

Photo by [CDC](#) on [Unsplash](#)

Project 1 – Pandemic Flu Spread

Table of Content

1.0	Abstract	3
2.0	Background & Problem Description	4
2.1	Problem Statement	4
2.2	Literature Review	4
2.3	Report Structure	5
3.0	Main Findings	6
3.1	What is a Pandemic?	6
3.2	Pandemics Spread Models & EpiModel	6
3.2.1	Susceptible-Infectious (SI)	7
3.2.2	Susceptible-Infectious-Susceptible (SIS)	8
3.2.3	Susceptible-Infectious-Recovered (SIR)	8
3.2.4	Stochastic Version of Models	9
3.2.5	Using EpiModel for School Flu Question	9
3.3	Analysis for the Disease Spread Modeling	10
3.4	Solution Based on Deterministic Spread Models	11
3.4.1	Deterministic SI Model Solution	11
3.4.2	Deterministic SIR Model Solution	15
3.5	Solution Based on Stochastic Spread Models	18
3.5.1	Stochastic SI Model Solution	18
3.5.2	Stochastic SIR Model Solution	22
4.0	Conclusion	28
4.1	Learnings from the Output	28
4.2	Learnings from Using EpiModel Library in R	28
	Appendix	29

Table of Figures

Figure 1 - Diagram of EpiModel Design	10
Figure 2 - Code block for deterministic SI model.....	12
Figure 3 - Deterministic SI model parameters	12
Figure 4 - Plot of susceptible and infected over time for deterministic SI model	13
Figure 5 - Table of susceptible & infected over time for deterministic SI model	14
Figure 6 - Code block for deterministic SIR model.....	15
Figure 7 - Plot of disease spread over time for deterministic SIR model	16
Figure 8 - Table of susceptible, infected & recovered over time for deterministic SIR model.....	17
Figure 9 - Code block for stochastic SI model	18
Figure 10 - Plot of disease spread over time for stochastic SI model.....	19
Figure 11 - Plot of disease spread with lines for each simulation for Stochastic SI model.....	19
Figure 12 - Table of Avg values of susceptible & infected for stochastic SI model	20
Figure 13 - Distribution of day 1 infections for stochastic SI model.....	21
Figure 14 - Histogram of epidemic duration for stochastic SI model	22
Figure 15 - Code block for stochastic SIR model.....	23
Figure 16 - Plot of disease spread over time for stochastic SIR model.....	23
Figure 17 - Plot of disease spread with lines for each simulation for Stochastic SIR model	24
Figure 18 - Table of avg values of susceptible & infected for stochastic SIR model.....	25
Figure 19 - Distribution of day 1 infections for stochastic SIR model.....	26
Figure 20 - Histogram of epidemic duration for stochastic SIR model	27
Figure 21 - Summary statistics for all models.....	28

1.0 Abstract

The problem at hand is a study of how flu virus spreads in a small classroom of 21 students. 1 kid gets infected with flu and walks into the class on day 1, leaving remaining 20 kids susceptible to catching the virus. The infectious period is 3 days for each kid, starting the next day from contracting the virus. The probability of transmissions is 2% on each of the 3 infectious day (as i.i.d. Bern(p) trials) and is defined by the Binomial distribution. The solution revolves around finding out the infected kids on days 1 and 2 and distributions of infection on subsequent days. Another key aspect of the analysis is finding out the duration of epidemic and its distribution across many simulations.

A spread of disease in a population is typically well defined by Ordinary Differential Equations (ODE) and multiple mathematical models exist to define it (e.g. SI, SIS, SIR). 4 different models, 2 deterministic & 2 stochastic, were built using an R library *EpiModel*. One of the major findings from the modeling & analysis was that the number of infected kids on days 1 and 2 did not change much across all 4 models (1 and 2 kids respectively). Other major finding was around the duration of the epidemic, all 4 models showed that this epidemic will last around 9-11 days in the classroom before, either fizzling out (everybody getting infected or recovered) or stabilizing (with no more infections after a certain point).

2.0 Background & Problem Description

2.1 Problem Statement

This report documents the solution and accompanying modeling & analysis for the selected problem for project 1 of ISyE 6644 Simulation course. The selected problem (question #4) is related to the spread of a pandemic flu in a classroom of elementary students. Following are the fact laid out in the problem statement of this question.

- The elementary school kids classroom comprises of 21 students.
- 20 kids are healthy (and susceptible to the flu disease).
- 1 kid (Tommy) walks in with flu on day 1 (let's call whatever day he first joins the class with flu as day 1). He starts interacting with remaining 20 students straight away.
- Tommy comes to school every day i.e. total number of kids in class stays 21.
- Tommy stays infectious for 3 days i.e. there are 3 chances for Tommy to infect any of the remaining susceptible students.
- Probability of spreading flu to any susceptible kid on any of the 3 days is $p = 0.02$ (2%).
- All kids and days are independent & identically distributed (i.i.d.).
- The infection event on any of the 3 infectious days is a Bernoulli(p) trial defined by the Binomial distribution of i.i.d. Bernoulli events.
- If a susceptible kid gets infected by Tommy, he will become infectious for 3 days as well from the next day.

Expected solutions include:

1. Distribution of the number of kids that Tommy infects on Day 1.
2. Expected (average) number of kids that Tommy infects on Day 1.
3. Expected number of kids that are infected by Day 2 (you can count Tommy if you want).
4. Simulate the number of kids that are infected on Days 1, 2... so on.
 - a. Repeating the simulation many times.
 - b. Calculating estimated expected (average) numbers of kids that are infected by Day i , $i = 1, 2, \dots$ so on.
 - c. Producing a histogram of number of days, the "epidemic" will last.

It appears that questions 2 & 3 could be solved by deterministic models whereas all questions, especially questions 1 and 4 could be solved with stochastic modeling.

2.2 Literature Review

The goal of the literature review was twofold:

- Find the mathematical foundation to understand, build and solve disease spread models.
- Find a suitable programming library to build and solve these models.

Given comfort in R programming language, the research was focused on R libraries mainly. The research led to the following three (3) R packages that help solve the disease spread models (details of types of models for disease spread are in section 3.2):

- **EpiModel**¹ - an R library to build deterministic and stochastic disease spread models. Models are based on exponential models like SI, SIR and SIS and have options built in for deterministic compartmental models, stochastic individual contract models and network models.
- **SimInf**² - an R library to build stochastic simulations for disease spread using continuous time markov chain³ using Gillespie algorithm.
- **Epinet**⁴ - an R library to build models for disease spread using exponential graph models like SI, SIR and SIS. The simulations utilize the Bayesian approach and Markov Chain Monte Carlo algorithm.

The 3 libraries have several similarities like they use some form of exponential growth models and differential equations-based models like SI, SIS and SIR as well as same algorithm (like Monte Carlos and/or Gillespie) to build the simulations. I selected the **EpiModel** library for this project due to availability of both deterministic and stochastic approaches as well as ease of use.

2.3 Report Structure

The report is structured in a traditional format with the report starting with a title page followed by an Abstract (short summary of the problem and major findings). Next, a bit of background of the problem is provided supported by literature review of different R libraries (this section). The key content is in the 'Main Section' which includes:

- Typical classifications of disease spread
- A quick overview of different models used to measure & simulate disease spread
- Analysis of the problem parameters in preparation of modeling in R
- Two different solutions accompanied by code snippets and explanation.

The report is closed out with a 'Conclusion' section with commentary on the learnings from the solution output as well as using R library to build & simulate models is provided.

¹ Jenness SM, Goodreau SM and Morris M. EpiModel: An R Package for Mathematical Modeling of Infectious Disease over Networks. *Journal of Statistical Software*. 2018; 84(8): 1-47.

² Widgren S., Bauer P., Eriksson R. and Engblom S. SimInf: An R Package for Data-Driven Stochastic Disease Spread Simulations. *Journal of Statistical Software*. 2019; 91(12): 1-42.

³ Powell V. "Markov Chains Explained Visually." *Setosa*, <https://setosa.io/ev/markov-chains/>

⁴ Groendyke C. and Welch D. epinet: An R Package to Analyze Epidemics Spread across Contact Networks. *Journal of Statistical Software*. 2018; 83(11): 1-22.

3.0 Main Findings

3.1 What is a Pandemic?

There are typically 4 classifications⁵ of infectious disease spread. The key factors in determining the type of spread is the duration as well as the geographical parameters of the spread.

- **Outbreak** – is an occurrence of disease spread more than normal in a certain area and at a certain time.
- **Epidemic** – an outbreak in a certain area is classified as an epidemic.
- **Pandemic** – an epidemic becomes a pandemic when disease spreads to a wider area over time.
- **Endemic** – an epidemic becomes an endemic when the disease spread in a certain area become perpetual.

For example, let's take COVID19 example. When it first started in Wuhan, China it was an outbreak. When it outlasted for some time within Wuhan, China and neighboring regions, it was classified as an epidemic. When cases started spreading to other countries and become a global phenomenon, it was classified as a pandemic. Flu could be classified as an endemic because it spreads every winter in a certain region like a US. Whether COVID will become an endemic is not fully known yet. Other examples of endemic are malaria in Africa and Dengue in South Asia.

Interesting observation from our project question is that the question header declares it a flu pandemic but in the last question, it is called an epidemic. The spread of flu within a classroom over days is technically an epidemic.

3.2 Pandemics Spread Models & EpiModel

There are several disease models in epidemiology that are called compartmental models because they model the flow of people from one compartment to another. These models are Ordinary Differential Equation (ODE) models. The main ones are⁶:

- **Susceptible-Infectious (SI)** – a two state model where transition is one-way from susceptible population to infected, with no recovery after infection. Examples include HIV or herpes.
- **Susceptible-Infectious-Susceptible (SIS)** – a two state model where transition is from susceptible to infected and recovery doesn't induce immunity. Thus, individuals move to susceptible state after infection. Examples include common cold and flu.
- **Susceptible-Infectious-Recovered (SIR)** – a three state model where transition is from susceptible to infected and from infected to recovered. Cured individuals achieve an immunity and are not included as part of the susceptible population (i.e. can't get sick again). For example measles.

⁵ Stanborough RJ. and Biggers A. "How Is a Pandemic Different from an Epidemic?" *Healthline*, Healthline Media, April 30, 2020, <https://www.healthline.com/health/pandemic-vs-epidemic#what-is-an-epidemic>

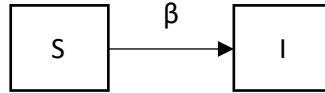
⁶ Jenness SM, Goodreau SM and Morris M. EpiModel: An R Package for Mathematical Modeling of Infectious Disease over Networks. *Journal of Statistical Software*. 2018; 84(8): 6.

In the section below, a basic mathematical overview of each model is provided for context. These mathematical models are implemented in the R package EpiModel that is used to solve the project question.

3.2.1 Susceptible-Infectious (SI)

The SI model is based on two states: susceptible (individual who can get sick) and infected (individuals who have been transmitted the disease). Following are some equations that represent this model⁷.

$S(t)$ is susceptible individuals at time t ; $I(t)$ is the infected at time t ; N is the total population.



$$S \rightarrow I \quad 1$$

$$S(t) + I(t) = N \quad 2$$

β is the transmission rate i.e. number of transmitting contacts per unit time; $T_c = 1 / \beta$ is time between transmitting contact. Thus, the infection equation is:

$$I(t + \delta t) = I(t) + \beta \frac{S(t)}{N} I(t) \delta t \quad 3$$

$$\frac{dI(t)}{dt} = \beta \frac{S(t)}{N} I(t) \quad 4$$

Converting the equation into fractions: $i(t) = I(t) / N$ and $s(t) = S(t) / N$. Thus,

$$\frac{di(t)}{dt} = \beta s(t) i(t) \quad 5$$

$$\frac{ds(t)}{dt} = -\beta s(t) i(t) \quad 6$$

$$s(t) + i(t) = 1 \quad 7$$

$$\frac{di(t)}{dt} = \beta (1 - i(t)) i(t) \quad 8$$

$i(t = 0) = i_0$ and solving for $i(t)$:

$$i(t) = \frac{i_0}{i_0 + (1 - i_0)e^{-\beta t}} \quad 9$$

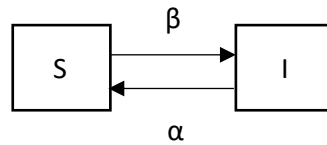
Equation 9 is the final solved deterministic equation for infected individuals at time t if $i(t = 0)$ is known. With this equation, we can keep incrementing time and find infected individuals. From $i(t)$, $I(t)$ as well as $s(t)$ and $S(t)$ could be calculated. The time at which all susceptible are infected will become the total duration of the epidemic.

