

A Mathematical look at the Ebola Virus

Zach Yarus

May 11, 2012

Abstract

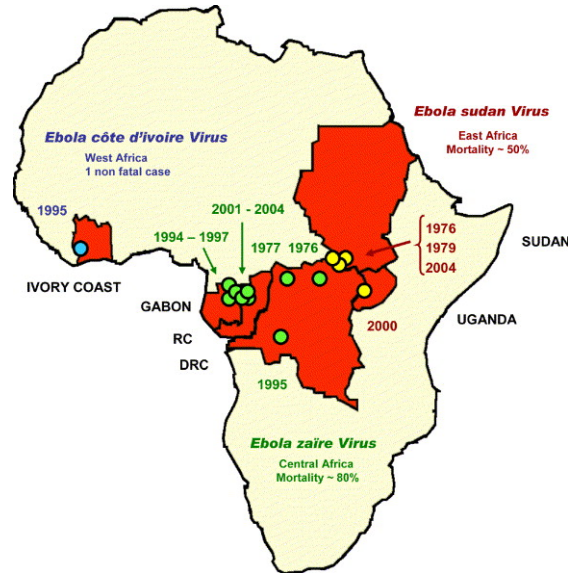
There are many fascinating things that can be described by differential equations. This paper will analyze an epidemic model created from first order differential equations to describe how the Ebola virus could potentially ravage a population. The population that will be used for this model will have roughly the same number of people as the number currently living in Kalispell, Montana.

Contents

| | | |
|---|---------------------------------------------|----|
| 1 | Choosing a Strain | 4 |
| 2 | Using this Virus in a Differential Equation | 5 |
| 3 | Our Assumptions | 6 |
| 4 | Creating the Differential Equation | 7 |
| 5 | Graphing the solution using Matlab | 9 |
| 6 | Conclusion | 11 |
| 7 | References | 12 |

1 Choosing a Strain

The Ebola virus from central Africa ranges from 974 to 1086 nano meters in length and kills 50 to 90 percent of the people it infects. First recognized in 1976, various strains of the virus have emerged and outbreaks have occurred, the latest occurring in 2011. The majority, as well as the largest outbreaks, have occurred only in Africa to date.



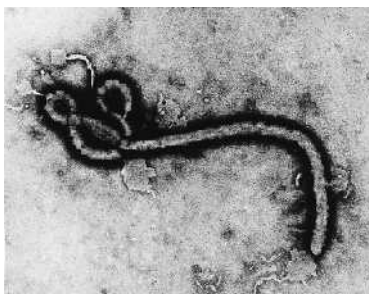
The Ebola virus is a member of the virus family Filoviridae, the viruses in the family, called Filoviruses, cause severe hemorrhagic fever in humans and primates. Hemorrhagic fever is characterized by high fever, internal bleeding, hypotension and shock. There are five known strains of Ebola virus, each named after its location where it was first recognized. These five strains are Ebola-Zaire, Ebola-Sudan, Ebola-Ivory Coast, Ebola-Bundibugyo and Ebola-Reston. Ebola-Reston was discovered in Reston, Virginia and only causes disease in primates. A virus called the Marburg virus was once considered an Ebola strain but is only a close cousin to Ebola.

A fact that is interesting with the Ebola Virus is that the exact origin, locations and natural habitat, known as the "natural reservoir", remains unknown. However, based on the nature of similar viruses, researchers believe that the virus is zoonotic. This means the virus is animal-borne, so there is likely an animal host native to Africa where the virus naturally lives. It is un-probable that the virus is native to anywhere but Africa. The Ebola-Reston strain was discovered in the United States but the monkeys were brought from Africa. What eliminates humans and primates from being the natural host is that it destroys those it infects far too quickly.

The virus is spread through close contact and contaminated medical equipment. When a person comes into contact and gets infected by the Ebola virus

the virus incubates within the body for a period of time, around 2 to 21 days. The onset of the illness is then abrupt and includes fever, headache, joint and muscle aches, sore throat and weakness, followed by diarrhea, vomiting and stomach pain. The condition is called hemorrhagic fever and it is often fatal. Reseachers do not understand why some people that become infected are able to recover while others are not, it is known that patients who die usually have not developed a significant immune response to the virus at the time of death.

The strain of Ebola virus this paper will use for the model is Ebola-Zaire. This strain has had the most outbreaks out of the four strains that cause disease in humans and it is also the deadliest. Ebola-Zaire was the first to emerge from the jungle in 1976 and has killed a total of 1084 people in Africa since its debut. The percent of death among human cases is 73.61 percent.



