Clovis Hatungimana
Tutor: Jacob Cook
MAT 3102 Calc 2: The Practice of Calculus
Block 4, Spring 2022
Quest University
April 11th

### The SIR Model

Human epidemics are often spread by contact with carriers of the disease often referred to as vectors. These vectors are infective carriers and can range from mosquitos to infective people. Mediums of spread can be airborne, casual or intimate contact. Smallpox, polio, measles, and covid-19 are examples of epidemics. Over the years, human attempts have been made to contain the spread of epidemics, some of which have succeeded while others have not fully eradicated the targeted diseases.

In this paper, we will be investigating more about the SIR model, which is one of the most famous mathematical models that scientists have relied on to "predict different scenarios related to epidemiologic factors and possible outcomes to assess epidemic spread" (Espinosa et al., 2020). By leveraging mathematical concepts in Calculus, we will be exploring how the SIR model works and how to calculate parameters for a specific disease. Then, for covid-19, we will look into how different covid responses (such as social distancing and vaccination) impact the model.

A basic model with very basic assumptions is first defined, and from there, depending on the real-world factors of the epidemic, the parameters and assumptions of the SIR model can be adapted to the context and real-life situation; which makes the model more complex and more accurate in modeling the real-world situation of the epidemic at hand.

### 1. Basic assumptions of the model



- The size of the population is fixed and confined to a well-defined region such as a city, organization, country, or the whole world, without external contact with the concerned region.
- In the SIR model, the population is broken up into three categories:
  - **Susceptible**: those who are susceptible to catching the disease. Here, the assumption is that every individual who has never got the disease can catch it.
  - **Infected (vectors)**: those who can spread the disease. Here, we assume that whoever gets the disease is capable of spreading it.
  - **Removed**: those who are immune and cannot spread the disease. This category includes those who have recovered and those who have died from the disease.

However, the most basic assumption here is that those who are recovered can no longer get or spread the disease.

- The size of the population is fixed and confined to a well-defined region such as a city, organization, country, or the whole world, without external contact with the concerned region. This means that there are no newborns or travelers (immigrants or emigrants).
- The population is well mixed that people come into contact with others from the other categories every day. If you are infective and you come into contact with someone who is susceptible, then you spread the disease to that person. If you are come into contact with an immune person, you will not spread the disease.
- We ignore factors (such as diet, weather and other diseases) that have small effects on the spread of the disease.

As mentioned earlier, these assumptions can be tweaked and adapted to best account for the real-life situation. For example, the notion of contact for covid can be redefined as follows: if you touch a contaminated surface and touch your mount or spends more than 15 minutes with someone who has had covid for less than 5 days, you get the disease but only two days later.

#### 2. Variables of the SIR model

Let t = the time in days, with t=0 at the start of observation of the epidemic.

S = the number of susceptible individuals.

I = the number of infective individuals.

R = the number of removed individuals.

N = the population size (number of people in a certain region).

The population size is N = S + I + R

# 3. Parameters of the SIR model

Each epidemic has three mainly basic parameters. Let:

- d = the average number of days a sick individual is contagious. We assume that this period is the same for all individuals across the entire population. We also assume that, after this period, the individual transfers from infected (I) to removed (R) individuals.
- b = the rate at which an individual transfers from infected (I) to the removed (R) population. Thus, b = 1/d.
- c = the average total number of contacts per infective person.
- a = the daily rate of contacts per infective individual. Thus, a = c/d which is also  $a = b \cdot c$

In real life, often the epidemic spreads faster during the first days when someone is sick than in the last days of their sickness. However, we will assume that each person in each category (S, I, R) has the same capacity to catch, recover or spread the disease independently of how long they are in that category. Now, we can write differential equations that model the daily changes in the number of our populations in the three categories (S, R, and I).

### 4. Equations of change

Let  $\Delta S$ ,  $\Delta I$ , and  $\Delta R$  represent respectively average daily changes in the number of people in susceptible (S), infected (I) and removed (R) categories of our total population N.