⁷ Zhukov L. "Mathematical Modeling of Epidemics. Lecture 1: basic SI/SIS/SIR models explained." YouTube, Google Corporation, Mar 16, 2020, <https://www.youtube.com/watch?v=IXkrOAsEh1w>

3.2.2 Susceptible-Infectious-Susceptible (SIS)

The SIS model is also based on two states but the infected move back to Susceptible state after being cured without immunity. Following equations represent this model⁸:

$S(t)$ is susceptible individuals at time t ; $I(t)$ is the infected at time t ; N is the total population.



$$S \rightarrow I \rightarrow S \quad 1$$

$$S(t) + I(t) = N \quad 2$$

β is the transmission rate i.e. number of transmitting contacts per unit time; α is the recovery rate. Thus, the equations are:

$$\frac{dS}{dt} = \alpha I - \beta SI \quad 10$$

$$\frac{dI}{dt} = \beta SI - \alpha I \quad 11$$

$$\frac{dS}{dt} + \frac{dI}{dt} = \alpha I - \beta SI + \beta SI - \alpha I \quad 12$$

$$\frac{dN}{dt} = 0 \quad 13$$

$dN / dt = 0$ because N (total population) should be a constant. After rearranging the dI / dt equation and using integral to solve for I , we get:

$$I = \frac{(\beta N - \alpha) I_0 e^{(\beta N - \alpha)t}}{(\beta N - \alpha) + \beta I_0 [e^{(\beta N - \alpha)t} - 1]} \quad 14$$

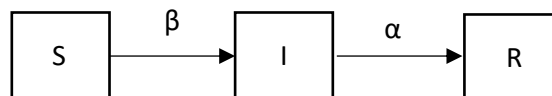
I_0 is the $I(t=0)$. Using the I equation, S can be calculated as well since $N = S + I$.

3.2.3 Susceptible-Infectious-Recovered (SIR)

The SIR model is based on 3 states where susceptible move to infected state and the infected move to a recovered state with immunity. Thus,

$$N = S + I + R \quad 15$$

And



⁸ Smith SR. "Simple Epidemic Models" Stacey Smith Homepage, University of Ottawa, <https://mysite.science.uottawa.ca/rsmith43/MAT4996/Epidemic.pdf>

Where β and α are the infection and recovery rates respectively. The system of equations for SIR is⁹:

$$\frac{dS}{dt} = -\beta SI \quad 16$$

$$\frac{dI}{dt} = \beta SI - \alpha I \quad 17$$

$$\frac{dR}{dt} = \alpha I \quad 18$$

These equations can be solved using the same integration technique used for SIS to get a deterministic formula for infected I individuals.

3.2.4 Stochastic Version of Models

The mathematical models shown in sections above are deterministic with rate of infection and recovery used for calculations. In stochastic versions of these models i.e. SI, SIS and SIR, the continuous variables are replaced by discrete numbers and the rates are replaced by probabilities. In order to simulate the stochastic models, Gillespie algorithm is used.

The Gillespie algorithm dictates two things: when the next event happens (e.g. susceptible being infected, infected being recovered etc.) and then it determines what happens exactly. The next time interval when something happens is exponentially distributed random numbers scaled by sum of all process rates. And what happens next is determined by drawing an event randomly from all possible events according to their respective probabilities¹⁰.

3.2.5 Using EpiModel for School Flu Question

The EpiModel library has the following models available¹¹:

⁹ Smith SR. "Simple Epidemic Models" Stacey Smith Homepage, University of Ottawa, <https://mysite.science.uottawa.ca/rsmith43/MAT4996/Epidemic.pdf>

¹⁰ Regoes RR., Schaefroth S. "Stochastic Simulation of simple epidemic" Biosym, http://www.biosym.uzh.ch/modules/models/ETHZ/StochasticSimulation/sir_stoch.xhtml

¹¹ Jenness SM, Goodreau SM and Morris M. EpiModel: An R Package for Mathematical Modeling of Infectious Disease over Networks. *Journal of Statistical Software*. 2018; 84(8): 6.

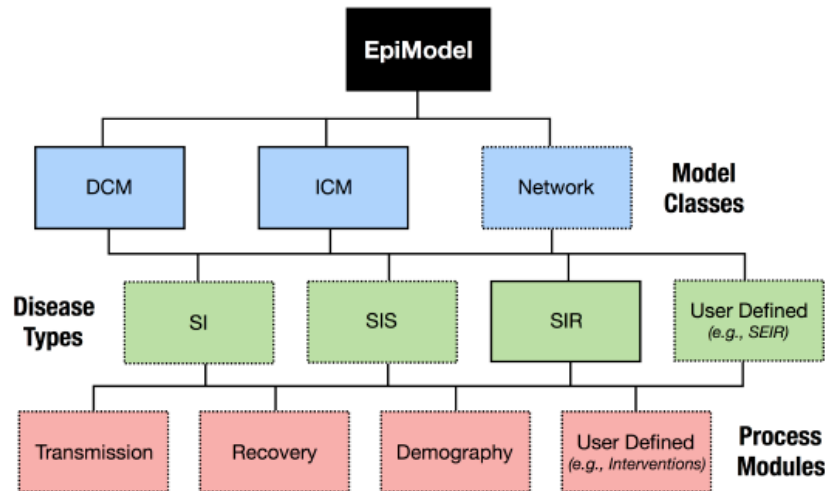


Figure 1 - Diagram of EpiModel Design

DCM means deterministic compartmental models and ICM means stochastic individual contract models (simulations). This aligns with the mathematical models review in sections above. Thus, the question 4 pandemic flu in school is solved using the DCM and ICM models in EpiModel (in section 3.4 and 3.5). SI and SIR models are used with some assumptions.

3.3 Analysis for the Disease Spread Modeling

Before we jump into the solution, some analysis of the parameters provided in the question is needed. Let's list out the known parameters first:

- Total population in class, $N = 21$
- Total susceptible at $t=0$, $S_0 = 20$
- Total infected at $t=0$, $I_0 = 1$
- Infectious period = 3 days
- Recovery rate, avg rate of recovery = $1 / 3 = 0.33$
 - Recovery rate is the reciprocal of disease duration¹², which in this case is assumed to be the infectious period. Technically, person might still be sick after 3 days, just not infectious, but this assumption that recovery and infectious period are the same is being made here.
- Probability of transmission = TBD
 - Provided probability 0.02 is the Bern(p) i.i.d. event for all kids and each of the 3 infectious days.

¹² Jenness SM, Goodreau SM and Morris M., Beylerian E. "Package EpiModel". CRAN, r-project organization, Nov 9, 2020, <https://cran.r-project.org/web/packages/EpiModel/EpiModel.pdf>; 6

- Thus, we need the probability of transmission occurring on any of these 3 Bernoulli events¹³. This can be done using the Binomial distribution (Bi (n, p)) probability mass function, since Binomial distribution is the probability distribution of x number of successful trials of n Bernoulli trials with probability p.

$$f(x) = p^x (1 - p)^{n-x} \binom{n}{x}$$

Where x are successful trials and n are total number of Bernoulli trials.

Thus, in our question, we want to calculate probability of at least 1 (=x) successful transmission out of 3 (=n) i.i.d. Bernoulli trials with probability of each (=p) 0.02. Plugging it in give us **0.0576** as probability of at least 1 infection occurring on a 3-day infectious period.

- Thus, infection transmission probability = 0.06

We are now ready to dive into the solution.

3.4 Solution Based on Deterministic Spread Models

As noted in the sections above, the school flu spread question will be solved using deterministic as well as stochastic SI and SIR models in EpiModel.

3.4.1 Deterministic SI Model Solution

Assumption for SI model:

- Students are not getting recovered and epidemic ends when everybody is infected.

This will be a simpler version of the solution and recovery rate will not be utilized in modeling. Using model parameters:

- inf.prob (infection probability) = 0.06
- act.rate (acts per person per unit time) = 20
 - Since a kid could interact with 20 other kids in the class at any given time.
- s.num (susceptible) = 20
- i.num (infected initially) = 1
- model type = SI
- nsteps = 50 (number of days)

The code looks like this (full code is in the appendix):

¹³ Sugiyama M. "Examples of Discrete Probability Distributions". Science Direct, 2016, <https://www.sciencedirect.com/topics/computer-science/bernoulli-trial>

```

```{r deterministic SI model}

require(EpiModel)
set.seed(101)

param <- param.dcm(inf.prob = 0.06, act.rate = 20)
init <- init.dcm(s.num = 20, i.num = 1)
control <- control.dcm(type = "SI", nsteps = 50)
mod <- dcm(param, init, control)

mod
plot(mod)

#summary function shows the state of compartments at a given time
summary(mod, at = 25)

```

```

Figure 2 - Code block for deterministic SI model

And the output looks like:

```

EpiModel Simulation
=====
Model class: dcm

Simulation Summary
-----
Model type: SI
No. runs: 1
No. time steps: 50
No. groups: 1

Model Parameters
-----
inf.prob = 0.06
act.rate = 20

Model Output
-----
Variables: s.num i.num si.flow num

```

Figure 3 - Deterministic SI model parameters

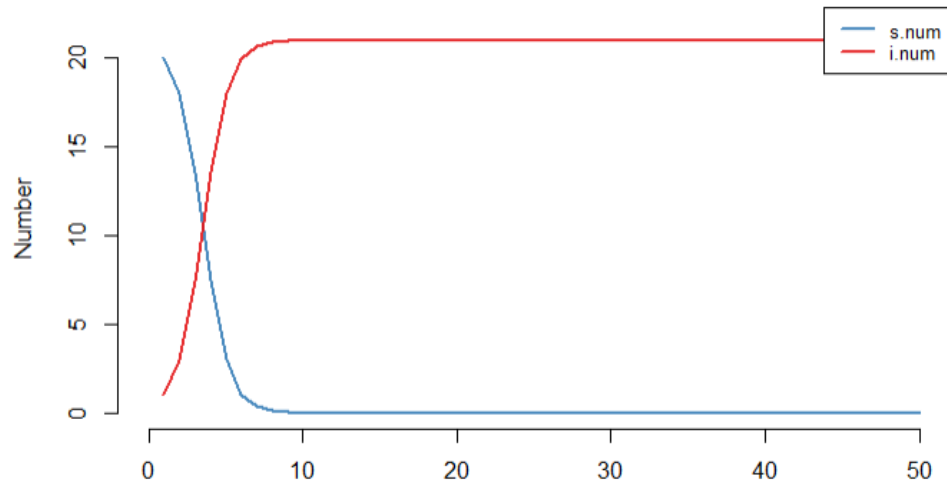


Figure 4 - Plot of susceptible and infected over time for deterministic SI model

s.num and i.num represent the progression of susceptible and infected kids over time in the graph above. The same information is shown in a clean table format below (only till 20 days since epidemic seems to end after 10 days when everybody gets infected):

| | Susceptible | Infected | SI Flow | Total |
|----|-------------|----------|---------|-------|
| 1 | 20.00 | 1.00 | 1.98 | 21 |
| 2 | 18.02 | 2.98 | 4.44 | 21 |
| 3 | 13.58 | 7.42 | 6.11 | 21 |
| 4 | 7.47 | 13.53 | 4.46 | 21 |
| 5 | 3.02 | 17.98 | 1.98 | 21 |
| 6 | 1.04 | 19.96 | 0.70 | 21 |
| 7 | 0.34 | 20.66 | 0.23 | 21 |
| 8 | 0.11 | 20.89 | 0.07 | 21 |
| 9 | 0.03 | 20.97 | 0.02 | 21 |
| 10 | 0.01 | 20.99 | 0.01 | 21 |
| 11 | 0.00 | 21.00 | 0.00 | 21 |
| 12 | 0.00 | 21.00 | 0.00 | 21 |
| 13 | 0.00 | 21.00 | 0.00 | 21 |
| 14 | 0.00 | 21.00 | 0.00 | 21 |
| 15 | 0.00 | 21.00 | 0.00 | 21 |
| 16 | 0.00 | 21.00 | 0.00 | 21 |
| 17 | 0.00 | 21.00 | 0.00 | 21 |
| 18 | 0.00 | 21.00 | 0.00 | 21 |
| 19 | 0.00 | 21.00 | 0.00 | 21 |
| 20 | 0.00 | 21.00 | 0.00 | 21 |

Figure 5 - Table of susceptible & infected over time for deterministic SI model

We have enough information to answer the questions for the problems.

- a) *Distribution of kids infected on day 1*
 - a. 1 kid is infected on day 1 (figure 5)
 - b. Distribution can't be created for deterministic solution.
- b) *Expected number of kids Tommy infects on day 1*
 - a. 1 kid is infected on day 1 (figure 5)
- c) *Expected number of kids infected on day 2 including Tommy*
 - a. 2.98 (or rounded to 3 kids from figure 5)
- d) *Expected number of kids infected on each day based on simulation.*
 - a. Deterministic solution doesn't have simulation, but the number of kids infected on each day is shown in figure 5.
- e) *Histogram of how long epidemic lasts*
 - a. Figure 4 shows plots the number of days with infections.
 - b. Figure 4 and 5 both show that epidemic lasts around 11 days (when everybody gets infected).