2 Using this Virus in a Differential Equation

This project will use an epidemic model, a common epidemic model is the SIR model. How the SIR model works is it divides a population into groups, the susceptible group, the infected group and the recovered group. The susceptible are described by $S(t)$, infected are decribed by $I(t)$ and the recovered group is described by $R(t)$. The sum of $S + I + R$ is equal to N , which is the total population. Most models assume N is a constant value, this means there are no births or deaths.

The disease will spread through contact between the infected and the susceptible, the rate of change of the populations is proportional to the number of contacts. By assuming the populations of susceptible and infected are randomly distributed over area, the number of contacts between the two populations is proportional to the product of S and I , this means there is a positive constant "a", that defines the rate at which susceptible become infected. The infected population will relate to the recovered group in the same way. The model that will be created for the Ebola virus is going to work based on this.

Creating an SIR model using equations is quite easy. One can begin by describing the rate of change of the susceptible population over time by multiplying $S(t)$, $I(t)$ and the constant "a". The population of the susceptible group will be reduced as the infected come into contact with the infected. This can

be show as:

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

The population of the infected group changes in two ways. One, individuals leave the susceptible group and join the infected group, adding to the total population of infected people. And two, individuals leave the infected group and join the recovered group, reducing the infected population. Let's say the rate at which the infected join the recovery group is the constant "b". To build a model for the change in population as it changes with time one can multiply "a", S(t) and I(t) and subtract "b" multiplied with I(t). This looks like:

$$\frac{dI(t)}{dt} = aS(t)I(t) - bI(t)$$

And finally the recovery population can be created. The only factor that changes the population of the recovery group is the addition of the newly recovered infected. This can be shown as the multiplication of "b" with I(t):

$$\frac{dR(t)}{dt} = bI(t)$$

This graphic demonstrates the relationship between variables for an SIR model.



While the epidemic model created for Ebola-Zaire will be similar to the SIR model in the text book Differential Equations with boundary values problems second edition, the model will have two differences. The first difference is that the SIR model is non-lethal and the model that will be created for Ebola-Zaire will be fatal, the infected will die at a certain rate. In the SIR model, the recovery group obtains immunity from the disease after they become infected. In the Ebola-Zaire model the recovered group will not remain immune to the infection. The Ebola-Zaire model will have the recovered group joining the susceptible at a certain rate.

3 Our Assumptions

A few interesting assumptions are made for this model.

1. The total population of people used in this model are randomly distributed over area, allowing for a constant to be defined for the contact made between

the susceptible and the infected.

2. The virus always kills 74 percent of the people it infects, the survivors will become the recovered group.
3. Individuals that recover are given no immunity.
4. The population involved stays constant (no births or unrelated deaths).
5. Another assumption that will be made that cannot be placed in the model is that people do not panic completely and that they do not treat the infected with quarantine procedures. In other words, the susceptible and the infected go about their daily lives.

4 Creating the Differential Equation

As was stated, the differential equation that will be used is a modified epidemic model. Essentially a SIR model with the difference that the recovered can become infected again and the infected die at a certain rate.

The susceptible are described by $S(t)$, the infected by $I(t)$, the recovered by $R(t)$, and people that are killed by the virus are described by $D(t)$. To create an equation that describes the population of the susceptible group with respect to time, start with the fact that the susceptible become infected at rate "a". This means the change in population of the susceptible group is equal to the negative product of a, $S(t)$ and $I(t)$. This can be written as:

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

This equation isn't complete yet. Individuals from the recovery group become susceptible again at a certain rate "c". This can be multiplied by $R(t)$ and added to the previous equation and written as:

$$\frac{dS(t)}{dt} = -aS(t)I(t) + cR(t)$$

This is the complete equation that describes the change in population of the susceptible group over time.

The equation that describes the population of the infected group begins with adding what was just removed from the susceptible population, $aS(t)I(t)$. This is written as:

$$\frac{dI(t)}{dt} = aS(t)I(t)$$

The population of the infected group is reduced in two ways, people can either recover or they are killed by the virus. Both options remove people from the

infected group. When the infected recover they join the recovery group at rate "b" and when infected die they join the deceased group at rate "e". This is written as:

$$\frac{dI(t)}{dt} = aS(t)I(t) - bI(t) - eI(t)$$

This is the complete equation that describes the change in population of the infected group over time.