A person in the susceptible category (S) immediately moves to the infective one (I) right after getting into contact with someone who has the disease. The number of infected people on a daily basis is calculated by (a = the daily rate of contacts per infective individual) x (I = the number of infective individuals), which is  $a \cdot I$ . However, we multiply by the fraction S/N (number of susceptible people divided by the total population) because people in I can only infect those in S. Thus,  $a \cdot I \cdot \frac{S}{N}$  is the number of infected people on a daily basis, and this is the rate by which the number of individuals in S is decreasing by. Thus,

$$\Delta S = -a \cdot I \cdot \frac{S}{N}$$

The number in category (R) receives infected people by the rate  $a \cdot I \cdot \frac{S}{N}$  daily. However, at the same time, infected people have to spend d days in the infective category (I), after which period they immediately move to the removed category (R). So, the number of people who leave this category on a daily basis is obtained by (b = the rate at which an individual transfers from infected (I) to the removed (R) population) x (I = the number of infective individuals), which is  $b \cdot I$ . Thus, the change in the number of people in this category (I) is obtained by

$$\Delta I = a \cdot I \cdot \frac{S}{N} - b \cdot I$$

People in the removed category (R) increase by

$$\Lambda R = b \cdot I$$

Those who are already in this category (R) stay there indefinitely because they can no longer catch the disease.

Notice: The above expressions of change (in S, I and R) are average rates of change per day. Let  $\Delta t$  represent the change in time from  $t_0$  to  $t_1$ . The average rates of change  $\Delta S$ ,  $\Delta I$  and  $\Delta R$  are obtained by the difference between our initial values of S, I and R (at time  $t_0$ ) and our final values (at time  $t_1$ ) divided by the change in time (from time  $t_0$  to time  $t_1$ ). Let  $S(t_0)$ ,  $I(t_0)$  and  $R(t_0)$  represent our initial values and  $S(t_1)$ ,  $I(t_1)$  and  $R(t_1)$  our final values of the population in each of our categories. Thus, the expressions for the numbers of people in each category (S, I and R) are:

- Rate of change in the number of the susceptible population on interval  $\Delta t$ :

$$\frac{S(t_1) - S(t_0)}{t_1 - t_0} = -a \cdot I \cdot \frac{S}{N} \quad or \quad S(t_1) - S(t_0) = \Delta S \quad and \quad t_1 - t_0 = \Delta t$$
 We get:

$$\frac{\Delta S}{\Delta t} = -a \cdot I \cdot \frac{S}{N}$$

- Rate of change in the number of the infected population on interval  $\Delta t$ :

$$\frac{I(t_1) - I(t_0)}{t_1 - t_0} = a \cdot I \cdot \frac{S}{N} - b \cdot I \quad or \quad I(t_1) - I(t_0) = \Delta I \quad and \quad t_1 - t_0 = \Delta t$$

We get:

$$\frac{\Delta I}{\Delta t} = a \cdot I \cdot \frac{S}{N} - b \cdot I$$

- Rate of change in the number of the removed population on interval  $\Delta t$ :

$$\frac{R(t_1) - I(t_0)}{t_1 - t_0} = a \cdot I \cdot \frac{S}{N} - b \cdot I \quad or \quad R(t_1) - R(t_0) = \Delta I \quad and \quad t_1 - t_0 = \Delta t$$

We get: 
$$\frac{\Delta R}{\Delta t} = b \cdot I$$

From the above expression, we want to derive our differential equations. Thus, we need to calculate the rates of changes in S, I and R on much smaller intervals of time  $\Delta t$ . This is the instantaneous rate of change (which is calculated by evaluating the limit of each of the above expressions as  $\Delta t$  equals almost 0). Because the above expressions don't explicitly depend on time t, the evaluations of the limits of our expressions will yield the exact expressions like the above ones, except for d symbols in place of  $\Delta$  symbols to signal that the rate of change was taken on much smaller intervals of time. Let dt represent the smallest interval of time possible that is closer to 0. Similarly, dS, dI, and dR will represent the smallest values of the differences  $S(t_1) - S(t_0)$ ,  $I(t_1) - I(t_0)$  and  $R(t_1) - R(t_0)$ . Our differential equations are:

- Instantaneous rate of change in the number of the susceptible population:

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

- Instantaneous rate of change in the number of the infected population:

$$\frac{dI}{dt} = a \cdot I \cdot \frac{S}{N} - b \cdot I$$

- Rate of change in the number of the removed population on interval  $\Delta t$ :

$$\frac{dR}{dt} = b \cdot I$$

<u>Note</u>: We could not think of an algebraic way of finding solutions to the above differential equations. We were unable to find expressions of S(t) and I(t), and R(t) that would allow us to continuously predict values of S, I, and R at any time t. The main reason, in our understanding, is that the above expressions have more than two variables (t, S, I and R). Dealing with such expressions requires more advanced knowledge than the one we have from our Calculus 2 class. We suspect this is for multivariable Calculus.