3.4.2 Deterministic SIR Model Solution

Assumption for SI model:

- Students are getting recovered and epidemic ends when everybody is recovered or spread stabilizes.
- The recovery rate is determined from infectious period and kid is considered recovered after 3 days of infectious period.

Using model parameters:

- inf.prob (infection probability) = 0.06
- act.rate (acts per person per unit time) = 20
 - Since a kid could interact with 20 other kids in the class at any given time.
- Rec.rate (recovery rate) = 1/3
- s.num (susceptible) = 20
- i.num (infected initially) = 1
- model type = SIR
- nsteps = 50 (number of days)

The code looks like this (full code is in the appendix):

```
```{r deterministic SIR model}

set.seed(101)

param1 <- param.dcm(inf.prob = 0.06, act.rate = 20, rec.rate = 1/3)
init1 <- init.dcm(s.num = 20, i.num = 1, r.num = 0)
control1 <- control.dcm(type = "SIR", nsteps = 50, dt = 1)
mod1 <- dcm(param1, init1, control1)

#plotting the model
par(mar = c(3.2, 3, 2, 1), mgp = c(2, 1, 0), mfrow = c(1, 2))
plot(mod1, popfrac = FALSE, alpha = 0.5,
 lwd = 4, main = "Flow Histograms")
plot(mod1, y = "si.flow", lwd = 4, col = "firebrick",
 main = "Disease Incidence", legend = "n")

```
```

Figure 6 - Code block for deterministic SIR model

And the output looks like:

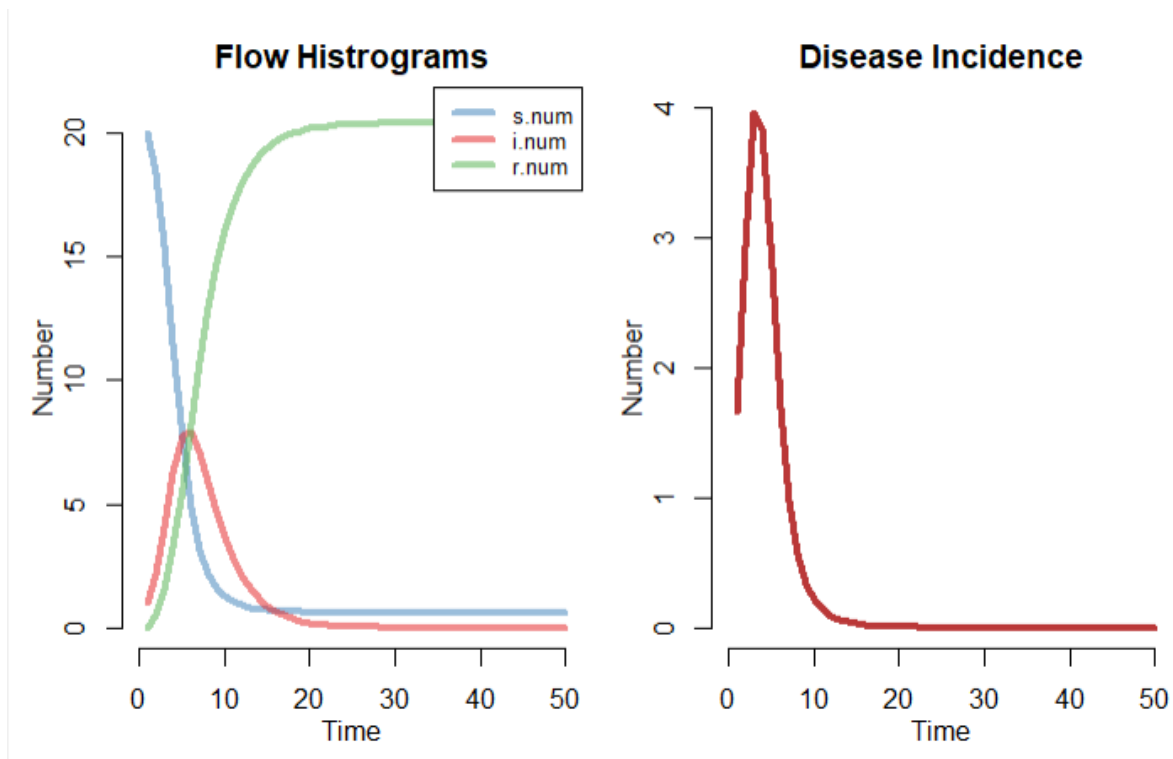


Figure 7 - Plot of disease spread over time for deterministic SIR model

s.num and i.num represent the progression of susceptible and infected kids over time in the graph above and r.num is the number of kids recovered over time. Disease incidence is the si.flow which is the rate at which susceptible kids become infected.

The same information is shown in a clean table format below (only till 20 days since epidemic seems to slow down after that):

| | Susceptible | Infected | Recovered | SI Flow | IR Flow | Total |
|----|-------------|----------|-----------|---------|---------|-------|
| 1 | 20.00 | 1.00 | 0.00 | 1.66 | 0.50 | 21 |
| 2 | 18.34 | 2.15 | 0.50 | 2.93 | 1.02 | 21 |
| 3 | 15.41 | 4.07 | 1.52 | 3.96 | 1.73 | 21 |
| 4 | 11.45 | 6.30 | 3.26 | 3.83 | 2.38 | 21 |
| 5 | 7.62 | 7.75 | 5.63 | 2.77 | 2.64 | 21 |
| 6 | 4.85 | 7.88 | 8.27 | 1.69 | 2.51 | 21 |
| 7 | 3.15 | 7.07 | 10.78 | 0.98 | 2.16 | 21 |
| 8 | 2.18 | 5.88 | 12.94 | 0.57 | 1.76 | 21 |
| 9 | 1.61 | 4.69 | 14.70 | 0.34 | 1.38 | 21 |
| 10 | 1.27 | 3.64 | 16.09 | 0.21 | 1.07 | 21 |
| 11 | 1.06 | 2.79 | 17.15 | 0.14 | 0.81 | 21 |
| 12 | 0.92 | 2.11 | 17.97 | 0.09 | 0.61 | 21 |
| 13 | 0.83 | 1.59 | 18.58 | 0.06 | 0.46 | 21 |
| 14 | 0.77 | 1.19 | 19.04 | 0.04 | 0.35 | 21 |
| 15 | 0.72 | 0.89 | 19.39 | 0.03 | 0.26 | 21 |
| 16 | 0.69 | 0.67 | 19.64 | 0.02 | 0.19 | 21 |
| 17 | 0.67 | 0.50 | 19.84 | 0.02 | 0.14 | 21 |
| 18 | 0.65 | 0.37 | 19.98 | 0.01 | 0.11 | 21 |
| 19 | 0.64 | 0.27 | 20.09 | 0.01 | 0.08 | 21 |
| 20 | 0.63 | 0.20 | 20.17 | 0.01 | 0.06 | 21 |

Figure 8 - Table of susceptible, infected & recovered over time for deterministic SIR model

Answer the questions for the problems.

- f) *Distribution of kids infected on day 1*
 - a. 1 kid is infected on day 1 (figure 8)
 - b. Distribution can't be created for deterministic solution.
- g) *Expected number of kids Tommy infects on day 1*
 - a. 1 kid is infected on day 1 (figure 8)
- h) *Expected number of kids infected on day 2 including Tommy*
 - a. 2.15 (or rounded to 2 kids from figure 8)
- i) *Expected number of kids infected on each day based on simulation.*
 - a. Deterministic solution doesn't have simulation, but the number of kids infected on each day is shown in figure 8.
- j) *Histogram of how long epidemic lasts*
 - a. Figure 7 shows plots the number of days with infections.
 - b. Defining the end of epidemic is not a straightforward thing in epidemiology studies. In this model, given it's an SIR model, we can see that epidemic ends when:

- i. Everybody is infected i.e. susceptible kids is 0
- ii. Or, infections have dropped to 0, even if there are still susceptible kids.
- c. Figure 7 and 8 both show that epidemic lasts around 18 days because infections are dropped below 1 (rounded off) and almost everybody is infected. It might be possible by chance at 1 kid does not gets infected at all at the end of the epidemic.

3.5 Solution Based on Stochastic Spread Models

3.5.1 Stochastic SI Model Solution

Assumption for stochastic SI model are the same as deterministic model.

Using model parameters:

- inf.prob (infection probability) = 0.06
- act.rate (acts per person per unit time) = 20
- s.num (susceptible) = 20
- i.num (infected initially) = 1
- model type = SI
- nsims (number of simulations) = 1000
- nsteps (number of days) = 50

The code looks like this (full code is in the appendix):

```
```{r stochastic SI model}

set.seed(101)

param2 <- param.icm(inf.prob = 0.06, act.rate = 20)
init2 <- init.icm(s.num = 20, i.num = 1)
control2 <- control.icm(type = "SI", nsims = 1000, nsteps = 50)
mod2 <- icm(param2, init2, control2)
mod2

#plotting the model
plot(mod2)

#with each simulation plot & average
plot(mod2, sim.lines = TRUE, mean.smooth = TRUE, qnts.smooth = TRUE)

```
```

Figure 9 - Code block for stochastic SI model

And the output looks like:

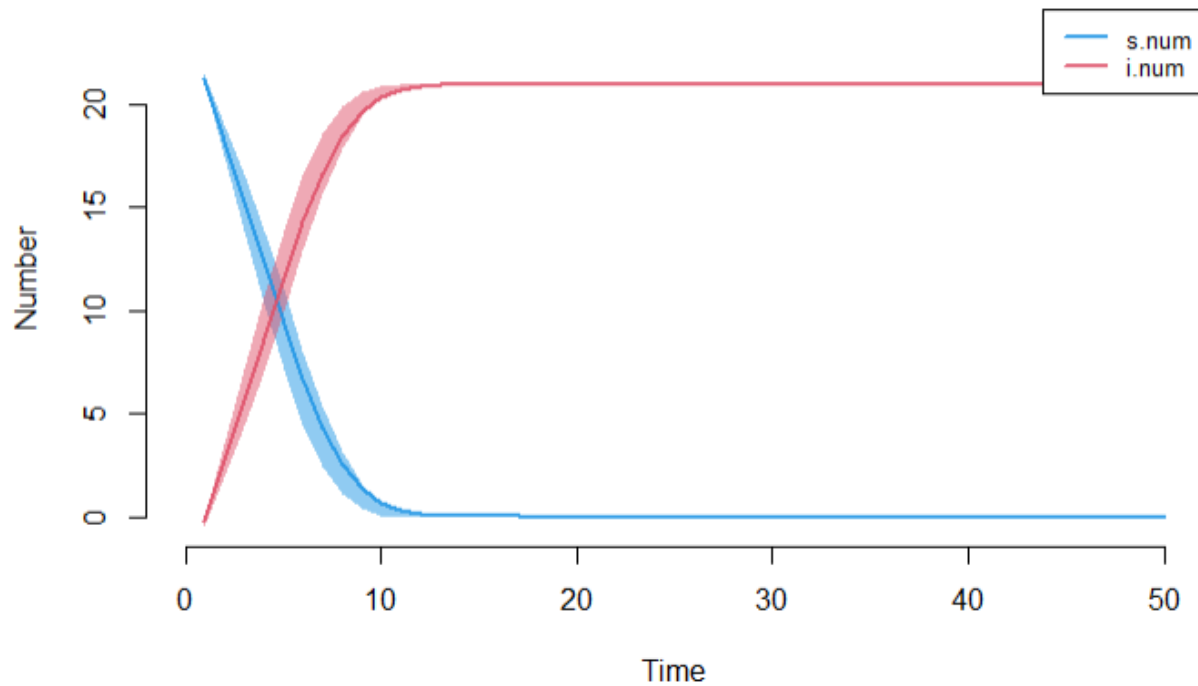


Figure 10 - Plot of disease spread over time for stochastic SI model

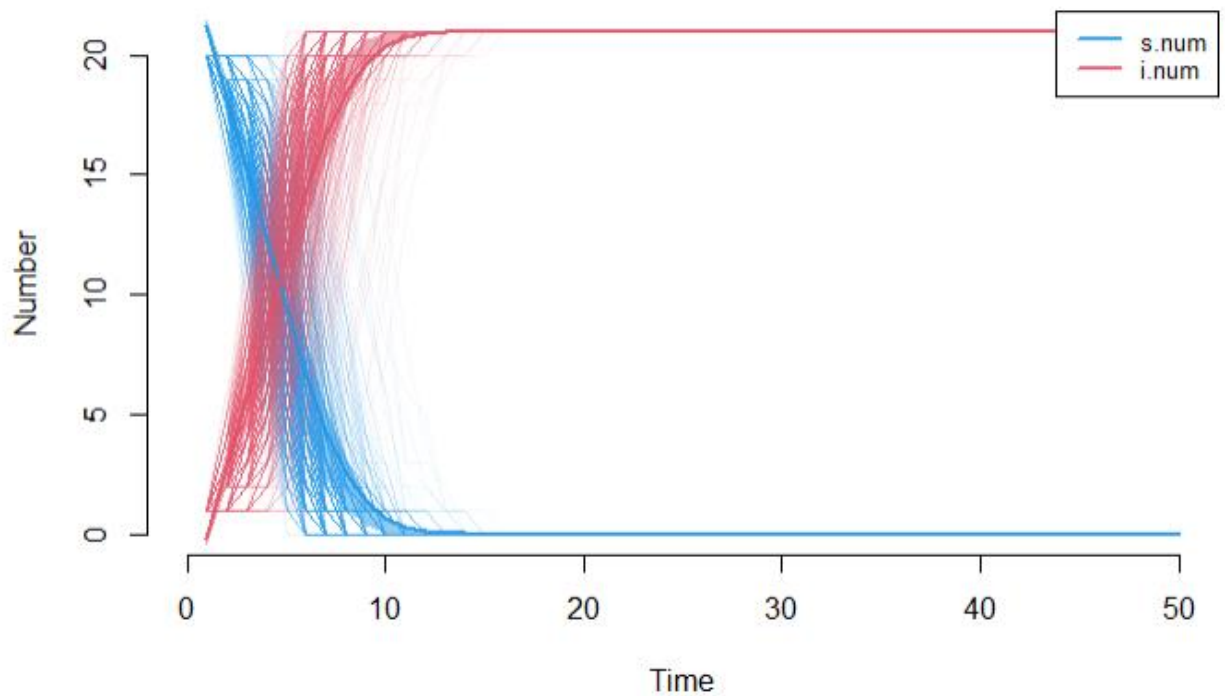


Figure 11 - Plot of disease spread with lines for each simulation for Stochastic SI model

