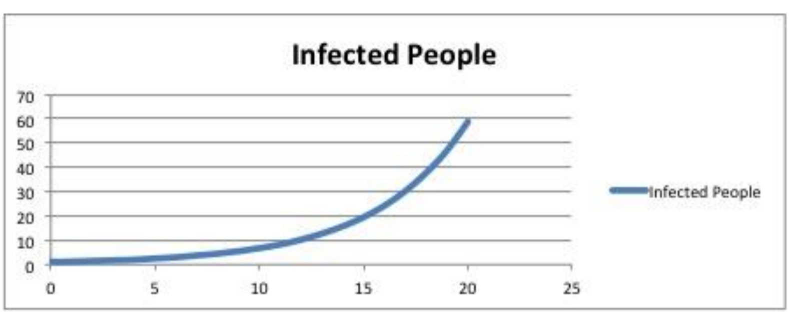
# SIR: Modeling an Epidemic, Part I [for BLC-150]

Let's say that you start with a Total population of 100 people... and of that total, one person, [Patient Zero, $P_0$] becomes infected by some sort of crossover event... maybe the he gets infected by a bat-born germ (we'll assume it's a "he" for no particular reason).

Let's say that the infection rate once unleashed in the human population is 25% per day, meaning that each day a single infected person has a 25% *chance* of infecting someone else. For simplicity, instead of *chance*, we'll use a rate. So a single infected person with a 25% infection rate per day will on average infect 0.25 people per day.... or 1 full person every 4 days. Follow? In this particular model on Day-1 Patient-Zero, $P_0$, gives his disease to 0.25 people. Then on Day-2 he has spread the disease to 0.5 people... Day-3 to 0.75 and finally, as you can deduce, by Day-4 he has given it to a full person.

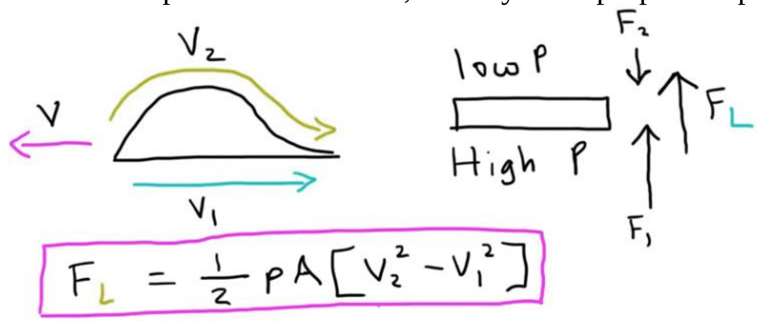
On Day-4 there are 2 infected people. That's going to double the rate of infection as it relates to the Total population. There are now 2 people passing on the disease at a rate of 0.25 per day. So, on Day-5 these two infected people will have given the disease to 0.25 + 0.25 = 0.5 people and on Day-6 to another full person. So on Day-6 we have 3 Infected people. Day-7 these 3 infected people pass it on to 0.75 people, and on Day-8 to an additional 1.5 people. The total on Day-8 is now 4 full Infected people (and a half of a person). Day-9 there will be 5 Infected people and a half person, Day-10 will have 6 and ¾ of a person... etc.

| Day # | Fractional Infected People | Number of Fully Infected People | + | Fractional Remainder of Infected People |
|---|---|---|---|---|
| 0 | 1 | 1 | + | 0 |
| 1 | 1.25 | 1 | + | 0.25 |
| 2 | 1.5 | 1 | + | 0.5 |
| 3 | 1.75 | 1 | + | 0.75 |
| 4 | 2 | 2 | + | 0 |
| 5 | 2.5 | 2 | + | 0.5 |
| 6 | 3 | 3 | + | 0 |
| 7 | 3.75 | 3 | + | 0.75 |
| 8 | 4.5 | 4 | + | 0.5 |
| 9 | 5.5 | 5 | + | 0.5 |
| 10 | 6.75 | 6 | + | 0.75 |
| 11 | 8.25 | 8 | + | 0.25 |
| 12 | 10.25 | 10 | + | 0.25 |
| 13 | 12.75 | 12 | + | 0.75 |
| 14 | 15.75 | 15 | + | 0.75 |
| 15 | 19.5 | 19 | + | 0.5 |
| 16 | 24.25 | 24 | + | 0.25 |
| 17 | 30.25 | 30 | + | 0.25 |
| 18 | 37.75 | 37 | + | 0.75 |
| 19 | 47 | 47 | + | 0 |
| 20 | 58.75 | 58 | + | 0.75 |



