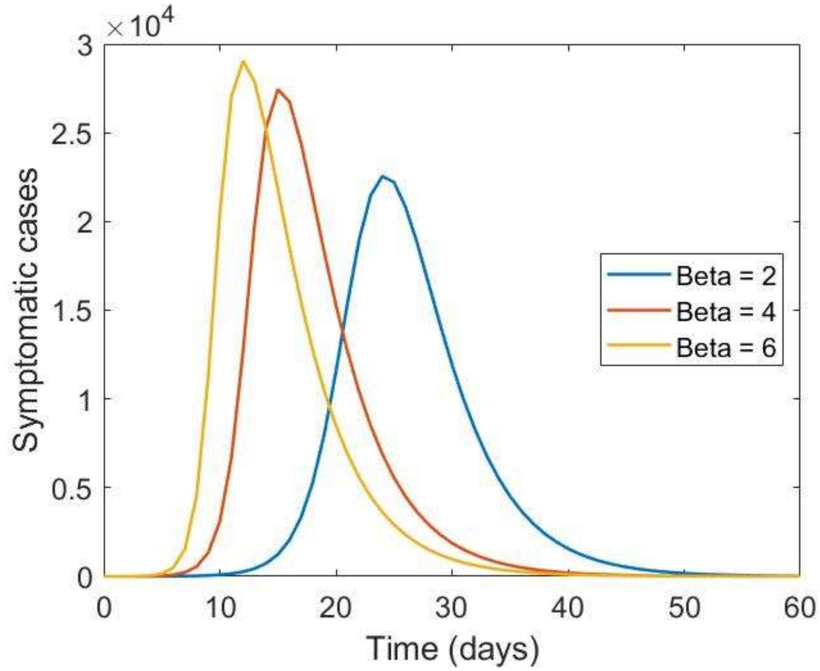


Fitting with data



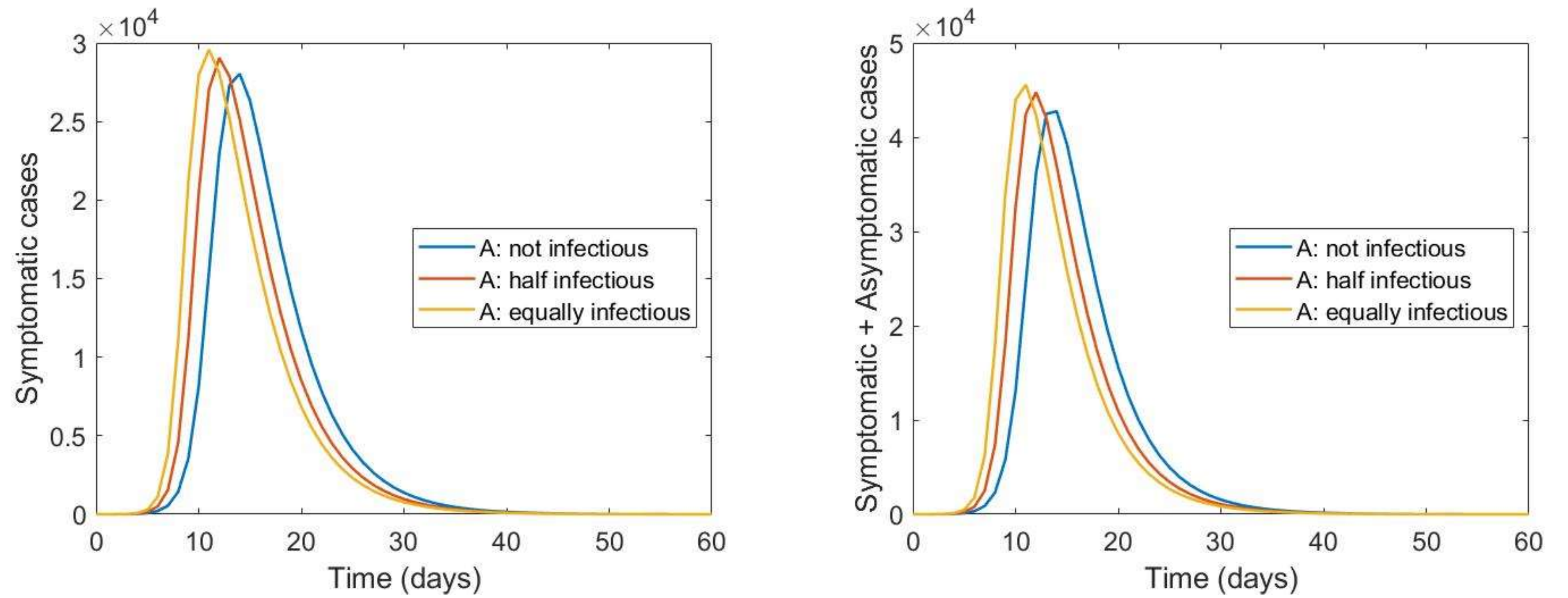
Manual fitting

5. Describe how the epidemic curve varies by varying transmission rate.



- For low transmission rate peak height is lower
- To end epidemic, it takes longer time for low beta value
- For high beta, peak appears early