Figure 10 shows the mean values of susceptible (s.num) and infected (i.num) along with the smoothed inter-quartile range of values. Figure 11 shows the actual simulation lines for susceptible and infected.

The average value for susceptible & infected for 1000 simulations is shown in the table below (only till 20 days):

| | time | s.num | i.num | num | si.flow |
|----|------|-------|-------|-----|---------|
| 1 | 1 | 20.00 | 1.00 | 21 | 0.00 |
| 2 | 2 | 18.83 | 2.17 | 21 | 1.17 |
| 3 | 3 | 16.66 | 4.34 | 21 | 2.17 |
| 4 | 4 | 13.27 | 7.73 | 21 | 3.39 |
| 5 | 5 | 9.15 | 11.85 | 21 | 4.12 |
| 6 | 6 | 5.31 | 15.69 | 21 | 3.84 |
| 7 | 7 | 2.62 | 18.38 | 21 | 2.69 |
| 8 | 8 | 1.14 | 19.86 | 21 | 1.48 |
| 9 | 9 | 0.46 | 20.54 | 21 | 0.69 |
| 10 | 10 | 0.17 | 20.83 | 21 | 0.29 |
| 11 | 11 | 0.05 | 20.95 | 21 | 0.12 |
| 12 | 12 | 0.02 | 20.98 | 21 | 0.04 |
| 13 | 13 | 0.00 | 21.00 | 21 | 0.01 |
| 14 | 14 | 0.00 | 21.00 | 21 | 0.00 |
| 15 | 15 | 0.00 | 21.00 | 21 | 0.00 |
| 16 | 16 | 0.00 | 21.00 | 21 | 0.00 |
| 17 | 17 | 0.00 | 21.00 | 21 | 0.00 |
| 18 | 18 | 0.00 | 21.00 | 21 | 0.00 |
| 19 | 19 | 0.00 | 21.00 | 21 | 0.00 |
| 20 | 20 | 0.00 | 21.00 | 21 | 0.00 |

Figure 12 - Table of Avg values of susceptible & infected for stochastic SI model

Answer the questions for the problems.

a) *Distribution of kids infected on day 1*

- 1 kid is infected on day 1 (figure 12)
- Distribution of infections on day 1 across 1000 simulations is shown below. It shows that all values for 1000 simulations for day 1 infected are 1.

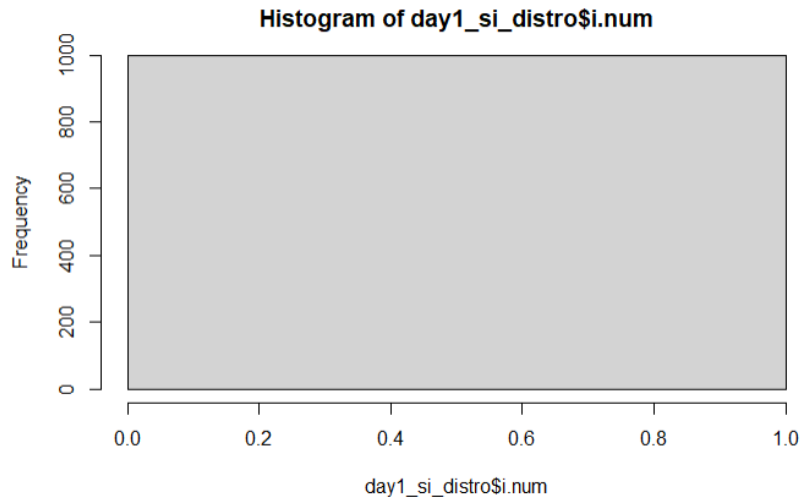


Figure 13 - Distribution of day 1 infections for stochastic SI model

- b) *Expected number of kids Tommy infects on day 1*
 - a. Expected (average) value of infected on day 1 is 1 (figure 12)
- c) *Expected number of kids infected on day 2 including Tommy*
 - a. 2.17 (or rounded to 2 kids from figure 12)
- d) *Expected number of kids infected on each day based on simulation.*
 - a. The expected (average) number of kids infected on each day is shown in figure 12.
 - b. If numbers are rounded off, it appears that epidemic ends (everybody infected) by day 9 and if numbers are not rounded off, 13 days.
- e) *Histogram of how long epidemic lasts*
 - a. In this model, given it's a simple SI model, we can see that epidemic ends when everybody is infected i.e. susceptible kids is 0 or infected kids is 21.
 - b. Based on this criterion, a summary dataframe is prepared with a log of duration of epidemic for each of 1000 simulations. A histogram of this table is shown below.

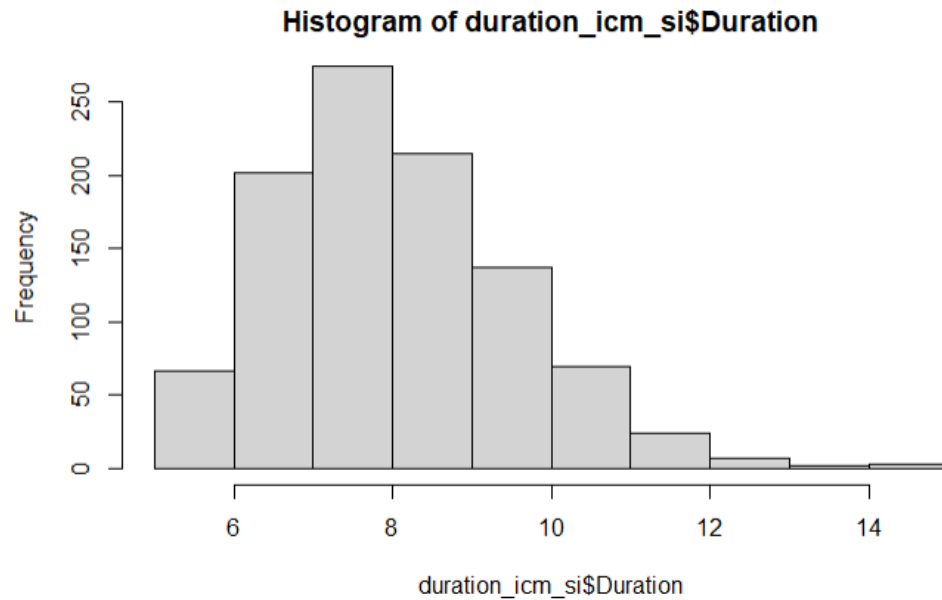


Figure 14 - Histogram of epidemic duration for stochastic SI model

- c. This distribution is right-skewed with median (8 days) lower than the mean (8.52 days).

3.5.2 Stochastic SIR Model Solution

Assumption for stochastic SIR model are the same as deterministic model.

Using model parameters:

- inf.prob (infection probability) = 0.06
- act.rate (acts per person per unit time) = 20
- s.num (susceptible) = 20
- i.num (infected initially) = 1
- model type = SIR
- nsims (number of simulations) = 1000
- nsteps (number of days) = 50

The code looks like this (full code is in the appendix):

```

```{r Stochastic SIR Model}

set.seed(101)

param3 <- param.icm(inf.prob = 0.06, act.rate = 20, rec.rate = 1/3)
init3 <- init.icm(s.num = 20, i.num = 1, r.num = 0)
control3 <- control.icm(type = "SIR", nsteps = 50, nsims = 1000)
mod3 <- icm(param3, init3, control3)

#plotting the model
plot(mod3)

#each simulation plot with average
plot(mod3, sim.lines = TRUE, mean.smooth = TRUE, qnts.smooth = TRUE)
```

```

Figure 15 - Code block for stochastic SIR model

And the output looks like:

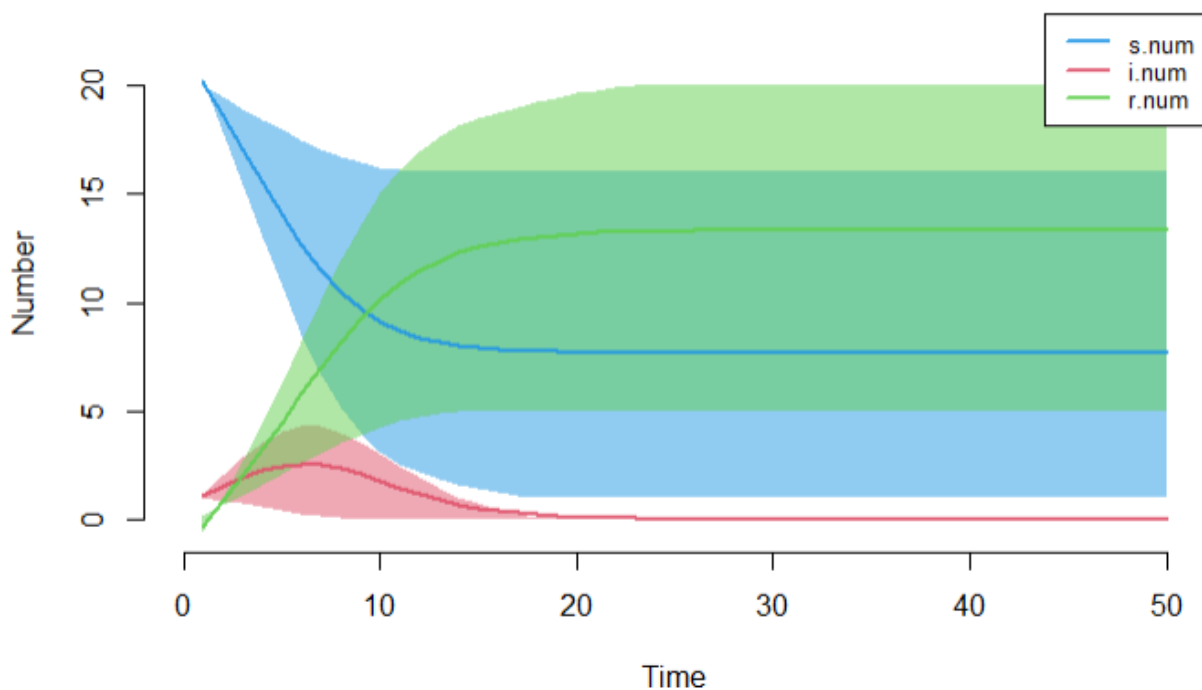


Figure 16 - Plot of disease spread over time for stochastic SIR model

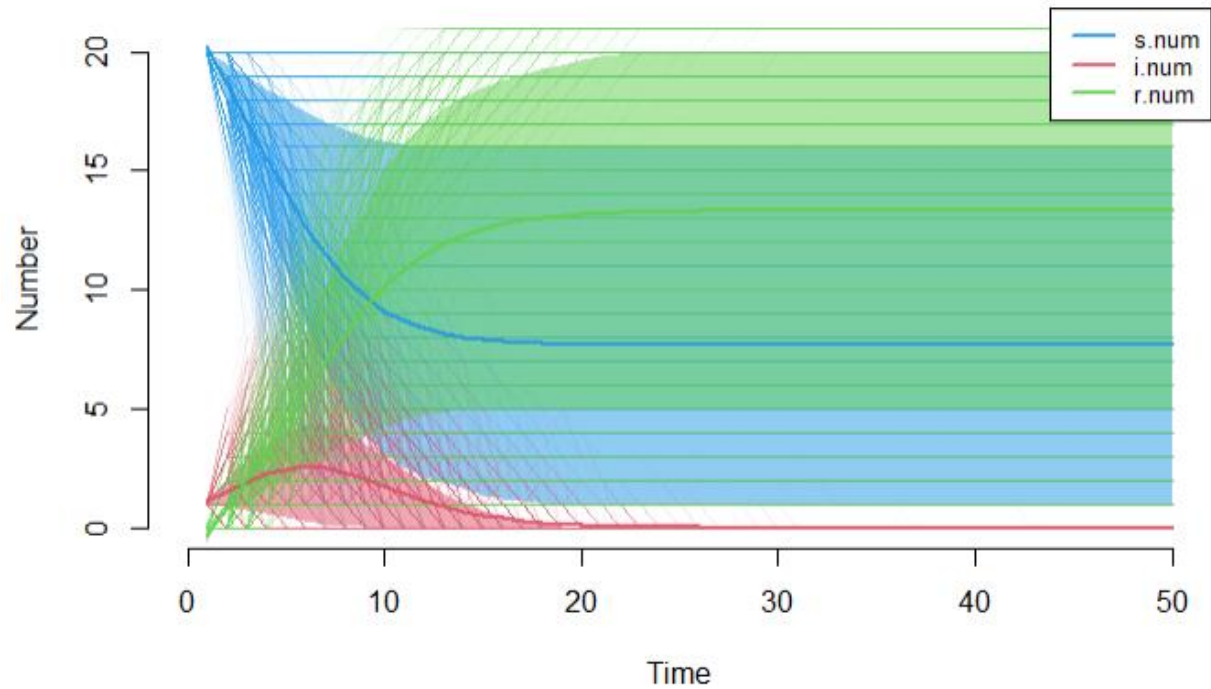


Figure 17 - Plot of disease spread with lines for each simulation for Stochastic SIR model