However, we can only find numerical or graphical solutions to our differential equations through recursion. This means that, if we know our initial values of *S*, *I* and *R*, we can calculate the next values based on the previous ones. First, we need to transform our differential equations as follows:

- The number of the susceptible population at:

$$\frac{dS}{dt} = -a \cdot I \cdot \frac{S}{N} \quad or \quad dS = S(t_1) - S(t_0)$$

$$\frac{S(t_1) - S(t_0)}{dt} = -a \cdot I \cdot \frac{S}{N}$$

$$\Rightarrow \qquad S(t_1) = S(t_0) - \left(a \cdot I \cdot \frac{S}{N}\right) \cdot dt$$

- The number of the infected population:

$$\begin{split} \frac{dI}{dt} &= a \cdot I \cdot \frac{S}{N} - b \cdot I \quad or \quad dI = I(t_1) - I(t_0) \\ \frac{I(t_1) - I(t_0)}{dt} &= a \cdot I \cdot \frac{S}{N} - b \cdot I \\ \Rightarrow \qquad I(t_1) &= I(t_0) + \left( a \cdot I \cdot \frac{S}{N} - b \cdot I \right) \cdot dt \end{split}$$

- The number of the removed population on interval  $\Delta t$ :

$$\frac{dR}{dt} = b \cdot I \quad or \, dR = R(t_1) - R(t_0)$$

$$\frac{R(t_1) - R(t_0)}{dt} = b \cdot I$$

$$\Rightarrow \qquad R(t_1) = R(t_0) + (b \cdot I) \cdot dt$$

The graphical representation of the general solutions to the above equations can take many different shapes, depending on the values of variables and parameters. However, we will go straight into examples of covid-19 and look at how different responses can affect the spread of the epidemic; thus, affecting the graphs of the model.

We know that the number of infective people should be affected by quarantine (because those are isolated from the population) and that the rate of spread of covid should be affected by social

distancing, hygiene, mask mandate and lockdown. So, we will look at three cases, where we vary the rate of spread, then the number of infective people and investigate who the SIR model for covid changes.

# 5. The spread of covid-19 versus human responses

# Case 1

Through modeling, we will be investigating the spread of covid-19 in a population partitioned as follows: 30,000 susceptible individuals; 1,000 infective and 0 removed individuals (recovered). According to the WHO, each individual with covid-19 is contagious between 2 and 14, days after which they no longer have the disease (they recover). In this example, we will assume that each infective person is contagious for 14 days. The average total number of contacts per infective person is 6.8.

# Our basic assumptions are:

- every individual who has never got the disease can catch it.
- whoever gets the disease is capable of spreading it.
- those who are recovered can no longer get or spread the disease.
- The size of the population is fixed at 30,000 people (no newborn travelers or visitors from outside this population of our interest) for the time of our interest.
- The population is well mixed that people come into contact with others from the other categories every day.
- If you are infective and you come into contact with someone who is susceptible, then you spread the disease to that person. If you come into contact with a recovered person, you will not spread the disease.
- the notion of contact for covid in our case is defined as spending more than 15 minutes with someone who has covid, you directly get the disease.
- We ignore factors (such as diet, weather and other diseases) that have small effects on the spread of the disease.
- each person in each category (S, I, R) has the same capacity to catch, recover or spread the disease independently of their immunity or other factors that could confer accelerated recovery or protection.

#### Our variables are:

- At time t=0:
  - S = 30,000 (susceptible individuals)
  - I = 1,000 (infective individuals)
  - R = 0 (removed individuals: recovered or dead)
- d = 14 (the average number of days a sick individual is contagious).
- b = 1/14 (the rate at which an individual transfers from infected (*I*) to the removed (*R*) population).
- c = 6.8 (the average total number of contacts per infective person).
- $a = b \cdot c$  (the daily rate of contacts per infective individual). Thus, a = 1/14 \* 6.8 = 6.8/14.

To model the situation described above, we will calculate the numbers of individuals for S, R, and I by regression because that's what we can do for now. Then we will make graphs of S, R, and I as functions of t, where t is in days (from t=0 to t=150 days).