I've made some assumptions in this model–assumptions about how this world works. It's not a very realistic world. That's what models are. They are simplified worlds. A model airplane represents what a real airplane looks like, but it doesn't represent how a real airplane flies. The real world is way too complicated. That's why we make models. We make them in order to isolate what we think matters in a particular situation. For the model airplane, what matters is modeling its appearance.


But a mathematical model of its aerodynamics will not give you much information about what the plane looks like. It won't include the teeth painted on the side, or the yellow propeller tips, or its overall color.




The infection model described above represented how a single Infected person spread a virus. But the mechanism for infection was quite simple. I assumed that only fully infected people actually have the virus and can spread it. The partially infected (0.25s, 0.5s and 0.75s) are treated as uninfected. Only when they are a whole number do they become Infected and infectious. The above model tells a story of sorts: How an infected population grows. It has nothing to say about the world of the uninfected.

---

The model is rather 1-dimensional. I imagine an infected person as a drunk bumbling idiot bouncing around in a crowd of mindless people in a trance. Every once in a while he transmits his germs to one of them.... on average... this occurs once every 4 days (for an infection rate of ¼ per day.) As the days go by you get more and more bumbling idiots bouncing off the unlimited supply of the hapless healthy.

A big assumption in this model is that all infected people are equal: equally infectious and have equal access to the healthy population. This assumption is represented by the 25% infection rate. There is no variation. For example, a 90-year-old Infected person and a 20-year-old Infected person are treated as equals. Both can spread the virus with equal efficacy. This is patently absurd. A 20-year-old will most likely be much more mobile and meet more people in a day. This model averages out these details. Maybe the typical Infected 20-year-old has a 35% infection rate per day and the 90-year-old has a 15% rate. These rates average to 25%. If we were studying infection rates as a function of age, we would want to incorporate these details. But details complicate models and models are about simplification.

Q: Why simplification, you ask?
A: Because we already have the ultimate model: REALITY. Reality is too complicated to understand.

---

Another assumption in the infection model described above was that the healthy population was always available to be infected. There was no limit to the supply of victims. The Infected population could grow without limit. In the real world this is not typically the case. Eventually Infected people can't find new victims because everybody they meet is Infected (or immune after recovering from an infection). These details are absent from the above model. Why? For simplicity. But simplicity is... well... simple. I'd like to know how an infection spreads and accelerates and also how it slows down. From experience and history we know that infections, (epidemics, pandemics, plagues, etc.) eventually end. They die out. To be more realistic (and optimistic) we need more detail. We need something to push back against the ever-rising Infected population.

---

Enter the ***S-I*** **Model**. In this model we'll have Infected people, *I*, as before, but this time we'll have a new population, the Susceptibles, *S*. The previous model assumed an unlimited supply of people who could be infected. This new model will limit the population of victims... what we'll call the Susceptibles. The effect will be that as the Infected numbers rise, the Susceptible numbers will fall.

To do this we will put a limitation on the overall population. This is more realistic in terms of the dynamics of an infection in reality. This will allow us to put a break on infections as the numbers of Susceptibles fall. This adds a level of realism to the model. Susceptibles can become Infecteds. However, Infecteds cannot become Susceptibles. The Infected population goes up, which means that the Susceptible population goes down. The same Total number of people, but their categories switch. Simple.