6. If asymptomatic infections are not infectious then what would be the change in disease pattern.



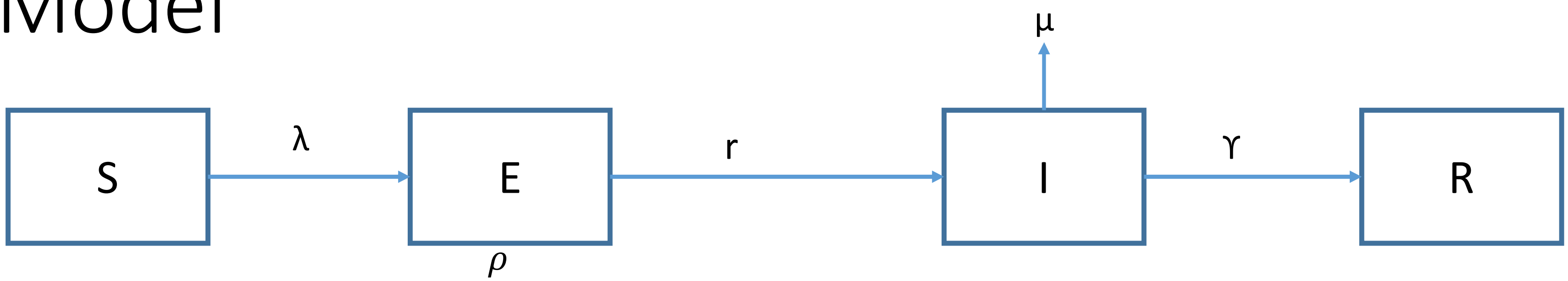
Problem 2

Develop a simple epidemic model (SEIR) to see the impact of quarantine of infectious individuals using the following assumptions.

- Average duration of incubation period and the infectious period are 3 days
- Case fatality rate = 2% (dies only during symptomatic infection)
- During incubation period, individuals are half infectious than the symptomatic individuals.
- Basic reproduction number R0 = 2

1. Show the percentage reduction of peak prevalence by isolating symptomatic individuals from 0% - 50% in average 4 days time period.
2. Do the same as above by isolating symptomatic individual in average 2 days and 3 days and describe the difference.

Model



Average duration of incubation period and the infectious period are 3 days ➔ r = 1/3 day-1 and ϒ = 1/3 day-1

Case fatality rate = 2% (dies only during symptomatic infection) ➔ CFR = 0.02; $\frac{\mu}{\mu+\Upsilon} = 0.02;$ ⇒ μ = 0.0068

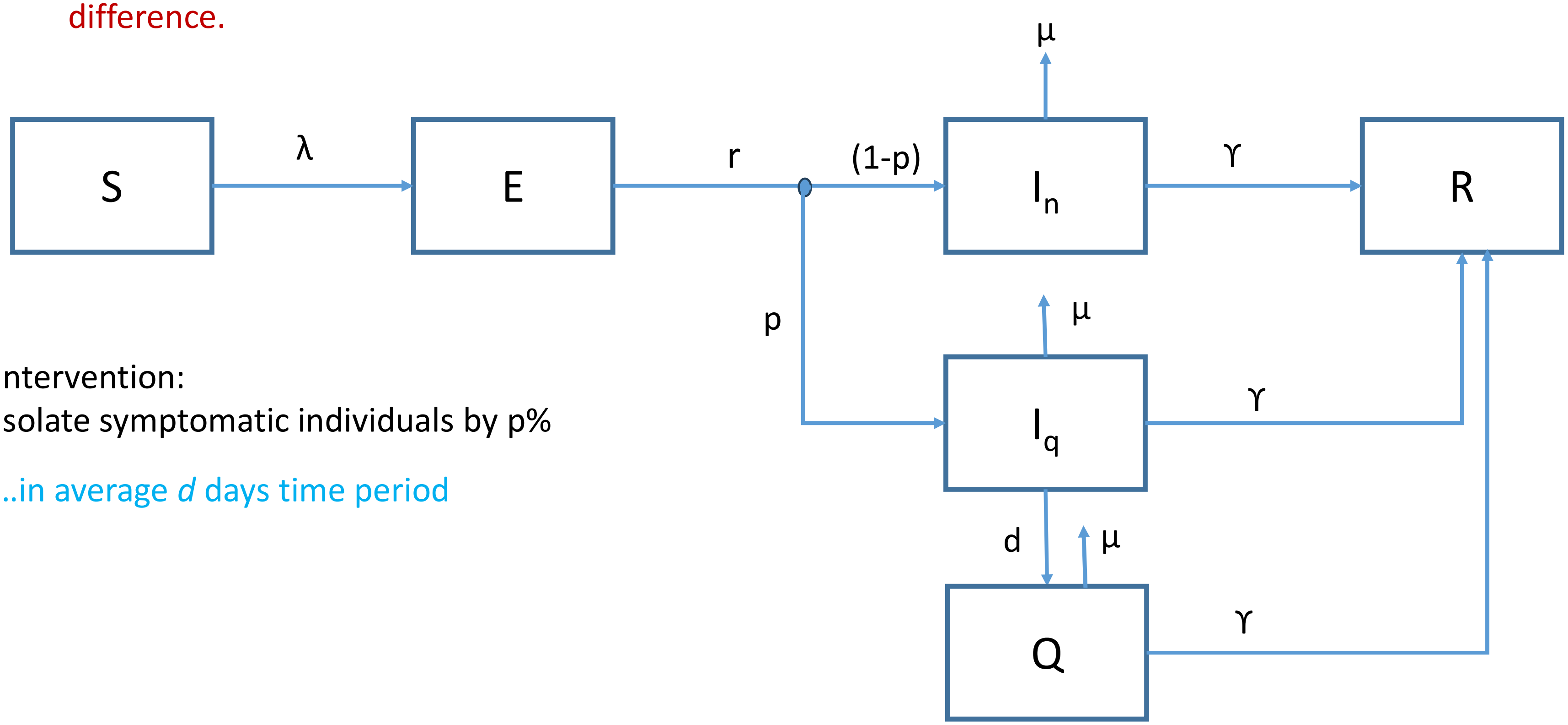
During incubation period, individuals are half infectious than the symptomatic individuals

Infectiousness of E, relative to I compartment ρ = 0.5; λ = β *(I + 0.5 * E)/N

Basic reproduction number R0 = 2;

