On a Monte Carlo Epidemics Numerical Simulations

João Luiz Kohl Moreira

February 9, 2022

Abstract

A numerical simulation is being developed to test different quantities: social distancing, government policy, social mobility, and demographic density, considering the actual global situation, with the Coronavirus spreading out worldwide. Each of them is put in numerical parameters, and the results are inspected under their variations. Numerical simulations are placed side-by-side with data given by officials about the real values.

1 User's Guide

I developed the application in Java, improved by some public domain libraries, namely SQLite library set, jFreeChart for graphics, etc. So, it is easy to launch it, and for that, the procedure changes according to the platform you are applying, suffice to have Java-8 or later. The package comes from the "`Epidemia`" path tree. It is supposed to work with Java-8 JRE-edition. With Windows-10 or later, and Mac OS 10 or later, you create a shortcut or an alias (the word depends on being Windows or Mac) of the file "Epidemia/Epidemia.jar" and put it in the folder Desktop (don't move the original jar-file from its place). Then, you can double click on the icon so created on the Desktop. For the multiple distros of Linux, the general rule is to create a command on the Desktop or Main Menu with an alias to the command "`java -jar {path}/Epidemia/Epidemia.jar.`". There is a way to straightforward change the icon for Windows or Linux from the png-file stored in the folder "icons". For the Mac, this procedure is a little more complicated, but it is possible.

Once the installation is complete, you can launch the app. You should have a window on the screen similar to the Figure 1. An initial set of parameters is set by default, and the program already proceeds a simulation. The units (of space and time) are relative. To compare the results with real data, one should adapt the parameters to values such that the aspect of the plot would be comparable.

When starting, the program attempts to read a file Json-type to retrieve the parameters stored in it. If, for some reason, namely, for not the existence of this file, the program assumes a set of parameters taken by default. This file may be generated by clicking on the button "Json" on the panel's lower-right side.

The plot you see is of the number of active infections per step (not the number of new ones). This quantity indicates the potential for contamination. Looking at the upper-side plot, we see five step-looking curves in different colors. The black curve is the total of infected individuals per step; the blue one is the observed number of infected. Since we have asymptomatic cases, the observed number is what we usually can count. The light-blue curve represents the asymptomatic cases. Additionally, we have the plot of the number of risky populations (who could die from the disease) in magenta and who effectively died (red curve).

A horizontal dashed red line marks the limit of the health system capacity. This threshold will matter to define the number of who dies in the absence of assistance. The counting process works this way: one takes into account the number of risky infected beings; if it happens that this number reaches the threshold of the health system, the probability of dying changes from the "normal" value
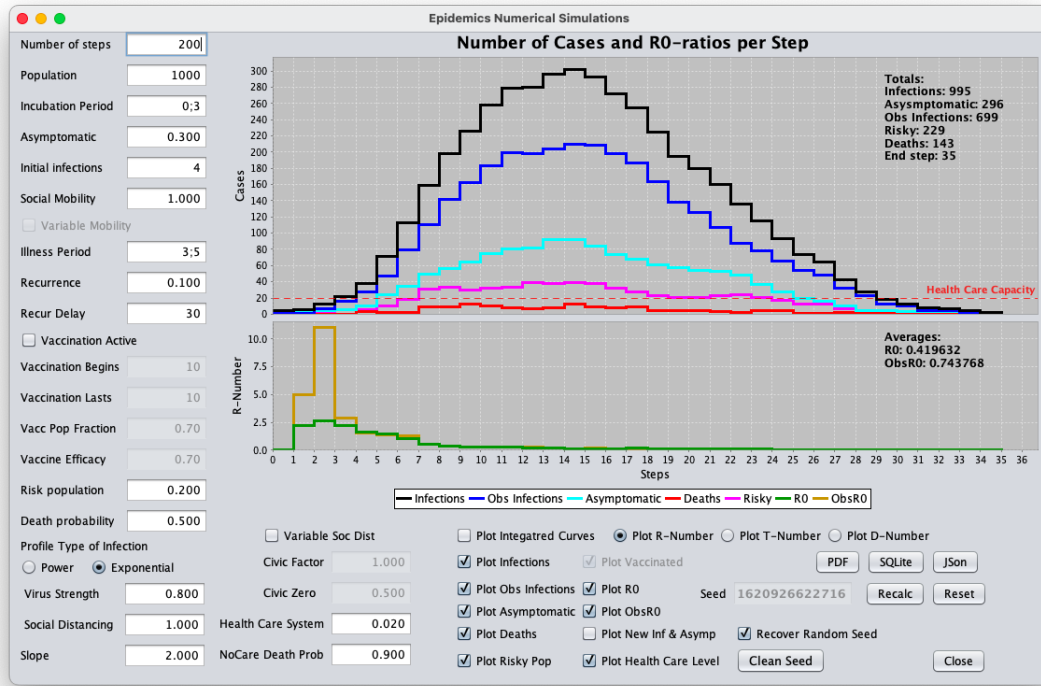
Figure 1: The initial window of the application for Macos.

to the one in the field text labeled "NoCare Death", and the "new" risky infected individual at this step now has this probability of dying.

The plot at the lower-side shows one of these three quantities: the R-number for each step, the T-number and the D-Number. There are two curves. For the R-number, the dark-green one is the total (real) value obtained from the number of newly infected beings taking into account the number of effectively infected ones. The other curve, in gold-yellow, is the observed R-number; the new infected subtracted the asymptomatic beings. You can toggle what is plotted on the graphic below by checking one of the boxes just on the lower-right side of the window.

Checking the "T-Number" box, you have the plot of the ("real" and "observed") ratio between the closed cases, recovered or dead, and the new infected per step. The name "T-number" is because Kermack & McKendrick (1927) introduced this quantity as a "threshold" for the population density as a condition for the epidemic to occur. They supposed the steady-state for the evolution of the illness in time. We can see that it barely happens in our case. Anyway, the average values are in the upper-right corner of the plot.

Finally, checking the "D-Number" box, you have the Death-to-Closed-cases ratio. This quantity appears on the WorldOMeter site (`https://www.worldometers.info/coronavirus/country/brazil/`) in the form of "Percent (%)Outcome of total closed cases (recovery rate vs. death rate)" with the same ratio relative to the recovered cases. They are complementary, so I think that the mention of both unnecessary. This number is important to know how much the illness is lethal.

Also, at the upper-right side of the plot above, you can see annotations giving the statistics of infections, observed ones, the asymptomatic total, and the number of deaths.