Restated: The Susceptible population will drop as people become Infected.
Let's say the infection rate, *i*, is ¼, like before. In our scenario this might mean that each infected person will infect another person, on average, once every 4 days. If we start on Day-1 with **4 infected people** in the population, then one additional person per day would become infected. However, this system is dynamic... On Day-2 there are 5 Infected people, on Day-3 there should be 6.25 people... etc. But there is another power at play. As more people become infected, there are less people to infect, so the ¼ rate is diminished because a day's worth of encounters will be with fewer and fewer Susceptible people. As the Infected population gets really high, the Susceptible population must become very low, and thus towards the end of a hypothetical outbreak an Infected person rarely meets a Susceptible person and the epidemic stops.

---

**For example: Everybody is Susceptible**. Let's say on a given day for an infection rate of ¼ per day, the Infected population is 8, and that the Susceptible population is 2000, and that the Total population, T, is 2000.

**S = 2000 and T = 2000**.

This means that everybody is Susceptible. The percentage of Susceptible out of the Total population is $\frac{2000}{2000}$, or 100%. The portion of the Total population who are Susceptible is 100%.

Here is the math for the situation where 8 Infected people are wandering around on Day-1:

$$I_{infected} = \frac{1}{4}(8)\left(\frac{S}{T}\right) = ? \;\rightarrow\; \frac{1}{4}(8)\left(\frac{2000}{2000}\right) = 2.$$

Eight people with an infection rate of ¼ per day, infect 2 whole people in 1 day.

---

**Another example: S = ½ T**: Now let's look at a situation in which only half of the Total population are Susceptible. **S = 1000 and T = 2000.**

[One thousand Susceptibles out of a Total Population of 2000]

The math...

$$\frac{1}{4}(8)\left(\frac{S}{T}\right) = ? \;\rightarrow\; \frac{1}{4}(8)\left(\frac{1000}{2000}\right) = 1$$

8 Infected people (carrying an infection rate of 1/4) only infect 1 person this time, because half the people they meet are not Susceptible.

The $\left(\frac{S}{T}\right)$ term is the percentage of the population an infected person meets who are Susceptible.

In this example S = T/2, which means that $\left(\frac{S}{T}\right)$ is 50%, and the infection rate is halved.

$$S = \frac{T}{2} \quad \rightarrow \quad S \times \frac{1}{T} = \frac{T}{2} \times \frac{1}{T} = \frac{S}{T} = \frac{1}{2}$$

It might make more sense to write it in this order, which shows the infection rate being halved to $\frac{1}{8}$.

$$\frac{1}{4}\left(\frac{1000}{2000}\right)(8) = \frac{1}{4}\left(\frac{1}{2}\right)(8) = \frac{1}{8} \cdot 8 = 1$$

---

**Another Example: S = 0.** Nobody is Susceptible.

If S = 0, then $\left(\frac{S}{T}\right)$ is 0% and nobody gets infected, no matter how many Infected people there are.

$$\frac{1}{4}(8)\left(\frac{S}{T}\right) = ? \;\rightarrow\; \frac{1}{4}(8)\left(\frac{0}{2000}\right) = 0$$

---

Restating:

When $\left(\frac{S}{T}\right) = \frac{1}{2}$ = 50% the infection rate was halved due to the fact that only half of the uninfected population was available for infection. If half were available, the other half must have been unavailable.

Q: What does it mean to be unavailable?

A: They might be immune, or are immune due to vaccination, or they had been **already infected**.

Oops. That's actually an issue. Being "already infected." We need to take into account that infected people might meet other infected people. Fortunately this can be easily dealt with. We simply need to account for the infected people in the overall Total population count. Infected people are people too. So in our example above, we should really have included the infected people in the overall total.

So, to add some more realism to this model, let's do that. Let's look at ¼ infection rate per day, a Susceptible population of 1000, an unSusceptible population of 1000, and 8 Infected people staggering around in and amongst the population at large. That puts the Total population, including the Infected, at 2,008.

This changes the results slightly.

$$\frac{1}{4}(8)\left(\frac{S}{T}\right) = ? \;\rightarrow\; \frac{1}{4}(8)\left(\frac{1000}{2008}\right) \approx 0.996.$$

With this added subtlety, only 99.6% of a person was infected on that day.