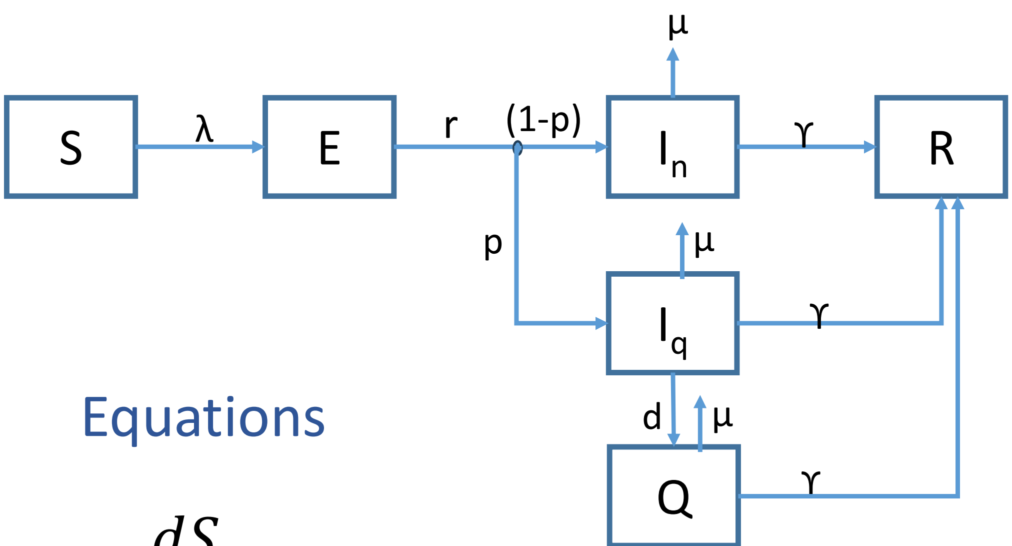
β = R0*Average recovery rate = R0/Effective infectious duration $\beta = \frac{R0}{\left(\frac{\rho}{r} + \frac{1}{\gamma + \mu}\right)}$ ⇒ β = 0.45

1. Show the percentage reduction of peak prevalence by isolating symptomatic individuals from 0% - 50% in average 4 days time period.
2. Do the same as above by isolating symptomatic individual in average 2 days and 3 days and describe the difference.



Intervention:
Isolate symptomatic individuals by p%

…in average *d* days time period



Equations

$$\frac{dS}{dt} = -\lambda S$$

$$\frac{dE}{dt} = \lambda S - rE$$

$$\frac{dI_n}{dt} = (1-p)\, r\, E - \gamma I_n - \mu I_n$$

$$\frac{dI_q}{dt} = p\, r\, E - \gamma I_q - \mathrm{d}\, I_q - \mu I_q$$

$$\frac{dQ}{dt} = \mathrm{d}\, I_q - \gamma Q - \mu Q$$

$$\frac{dR}{dt} = \gamma\, (I_n + I_q + \mathrm{Q})$$

Force of infection

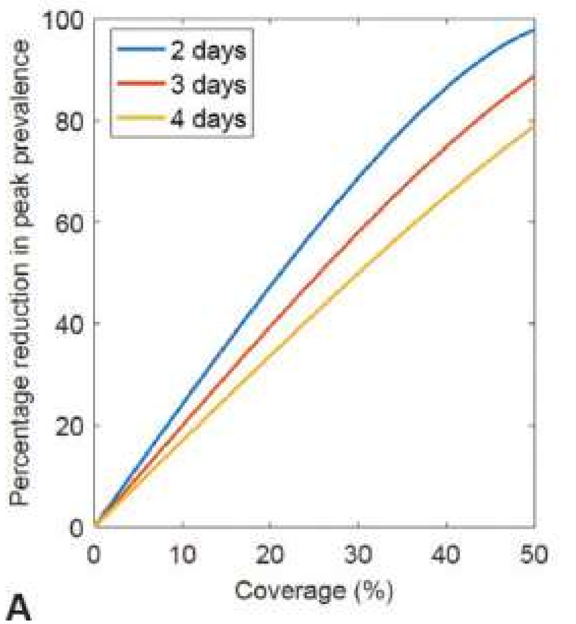
Before Isolating infected individuals

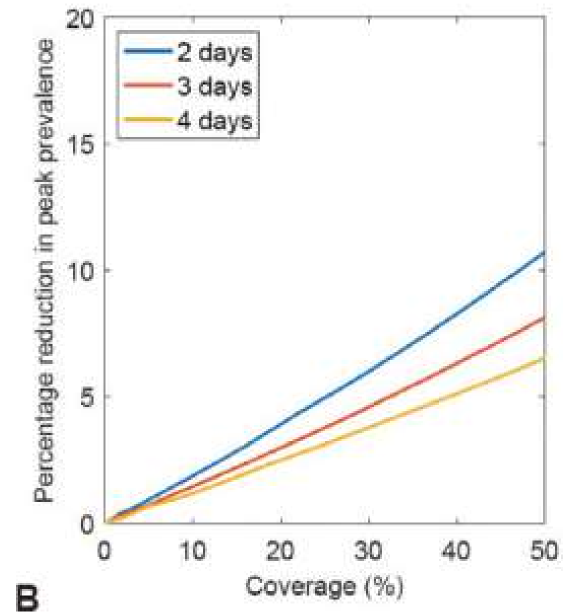
$\lambda = \beta * (0.5 * E + I)/N$

After isolating infected individuals

$\lambda = \beta * (0.5 * E + I_n + I_q)/N$

Example scenarios





Basic reproduction number R0 = 1.5

Exposed are non-infectious

Basic reproduction number R0 = 4

Exposed are half infectious than the symptomatic individuals

Problem 3

Construct a simple vaccination model assuming disease transmission rate =0.25 day-1 and average recovery period of 10 days. If the vaccine is leaky and efficacy is 70%, what proportion of the population would have to be vaccinated to prevent an epidemic over the next 2 years. Write a code and check your analytical result numerically.