When you span the cursor through the text fields in the panel, a banner window appears with a quick explanation of what that quantity means. The graphic that appears at the beginning is the result of a simulation, according to the input parameters given by default, and presents four distributions along the time, the horizontal axis representing it. The vertical axis shows the counts.

2

Health Care System, the Risk Population and Deaths

You can see the count of the individuals that belong to the risk populations, usually in magenta. It means that these individuals need to be hospitalized and be assisted in Intensive Care Unit. The text field "Risk Population" at the bottom-left corner of the panel determines the fraction of the population that is considered under risk. If the individual has the health system's attention, it has a chance of dying given by the text field "Death Probability", just below. The Health System level is provided by the text field "Care Health System" at the panel's left-bottom side, in a fraction of the population. Its absolute value is plotted on the graphics represented by a dashed red horizontal line, tagged with "Health Care Capacity", to indicate when the number of infected risk individuals goes over this level. When it is the case, the probability of dying is given by the text field named "NoCare Death Prob," which is usually greater than the initial "Deaths Probability."

The second distribution, in red, is the number of deaths per step. You see that if the number of risk individuals is under or over the red line of the Health System, the number of deaths reflects the correspondent dying probabilities.

Steps and Population

The first text field at the upper left corner gives the simulation's number of time binning, that I call it steps. The number of steps represents the total number of bins in time of the simulation. Still, it is only plotted where there effectively is something happening, either infection or dead individuals. The defined number of steps may be insufficient to represent the development of the illness all along the time. In this case, you should change it to a greater one.

Below the number of steps' text field, we find the population's definition. The greater this number, the lower is the noise level of the result of each simulation. However, the processing time depends strongly on this parameter. Since the simulation takes each individual to test and change its state each step, the processing time grows in a combinatorial way, not a polynomial or exponential one. Increasing the number of individuals in the population may overload the computer memory and extend to undesirable computation time.

Incubation

Below comes the incubation period definition. It is in the form "n;m", that is to say, two integer numbers separated by a semicolon. They define the range into where the infectious agent rest incubating in the individual. An example is the new-coronavirus, which may be resting incubating from two to fourteen days in the organism. The sign ";" used to separate two parameters in the same text field defines a list. In the case of the incubation period, it then establishes a list of two integer elements. Each individual is set to an intrinsic incubation time (if it is infected) that is drawn randomly from this range.

Asymptomatic Population

The text field that comes next gives the fraction of population that is asymptomatic. It corresponds to they who does not present any desease's symptom, or that, for some reason, its condition is not communicated to the authorities, distorting the total of official counted infections.

Initial Infections

Every epidemic begins with a certain number of infected persons. The text field that follows gives this number. Of course, giving zero will be no epidemic at all, so the minimum value is the unit. During generating the population, the infected individual(s) is (are) drawn randomly according to a coin-toss process. So, the infected individual(s) may be in any place in the area.

Social Mobility

The text field Social Mobility gives the value of σ as seen in Eq. (1). It gives the parameter that measures the population's compliance to the ordering "stay at home". Mobility causes a concentration of infections but does not increase them. Its units depend on the ones adopted in the population density. A condition for the social mobility is that, from the model's construction, it cannot be greater than $\sqrt{N}/6$, which guarantees that no more than 0.1% of the population "escapes" from the epidemic area (see below, in Numerical Model).

Mobility Variation

As we is going to see in the section 2, Numerical Model, social distancing and mobility may change over time depending on the population's capacity to organize itself before the Epidemic emergence. We could mark this option if we chose to tag the mobility's variation check box. We will see it later.

Illness Period

As the incubation period, we can choose the period that an individual may be ill. Each individual will have a value drawn randomly from this range. During this time (steps) the person will be able to infect others.

Recurrence

Here, I call recurrence the individual's condition of being infected more than once, independent of whether it can die or be vaccinated. There is no limit to the number of infections while the individual keeps alive. The number in the text field means the population's fraction of susceptible to recurrence individuals.

"Recur Delay" is the next text field, which is the number of steps an susceptible to recurrence individual may be infected with again when exposed to an infectious condition, that is to say, be close enough to an infected one.

Activate Vaccination

One can program one or more vaccination periods. If you check the "Vaccination Active" box, the fields concerned will be enabled to be modified at your will. The text fields are the beginning of the vaccination and the period it lasts, the fraction of the population to be vaccinated, and the vaccine efficacy. All these fields may be filled with a list of values, that is to say, many values separated by semicolons. If you put a list in the field text "Vaccination Begins", the program takes the correspondent values in the other field texts to build the process of vaccination. They may have the same number of values, or less, in their lists, but not more. If the number in the list is less than the given at the beginning of vaccination, the program takes the last value in the list to define the processes left. The vaccination periods may overlap. So do the fraction of the population to be vaccinated and the vaccine efficacy.

Risk Population

A fraction of the population may die when infected by the epidemic agent, defining the risk group. This fraction is defined in the text field "Risk population". If in the risk group, a probability of dying is in the field text below in "Death probability". As long the number of "risk" people stay under the "Health Care System" threshold, the probability of dying remains the same. But, if it exceeds this limit, then it will be the value in the field "NoCare Death Prob", just below at right.

Profile of Infection

When two individuals get closed, the other is under a probability of infection if one of them is ill. This probability has a profile, such that the closer the individuals are, the more likely the contagious may be. You can choose two types of profiles: a power-law or exponential law. Below, in the section Numerical Model, these profiles are discussed in more detail. The role of the parameters "Virus Strength", "Social Distance" and "Slope" depends on the chosen profile.

Social Distancing

In few words, social distancing and virus strength is associated with the illness's capacity to spread. The parameter "slope" is related to the model of the contagious process. If we want to alleviate the epidemic, we must modify them. In real life, the best we can do is to keep away from each other the most it gets to increase social distance. We should use a mask to decrease the virus's strength. These two quantities and social mobility are the ones we can change to avoid the epidemic.

Social Distancing Variation

We know that the people of a community tend to change their behavior according to the epidemics state. If there are few infected persons, people will be more relaxed in respecting the necessary social distancing. The state of "relaxation" depends on the civic condition of the community in question. Organized societies will respond to the epidemic increasing quicker than low organized ones. Checking the box "Social Distancing Variable" will make the program variate the social distancing from near zero to near the value of Social Distancing, according to the ill population's density, following the "logistic" function. This point is seen in the section below. The fields "Civic Factor" and "Civic Zero" are the parameters for the logistic function.