Of course, this is a weird characteristic of this model and of statistics in general.

What constitutes 99.6% of a person?

Unfortunately, making this model output whole numbers instead of these partial people adds unnecessary complexity to the execution of the model in Excel (or GoogleSheets or Apple Numbers). If this bothers you, just round it to the nearest whole number and call it a day. This model is already pretty simplified, so rounding a result to a whole number doesn't do much damage.

---

End of S-I model with a Total population restriction.

---

Now let's make a GoogleSheet that does this....

Search on "Google Sheets" and open up a spread sheet to work on.

Then, in another tab...

Go to video: Making a GoogleSheet Model.

GoogleSheet Tutorial-pts1-3 [video-49']

Link to the video is also on our class page: BLC150-2020Sp/FrontPorch150.html

Additional SIR material: Optional reading for BLC-150

Scroll Down for Part II

⇓

⇓

⇓

⇓

⇓

⇓

⇓

⇓

⇓

⇓

⇓

⇓

# SIR: Modeling an Epidemic, Part II - The Full S-I-R

In Part I we developed the idea of infection rates and populations. In the exercises we looked at how various parameters influenced infection over the time span of 1 day.

In this chapter we are going to make a fully functioning SIR Population Model.

Here are the basic formulas we worked with previously in a slightly more generalized form:
*Editor's Note: I used subscripts instead of functional notation in order to keep the formulas uncluttered.*

| | |
|---|---|
| $I_{n+1} = I_n + iI_n\left(\frac{S_n}{T}\right)$ | E.g. For *n = 0*. $I_1 = I_0 + iI_0\left(\frac{S_0}{T}\right)$<br>The Infected population on Day-1 equals the Infected population of Day-0 <u>plus</u> the infection rate, *i*, times the Infected population of Day-0 times the Susceptibles on Day-0 divided by the Total population.<br>Recall that $\left(\frac{S_n}{T}\right)$ is just the percentage of the Total population who can be infected. |
| $S_{n+1} = S_n - iI_n\left(\frac{S_n}{T}\right)$ | Similarly....<br>E.g. For *n = 0*. $S_1 = S_0 - iI_0\left(\frac{S_0}{T}\right)$<br>The Infected population on Day-1 equals the Infected population of Day-0 <u>minus</u> the infection rate, *i*, times the Infected population of Day-0 times the Susceptibles on Day-0 divided by the Total population.<br>Recall that $\left(\frac{S_n}{T}\right)$ is just the percentage of the Total population who can be infected. |
| Add these two equations together to get the Total population, *T*: | $T = I_{n+1} + S_{n+1} = I_n + iI_n\left(\frac{S_n}{T}\right) + S_n - iI_n\left(\frac{S_n}{T}\right)$ |
| Notice what cancels: | $T = I_{n+1} + S_{n+1} = I_n \cancel{+ iI_n\left(\frac{S_n}{T}\right)} + S_n \cancel{- iI_n\left(\frac{S_n}{T}\right)}$ |
| This means that ... | $T = I_{n+1} + S_{n+1} = I_n + S_n$ ...which makes sense. |
| The Total population remains the same, but the distribution between I and S changes.<br>The $iI_n\left(\frac{S_n}{T}\right)$ just transfers Susceptibes to Infecteds. | |

Now, finally, we are going to add the Recovered population to the model. The Recovered population is controlled by the recovery rate, *r*. This rate is very much like the infection rate, *i*. It is a rate per day.

E.g. A recovery rate of 0.1 means that an Infected person recovers at a rate of 0.1 per day, meaning that he fully recovers after 10 days. If you have 10 Infected people, one whole person will recover per day. I realize that this conceit is weird if you think about it, but it is part of how this model works. Just don't think about it.

The recovery rate, r, only affects the Infected population.

Q: Why?

A: Because uninfected people don't need to recover from anything.