Vaccination (perfectly effective)

Using the formula of Herd Immunity Threshold (HIT)

$$\boldsymbol{HIT} = \left(1 - \frac{1}{R_0}\right)$$

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

$$R_0 = \frac{0.25}{0.1} = 2.5$$

$$\boldsymbol{HIT} = \left(1 - \frac{1}{2.5}\right) = 0.6$$

We need a vaccine coverage of 60% to prevent the epidemic

Modelling a leaky vaccine

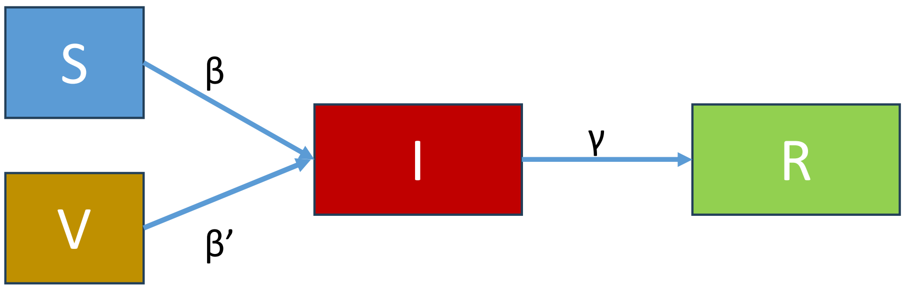
In the previous example, we need a vaccinate 60% to prevent the epidemic when the efficacy of the vaccine is 100%

$$\boldsymbol{HIT = \left(1 - \frac{1}{2.5}\right) = 0.6}$$

When the efficacy is less than 100%, we need to cover more people to prevent the epidemic. In this case the efficacy of the vaccine is 70%. Therefore,

$$\boldsymbol{HIT = \frac{0.6}{0.7} = 0.857}$$

Vaccination: model diagram



Total population S+V+I+R = N

Vaccine Coverage = *p* proportion of N

$$\frac{dS}{dt} = -\beta \ \frac{I}{N} S$$

$$\frac{dV}{dt} = -\beta' \ \frac{I}{N} V$$

$$\frac{dI}{dt} = \beta \ \frac{I}{N} S + \beta' \ \frac{I}{N} V - \gamma I$$

$$\frac{dR}{dt} = \gamma I$$

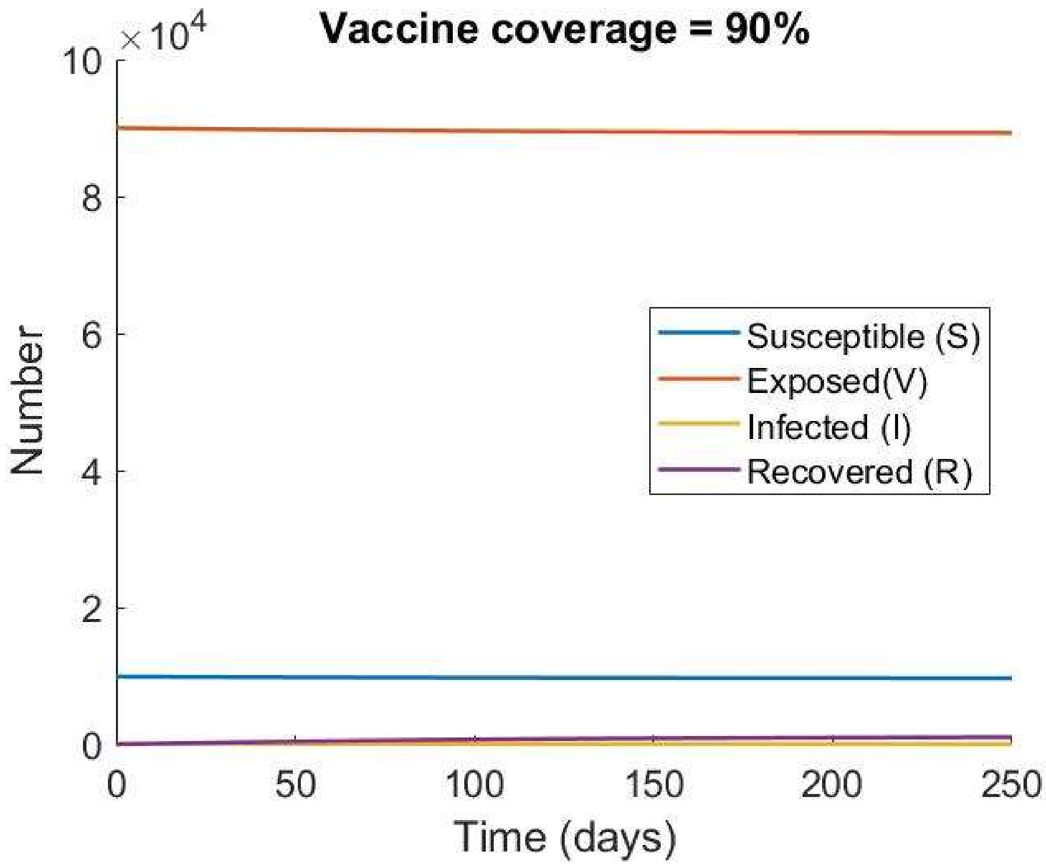
$$\beta' = \beta * Cs = \beta * (1 - \varepsilon)$$