Figure 16 shows the mean values of susceptible (s.num) and infected (i.num) along with the smoothed inter-quartile range of values. Figure 17 shows the actual simulation lines for susceptible and infected.

The average value for susceptible & infected for 1000 simulations is shown in the table below (only till 30 days):

| | time | s.num | i.num | num | r.num | si.flow | ir.flow |
|----|------|-------|-------|-----|-------|---------|---------|
| 1 | 1 | 20.00 | 1.00 | 21 | 0.00 | 0.00 | 0.00 |
| 2 | 2 | 18.80 | 1.53 | 21 | 0.67 | 1.20 | 0.67 |
| 3 | 3 | 17.22 | 2.08 | 21 | 1.70 | 1.58 | 1.03 |
| 4 | 4 | 15.44 | 2.56 | 21 | 3.00 | 1.78 | 1.31 |
| 5 | 5 | 13.72 | 2.88 | 21 | 4.39 | 1.71 | 1.39 |
| 6 | 6 | 12.19 | 2.95 | 21 | 5.86 | 1.53 | 1.47 |
| 7 | 7 | 10.94 | 2.75 | 21 | 7.31 | 1.25 | 1.45 |
| 8 | 8 | 10.00 | 2.48 | 21 | 8.52 | 0.94 | 1.21 |
| 9 | 9 | 9.26 | 2.14 | 21 | 9.59 | 0.74 | 1.07 |
| 10 | 10 | 8.76 | 1.74 | 21 | 10.50 | 0.50 | 0.91 |
| 11 | 11 | 8.43 | 1.37 | 21 | 11.20 | 0.33 | 0.69 |
| 12 | 12 | 8.21 | 1.07 | 21 | 11.72 | 0.22 | 0.53 |
| 13 | 13 | 8.05 | 0.81 | 21 | 12.14 | 0.16 | 0.41 |
| 14 | 14 | 7.93 | 0.63 | 21 | 12.44 | 0.12 | 0.30 |
| 15 | 15 | 7.86 | 0.46 | 21 | 12.68 | 0.07 | 0.24 |
| 16 | 16 | 7.79 | 0.34 | 21 | 12.87 | 0.07 | 0.19 |
| 17 | 17 | 7.75 | 0.26 | 21 | 12.98 | 0.04 | 0.11 |
| 18 | 18 | 7.72 | 0.19 | 21 | 13.09 | 0.03 | 0.11 |
| 19 | 19 | 7.71 | 0.13 | 21 | 13.16 | 0.01 | 0.07 |
| 20 | 20 | 7.70 | 0.10 | 21 | 13.20 | 0.01 | 0.05 |
| 21 | 21 | 7.69 | 0.07 | 21 | 13.23 | 0.00 | 0.03 |
| 22 | 22 | 7.69 | 0.05 | 21 | 13.26 | 0.00 | 0.02 |
| 23 | 23 | 7.69 | 0.03 | 21 | 13.28 | 0.00 | 0.02 |
| 24 | 24 | 7.69 | 0.02 | 21 | 13.29 | 0.00 | 0.01 |
| 25 | 25 | 7.69 | 0.02 | 21 | 13.29 | 0.00 | 0.00 |
| 26 | 26 | 7.69 | 0.01 | 21 | 13.30 | 0.00 | 0.01 |
| 27 | 27 | 7.69 | 0.01 | 21 | 13.30 | 0.00 | 0.00 |
| 28 | 28 | 7.69 | 0.01 | 21 | 13.30 | 0.00 | 0.00 |
| 29 | 29 | 7.69 | 0.01 | 21 | 13.31 | 0.00 | 0.00 |
| 30 | 30 | 7.69 | 0.00 | 21 | 13.31 | 0.00 | 0.00 |

Figure 18 - Table of avg values of susceptible & infected for stochastic SIR model

Answer the questions for the problems.

a) Distribution of kids infected on day 1

a. 1 kid is infected on day 1 (figure 18)

- b. Distribution of infections on day 1 across 1000 simulations is shown below. It shows that all values for 1000 simulations for day 1 infected are 1.

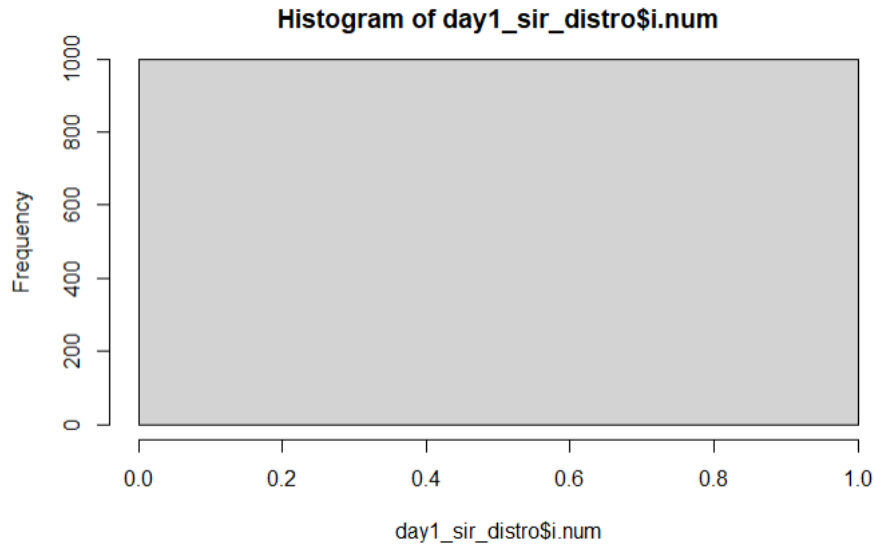


Figure 19 -Distribution of day 1 infections for stochastic SIR model

- b) *Expected number of kids Tommy infects on day 1*
- Expected (average) value of infected on day 1 is 1 (figure 18)
- c) *Expected number of kids infected on day 2 including Tommy*
- 1.53 (or rounded to 2 kids from figure 18)
- d) *Expected number of kids infected on each day based on simulation.*
- The expected (average) number of kids infected on each day is shown in figure 18.
 - If numbers are rounded off, it appears that epidemic ends (everybody infected) by day 15 and if numbers are not rounded off, 30 days.
- e) *Histogram of how long epidemic lasts*
- In this model, given it's an SIR model, we can see that epidemic ends when:
 - Everybody is infected i.e. susceptible kids is 0
 - Or, infections have dropped to 0, even if there are still susceptible kids.
 - Based on this criterion, a summary dataframe is prepared with a log of duration of epidemic for each of 1000 simulations. A histogram of this table is shown below.

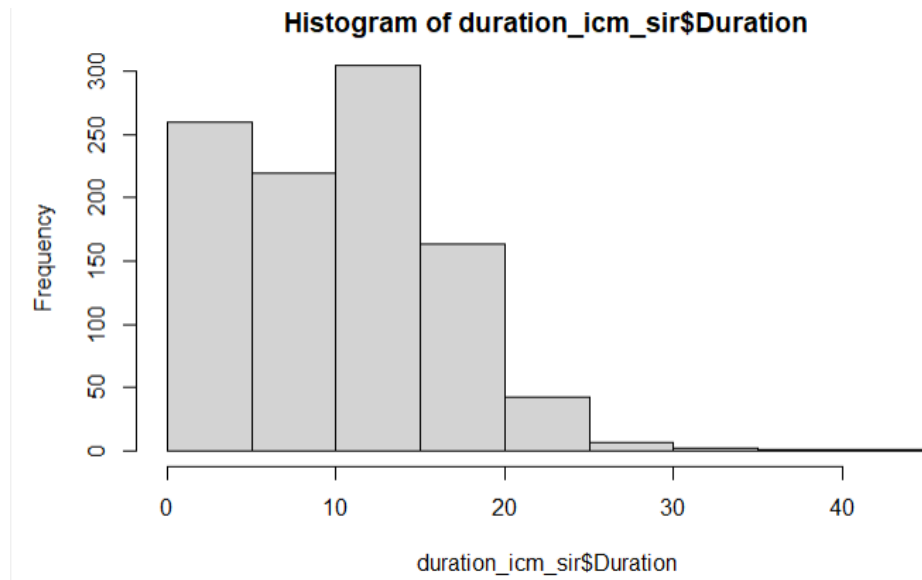


Figure 20 - Histogram of epidemic duration for stochastic SIR model

- c. This distribution is slightly left-skewed with mean (10.63 days) lower than the median (11 days).

4.0 Conclusion

4.1 Learnings from the Output

Let's compare the output for each of the 4 models:

| Model/Question | Day 1 Infections | Day 2 Infections | Epidemic Duration |
|--------------------------|--------------------------|--------------------------|---|
| Deterministic SI | 1.00 | 2.98 | 11 days |
| Deterministic SIR | 1.00 | 2.15 | 18 days
<i>(based on rounded infections)</i> |
| Stochastic SI | 1.00
<i>(average)</i> | 2.17
<i>(average)</i> | 9 days
<i>(mean of durations)</i> |
| Stochastic SIR | 1.00
<i>(average)</i> | 1.53
<i>(average)</i> | 11 days
<i>(mean of durations)</i> |

Figure 21 - Summary statistics for all models

It is interesting to note from the summary table above that the outputs from these 4 models are not very different. The day 1 and day 2 infections are consistently 1 and 2 kids (with exception of deterministic SI) and epidemic duration is in the same range of 9-11 days (with the exception of deterministic SIR). Though the stochastic models are preferred given these are based on many simulations and the variations of the occurrences are incorporated but interestingly the deterministic models were not very off as well.

4.2 Learnings from Using EpiModel Library in R

Using the EpiModel library made the analysis for this project much easier and manageable. As with all Python/R libraries, the benefit of having the models pre-coded and made available as parameterized functions is very efficient. But using these functions requires extensive studying of the library documentation to ensure the right usage. EpiModel provides nice functions for modeling, plotting and data output. That helped focus more on the theory and analysis rather than coding. Full code is attached (print out of R notebook) in the appendix of the report.

Appendix

The pages attached below are the output of the R Markdown Notebook with the full code solution for this project.