At last, here are the SIR equations in all of their glory:

| $S_{n+1} = S_n - iI_n\left(\frac{S_n}{T}\right)$ | $I_{n+1} = I_n + iI_n\left(\frac{S_n}{T}\right) - rI_n$ | $R_{n+1} = R_n + rI_n$ | Total Population |
|---|---|---|---|
| | Notice the $-rI_n$ and the $+rI_n$. They cancel each other out, but they send people from the Infected population into the Recovered population... ...just like how the $iI_n\left(\frac{S_n}{T}\right)$ sends Susceptibles into the Infected population. | | |

| $(S - \beta) + (I + \beta - \mu) + (R + \mu) = T$ or $S + I + R = T$ | This equation is sort of stupid when you look at the algebra. It just reduces to $S + I + R = T$. It sort of reminds me of *Completing the Square*... in that we add and subtract the same thing from it, maintaining equality, but producing interesting internal effects. |
|---|---|

Now let's take it out for a spin.
First we need to choose the initial conditions. Let's work with these numbers...

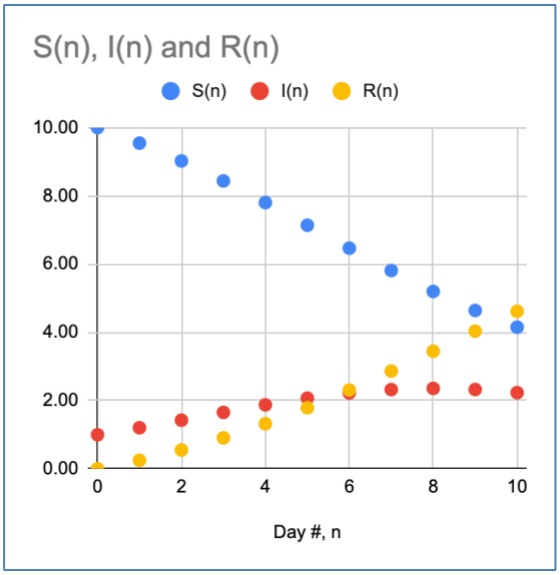
| infection rate, ***i*** | recovery rate, ***r*** | Initial Susceptible Population, $S_0$ | Initial Infected Population, $I_0$ | Initial Recovered Population, $R_0$ | Total Population, *S*+*I*+*R*= ***T*** |
|---|---|---|---|---|---|
| 50% | 25% | 10 | 1 | 0 | 11 |

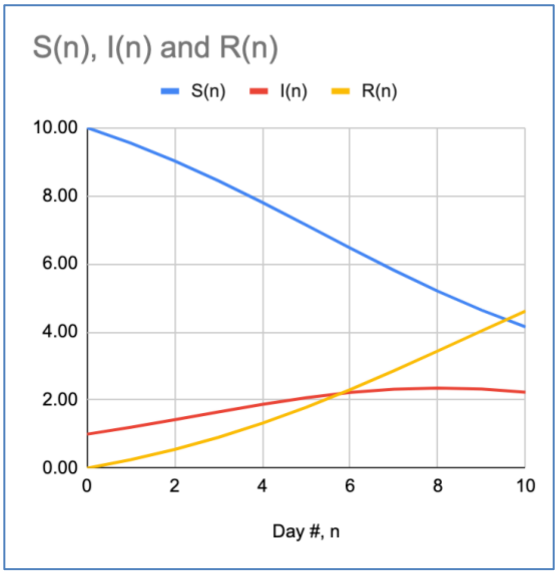
Once we have initial conditions we can let our SIR machine crank out new population distributions.
Editor's Note: We need initial conditions in order for our equations to work on something.

Below: I showed all the calculation for Day-1, some of it for Day-2,
and then I used a spreadsheet to work out the rest.