Checking the box "Social Distancing Variate" enables the checking box "Mobility Variate". The mobility parameter will variate in the same way the social distance parameter does but in the inversed way. If the social distance increase, the mobility decreases, obeying the same logic that the civic acts do: for a large number of infections, the society tends to move less, as to increase the social distancing. Once the box "Soc Dist Variable" is checked, the check box "Mobility Variable" is put to be enabled. Then, you can choose to mark it or not. As long as the social distancing increases, mobility decreases, doing what a reactive community is expected to do.

Plot Presentation

Aside from the calculation parameters, we have some options to plot the results. We can withdraw or add the plot of some results, like the number of active infected per step, the asymptomatic number of infected individuals, the observed active infected ones (the difference between the former and the latter), the number of risk group, the deaths and so on, just by tagging or not the related check box. By default, all the possible numbers are plotted, but the "New Infections" and "Observed New Infections", which correspond to the real and observed recorded communications of cases. They are not plotted by default for the sake of cleanness. By combining the plots on the graphics, you may have the presentation you want.

Yet, you can choose to plot or not the level of Health Care Capacity the system may offer by checking the related box at the panel's bottom-center.

The graphic on the panel below presents some parameters of the epidemic evolution. Firstly, the so-called R-Number is the ratio between the number of newly infected by the active individuals coming by default. Clicking on the related box, you may have other estimators: the T-Number (that I am explaining below), and so on. They are helpful to understand the evolution of the epidemics.

The check box "Plot Integrated Curves" allows you to visualize the curves' evolution in an integrated form.

Steps and Time Unit

How could we interpret each step and its relationship with the timeline? We have three parameters having a direct action on the number of steps the epidemic occurs: "Incubation," "Illness Period," and "Recurrence Delay." The unit in time you adopt for these three parameters will define the step-unit. The default values when the program starts: "0;3", "3;5" and "30", respectively, are the typical values for the coronavirus in weeks. So the steps in the plot correspond to weeks. But this issue does not mean a deterministic relationship for a generic situation about the coronavirus. The number of steps increases with the number of individuals in the population. This application is not supposed to be a forecasting tool.

Seed of Random Series

By default, when re-doing a simulation, the program takes the same seed to generate the random number series that the simulation needs. By doing so, the program lets you choose all the calculation parameters at will. Thus, you may compare how a particular population's configuration behaves according to some parameters changing. By returning to the default values, the result of the simulation will be the same.

You can change this condition by clicking on the button "Clean Seed" to change the "seed" once or by unchecking the box "Recover Random Seed" when you have a new "seed" generated each time you want to re-do the simulation.

When the program starts, it seeks the newest file named "Epidemia_*.json" in the user root folder to retrieve the "Seed" stored in it. If there is no such file, the program creates a new "seed" as it is done when clicking on the button "Clean Seed".

Recalculate

The button "Recalc" allows you to re-do the simulation. Each simulation generates a different result, even if all the parameters are preserved. The calculation is based on random-generated numbers that the seed is other for each run, by an internal Java library procedure.

The button "Reset" re-do the simulation with all the parameters reset to the default data.

Output

There are three types of output: the plot image in PDF and the data generated in the simulation in "SQLITE" and "JSON".

Clicking on the button "SQLite," the simulation data goes to a file called "Epidemia.sqlite". If this file does not exist, the program creates one in your root folder. The SQLite form generates a structure containing the data inside the program. There is the table "Urbe", its the sub-structure, the table "Individual", and for each "Individual," there is the derived table "Infection" and the table "Position". The table "Urbe" contains all the parameters you entered to generate the simulation, that is to say, the values in the text fields in the plot window. The table "Individual" contains each individual's data in the population that was somehow modified by the program: being infected, vaccinated, or dead. The table "Infection" has the data of the conditions an individual was modified, having information of when it happened and who did it. Finally, the table "Position" contains all the individual's positions from the first step until the last one. Because of the data volume, the loading process of the file ".sqlite" is long.

The JSON file contains the simulation's context, the counts resulting from the simulation, and the R,T,D-Numbers. It is conceived to give data to plot in graphics.

All the output procedures announce the created file in a message-window at the end of the process.

Close

The button "Close" allows you to terminate the program.

2 Numerical Model

Variables

It is considered a two-dimensional space with independent variables, the number of beings (N), and the mean density ρ (assumed to be unitary), so the mean distance is obtained from the square root of the number of beings multiplied by this average distance, so

$$L = \sqrt{N},$$

where L is taken as the "space" where the beings live.

After each iteration, the position of the being changes according to the linear transformation:

$$\begin{aligned} x_{i+1} &= x_i + r_i \cos \theta_i \\ y_{i+1} &= y_i + r_i \sin \theta_i \end{aligned},$$

where

$$\begin{aligned} r_i &= \mathcal{G}\{0, \sigma\} \\ &\text{and} \\ \theta_i &= \mathcal{R}\{0, 2\pi\} \end{aligned} \tag{1}$$

The operators $\mathcal{G}$ and $\mathcal{R}$ are the random number generations according to the Gaussian distribution and linear distributions, respectively. Otherwise saying, the parameter theta - the angle coordinate in the polar system - may take any value between $0 \ldots 2\pi$. Still, the parameter R takes values that are more concentrated near the center (supposed to be the original position of the individual in consideration), such that, according to the Gauss distribution, it has around chances of 84% to be inside the circle $r \leq \sigma$ and 98% of being inside 2σ. The parameter σ is associated to the social mobility. The conditions above mean that each individual may displace to a point in the neighborhood in any direction keeping a distance from the original point, scaled by σ.

Notice that the "space" variable x or y may not be outside the limits $[0, L]$, so, if it is the case, a "bounce" operator is applied, such that:

$$x_{+1} = \begin{cases} L - (x \text{ MOD } L) & x > L \\ -x & x < 0 \\ x & 0 \leq x \leq L \end{cases}.$$

Otherwise saying, x is forced to be always inside the domain $[0, L]$. It stands for either x or y. The operator MOD stands for "modulus", the rest of the division of x by L ($L(s - \text{int}(s))$, $s = x/L$. It prevents the possibility that an individual is found outside the square $L \times L$, that the population is considered to be in.