ISyE-6644 Project 1 - Pandemic Flu Spread

[Muhammad Omer Farooq](#)

03/09/2021

1. Deterministic SI Model

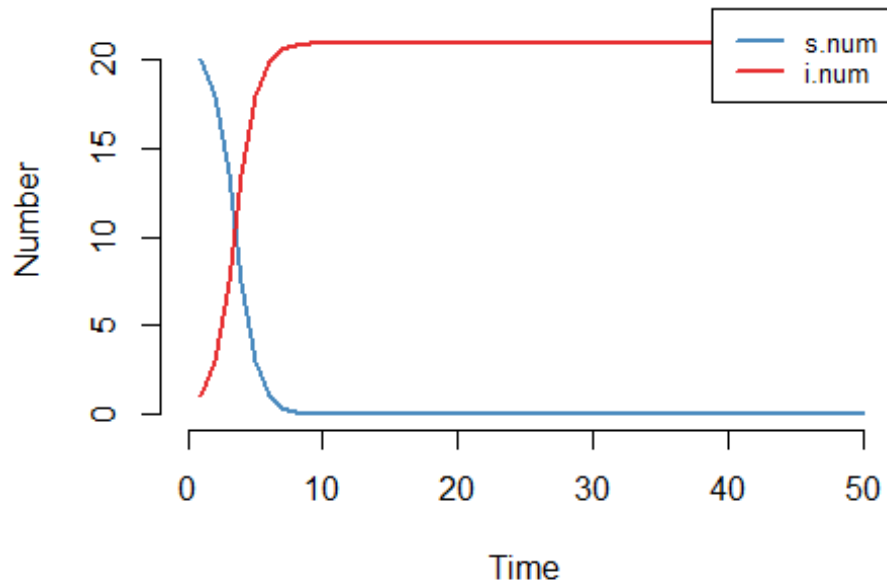
```
require(EpiModel)
set.seed(101)

param <- param.dcm(inf.prob = 0.06, act.rate = 20)
init <- init.dcm(s.num = 20, i.num = 1)
control <- control.dcm(type = "SI", nsteps = 50)
mod <- dcm(param, init, control)

mod

## EpiModel Simulation
## =====
## Model class: dcm
##
## Simulation Summary
## -----
## Model type: SI
## No. runs: 1
## No. time steps: 50
## No. groups: 1
##
## Model Parameters
## -----
## inf.prob = 0.06
## act.rate = 20
##
## Model Output
## -----
## Variables: s.num i.num si.flow num

plot(mod)
```



#summary function shows the state of compartments at a given time
 summary(mod, at = 25)

```
## EpiModel Summary
## =====
## Model class: dcm
##
## Simulation Summary
## -----
## Model type: SI
## No. runs: 1
## No. time steps:
## No. groups: 1
##
## Model Statistics
## -----
## Time: 25 Run: 1
## -----
##
```

| | n | pct |
|----------|----|-----|
| Suscept. | 0 | 0 |
| Infect. | 21 | 1 |
| Total | 21 | 1 |
| S -> I | 0 | NA |

```
## -----
##
```

#creating output dataframe
 s_num <- mod\$epi\$s.num
 i_num <- mod\$epi\$i.num
 si_flow <- mod\$epi\$si.flow

```

num <- mod$epi$num
dcm_si <- data.frame(s_num, i_num, si_flow, num)
colnames(dcm_si) <- c("Susceptible", "Infected", "SI Flow", "Total")
dcm_si <- round(dcm_si, digits = 2)
dcm_si

```

| ## | Susceptible | Infected | SI Flow | Total |
|-------|-------------|----------|---------|-------|
| ## 1 | 20.00 | 1.00 | 1.98 | 21 |
| ## 2 | 18.02 | 2.98 | 4.44 | 21 |
| ## 3 | 13.58 | 7.42 | 6.11 | 21 |
| ## 4 | 7.47 | 13.53 | 4.46 | 21 |
| ## 5 | 3.02 | 17.98 | 1.98 | 21 |
| ## 6 | 1.04 | 19.96 | 0.70 | 21 |
| ## 7 | 0.34 | 20.66 | 0.23 | 21 |
| ## 8 | 0.11 | 20.89 | 0.07 | 21 |
| ## 9 | 0.03 | 20.97 | 0.02 | 21 |
| ## 10 | 0.01 | 20.99 | 0.01 | 21 |
| ## 11 | 0.00 | 21.00 | 0.00 | 21 |
| ## 12 | 0.00 | 21.00 | 0.00 | 21 |
| ## 13 | 0.00 | 21.00 | 0.00 | 21 |
| ## 14 | 0.00 | 21.00 | 0.00 | 21 |
| ## 15 | 0.00 | 21.00 | 0.00 | 21 |
| ## 16 | 0.00 | 21.00 | 0.00 | 21 |
| ## 17 | 0.00 | 21.00 | 0.00 | 21 |
| ## 18 | 0.00 | 21.00 | 0.00 | 21 |
| ## 19 | 0.00 | 21.00 | 0.00 | 21 |
| ## 20 | 0.00 | 21.00 | 0.00 | 21 |
| ## 21 | 0.00 | 21.00 | 0.00 | 21 |
| ## 22 | 0.00 | 21.00 | 0.00 | 21 |
| ## 23 | 0.00 | 21.00 | 0.00 | 21 |
| ## 24 | 0.00 | 21.00 | 0.00 | 21 |
| ## 25 | 0.00 | 21.00 | 0.00 | 21 |
| ## 26 | 0.00 | 21.00 | 0.00 | 21 |
| ## 27 | 0.00 | 21.00 | 0.00 | 21 |
| ## 28 | 0.00 | 21.00 | 0.00 | 21 |
| ## 29 | 0.00 | 21.00 | 0.00 | 21 |
| ## 30 | 0.00 | 21.00 | 0.00 | 21 |
| ## 31 | 0.00 | 21.00 | 0.00 | 21 |
| ## 32 | 0.00 | 21.00 | 0.00 | 21 |
| ## 33 | 0.00 | 21.00 | 0.00 | 21 |
| ## 34 | 0.00 | 21.00 | 0.00 | 21 |
| ## 35 | 0.00 | 21.00 | 0.00 | 21 |
| ## 36 | 0.00 | 21.00 | 0.00 | 21 |
| ## 37 | 0.00 | 21.00 | 0.00 | 21 |
| ## 38 | 0.00 | 21.00 | 0.00 | 21 |
| ## 39 | 0.00 | 21.00 | 0.00 | 21 |
| ## 40 | 0.00 | 21.00 | 0.00 | 21 |
| ## 41 | 0.00 | 21.00 | 0.00 | 21 |
| ## 42 | 0.00 | 21.00 | 0.00 | 21 |
| ## 43 | 0.00 | 21.00 | 0.00 | 21 |
| ## 44 | 0.00 | 21.00 | 0.00 | 21 |
| ## 45 | 0.00 | 21.00 | 0.00 | 21 |
| ## 46 | 0.00 | 21.00 | 0.00 | 21 |
| ## 47 | 0.00 | 21.00 | 0.00 | 21 |
| ## 48 | 0.00 | 21.00 | 0.00 | 21 |

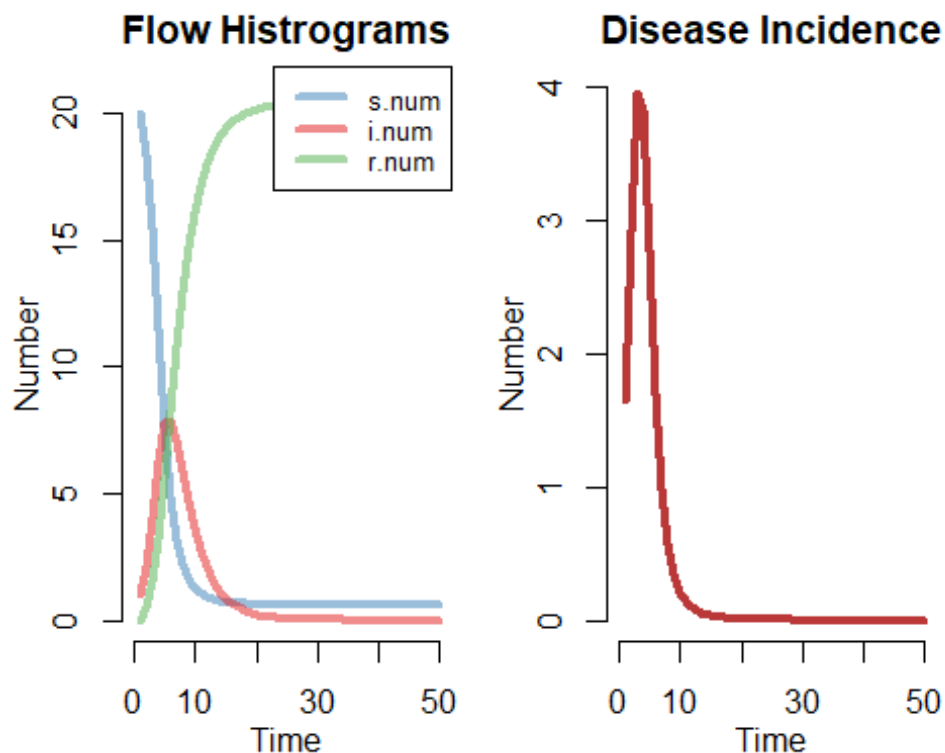

```
## 49      0.00    21.00    0.00    21
## 50      0.00    21.00    0.00    21
```

2. Deterministic SIR Model

```
set.seed(101)

param1 <- param.dcm(inf.prob = 0.06, act.rate = 20, rec.rate = 1/3)
init1 <- init.dcm(s.num = 20, i.num = 1, r.num = 0)
control1 <- control.dcm(type = "SIR", nsteps = 50, dt = 1)
mod1 <- dcm(param1, init1, control1)

#plotting the model
par(mar = c(3.2, 3, 2, 1), mgp = c(2, 1, 0), mfrow = c(1, 2))
plot(mod1, popfrac = FALSE, alpha = 0.5,
      lwd = 4, main = "Flow Histograms")
plot(mod1, y = "si.flow", lwd = 4, col = "firebrick",
      main = "Disease Incidence", legend = "n")
```



```
#creating output dataframe
s_num1 <- mod1$epi$s.num
i_num1 <- mod1$epi$i.num
r_num <- mod1$epi$r.num
si_flow1 <- mod1$epi$si.flow
ir_flow <- mod1$epi$ir.flow
num1 <- mod1$epi$num
dcm_sir <- data.frame(s_num1, i_num1, r_num, si_flow1, ir_flow, num1)
colnames(dcm_sir) <- c("Susceptible", "Infected", "Recovered", "SI Flow", "IR Flow", "Total")
```

```
dcm_sir <- round(dcm_sir, digits = 2)
dcm_sir
```

| ## | Susceptible | Infected | Recovered | SI | Flow | IR | Flow | Total |
|-------|-------------|----------|-----------|------|------|----|------|-------|
| ## 1 | 20.00 | 1.00 | 0.00 | 1.66 | 0.50 | | | 21 |
| ## 2 | 18.34 | 2.15 | 0.50 | 2.93 | 1.02 | | | 21 |
| ## 3 | 15.41 | 4.07 | 1.52 | 3.96 | 1.73 | | | 21 |
| ## 4 | 11.45 | 6.30 | 3.26 | 3.83 | 2.38 | | | 21 |
| ## 5 | 7.62 | 7.75 | 5.63 | 2.77 | 2.64 | | | 21 |
| ## 6 | 4.85 | 7.88 | 8.27 | 1.69 | 2.51 | | | 21 |
| ## 7 | 3.15 | 7.07 | 10.78 | 0.98 | 2.16 | | | 21 |
| ## 8 | 2.18 | 5.88 | 12.94 | 0.57 | 1.76 | | | 21 |
| ## 9 | 1.61 | 4.69 | 14.70 | 0.34 | 1.38 | | | 21 |
| ## 10 | 1.27 | 3.64 | 16.09 | 0.21 | 1.07 | | | 21 |
| ## 11 | 1.06 | 2.79 | 17.15 | 0.14 | 0.81 | | | 21 |
| ## 12 | 0.92 | 2.11 | 17.97 | 0.09 | 0.61 | | | 21 |
| ## 13 | 0.83 | 1.59 | 18.58 | 0.06 | 0.46 | | | 21 |
| ## 14 | 0.77 | 1.19 | 19.04 | 0.04 | 0.35 | | | 21 |
| ## 15 | 0.72 | 0.89 | 19.39 | 0.03 | 0.26 | | | 21 |
| ## 16 | 0.69 | 0.67 | 19.64 | 0.02 | 0.19 | | | 21 |
| ## 17 | 0.67 | 0.50 | 19.84 | 0.02 | 0.14 | | | 21 |
| ## 18 | 0.65 | 0.37 | 19.98 | 0.01 | 0.11 | | | 21 |
| ## 19 | 0.64 | 0.27 | 20.09 | 0.01 | 0.08 | | | 21 |
| ## 20 | 0.63 | 0.20 | 20.17 | 0.01 | 0.06 | | | 21 |
| ## 21 | 0.62 | 0.15 | 20.22 | 0.00 | 0.04 | | | 21 |
| ## 22 | 0.62 | 0.11 | 20.27 | 0.00 | 0.03 | | | 21 |
| ## 23 | 0.62 | 0.08 | 20.30 | 0.00 | 0.02 | | | 21 |
| ## 24 | 0.61 | 0.06 | 20.32 | 0.00 | 0.02 | | | 21 |
| ## 25 | 0.61 | 0.05 | 20.34 | 0.00 | 0.01 | | | 21 |
| ## 26 | 0.61 | 0.03 | 20.36 | 0.00 | 0.01 | | | 21 |
| ## 27 | 0.61 | 0.03 | 20.37 | 0.00 | 0.01 | | | 21 |
| ## 28 | 0.61 | 0.02 | 20.37 | 0.00 | 0.01 | | | 21 |
| ## 29 | 0.61 | 0.01 | 20.38 | 0.00 | 0.00 | | | 21 |
| ## 30 | 0.61 | 0.01 | 20.38 | 0.00 | 0.00 | | | 21 |
| ## 31 | 0.61 | 0.01 | 20.38 | 0.00 | 0.00 | | | 21 |
| ## 32 | 0.61 | 0.01 | 20.39 | 0.00 | 0.00 | | | 21 |
| ## 33 | 0.61 | 0.00 | 20.39 | 0.00 | 0.00 | | | 21 |
| ## 34 | 0.61 | 0.00 | 20.39 | 0.00 | 0.00 | | | 21 |
| ## 35 | 0.61 | 0.00 | 20.39 | 0.00 | 0.00 | | | 21 |
| ## 36 | 0.61 | 0.00 | 20.39 | 0.00 | 0.00 | | | 21 |
| ## 37 | 0.61 | 0.00 | 20.39 | 0.00 | 0.00 | | | 21 |
| ## 38 | 0.61 | 0.00 | 20.39 | 0.00 | 0.00 | | | 21 |
| ## 39 | 0.61 | 0.00 | 20.39 | 0.00 | 0.00 | | | 21 |
| ## 40 | 0.61 | 0.00 | 20.39 | 0.00 | 0.00 | | | 21 |
| ## 41 | 0.61 | 0.00 | 20.39 | 0.00 | 0.00 | | | 21 |
| ## 42 | 0.61 | 0.00 | 20.39 | 0.00 | 0.00 | | | 21 |
| ## 43 | 0.61 | 0.00 | 20.39 | 0.00 | 0.00 | | | 21 |
| ## 44 | 0.61 | 0.00 | 20.39 | 0.00 | 0.00 | | | 21 |
| ## 45 | 0.61 | 0.00 | 20.39 | 0.00 | 0.00 | | | 21 |
| ## 46 | 0.61 | 0.00 | 20.39 | 0.00 | 0.00 | | | 21 |
| ## 47 | 0.61 | 0.00 | 20.39 | 0.00 | 0.00 | | | 21 |
| ## 48 | 0.61 | 0.00 | 20.39 | 0.00 | 0.00 | | | 21 |
| ## 49 | 0.61 | 0.00 | 20.39 | 0.00 | 0.00 | | | 21 |
| ## 50 | 0.61 | 0.00 | 20.39 | 0.00 | 0.00 | | | 21 |