ε = Vaccine efficacy

Initial condition,
before introduction
of the disease

S(0) = N*(1 - p)
V(0) = N*p
I(0) = 0
R(0) = 0

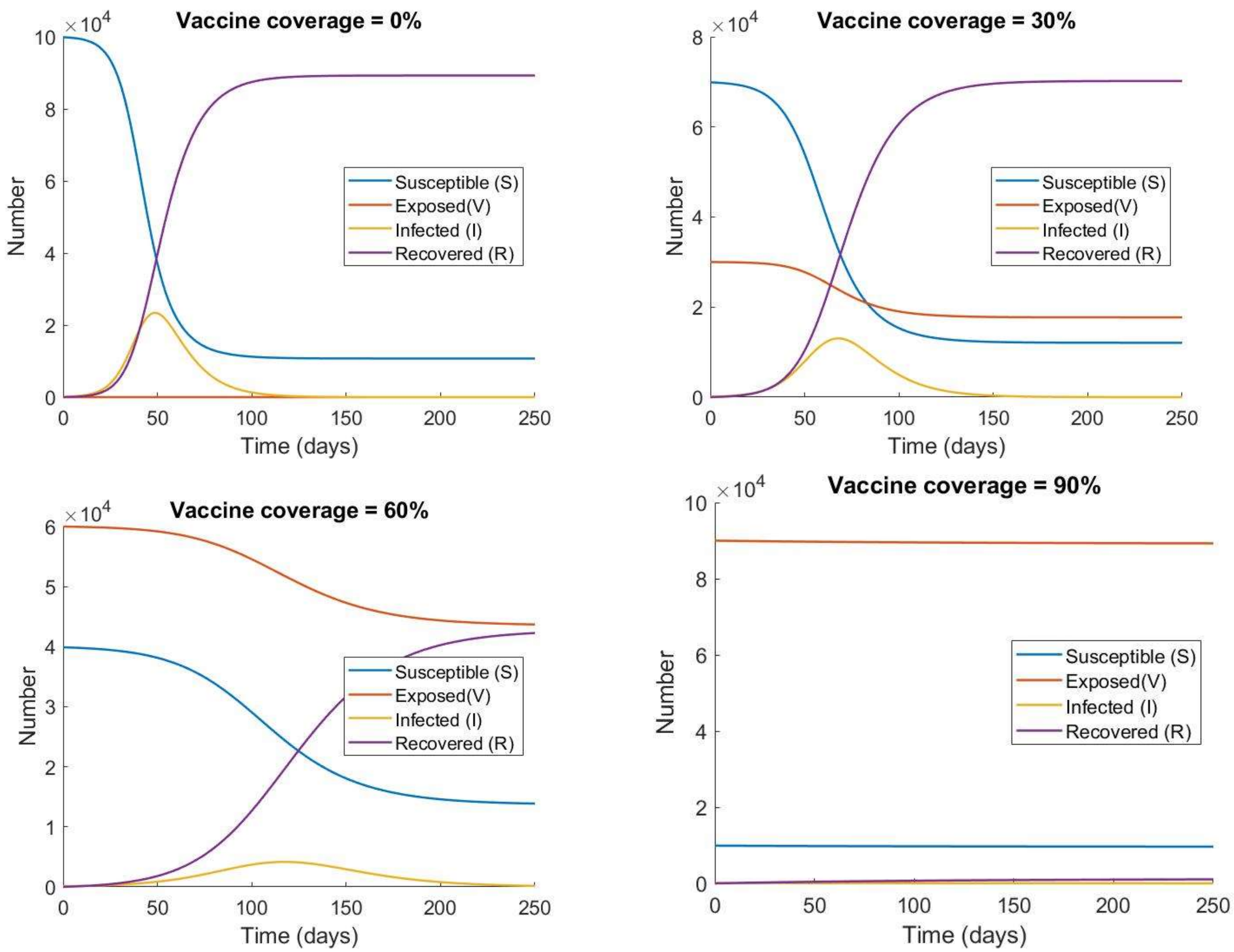
Estimation of threshold vaccine coverage



Theoretically, we estimated, heard immunity threshold (slide 20)

$$\boldsymbol{HIT = 0.857}$$

Estimation of threshold vaccine coverage



Theoretically, we estimated, heard immunity threshold (slide 20)

$$\boldsymbol{HIT = 0.857}$$

So far….

In all our previous examples, parameter values were given directly or indirectly

- For example, in the SIR model β and ϒ were given directly or,
- Value of R0 and average disease duration were given from which we could estimate all required parameter values

Challenges

In reality many parameters are unknown and difficult to estimate from the epidemiological observations

- For example, no direct method for knowing β
- For some diseases, it is difficult to know the duration between ‘onset of disease’ to ‘develop of symptom’

Model calibration

Revisiting problem 1

In a city of population one lakh an epidemic occurred last year and the number of symptomatic cases were recorded as below.