Contagious Probability

The vector of distance $\vec{r}_{ij}$ is calculated for every pair of beings in space. It will determine the probability of contagious if one of the beings in the considered couple is infected. The contagious probability may be placed in two forms. The first one:

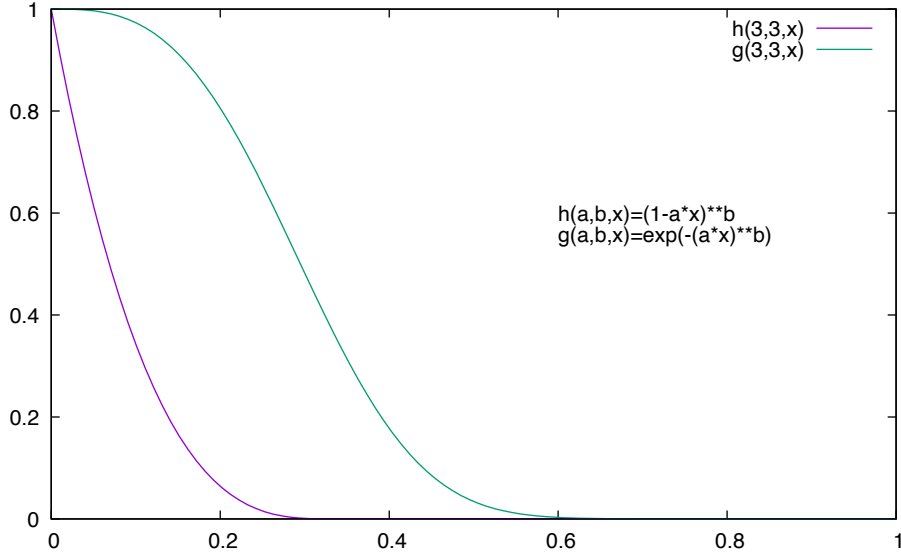$$p_c = \Phi \left(1 - \Lambda \left|r_{ij}\right|\right)^{\beta}, \tag{2}$$

Figure 2: Different slopes for the probability function p_c as given in Eq. (2) and Eq. (3).

where Φ is linked to the contagious power of the epidemic agent (for instance, the new coronavirus) and may take values in the limits $[0, 1]$, where 0 represents no contagious power (agent is innocuous) and 1 is the maximum contagious capability. β is the 'slope' of contagious. The parameter Λ gives the social distancing. It introduces a problem when it is > 1, so term inside the big parentheses becomes negative. Thus, there is an additional condition that states that if it is negative, it is replaced by zero:

$$Y = 1 - \Lambda \left| r_{ij} \right| = \begin{cases} Y & Y \geq 0 \\ 0 & Y < 0 \end{cases}.$$

In this way, the probability function p_c does the exact purpose we want for it, as seen in Fig. (2).

The second form is exponential:

$$p_c = \Phi \exp\left(-(\Lambda r)^\beta\right). \tag{3}$$

In this case no test for negative or crescent function is needed.

The maximum of the social distancing value will depend on the actual population density. If we adopt the average density as being the city of Rio de Janeiro's: $\sim 5300\,\text{hab.km}^{-2} = 5.3 \times 10^{-3}\text{hab.m}^{-2}$, the possible maximum social distancing is

$$\Lambda_{Max} = \frac{1}{\sqrt{5.3 \times 10^{-3}}} \simeq 13.7\text{m}.$$

This is the maximum possible distancing for a surface of 1200km^2, which is the Rio de Janeiro City's area.

Recurrence

One takes the "recurrence" parameter to stop the being's immunity to be infected by the illness and be susceptible to new contagious (t_N). It is the case of COVID-19. It is observed that the virus may re-infect those who were infected a time ago (let's say, three months, t_R). This re-infection possibility is individual, and its collective distribution is controlled by a parameter named "ethnic susceptibility" p_S. This probability is applied to choose some individuals with this "innate" condition to be re-infected at the population generation time.

Random Walk

To start the simulation process, one chooses the beings distributed randomly on the considered space to have the initial population spread with a given distribution. One may start with the uniform one. Other ones may be Gaussian $G[x,\sigma]$, Rect $R[x,\tau]$, or Teeth $T[x,\tau]$. Each individual gets an identification number that will be chosen to be firstly infected at the initial step of the simulation. The number of the initially selected individuals is previously defined by the parameter "Initial Infections". At each step (round) of the simulation the beings are sorted to be infected in a coin-toss simulation, with a bias that corresponds to the probability given by Eqs. (2) or (3) in function of the distance to the infected individual. The greater is p_c, the more likely the being will be infected, if one of the beings (i or j) is infected, of course. The scale σ is given in terms of the characteristic distance of the infecting agente, namely, the respiratory illness, like influenza or the new coronavirus, has the σ in the scale of meters.

Individuals Properties

Each individual should save in memory its contagious history, how many other individuals it had contaminated, the beings it met, if they were infected or not, and if it was infected or not at its own. It keeps the position at each iteration.

Asymptomatic Individuals

There are asymptomatic individuals that, even infected, don't count to the number of infections. However, they keep on infecting others. They are asymptomatic but may get into a severe condition, demanding Intensive Care, and even may die.

Social Distancing

The value of the social distancing Λ may change during the timeline. This variation is the sign of the "civic" capability watching the actual fraction of infected people. Setting Φ variation makes "people" react for greater social distancing if the number of infected individuals increases, relaxing the surveillance for it when decreasing. The best function that represents this behavior is the "logistic" function:

$$\Lambda(u) = \frac{(\Lambda_0 - \Gamma_0)}{1 + \exp\left(-a(u-1)\right)} + \Gamma_0, \ 0 \leq \Gamma_0 < \Lambda_0. \tag{4}$$