Our expressions for regression:

$$S(t_1) = S(t_0) - \left(a \cdot I \cdot \frac{S}{N}\right) \cdot dt$$

$$I(t_1) = I(t_0) + \left(a \cdot I \cdot \frac{S}{N} - b \cdot I\right) \cdot dt$$

$$R(t_1) = R(t_0) + (b \cdot I) \cdot dt$$

Using Python, we were able to calculate and plot values for S, I, and R for t values that range from 0 to 150 days. In the program source code, we can assign different values to our variables the parameters and graph our S, I, and R as functions of time for any epidemic of interest. However, below is the source code for our situation described above. For the sake of simplicity, we sed dt = 1 because we are calculating covid cases on a daily report.

```
#Code written by Clovis Hatungimana for the SIR model
import matplotlib.pyplot as plt
# defining variables and parameters
                    #susceptible individuals
S = [30000] #SUSCEPTIBLE INDUSTRIALS

I = [1000] #Infected individuals

R = [0] #Removed individuals

N = S[0]+I[0]+R[0] #Total population = S + I + R

d = 14  # number of days a person with covid is infective

b = 1/d # the rate at which an individual transfers from infected (I) to the removed (R) population
            # the average total number of contacts per infective person
# the daily rate of contacts per infective individual
#changes in the population are calulated per day.
period_of_observation=150
# setting the values for t as a list that will contain values from 0 to 150 days. t=[0]
 counter=0
 while counter < period_of_observation:</pre>
      #calculating by regression the number of remaining susceptible people over dt interval
      susceptible = round(S[counter], 2) - (a*round(I[counter],2)*round(S[counter],2)/N)*dt
      #calculating by regression the number of infected people over dt interval
infected = round(I[counter], 2) + ((a*round(I[counter],2)*round(S[counter],2)/N)-b*round(I[counter],2))*dt
      #calculating by regression the number of removed people over dt interval removed = round (R[counter], 2) + b*round(I[counter],2)*dt
      susceptible= round(susceptible, 2)
      infected= round(infected, 2)
removed = round(removed, 2)
      #Storing lists of values for S(t), I(t), and R(t) over our period of observation
      S.append(susceptible)
I.append(infected)
      R.append(removed)
       counter=counter+1
       t.append(counter)
```

Here is a table of values of S, I, and R for t values between 0 and 4 days.

t	0	1	2	3	4
S	30,000	29,529.95	28,882.83	28,002.26	26,823.17
I	1,000	1,398.62	1,945.83	2,687.41	3,674.54
R	0	71.43	171.33	310.32	502.28

### Graphical representation of S(t), I(t), and R(t).

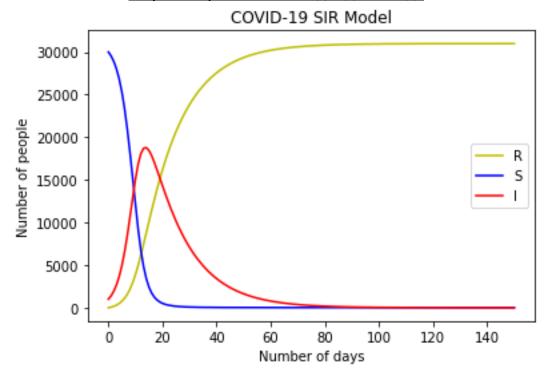


Fig.1. Covid-19 SIR model.  $S(t_0) = 30,000$ ;  $I(t_0) = 1,000$ ;  $R(t_0) = 0$ ; c=6.8; b = 1/14

<u>Discussion for this particular example</u> (the exact values used in this section were retrieved through computer-generated calculations).

From the graph above, we see that the number of S and I are inversely proportional (one decreases when the other is increasing, vice versa). S and R are also inversely proportional. However, I and R are directly proportional (one increases when the other is also increasing) up to a certain point on an interval and then become inversely proportional. This is because (from our differential equations as well as our equations for regression) S, I and R are related to one another (mostly through the rate of change and the fact that the expressions of S and R contain I). This means that the spread of the epidemic, covid-19 in this case, follows some pattern underlying in the structure of the equations of S(t), I(t) and R(t). I grows faster and peaks at t = 14 (where I = 18,781.83, S = 3,430.18). This value of S is about 11 percent of the initial susceptible population, meaning that 89 percent of them got sick in 14 days. The spread continues at a continuously decreasing rate until the disease dies out.