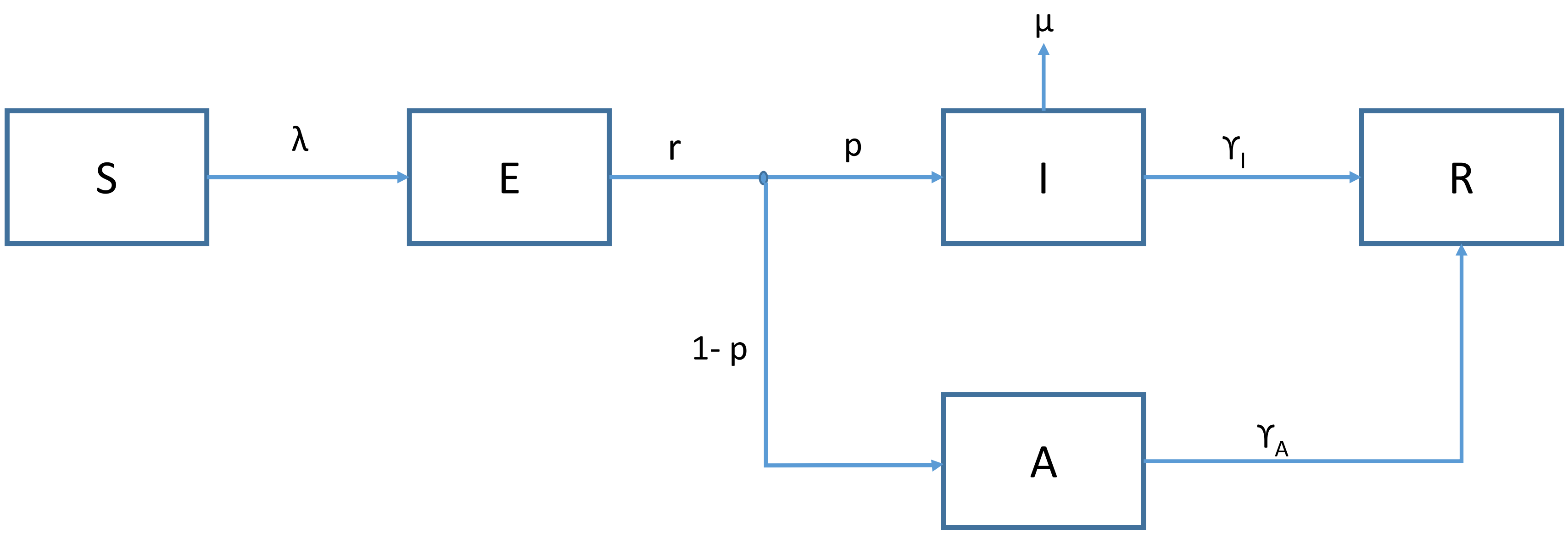
1 Mar	6 Mar	11 Mar	16 Mar	21 Mar	26 Mar	31 Mar	5 Apr	10 Apr	15 Apr
1	38	3085	27441	15302	5572	1880	623	206	68

The following factors are known about the disease:

- Upon infection, there is an *incubation period* of 2 days on average (non-infectious)
- After the incubation period, 60% develop symptoms, the rest remain asymptomatic
- Asymptomatic infections are *half* as infectious as symptomatic infection
- 10% of the symptomatic individual die due to the disease
- Average duration of symptomatic infection is 5 days
- Average duration of asymptomatic infection is 3 days

What is the value of beta (transmission rate) that gives the best fit with the observed data?

Model

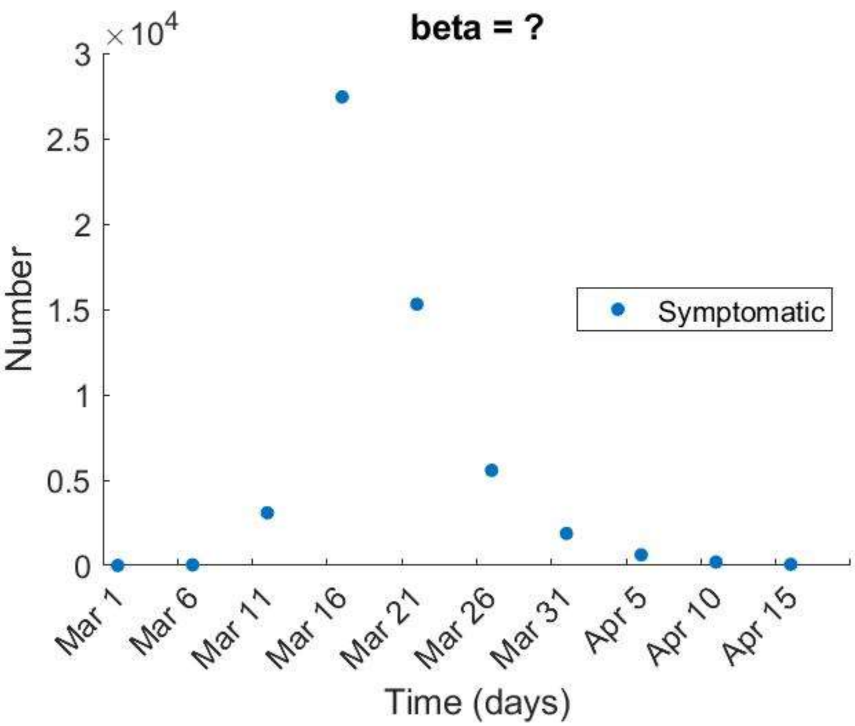


Incubation period of 2 days ➔ r = ½ day-1
60% develop symptoms ➔ p = 0.6
Average duration of symptomatic infection is 5 days ➔ Υ_I = 1/5 day-1
Average duration of asymptomatic infection is 3 days ➔ Υ_A = 1/3 day-1
10% of the symptomatic individual die due to the disease ➔ CFR = 0.1; $\frac{\mu}{\mu+\Upsilon_I} = 0.1; \Rightarrow \mu = 0.0022$

$\lambda = \beta *(I + 0.5 * A)/N$

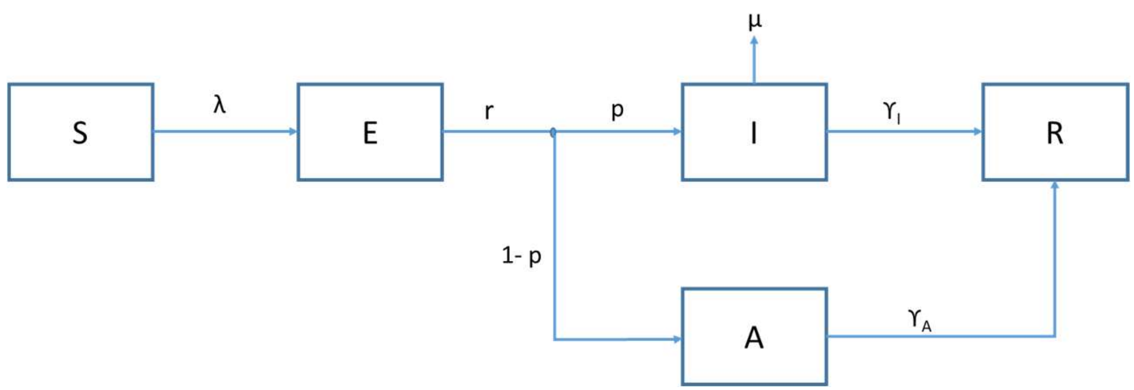
Observations

1 Mar	6 Mar	11 Mar	16 Mar	21 Mar	26 Mar	31 Mar	5 Apr	10 Apr	15 Apr
1	38	3085	27441	15302	5572	1880	623	206	68



What is the value of beta (transmission rate) that gives the best fit with the observed data?