The parameter u may represent the fraction n_R/H, where n_R is the actual number of risky infected individuals and H is the System Health Care capacity multiplied by the number of individuals in the population. Λ_0 and Γ_0 are the maximum and minimum social distancing a community may take and, so to speak, is associated with highly organized societies. On the other hand, the parameters $a = [1, \infty]$ and $\Lambda_0, \Gamma_0 = [0, \Lambda_0]$ also reflect the development of the society, such that little values appoint to the speed that the community may react to the increasing state of the infection, since the function is nearly linear: the more infected people we find, the faster is the increasing of the social distancing. Low organized societies act as a herd: reacts slowly to the infection increasing state, and when it reaches a dire condition, people go to a radical (as possible as it can be) distancing. It is the situation of high a values. The logistic function is also used to the social mobility variation, except that it is in the reversed way: when the social distancing increases, the mobility decreases. It suffices to invert the signal of a, inside the exponential term. We can see the plot of the logistic function, as defined above, in the Figure 3.
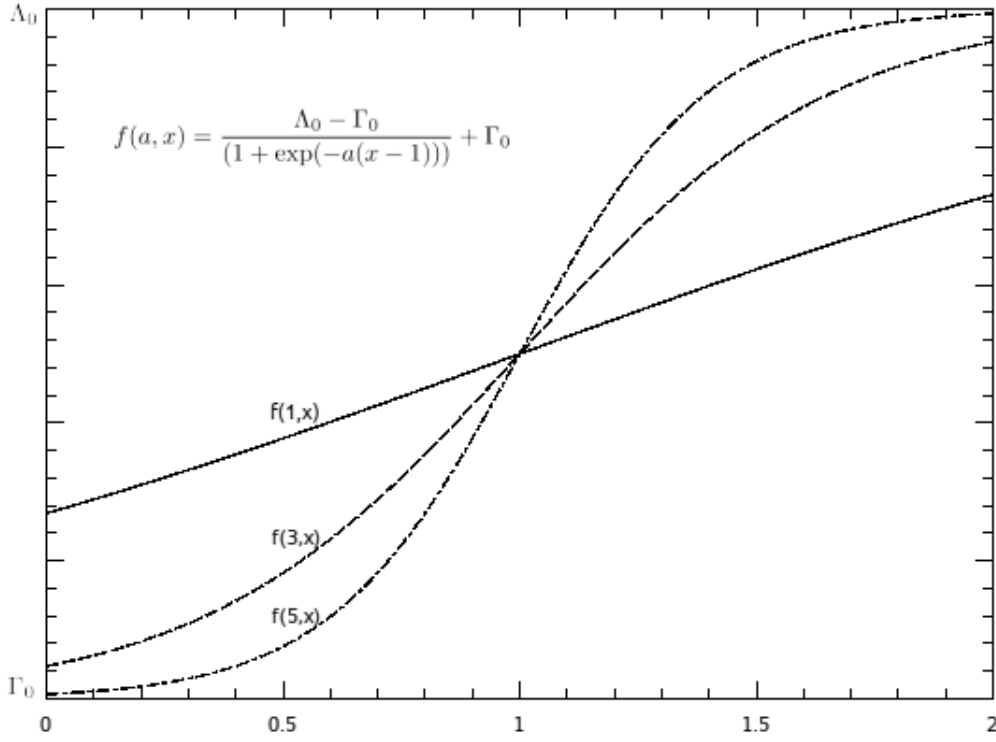
Figure 3: Logistic function for some values of a. Different values of Γ_0 and Λ_0 compress or stretch the logistic function.

Vaccination

We should consider a period of vaccination. The vaccination's role is to "infect" the individual with no risk of dying and no means for him to infect another one.

Efficacy

The immunity a vaccine offers is not a hundred percent of efficacy, so there is a probability the vaccine can immunize. Thus, we must consider an index of vaccination efficacy.

Deaths

If the individual is infected, he may die. This condition is dictated by the parameter the individual has, named "MayDie". A fraction of the population is set for this parameter to have the true value. If the infected individual may die, the coin is tossed for drawing if this is the case or not. Once he is dead, he interacts no more.

Health Care System

The locale where the individuals live should have a certain level of health care. Usually, the local officials install a hospital with a capacity of assistance compatible with the local population. When the community is under the attack of an epidemic agent, the health care system should be able to harbor its patients, and under these conditions, there will be some deaths that we would say are inevitable. However, it may happen that the number of patients needing care will overwhelm the system's capacity to receive them in the hospital. Then, the probability of deaths radically increases for individuals to be put out of the system.

Given the Care System Level, the program tests if the number of individuals needing Intensive Care is over this threshold level. If it is the case, the probability of death goes to a new value given in the text field's correspondent parameter. If not, it remains the usual value.

Discussion

First simulation: The first plot comes from a simulation done automatically from an initial set of parameters. These parameters come from the content of a JSON file saved when the app was applied or, in the absence of that file, a set of default parameter values. The graphic has the aspect seen in the Figure 1. From the figure, we can see that almost all susceptible individuals of the population were infected at the end of the process. Nearly 30% is asymptomatic, so the observed process is only about 70%. From the risk group, 21% were contaminated, and 15% (69% of the group) of the population died. The death toll is over 50% of the risk group because the health system's capacity has been exceeded, as seen in the figure. The probability of dying went from 50% to 90%, defined in the "DeadNoCare" text field. It is acceptable if you are dealing with lambs, pigs, cows. Not with humans. There is no worse situation if you, or your wife, or your son or parents need health care and you hear at the hospital entrance: "COMPLETE"! It is the collapse of the Health System. We saw it in Lombardia Region in Italy, in Manaus, Amazonas State in Brazil, and in some cities where this condition is very close to being reached, like Rio de Janeiro, São Paulo.

In the figure, you see the Health Care System-level plot, showing the regions where the system capacity is exceeded: the line in magenta (risk group) and the deaths (in red) coming from this condition. The Health Care level is defined in the text field "Health Care System".

The model for the simulations is as described by Kermack & McKendrick (1927). One takes the number of infected, step by step, one simulates contacts between these infected individuals with susceptible ones in the neighborhood. According to the distance between them, a probability is evaluated, following the law given in Eq (1) or Eq (3), depending on what of them you chose in the control panel (lower left side). When all the contact possibilities are tested, one counts the number of infected individuals for that step, as well as the asymptomatic cases, the observed number of infected individuals (the difference between the former and the latter), the number of individuals in the risk group and the dead ones. Then, one goes to the next step to perform the same operations until there are no more infected individuals. The result is seen in Figure 1. In the panel below the infections panel, we may observe three kinds of plots. The first one showed when the program starts is the ratio between the newly infected individuals and the infected ones in the last step. This ratio gives how many people an infected individual touch others to be infected too. This number is known as R-ratio or R-number. If it is greater than the unit, the epidemic is rising; otherwise, it points to declining.

Checking to box labelled "T-number", we have the second kind of plot, that is the relation between the removed people (closed cases: recovered plus dead) and the new infected ones for a single step. I call this ratio T-number because Kermack & McKendrick raised the hypothesis that the relation between the removed rate and the infection rate represents a threshold of the population density for the epidemic to happen. Rigorously, the infection and remotion rates, in a sense defined by the authors, are tough numbers to estimate since they combine the equations in a very complex way. The numbers showed in this panel are very different from the rates defined by the authors. Anyway, they may hint at some interesting conclusions, as we will see later on.