In 150 days, we see that covid-19 fades away, leaving only a susceptible population of 16.27, which is about 0.05 percent of the susceptible population unaffected by the disease. It is interesting to see that some people, despite lack of counter measures of the spread, will never get infected (which means they were never been exposed to infective people.

However, 0.05 percent of unaffected people is way less for a government that cares about its people. Using the SIR model for covid, if the government makes the above predictions at t=0, we would see, as was the case in 2020 and 2021, governments decree some policies such as enforcement of social distancing and hygiene (use of disinfectants) and masks. However, all this works on the premise that no one can get the disease more than once.

# Case 2: increase the rate of spread

Maybe, we considered a less realistic rate of spread of covid. In real life, for a period of 14 days, let's say that an infective person infects 1 person each day. This will make it 14 persons infected for a period of 14 days. We keep all the other variable same as previously.

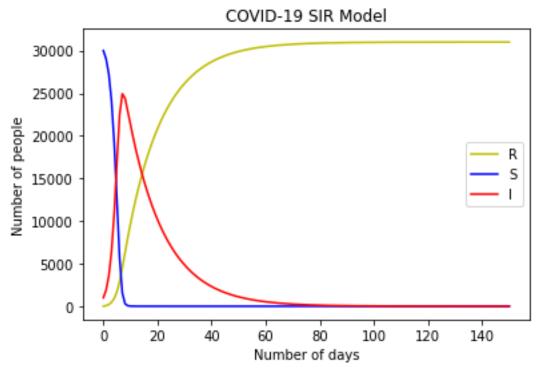


Fig.2. Covid-19 SIR model.  $S(t_0) = 30,000$ ;  $I(t_0) = 1,000$ ;  $R(t_0) = 0$ ; c = 14; b = 1/14

The infected population peaks on day 7, where I = 24,934.64 and S = 1,559.84. This is about 5.2 percent of the susceptible population, meaning that 94.8 percent of them got sick in only 7 days. After day 7, the rate of spread continuously drops, until everyone catches the disease in 12 days (I = 18,394.64 and S = 1.91). After 12 days, everyone of the susceptible population has got covid. However, covid lingers around until t = 145, time at which everyone who had been infected is recovered. The change in the rate of spread makes a big difference in how many people and how faster the susceptible population gets infected. To understand more about the impact of the rate of spread of covid, let's try another example in which we keep constant the

average total number of contacts per infective person (c=14) while we change the infective population to 10 individuals.

# Case 3: decrease the initial number of infective persons

In this case, of the 1000 infective population, 99 percent have been quarantined. For some reason, the government did not spot the 1 remaining percent of the infective population. So, there are only 10 individuals who can spread the disease at the average total number of contacts of 14 people per infective person. Here, we consider the 99 percent (or 990 individual in quarantine as if they are removed). Thus, R = 990 at t = 0.

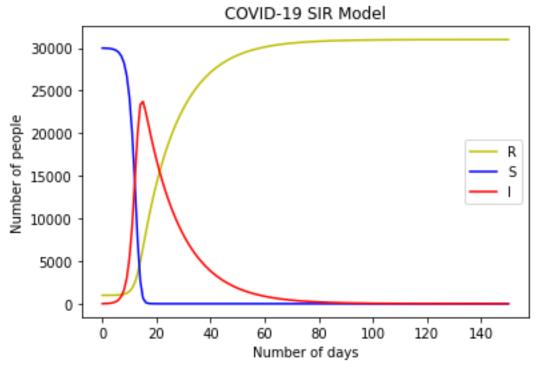


Fig.3. Covid-19 SIR model.  $S(t_0) = 30,000$ ;  $I(t_0) = 10$ ;  $R(t_0) = 990$ ; c = 14; b = 1/14

The infected population peaks on day 15, where I = 23,730.95 and S = 668.96. This is about 2.2 percent of the susceptible population, meaning that 97.8 percent of them got covid in 15 days. Afterward, covid keeps on spreading at a continuously decreasing rate, making everyone of the susceptible population sick at t=21. However, covid stays around until t=150 days, time at which it dies out.