Model equations



$$\frac{dS}{dt} = -\lambda S$$

$$\frac{dE}{dt} = \lambda S - rE$$

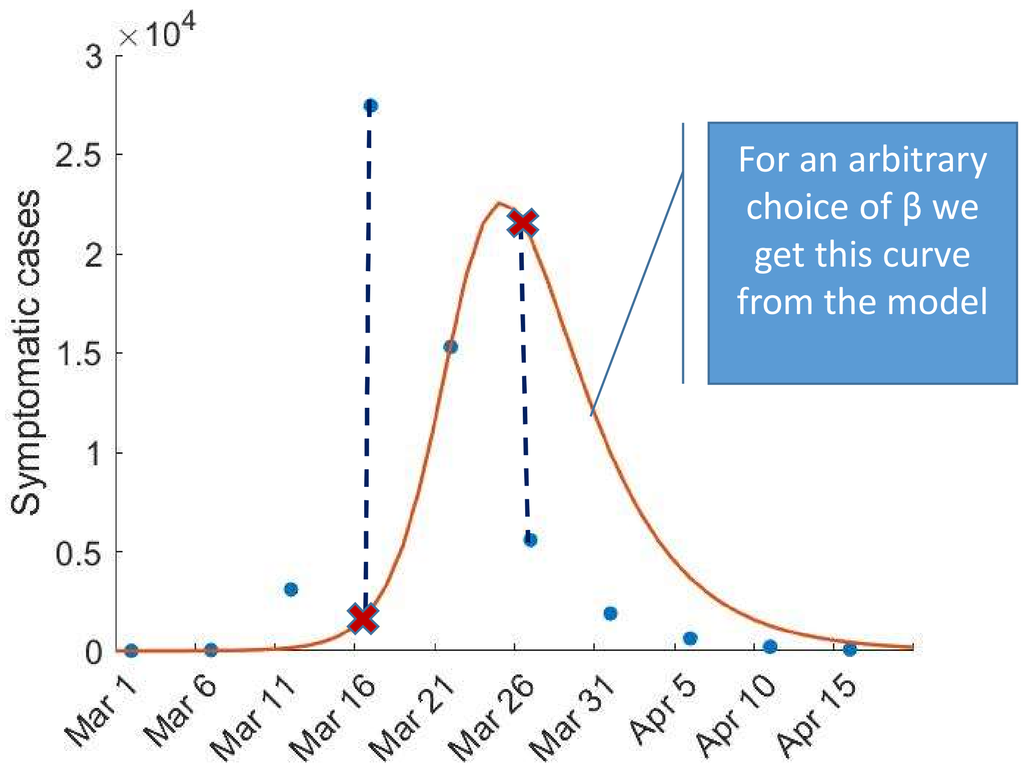
$$\frac{dI}{dt} = p\, r\, E - \gamma_I I \ - \mu\, I$$

$$\frac{dA}{dt} = (1 - p)\, r\, E - \gamma_A A$$

$$\frac{dR}{dt} = \gamma_I I + \gamma_A A$$

λ = β *(I + 0.5 * *A*)/N

Develop an algorithm to find the value of β by computer simulation



To measure how far the model result is from the data

Calculate Residuals

Residual (z_i) = Observed Value (y_i) - Predicted Value (m_i)

- Depending on the predicted values this could be either positive or negative.
- For the best-fit, model projected values and observed values are the same, means Residual = 0

For optimization of parameters, define a function

Sum of Squares of Residuals

Obj (β) $= \sum_i (y_i - m_i)^2$

Find the value of β for which this objective function is minimum.

This algorithm can be used to estimate multiple parameters together

Obj (β, ϒ, …) $= \sum_i (y_i - m_i)^2$

Parameter estimation: least squares method

The least squares method is a widely used technique in statistics and optimization for estimating the parameters of a mathematical model by minimizing the sum of the squared differences between observed and predicted values. It is commonly used in regression analysis, curve fitting, and various scientific and engineering applications.

Steps

1.Define the Model: First, you need to have a mathematical model that describes the relationship between the independent variables (predictors) and the dependent variable (the variable you want to predict or estimate). The model should have parameters that you want to estimate.
Example: For linear regression, the model might be: $y = mx + b$, where m and b are the parameters to be estimated.
2.Collect Data: You need a dataset with observations of both the independent and dependent variables. This dataset is used to estimate the model parameters.
3.Calculate Residuals: For each data point in your dataset, calculate the difference between the observed value and the value predicted by the model using the current parameter estimates. These differences are called residuals.
Residual = Observed Value - Predicted Value

4. Minimize the Sum of Squares of Residuals: The objective is to find the parameter values that minimize the sum of the squares of these residuals. The sum of squared residuals is often referred to as the "objective function" or the "loss function."

- Objective Function = Σ(residual^2) over all data points

5. Optimization: To find the parameter values that minimize the objective function, you can use various optimization techniques, such as gradient descent, least squares solvers, or closed-form solutions (if applicable).

  - **Gradient Descent**: Iteratively update the parameter values in the direction that reduces the objective function.
  - **Closed-Form Solution**: In some cases, you can solve for the parameter estimates directly using mathematical equations.

6. Parameter Estimates: Once the optimization process converges, you obtain the estimated parameter values that best fit the data.