The last kind of plot to be seen in the lower panel is the "D-number", the ratio of dead people to closed cases (removed: dead + recovered people). The site "Worldometer" (`worldometers. info/coronavirus/`) shows the plot of this number's evolution with its complement. This number's meaning is not clear to me yet, but I hope to figure it out sometime.
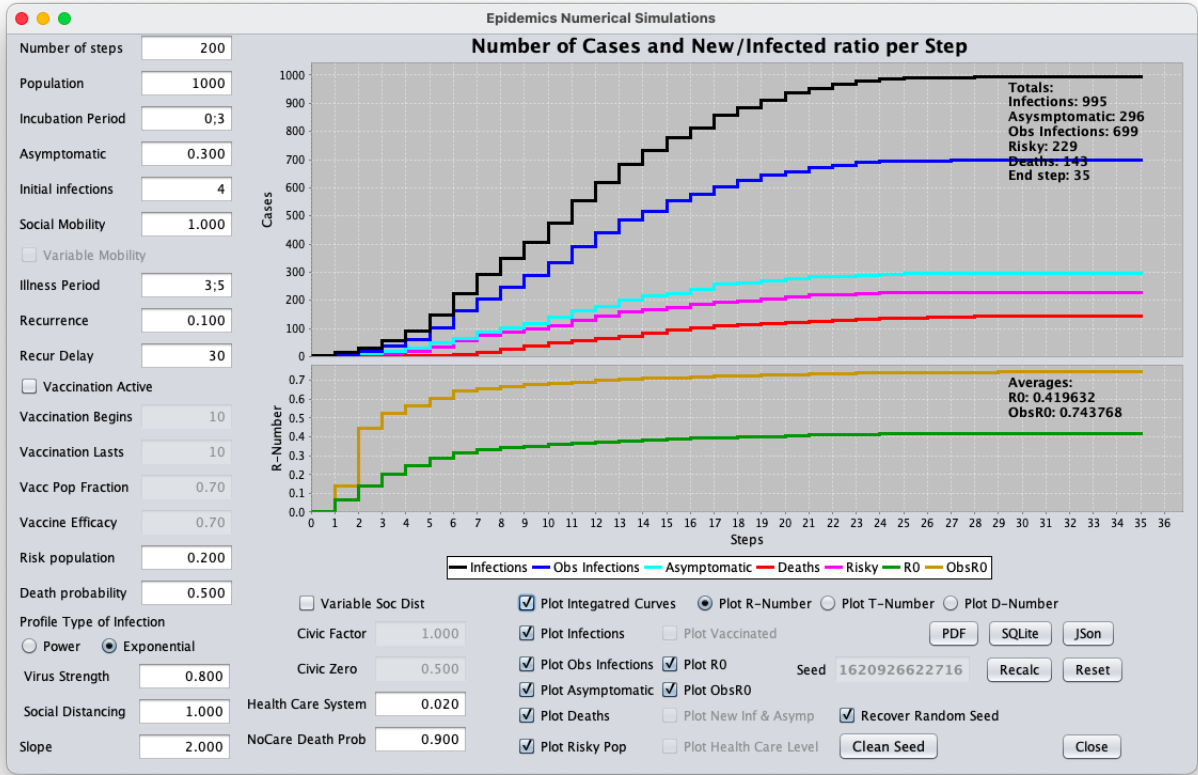
Figure 4: Integrated function curves.

Infections Panel (Upper Panel)

The curve in black represents the number of infected beings in the considered bin period. It is not the number of newly infected beings. More than the latter, it matters to know the actual state of a particular condition of the community's epidemic. It is known as the "active population". They are those who are infecting people. One wants to see the probability of being infected when sticking around the street, and this number may give it.

The curve in light-blue shows the evolution of the asymptomatic individuals that are being infected. The difference is the observed number of infections, that is, the curve in turquoise. It is important to note that this curve indicates what the health authorities have to analyze the epidemic evolution. It is especially sensitive in the coronavirus case. We see many people who don't manifest any symptom of the illness or the possibility that they are so weak that it does not communicate with the officials.

We also see the number of risk groups (magenta) and deaths (red) per round. A dashed, thin red line marks the threshold of the health system capacity. According to values in the concerned text field, if the risk group people exceeds this level, the probability of dying changes from 50% to 90%. You can change these values to fit the particular conditions of the infectious agent in the study.

Additionally, we can see the evolution of the newly infected individuals and the fraction of asymptomatic by checking the box "Plot New Inf & Asymp". They are not shown by default to not overload the figure with curves that could be some confusion.

Furthermore, we can turn the plots into their integrated form (Figure 4). These curves, especially the integrated infections, are sigmoid functions: to say, bound curves, having null derivatives in the domain's edges: $-\infty$ and $+\infty$. These curves may be model, under some simplifications, as a Gompertz-type function (PhysRevResearch.2.043381, 2020):

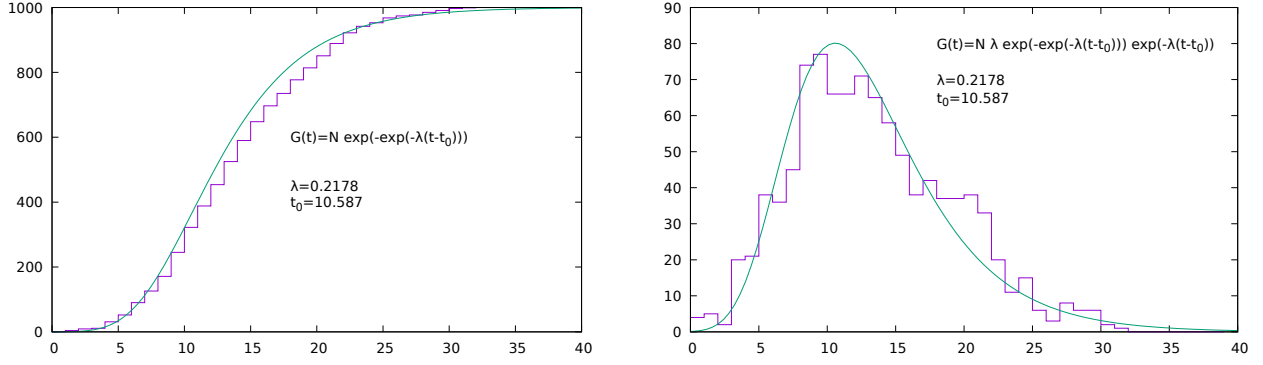$$I(t) = I_0 e^{-e^{-\lambda(t-t_0)}}, \tag{5}$$

12

Figure 5: Curves of infections, new (right) and integrated (left), with the fitted Gomperz-type curve.

where I_0 is the final number of infected beings, and λ and t_0 are parameter to be determined. The parameter λ is understood as the epidemic's speed, and t_0 is the inflection point, the maximum of the newly infected beings (the integrated curve derivative).

