

**I. (1 pt) mathbioLibrary/setABioc2Labs/Lab122.L3.SIR.pg**

Because of the accuracy of WebWork, you should use 4 or 5 significant figures on this problem.

Influenza or flu is a viral infection that is readily transmitted through the air and causes respiratory problems in humans and other animals. It occurs seasonally and results in an average of 30,000 deaths in the U. S. annually primarily among the young and old. There have been some epidemics, such as the pandemic of 1918, which resulted in over 30 million deaths around the world.

There are two primary types, A and B of influenza, which appear seasonally, usually starting in the fall and continuing through the winter months. These types mutate creating new strains each year. Flu is an RNA virus with 11 genes that mutate regularly, making it difficult to create general vaccines. Each year scientists anticipate the most likely dominant strain and determine which mutant is most likely to spread in order to create a vaccine to protect the population. In addition to the many deaths, this disease results in huge economic losses from lost work and treatment of fragile patients.

The disease follows a classic mathematical model known as an SIR model. The population can be divided into susceptible, infected, and recovered individuals. Once a person has a particular strain of one of the types of the influenza virus, then that individual develops immunity to that specific strain, preventing reinfection. (This is why flu viruses need to mutate to keep finding new hosts.) Since flu acts over a very short period of time, we will simplify our model by assuming that the population remains constant with a size of  $N$ . This means that it is sufficient to simply keep track of two classes of individuals, the susceptibles,  $S_n$ , and the infected,  $I_n$ , as the recovered satisfy

$$R_n = N - S_n - I_n.$$

Since we are ignoring births and deaths in the population, we only consider the spread of the disease, which is based on successful contacts (such as inhaling aerosol infected particles from the sneeze of an infected individual into the lungs of a susceptible host). The discrete mathematical model is given by the system of equations:

$$S_{n+1} = S_n - \frac{\beta}{N} S_n I_n,$$

$$I_{n+1} = I_n(1 - \gamma) + \frac{\beta}{N} S_n I_n,$$

where  $\beta S/N$  represents the proportion of contacts by an infected individual that result in the infection of a susceptible individual. The parameter  $\gamma$  is the probability that an infected person recovers (enters class  $R$  of the SIR model). The ratio  $1/\gamma$  is the average length of the infectious period of the disease.

Below we provide data from the Center of Disease Control for one season of flu. The table below gives the number of cases week by week for a control set of individuals. In particular, we consider the data for a particular strain of Type A for the flu season in 2005 and 2006. Assume that the control population

consists of  $N = 148636$ . The first week, ( $n = 0$ ), corresponds to the last week in September, which is near the seasonal start of the flu season in the U. S.

$n$ (wk)	$I_n$	$n$ (wk)	$I_n$	$n$ (wk)	$I_n$
0	4	17	468	34	28
1	4	18	508	35	19
2	4	19	578	36	8
3	6	20	730	37	11
4	5	21	812	38	2
5	4	22	781	39	2
6	13	23	784	40	4
7	16	24	551	41	6
8	17	25	456	42	9
9	36	26	301	43	8
10	69	27	180	44	16
11	106	28	128	45	9
12	328	29	67	46	7
13	343	30	43	47	9
14	361	31	49	48	8
15	353	32	45	-	-
16	375	33	41	-	-

a. We want to simulate the SIR model above, finding the best parameters,  $\beta$  and  $\gamma$ , that match the data from the CDC. Insert the week number,  $n$ , and the data on infected individuals in the first two columns. Use Columns C and D for the simulation of the discrete model,  $S_n$ , and  $I_n$ . For initial values, take  $I_0 = 4$  and  $S_0 = N - I_0 = 148632$ . Define your parameters,  $\beta$  and  $\gamma$ , and take your initial guesses for these parameters to be  $\beta = 3.7$  and  $\gamma = 3.3$ . In Column E, compute the square error between the number of infected individuals in the data and those found by the model. Use Excel's solver to find the best fitting parameters  $\beta$  and  $\gamma$  that minimize the sum of square errors of the infected individuals. Also, give the least sum of square errors.

$\beta =$  \_\_\_\_\_

$\gamma =$  \_\_\_\_\_

$SSE =$  \_\_\_\_\_

Give the model prediction for people infected with influenza at  $n = 15$  and  $n = 25$  and find the percent error at each of these times from the actual CDC data given:

$I_{15} =$  \_\_\_\_\_

Percent Error at  $n = 15$  \_\_\_\_\_

$I_{25} =$  \_\_\_\_\_

Percent Error at  $n = 25$  \_\_\_\_\_

As noted above, the average length of the infectious period is equal to  $1/\gamma$ . Find this period in units of days. (Recall that  $n$  is in weeks.)

Infectious Period = \_\_\_\_\_ days

Epidemiologists often examine what is called the *basic reproduction ratio* given by

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

which provides a measure of how rapidly a disease will spread and how much of the population will be affected by a particular disease. Use your values of  $\beta$  and  $\gamma$  to find  $R_0$ .

$R_0 =$  \_\_\_\_\_

To determine the impact of a particular flu season, we want to know the total number of individuals who were infected by the influenza virus. Since we are assuming a constant population  $N$  and because the number of infected individuals is essentially zero at the end of the simulation. We estimate the total number of cases of flu by computing the number in the recovered class, so

$$R_n = N - S_n,$$

for  $n$  large. Estimate the number of Influenza Type A cases for the 2005 - 2006 flu season. What percent of the original population ultimately got this strain of influenza?