The recovery group's population is increased by those that recover from the virus. People recovered from the virus as rate "b", this means the population of the recovery group is increased by "b" multiplied by I(t), this is written as:

$$\frac{dR(t)}{dt} = bI(t)$$

The recovery group's population is reduced by the number of people that join the susceptible group, the recovered joined the susceptible group at rate "c". Subtracting cR(t) from the previous equation is written as:

$$\frac{dR(t)}{dt} = bI(t) - cR(t)$$

This is the complete equation that describes the change in population of the infected group over time. The last equation to build is the for the deceased group. The population of the deceased group is defined by the number of people that are killed by the virus within the infected group. People were killed at rate "e", so the equation for the population of the deceased group is written as:

$$\frac{dD(t)}{dt} = eI(t)$$

The sum of S(t), I(t), R(t) and D(t) will always remain constant. The four complete equations are:

$$\frac{dS(t)}{dt} = -aS(t)I(t) + cR(T) \quad (1)$$

$$\frac{dI(t)}{dt} = aS(t)I(t) - bI(t) - eI(t) \quad (2)$$

$$\frac{dR(t)}{dt} = bI(t) - cR(T) \quad (3)$$

$$\frac{dD(t)}{dt} = eI(t) \quad (4)$$

To summarize the coefficients:

a = the rate of infection

b = the rate of recovery

c = the rate of susceptibility

e = the rate of death

5 Graphing the solution using Matlab

Because the Ebola-Zaire model uses four separate differential equations one must use a numerical solver to plot the solution. Matlab makes this easy. Perhaps the easiest way to plot a solution is to create a function m-file with the equations and constants a , b , c , and e defined. Next, run that function file with the solver `ode45` and plot.

Two scenarios will be plotted for this project. The first scenario uses the current population of Kalispell, Montana as the initial susceptible group, this value is 22,000. The susceptible group was started at 21,950. The initial infected group started out at 50. The deceased and the recovery group were both started at 0.

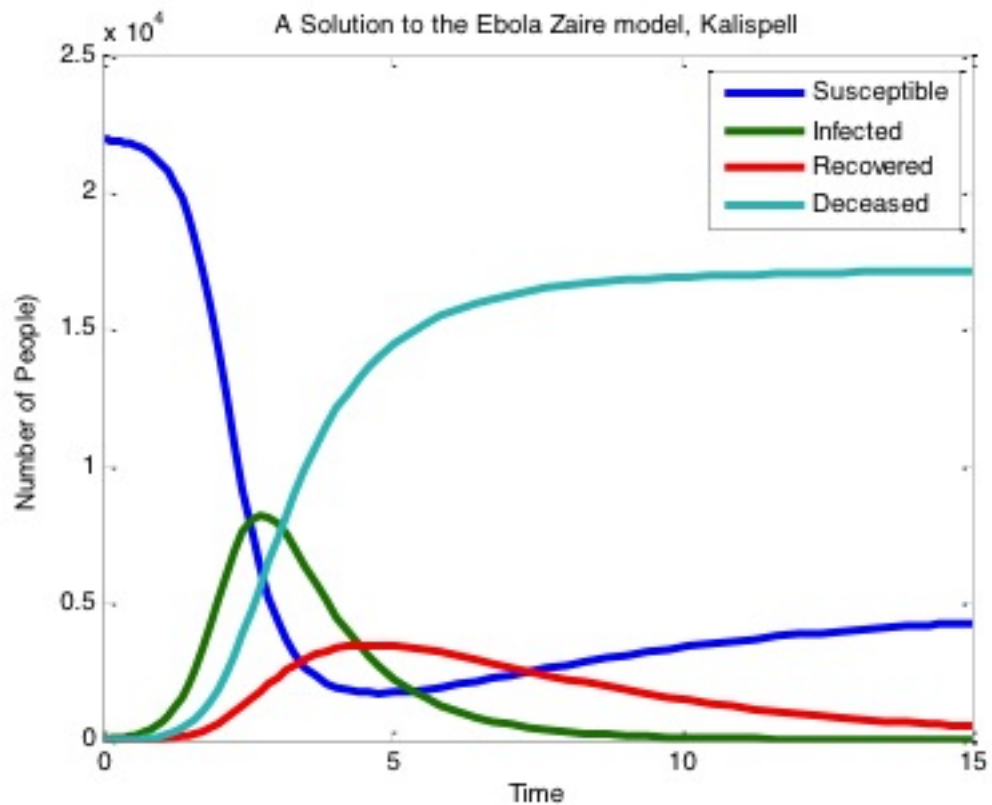
The constants were as follows:

a = rate of infection = .000165

b = rate of infection = .27

c = rate of susceptibility = .23

e = rate of death = .73



According to the graph the susceptible group immediately begins to plummet due to how infectious the virus is and at the same time the infected group's numbers begin to rise. The number of the dead begins to rise following the infected groups climb and continues to climb until nearly seventy percent of the population is dead. The recovery group's peak is around 4,000 people, around the same time when the deceased group's curve is hitting a plateau.

This solution results in the susceptible group's population leveling out just under 5,000 people. The recovery group and the infected group both head to zero as the last of the infected die off or join the recovery group. The recovery group joins with the susceptible group until no one remains in the group. The deceased group reaches equilibrium around 17,500. In total 79 percent of the population is wiped out by the Ebola-Zaire virus leaving less than 5,000 residents in a city that once contained 22,000.