Understanding the Integrated Curve

The curve of integrated number of infections is a solution to a set of ODE introduced by Kermack & McKendrick (1927), and since then, their theory gained complexity and sophistication. They considered the case of the epidemic taken the totality of the population. The equations are:

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

$$\frac{dI}{dt} = \phi I(t) S(t) - \psi I(t) \ .$$ (6)

$$\frac{dR}{dt} = \psi I(t)$$

Where $S(t)$ is the function of the number of susceptible individuals to the illness, $I(t)$, as seen before, the number of newly infected ones, and $R(t)$ is the number of the removed ones from the set of infected. The (supposedly) constants ϕ and ψ are the infection rate and remotion rate, respectively. As the authors show, by developing the solutions (they stay in solving only for $R(t)$), the relation between the pairs λ, t_0 with ϕ, ψ is very complex, and there is no point in trying to find it. Let's see what the pair λ, t_0, gives some clues about the epidemic's evolution. Thus, we can fit the function $I(t)$ by simple regression from the plot of the integrated number of infections. The result, we see in Figure 5.

The fitting is not perfect because the simulation does not follow precisely the assumptions the authors have made. However, for the conditions we have, it is in a good approximation. The plot on the right is the derivative for the same parameter values. The I-derivative is in the form:

$$G(t) = I_0 \lambda e^{-e^{-\lambda(t-t_0)}} e^{-\lambda(t-t_0)}.$$ (7)

Variating the simulation parameters "Mobility", "Virus Strength", and "Social Distancing", as well as the parameters of social distancing and mobility variations, results in changing of I_0, λ and t_0.

Determining Quantities

Once we have the theoretical function and simulation values in good accordance, we may deduce some of the quantities above described from the data. Constraint conditions may help: at $t = 0$, we have the initial number of infected beings:

$$N_0 = I_0 e^{-e^{\lambda t_0}},$$

13

or
$$\lambda t_0 = -\log(\log I_0 - \log N_0).$$

On the limit, every epidemic begins with at least one infected, so $N_0 = 1$, and

$$\lambda t_0 = \log \log I_0,$$

so, the parameters λ, t_0 and I_0 are tied by one condition. In practice, the evolution of a epidemic is defined by only two parameters: two of these three.

Rodriguez & Helene (2020)[1] estimated λ and t_0 for some countries and, from them, deduced I_0 the total number of infected people for the Covid-19, taking data until Aug, 2020, and for some cases, considering two or even three "waves". Actualized data show that these forecasts are far from reality. We see how unpredictable are all anticipation trials, but there is a reason that I discuss later.

The epidemic's end's theoretical determination is challenging since the Gompertz function reaches the total number of infections at infinity. Anyway, we can determine when there is one infected left:

$$e^{-\lambda(t_1-t_0)} = -\log \frac{I_0}{I_0 - 1},$$
or
$$\lambda t_1 = \lambda t_0 - \log\log \frac{I_0}{I_0 - 1} \qquad .$$
and
$$t_1 = \frac{1}{\lambda}\log\log I_0 - \log\log \frac{I_0}{I_0 - 1}$$

Let's take the last "log" on the right hand of the previous expression:

$$\log \frac{I_0}{I_0 - 1}.$$

Considering the epidemic, we could say that $I_0 \gg 1$, so we could write:

$$\log \frac{I_0}{I_0 - 1} = \log I_0 - \log(I_0 - 1) \approx \log I_0 - \log I_0 + \frac{1}{I_0} = \frac{1}{I_0}.$$

Putting it on the equation:
$$t_1 = \frac{1}{\lambda}\log\log I_0 + \frac{1}{\lambda}\log I_0.$$

We know that the time of the epidemic, T is a little greater than t_1, but we don't know how much it is longer. As a rule of thumb, we would say that

$$T \gtrapprox \frac{1}{\lambda}\log I_0.$$

Having T, we deduce a rough approximation of λ:

$$\lambda \gtrapprox \frac{1}{T}\log I_0.$$

I use to call $\tau = 1/\lambda$, the inverse of the epidemic "speed", the epidemic's characteristic time. So

$$\tau \lessapprox \frac{T}{\log I_0}.$$

Now, from what we have seen, it is time for a little exercise:

---

[1] PhysRevResearch.2.043381

- A city has a population of 10 thousand people. When the mayor knows a coronavirus epidemic is coming on, he decides to close the town: nobody gets in or out. But it is too late. Someone infected has been brought to the city before. When people figure out what is happening, the epidemic already has taken the population. Considering that for a population of one thousand people, the characteristic epidemic time is five weeks:

  1. How long does the epidemic take place in this city?

  2. When the infection level goes to maximum?

  3. Use the app Epidemia to verify your conclusions.

The Maximum of Infections

We see that the curve of newly infected beings is different of the curve of the number of infected ones per round. We can understand that there are some important differences between these two curves. The main reason is that an infected individual takes a time to recover (or die) and stay in the place infecting others. The program draws a number between 0 and the illness period for each individual to set the date it eventually dies, according to the probability it is set to him. On the other hand, when the illness period is reached, he gets recovered. Then, the number of infected people per round is greater than the new ones count. The maximum is also expected to be a bit later.

To figure out the maximum number of infected people per round, we need to sum the function through an average of these infected people's illness period. A reasonable estimator is

$$T = \frac{n_{min} + n_{max}}{2}.$$

The summation would be well approximated in the integral:

$$N(t) = \int_{t-T}^{t+T} G(t)dt = I_0 \left[\exp\left(-e^{-\lambda(t+T-t_0)}\right) - \exp\left(-e^{-\lambda(t-T-t_0)}\right) \right].$$

To find the maximum of N, we derivate it, put it equal to zero and solve the equation:

$$t_M = (t_0 - T) + \frac{1}{\lambda} \log\left(\frac{e^{2\lambda T} - 1}{2\lambda T} \right),$$

where t_M is the point of maximum number of infections. At this point, it is

$$N_{max} = \left\{ \exp\left(-\frac{2\lambda T}{e^{2\lambda T} - 1} \right) - \exp\left(-\frac{2\lambda T e^{2\lambda T}}{e^{2\lambda T} - 1} \right) \right\}.$$

Units: Time and Space

Units of time and space are not defined in the Epidemia simulation. They depend on the values adopted in the input parameters for the calculations. The default parameters define the new coronavirus's characteristic properties: incubation from 0 to 3 (for the coronavirus, the time unit is a week). Also, the illness period (3 to 5) is characteristic of new-coronavirus if it is in weeks. The infection recurrence of 30 (weeks) is compatible with the recurrence for it (7 to 8 months). So the abscisses adopted in the plot are supposed to be in weeks.