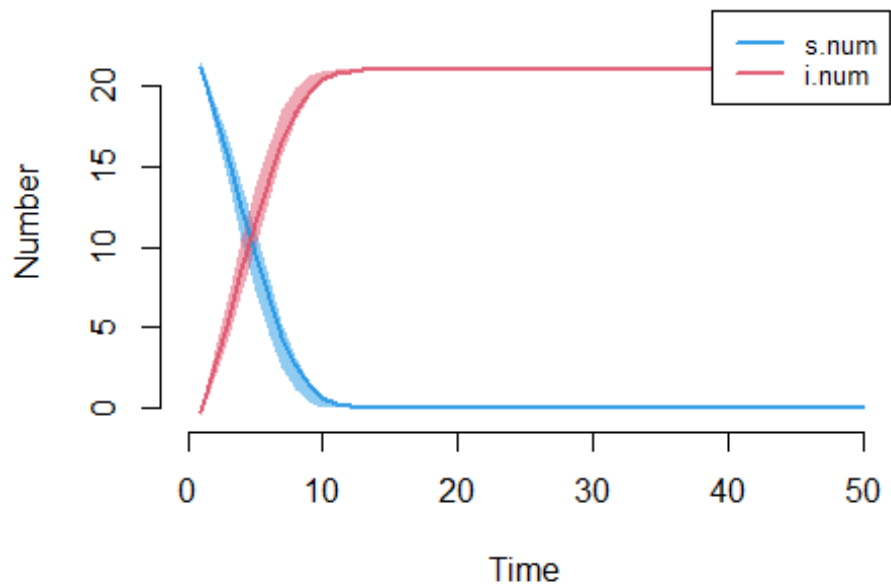
3. Stochastic SI Model

```
set.seed(101)

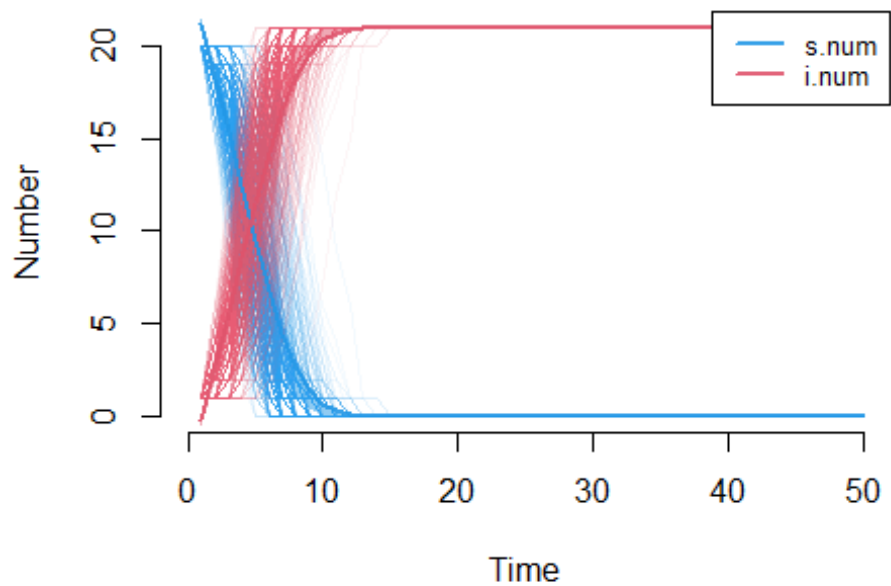
param2 <- param.icm(inf.prob = 0.06, act.rate = 20)
init2 <- init.icm(s.num = 20, i.num = 1)
control2 <- control.icm(type = "SI", nsims = 1000, nsteps = 50)
mod2 <- icm(param2, init2, control2)
mod2

## EpiModel Simulation
## =====
## Model class: icm
##
## Simulation Summary
## -----
## Model type: SI
## No. simulations: 1000
## No. time steps: 50
## No. groups: 1
##
## Model Parameters
## -----
## inf.prob = 0.06
## act.rate = 20
##
## Model Output
## -----
## Variables: s.num i.num num si.flow

#plotting the model
plot(mod2)
```



```
#with each simulation plot & average
plot(mod2, sim.lines = TRUE, mean.smooth = TRUE, qnts.smooth = TRUE)
```



```
icm_si <- round(as.data.frame(mod2, out = "mean"), digits = 2)
```

```
#average of 1000 simulations
```

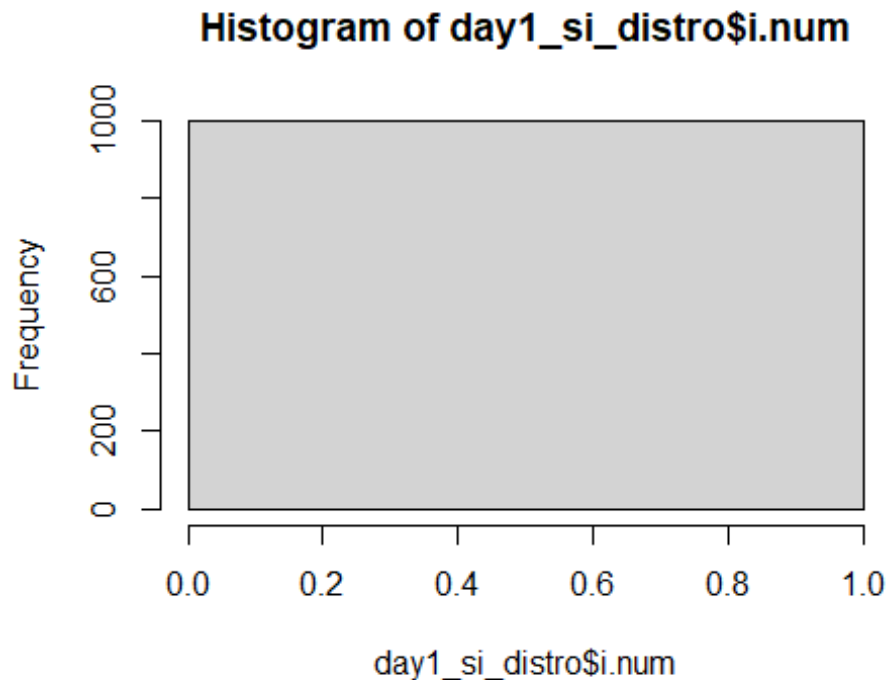
```
icm_si
```

```
##      time s.num i.num num si.flow
## 1      1 20.00  1.00  21    0.00
## 2      2 18.83  2.17  21    1.17
## 3      3 16.66  4.34  21    2.17
## 4      4 13.27  7.73  21    3.39
## 5      5  9.15 11.85  21    4.12
## 6      6  5.31 15.69  21    3.84
## 7      7  2.62 18.38  21    2.69
## 8      8  1.14 19.86  21    1.48
## 9      9  0.46 20.54  21    0.69
## 10     10 0.17 20.83  21    0.29
## 11     11 0.05 20.95  21    0.12
## 12     12 0.02 20.98  21    0.04
## 13     13 0.00 21.00  21    0.01
## 14     14 0.00 21.00  21    0.00
## 15     15 0.00 21.00  21    0.00
## 16     16 0.00 21.00  21    0.00
## 17     17 0.00 21.00  21    0.00
## 18     18 0.00 21.00  21    0.00
## 19     19 0.00 21.00  21    0.00
## 20     20 0.00 21.00  21    0.00
## 21     21 0.00 21.00  21    0.00
## 22     22 0.00 21.00  21    0.00
## 23     23 0.00 21.00  21    0.00
## 24     24 0.00 21.00  21    0.00
## 25     25 0.00 21.00  21    0.00
## 26     26 0.00 21.00  21    0.00
## 27     27 0.00 21.00  21    0.00
## 28     28 0.00 21.00  21    0.00
## 29     29 0.00 21.00  21    0.00
## 30     30 0.00 21.00  21    0.00
## 31     31 0.00 21.00  21    0.00
## 32     32 0.00 21.00  21    0.00
## 33     33 0.00 21.00  21    0.00
## 34     34 0.00 21.00  21    0.00
## 35     35 0.00 21.00  21    0.00
## 36     36 0.00 21.00  21    0.00
## 37     37 0.00 21.00  21    0.00
## 38     38 0.00 21.00  21    0.00
## 39     39 0.00 21.00  21    0.00
## 40     40 0.00 21.00  21    0.00
## 41     41 0.00 21.00  21    0.00
## 42     42 0.00 21.00  21    0.00
## 43     43 0.00 21.00  21    0.00
## 44     44 0.00 21.00  21    0.00
## 45     45 0.00 21.00  21    0.00
## 46     46 0.00 21.00  21    0.00
## 47     47 0.00 21.00  21    0.00
## 48     48 0.00 21.00  21    0.00
```

```
## 49  49  0.00 21.00  21    0.00
## 50  50  0.00 21.00  21    0.00

#day 1 data
icm_si1 <- round(as.data.frame(mod2, out = "vals"), digits = 2)
day1_si_distro <- icm_si1[icm_si1$time==1,]

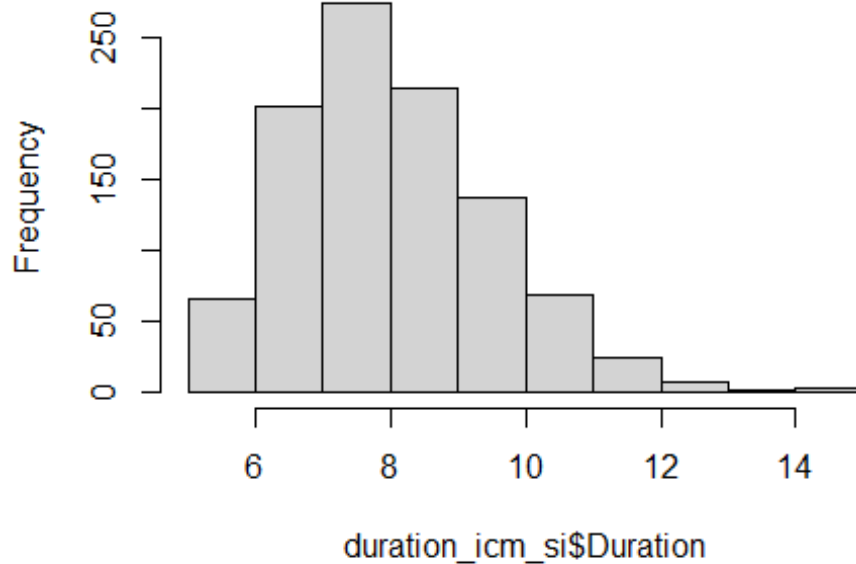
#day 1 infections histogram
hist(day1_si_distro$i.num)
```



```
#extracting epidemic duration for each simulation
duration_si_distro <- icm_si1[icm_si1$s.num>0,]
duration_icm_si <- data.frame(table(duration_si_distro$sim))
colnames(duration_icm_si) <- c("Simulation", "Duration")
duration_icm_si$Duration <- duration_icm_si$Duration+1

#epidemic duration histogram, mean & median
hist(duration_icm_si$Duration)
```