Never in history has this many people been killed by the Ebola virus. Africa is a populated place so there is no shortage of people that could fall ill, but the fact that the Ebola virus is spread primarily through contact and contaminated medical equipment limits its ability to spread. People afflicted with Ebola cannot do much on their own and require medical staff to care for them. Or at least family members and friends. Medical staff need to be equipped properly to safely deal with Ebola patients and many hospitals in Africa are poorly supplied. Medical staff are at highest risk of getting the virus when an outbreak happens.

This next scenario will take place in a theoretical hospital, the location where the majority of people that get the Ebola virus are exposed to it (in Africa). The same m-file that was used in the previous solution will be used once again but with the constant "a" altered and with different initial conditions.

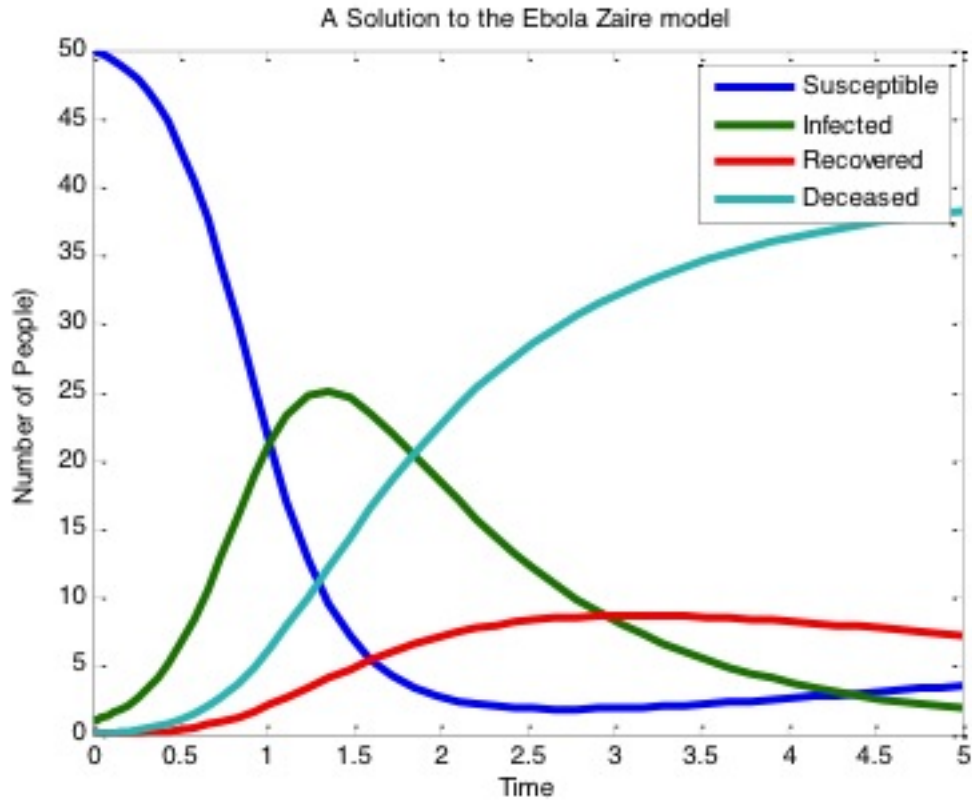
The initial conditions for the hospital model are 50 for hospital staff and 2 for infected, zero for both dead and recovered.

a = rate of infection = .1

b = rate of infection = .27

c = rate of susceptibility = .23

e = rate of death = .73



This solution is very similar to the previous solution, however upon comparison this solution **is more realistic**. It is more realistic because the virus likely wouldn't spread over thousands of people in such a manner even **though the virus is extremely contagious**. In a hospital setting the realistic chances of spreading are higher.

6 Conclusion

While this paper's equations and diagrams model an event that never occurred, the nature of the Ebola-Zaire virus makes a scenario like this frighteningly possible. Both scenarios didn't treat the infected with quarantine procedures as hospitals around the world have learned to do with highly infectious diseases and both populations ended up losing over half of their numbers as a result. The equations used in this model could be altered to account for other factors such as the population gaining immunity after infection or a theoretical vaccination being administered but history has shown the virus is very unforgiving. Other than the assumption that the susceptible don't treat the infected with quarantine procedures the models created represent the destruction of the virus

accurately.

7 References

1. <http://www.cdc.gov/ncidod/dvrd/spb/mnpages/dispages/ebola.htm>
2. <http://students.cis.uab.edu/sajayi11/Ebola>,
3. <http://www.who.int/mediacentre/factsheets/fs103/en/>