Total Cases = \_\_\_\_\_

Percent Infected = \_\_\_\_\_

b. In your Lab report, create one graph with the data and the model of the individuals infected with influenza. Create another graph showing the number of susceptible individuals for this particular strain of influenza over the 48 weeks. Write a brief paragraph discussing how well the model fits the data. Also, discuss if the predicted value of  $\gamma$  gives a reasonable estimate of the infectious period for the flu.

c. The CDC is interested in minimizing the impact of flu on the population, so uses a number of different controls. We will examine three different controls that are employed to fight outbreaks of the flu.

The first line of defense of which you are undoubtedly aware is the annual flu vaccine. The effect of a vaccine is to lower the population of susceptible individuals, which in terms of our model is simply to lower  $S_0$  by moving a number of individuals to  $R_0$ . (In fact, many individuals already have immunity to a given strain of the flu because of earlier contact with a related strain.)

Suppose we vaccinate 5 percent of the total population at the very beginning of the flu season. This immediately shifts some of the population to the recovered individuals in the model. We calculate the initial susceptible population by

$$S_0 = 0.95N - I_0,$$

so

$$R_0 = 0.05N.$$

Assuming no other changes for this epidemic, then the values of  $\gamma$  and  $\beta$  remain the same as above. Simulate the SIR model with this new initial condition. Find the number of infecteds and susceptibles at  $n = 25$  and  $n = 32$ . Also, simulate the model sufficiently long, so there are essentially no new infecteds and determine the total number of people who would have suffered from the flu if 5 percent of the population is vaccinated. This is given by the number of individuals in the recovered class (subtracting out the ones protected by the vaccine). Since the flu

epidemic has faded out ( $I = 0$ ),  $R_n \approx N - S_n$  for  $n$  large. An estimate of the number of people getting the flu is given by  $R_n - R_0$ . What percent of the population becomes infected?

$I_{25} =$  \_\_\_\_\_

$S_{25} =$  \_\_\_\_\_

$I_{32} =$  \_\_\_\_\_

$S_{32} =$  \_\_\_\_\_

Total Cases = \_\_\_\_\_

Percent Infected = \_\_\_\_\_

d. A second method of control is the quarantine of infected individuals or the education of the public on how to lessen the contact between infected and susceptible individuals. This lowers the value of  $\beta$ . Suppose we lower the value of  $\beta$  by 5 percent, which is equivalent to multiplying the value of  $\beta$  found in Part a by 0.95. Use the original initial conditions and best fitting value of  $\gamma$  and this new  $\beta$  value to simulate the SIR model. Find the number of infecteds and susceptibles at  $n = 25$  and  $n = 32$ . Also, simulate the model sufficiently long, so there are essentially no new infecteds and determine the total number of people who would have suffered from the flu if this control was used. (Recall that since the flu epidemic has faded out ( $I = 0$ ), the people who got the flu are the people in the recovered class. Thus,  $R_n \approx N - S_n$  for  $n$  large. What percent of the population becomes infected?

$I_{25} =$  \_\_\_\_\_

$S_{25} =$  \_\_\_\_\_

$I_{32} =$  \_\_\_\_\_

$S_{32} =$  \_\_\_\_\_

Total Cases = \_\_\_\_\_

Percent Infected = \_\_\_\_\_

e. Oseltamivir or Tamiflu is a drug that shortens the symptoms of flu for many people. If we assume that this drug shortens the length of the period of infectivity of the infected individuals, then this can be modeled by increasing  $\gamma$ . Suppose we increase the value of  $\gamma$  by 3 percent, which is equivalent to multiplying the value of  $\gamma$  found in Part a by 1.03. Use the original initial conditions and best fitting value of  $\beta$  to simulate the SIR model. Find the number of infecteds and susceptibles at  $n = 25$  and  $n = 32$ . Also, simulate the model sufficiently long, so there are essentially no new infected individuals and determine the total number of people who would have suffered from the flu if this drug was used. What percent of the population becomes infected?

$I_{25} =$  \_\_\_\_\_

$S_{25} =$  \_\_\_\_\_

$I_{32} =$  \_\_\_\_\_

$S_{32} =$  \_\_\_\_\_

Total Cases = \_\_\_\_\_

Percent Infected = \_\_\_\_\_

f. In your Lab report, reproduce the graph with the original data and the best fitting SIR model. Extend all models to 100 weeks. Add graphs of the infected individuals from the SIR model using the three means of controlling the disease discussed above for  $t \in [0, 100]$ . Be sure to label each type of control on

the graph. Describe what you observe in your graphs both quantitatively and qualitatively. Compare and contrast the different approaches to controlling a flu outbreak. Give advantages and disadvantages of each of the controls. Include a discussion of the practicality and financial burden of each of these approach. Give a couple of strengths and weaknesses for this SIR model. Select another disease that satisfies the SIR model and discuss

how this lab applies to treatment of the disease you are considering. How does this study apply to elimination of other serious diseases?

3

[1] CDC Flu website - [www.cdc.gov/flu/](http://www.cdc.gov/flu/) (last visited Sept. 2009).

[2] Wikipedia - [en.wikipedia.org/wiki/Influenza](http://en.wikipedia.org/wiki/Influenza) (last visited Sept. 2009).