Histogram of duration_icm_si\$Duration



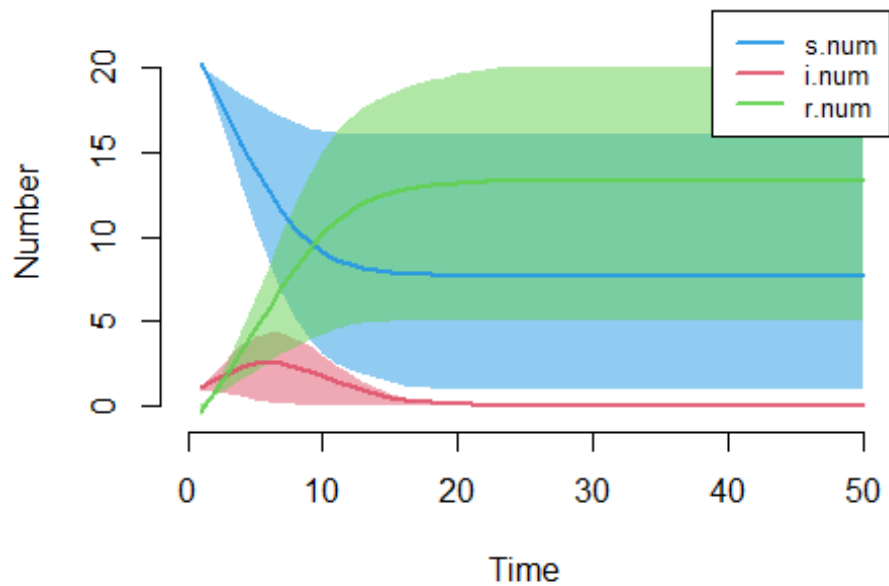
```
mean(duration_icm_si$Duration)
## [1] 8.521
median(duration_icm_si$Duration)
## [1] 8
```

4. Stochastic SIR Model

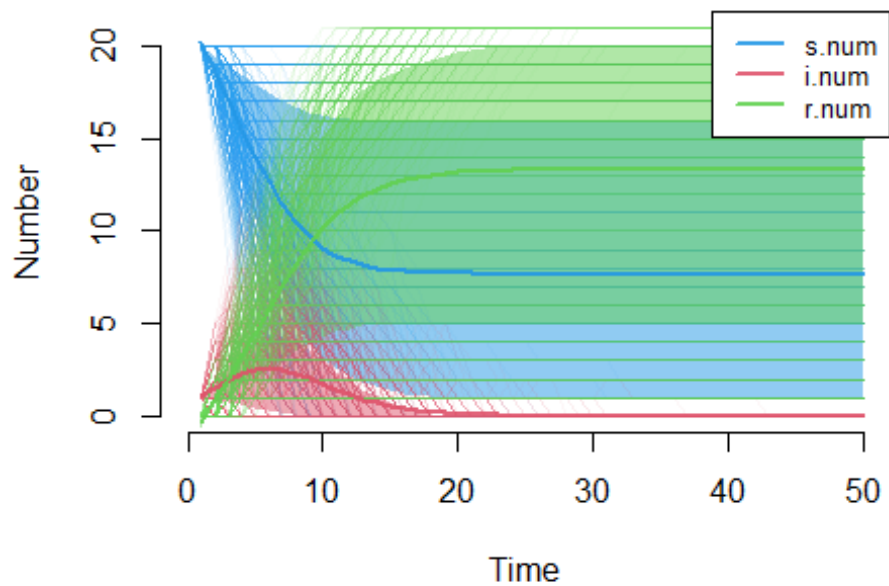
```
set.seed(101)

param3 <- param.icm(inf.prob = 0.06, act.rate = 20, rec.rate = 1/3)
init3 <- init.icm(s.num = 20, i.num = 1, r.num = 0)
control3 <- control.icm(type = "SIR", nsteps = 50, nsims = 1000)
mod3 <- icm(param3, init3, control3)

#plotting the model
plot(mod3)
```



```
#each simulation plot with average
plot(mod3, sim.lines = TRUE, mean.smooth = TRUE, qnts.smooth = TRUE)
```




```
icm_sir <- round(as.data.frame(mod3, out = "mean"), digits = 2)
```

```
#average of 1000 simulations
```

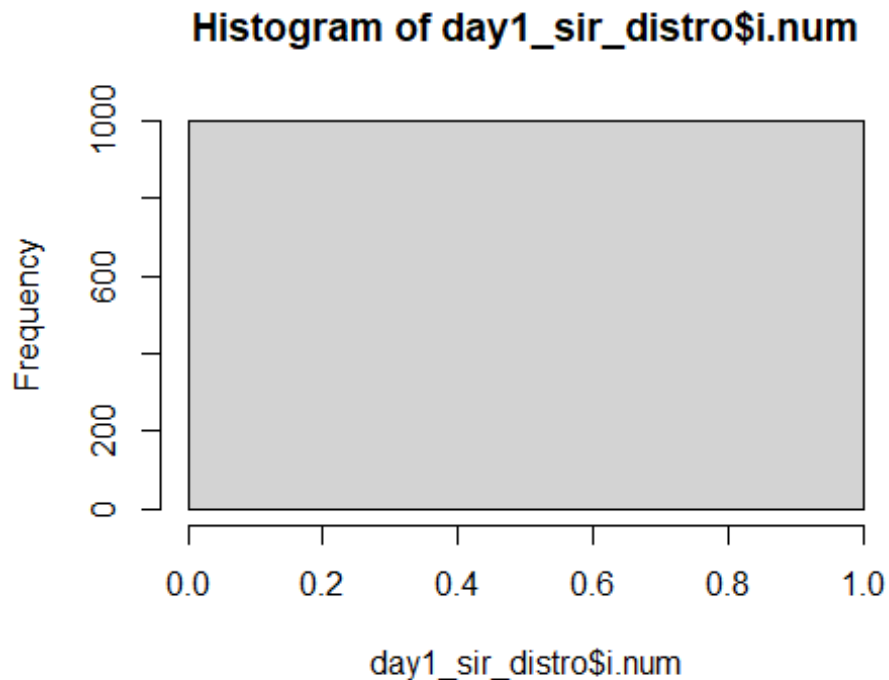
```
icm_sir
```

```
##      time s.num i.num num r.num si.flow ir.flow
## 1      1 20.00  1.00  21  0.00   0.00   0.00
## 2      2 18.80  1.53  21  0.67   1.20   0.67
## 3      3 17.22  2.08  21  1.70   1.58   1.03
## 4      4 15.44  2.56  21  3.00   1.78   1.31
## 5      5 13.72  2.88  21  4.39   1.71   1.39
## 6      6 12.19  2.95  21  5.86   1.53   1.47
## 7      7 10.94  2.75  21  7.31   1.25   1.45
## 8      8 10.00  2.48  21  8.52   0.94   1.21
## 9      9  9.26  2.14  21  9.59   0.74   1.07
## 10     10  8.76  1.74  21 10.50   0.50   0.91
## 11     11  8.43  1.37  21 11.20   0.33   0.69
## 12     12  8.21  1.07  21 11.72   0.22   0.53
## 13     13  8.05  0.81  21 12.14   0.16   0.41
## 14     14  7.93  0.63  21 12.44   0.12   0.30
## 15     15  7.86  0.46  21 12.68   0.07   0.24
## 16     16  7.79  0.34  21 12.87   0.07   0.19
## 17     17  7.75  0.26  21 12.98   0.04   0.11
## 18     18  7.72  0.19  21 13.09   0.03   0.11
## 19     19  7.71  0.13  21 13.16   0.01   0.07
## 20     20  7.70  0.10  21 13.20   0.01   0.05
## 21     21  7.69  0.07  21 13.23   0.00   0.03
## 22     22  7.69  0.05  21 13.26   0.00   0.02
## 23     23  7.69  0.03  21 13.28   0.00   0.02
## 24     24  7.69  0.02  21 13.29   0.00   0.01
## 25     25  7.69  0.02  21 13.29   0.00   0.00
## 26     26  7.69  0.01  21 13.30   0.00   0.01
## 27     27  7.69  0.01  21 13.30   0.00   0.00
## 28     28  7.69  0.01  21 13.30   0.00   0.00
## 29     29  7.69  0.01  21 13.31   0.00   0.00
## 30     30  7.69  0.00  21 13.31   0.00   0.00
## 31     31  7.68  0.00  21 13.31   0.00   0.00
## 32     32  7.68  0.00  21 13.31   0.00   0.00
## 33     33  7.68  0.00  21 13.31   0.00   0.00
## 34     34  7.68  0.00  21 13.31   0.00   0.00
## 35     35  7.68  0.00  21 13.31   0.00   0.00
## 36     36  7.68  0.00  21 13.31   0.00   0.00
## 37     37  7.68  0.00  21 13.31   0.00   0.00
## 38     38  7.68  0.00  21 13.31   0.00   0.00
## 39     39  7.68  0.00  21 13.31   0.00   0.00
## 40     40  7.68  0.00  21 13.31   0.00   0.00
## 41     41  7.68  0.00  21 13.31   0.00   0.00
## 42     42  7.68  0.00  21 13.31   0.00   0.00
## 43     43  7.68  0.00  21 13.32   0.00   0.00
## 44     44  7.68  0.00  21 13.32   0.00   0.00
## 45     45  7.68  0.00  21 13.32   0.00   0.00
## 46     46  7.68  0.00  21 13.32   0.00   0.00
## 47     47  7.68  0.00  21 13.32   0.00   0.00
## 48     48  7.68  0.00  21 13.32   0.00   0.00
```

```
## 49    49  7.68  0.00  21 13.32    0.00    0.00
## 50    50  7.68  0.00  21 13.32    0.00    0.00

#day 1 data
icm_sir1 <- round(as.data.frame(mod3, out = "vals"), digits = 2)
day1_sir_distro <- icm_sir1[icm_sir1$time==1,]

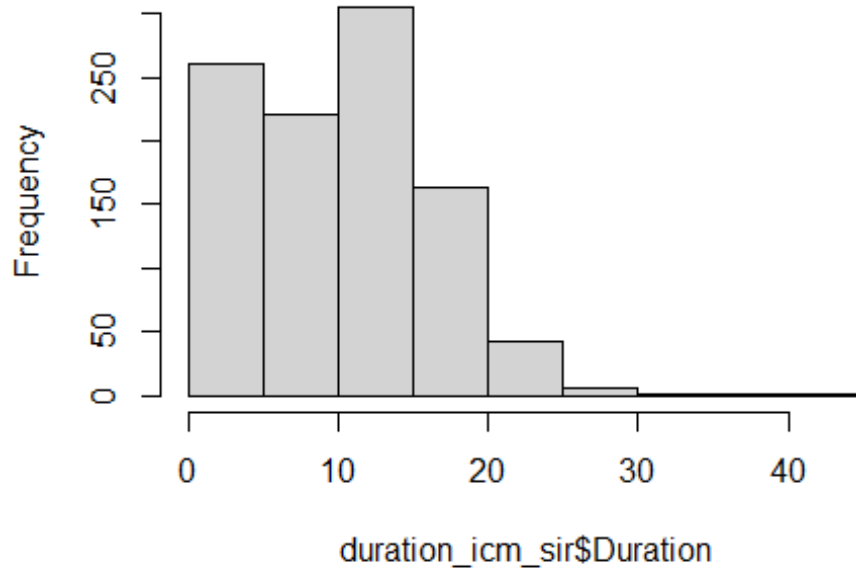
#day 1 infections histogram
hist(day1_sir_distro$i.num)
```



```
#extracting epidemic duration for each simulation
duration_sir_distro <- icm_sir1[icm_sir1$s.num>0 ,]
duration_sir_distro <- duration_sir_distro[duration_sir_distro$i.num>0 ,]
duration_icm_sir <- data.frame(table(duration_sir_distro$sim))
colnames(duration_icm_sir) <- c("Simulation", "Duration")
duration_icm_sir$Duration <- duration_icm_sir$Duration+1

#epidemic duration histogram, mean & median
hist(duration_icm_sir$Duration)
```

Histogram of duration_icm_sir\$Duration



```
mean(duration_icm_sir$Duration)
## [1] 10.632
median(duration_icm_sir$Duration)
## [1] 11
```