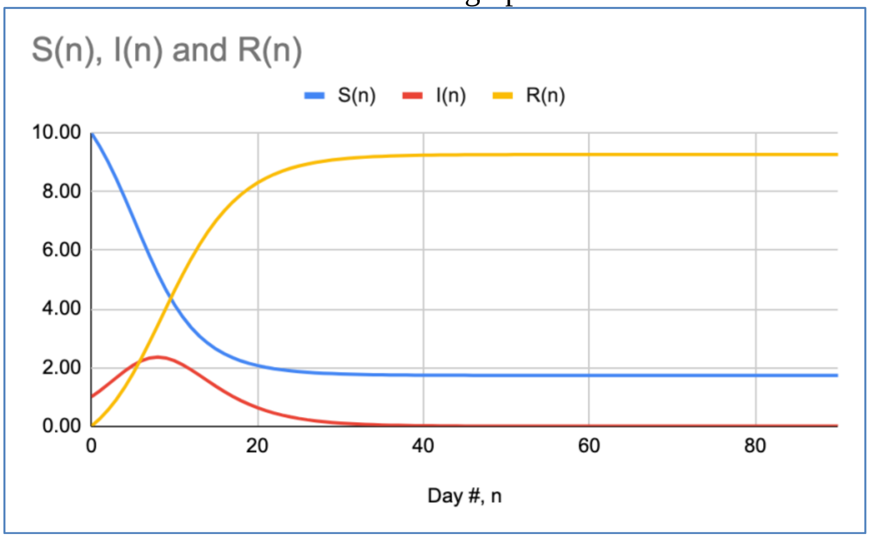
| Day Number, *n* | $S_{n+1} = S_n - iI_n\left(\frac{S_n}{T}\right)$ | $I_{n+1} = I_n + iI_n\left(\frac{S_n}{T}\right) - rI_n$ | $R_{n+1} = R_n + rI_n$ | Total Population |
|---|---|---|---|---|
| 0 | $S_0 = 10$ | $I_0 = 1$ | $R_0 = 0$ | 11 |
| 1 | $S_{0+1} = S_0 - iI_0\left(\frac{S_0}{T}\right) =$<br>$S_1 = 10 - 0.5(1)\left(\frac{10}{11}\right) =$<br>$\boldsymbol{S_1 \approx 9.545}$ | $I_{0+1} = I_0 + iI_0\left(\frac{S_0}{T}\right) - rI_0 =$<br>$I_1 = 1 + 0.5(1)\left(\frac{10}{11}\right) - 0.25(1) =$<br>$\boldsymbol{I_1 \approx 1.205}$ | $R_{0+1} = R_0 + rI_0$<br>$R_1 = 0 + 0.25(1) =$<br>$\boldsymbol{R_1 = 0.250}$ | 9.545<br>1.205<br>+ 0.250<br>11 |
| 2 | $S_2 = 9.545 - 0.5(1.205)\left(\frac{9.545}{11}\right) =$<br>$S_2 \approx 9.023$ | $I_2 = 1.205 + 0.5(1.205)\left(\frac{9.545}{11}\right) - 0.25(1.205) =$<br>$I_2 \approx 1.426$ | $R_2 = 0.25 + 0.25(1.205) =$<br>$R_2 \approx 0.551$ | 9.023<br>1.426<br>+ 0.551<br>11 |
| 3 | 8.438 | 1.654 | 0.908 | 11 |
| 4 | 7.803 | 1.875 | 1.321 | 11 |
| 5 | 7.138 | 2.072 | 1.790 | 11 |
| 6 | 6.466 | 2.226 | 2.308 | 11 |
| 7 | 5.812 | 2.324 | 2.864 | 11 |
| 8 | 5.198 | 2.357 | 3.445 | 11 |
| 9 | 4.641 | 2.324 | 4.035 | 11 |
| 10 | 4.151 | 2.234 | 4.616 | 11 |

Here are the data plotted in two different ways.




After 3 months of data the graph looks like this...




This is a pretty standard looking SIR graph.

-On Day-0 you can see the 1 Infected person, the 10 Susceptible, and the 0 Recovered.

-Approximately speaking only 2 Susceptibles never got the infection, as you can see by where the the blue line ends up.

-Judging from the yellow line about 9 appear to have become infected (including the original Infected person), but with time they Recovered.

-By Day-32 or so the infection died out as can be seen by the Infected population petering out to about 0 and the other lines flattening out.

---

There you have it.  The SIR Population Model.