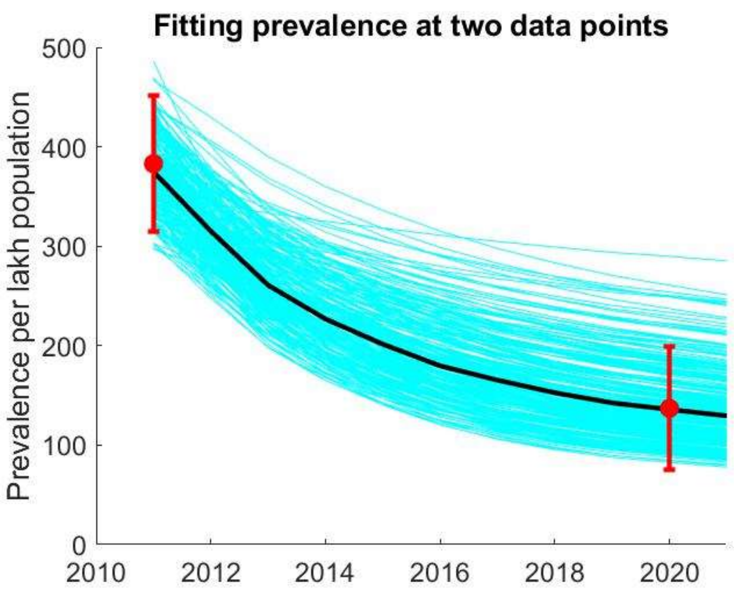
7. Evaluate Model Fit: It's essential to assess how well your model fits the data by looking at goodness-of-fit measures like R-squared, residual plots, and hypothesis tests for parameter significance.

Several built-in optimization functions are available

You can use built-in function to perform a least squares optimization to fit a model to your data.

- The **FindFit** function in Mathematica
- The **curve_fit function** in Python
- The **lsqcurvefit / fminsearch** in MATLAB

Model calibration & parameter estimation

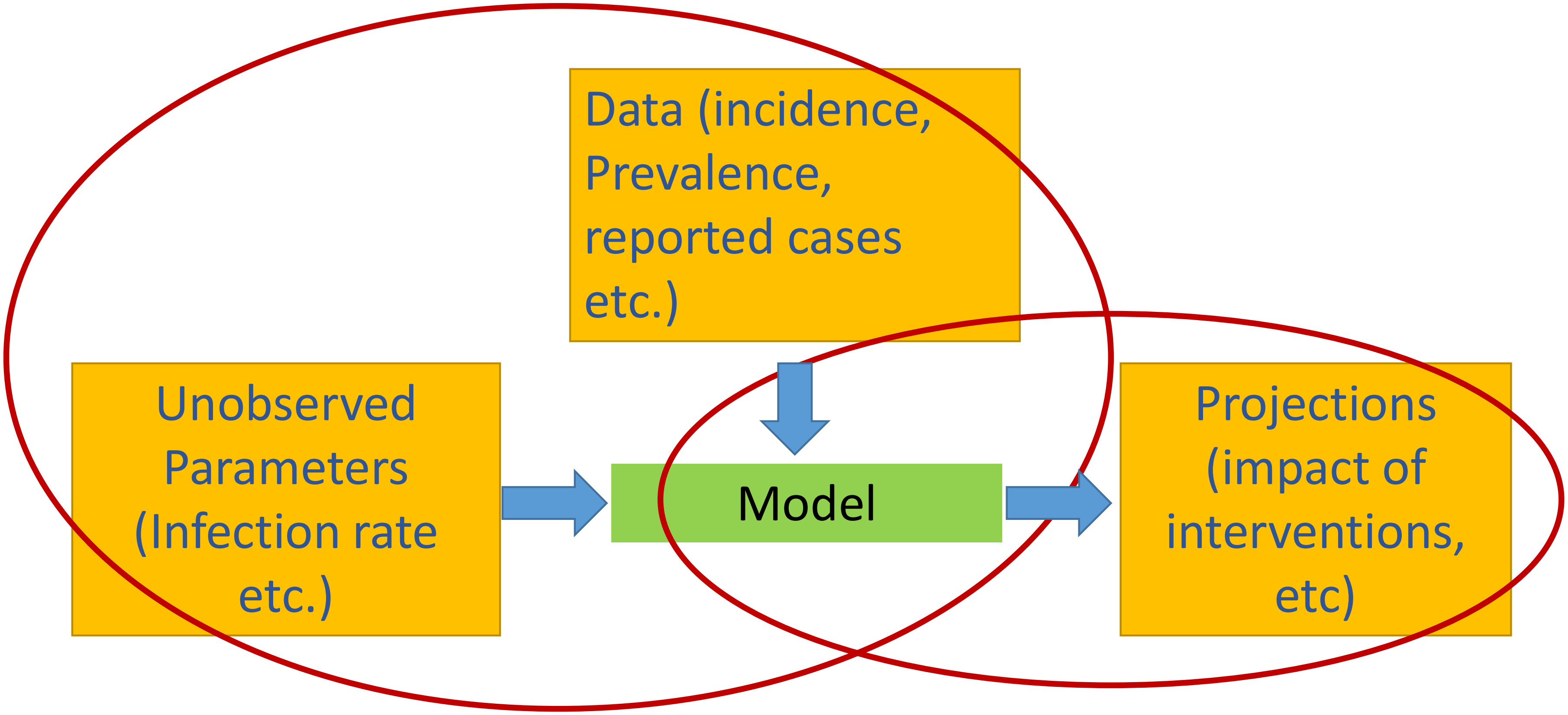


Important techniques

- Least square method
- Likelihood method
- Sensitivity analysis
- Propagating Uncertainty (Bayesian methods)

- Multiple data targets to meet
- Multiple parameters to estimate

Bayesian methods to propagate uncertainty



Thank you

Questions?