### Impact of human responses on the spread of covid

From Fig.1 to Fig.2, only the average total number of contacts per infective person was increased by a factor of 2.05 (c=6.8 versus c=14). This ended up in a change in the peak values of infected people as well as how long it takes covid to die away. In Fig.1; 89 percent (in 14 days) of the susceptible people got sick whereas in Fig.2; 94.8 percent got sick (in 7 days). This increase in percentage of people who got sick (from Fig.1 to Fig.2) is by a factor of 1.06. In Fig.1, covid dies away after 150 days, leaving 0.05 percent of the population unaffected by the disease. In

Fig.2, covid dies away after 145 days, after affecting everyone of the susceptible population. Thus, an increase in the rate of change of the spread of covid causes a noticeable change in how many people get affected and in how long the disease stays around. Thus, it makes sense why governments would strive to reduce the rate of spread by taking measures such as social distancing, lockdown, and hygiene.

From Fig.2 to Fig.3, only the initial number of infective individuals decreased by a factor of 1/100 (I=1,000 versus I=10). This resulted in changes in the peak values of infected people as well as how long it takes a disease to die away. In Fig.2, 94.8 percent of the susceptible people got sick (in 7 days) whereas in Fig.3; 97.8 percent got sick (in 15 days). This change in percentage of people who get sick (from Fig.2 to Fig.3) is by a factor of 1.03. In Fig.2, covid dies away after 145 days, leaving everyone of the susceptible population affected by the disease. However, in Fig.3, covid dies away in 150 days, leaving everyone of the susceptible population affected by the disease. Thus, measures that aim at decreasing the number of the infective population are less effective compared to those aimed at decreasing the rate of spread. This explains why governments didn't only opt to isolate sick people. They rather simultaneously adopted combined measures such as quarantine, mask mandate, lockdown, hygiene and vaccination to reduce both the rate of spread and the number of infective population, which sounds more effective.

#### 6. Conclusion

The SIR model is a mathematical model that serves scientists (especially epidemiologists) in predicting different scenarios related to epidemiologic factors and possible outcomes of an epidemic spread. We used mathematical concepts in Calculus to write the following differential

equations: 
$$\frac{dS}{dt} = -a \cdot I \cdot \frac{S}{N}$$
,  $\frac{dI}{dt} = a \cdot I \cdot \frac{S}{N} - b \cdot I$ ,  $\frac{dR}{dt} = b \cdot I$ .

dS/dt, dI/dt and dR/dt respectively represent the instantaneous rates of change in the number of the susceptible, infected and removed population. However, we found that the above equations don't have neatly written solutions S(t), I(t) and R(t) as functions of time. nevertheless, we wrote a computer program that allowed us to calculate by regression values of S(t), I(t) and R(t) at different times and then graph them. Through explorations of examples of the covid epidemic, we analysed graphical representations of S(t), I(t) and I(t) and found that an increase in the rate of change of the spread of covid causes a noticeable change in how many people get affected and in how long the disease stays around (Fig.1 versus Fig.2). On the other hand, keeping the rate of spread constant, we found that a decrease in the number of infective people causes small changes in how many people get infected and how long the disease stays around. Thus, striving to reduce the rate of spread (by taking measures such as social distancing, lockdown, and hygiene, etc.) can be more effective compared to reducing the number of infected people (through quarantine, etc.). However, for extra caution and for better results, it makes sense why governments simultaneously adopted multiple measures such as quarantine, mask mandate, lockdown, hygiene and vaccination to counteract the spread of covid in 2020 and 2021.

The SIR model is really useful in predicting different scenarios related to epidemiologic factors and possible outcomes to assess epidemic spread (Espinosa et al., 2020). However, there are

limitations to the SIR model. The model doesn't tell us when an individual would get sick. It doesn't tell us who will get sick and who will not get sick as we saw in Fig.1 (some people never got sick). The model make very basic assumptions to allows us to investigate and assess different outcomes of the overall spread of an epidemic.

### References

Espinosa, P., Quirola-Amores, P., & Teran, E. (2020). Application of a Susceptible, Infectious, and/or Recovered (SIR) Model to the COVID-19 Pandemic in Ecuador. *Frontiers in Applied Mathematics and Statistics*, 6.

https://www.frontiersin.org/article/10.3389/fams.2020.571544

Stroyan, K. Calculus (n.d). The Language of Change. *The University of Iowa*. https://homepage.divms.uiowa.edu/~stroyan/CTLC3rdEd/3rdCTLCText/Chapters/Ch2.pdf