For space units, since the population density is forced to 1, we must focus on social distancing. If we adopt the contagious model to be a Gaussian (exponential function type with slope = 2), a social distance of 1 is proper of the new coronavirus (as all respiratory illnesses like influenza, HxNx, etc.) when we adopt the meter as space units. So the default values of Epidemia for social distancing and mobility are compatible with distance units of meters.

The Meaning of Social Distancing

Though population density and social distancing are quantities that point out to the same concept, that is to say, an average distance between people, they refer to different conditions. If we take, namely, the city of Rio de Janeiro, the population density is about $5300\,\text{hab/km}^2$, according to IBGE. It is the same as $\sim 1\text{hab/200m}^2$. It is equivalent to having people at a mean distance of $\sim 13\,\text{m}$ one to another. If we adopt this value for social distancing, the epidemic does not take place. Rio de Janeiro's people may be at a maximum distance of 13 m, but they don't. People concentrate on very dense regions, like Copacabana, and leave vast areas empty. Thus, the population density is taken as a social distancing limit. In Rio de Janeiro City, one cannot go over it.

How Many Individuals to Take?

A population of 1000 individuals maybe not representative of the actual cases to be examined. Taking the simulation of 1000 beings' population has nothing to do with Rio de Janeiro or any other community in the world. Furthermore, cities are hardly closed systems, as we assume in the simulation. We must understand that the simulations in Epidemia are taken nothing more than what they are: simulations. They cannot be used as predictors. All they may give is qualitative results. They would be helpful for didactic purposes, not analytical studies.

Taking values much greater than 1000 may be very hard for the computer to run. Greater values increase confidence in the results, but it is useless, as we will see ahead. It is helpful if we keep in mind that the infection curves' peak increases proportionally when the population increases. However, the instant of the peak goes with $\log(\log(N))$, and the epidemic's duration increases with $\log(N)$. If we take Rio's population: 6 million people, so, a epidemic would last a couple of weeks more. Taking a population of 1000 people, or so, to study the epidemic would not be very far from the reality, at least on qualitative terms.

Flattening the Curve

If an epidemic is not acceptable (for humans, especially if its lethality is such that the health care capacity is exceeded), we must take measures to defuse it. We may modify three parameters: social distancing, virus strength, and mobility. They are, at some sort, connected, but they must be, at least, controlled. To increase social distancing is to tell people to keep the greatest distancing from each other as they can. People must guard distance in lines, malls, pharmacies, and so on. To diminish the virus strength is to use a mask to protect respiratory body entrances. To reduce mobility is to tell people to stay at home, to close theaters and sports stadiums. It helps the social distancing too.

There are numerical estimators for all these three parameters. They are not necessarily the same applied in Epidemia. If you are interested in make the correspondence between the officially adopted parameters and the used ones in Epidemia, go ahead, and share your conclusions, that will be welcome.

Running Simulations

Some exercises

The mayor of our fair 1000-people-village, knowing about the Covid-19, closes all public locations, theater and stadium, imposes a curfew and tells people to stay at home as much as they can. The Epidemia panel corresponds to cut the mobility by 50% and the virus strength to half and increases social distancing by 50%. Let's see what happens in Figure 6.

The measures taken to reduce the impact of the epidemic were effective. The measures taken to reduce the impact of the epidemic were effective. Infections dropped by 90%, and, most importantly,
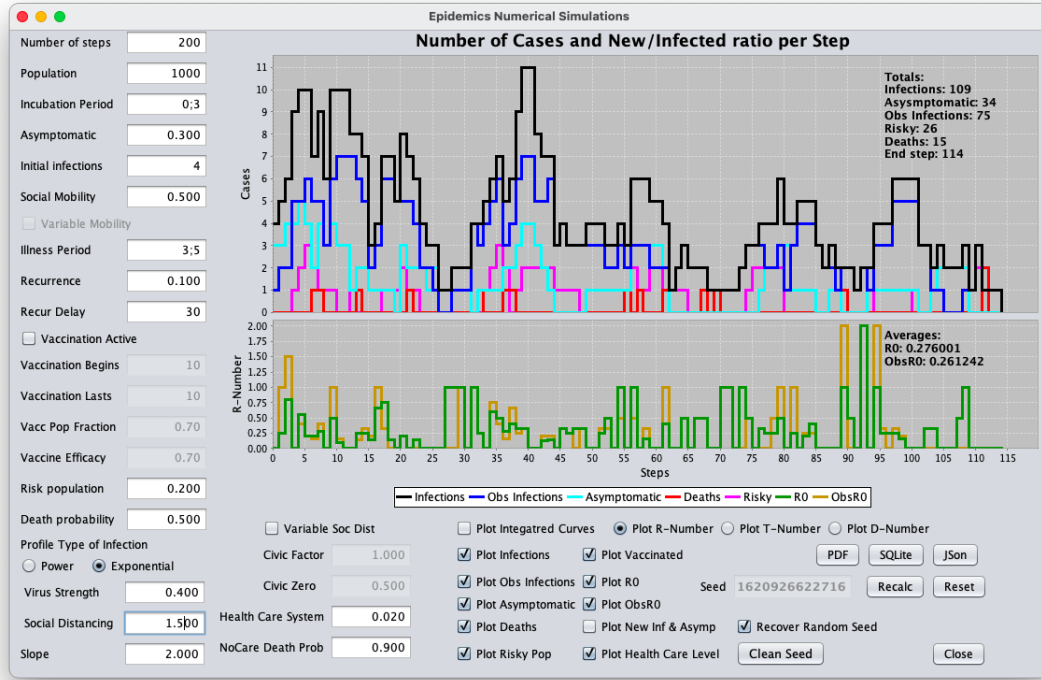
Figure 6: Lock-down in the little village.

the health care was sufficient to attend to all who needed it. There is a cost. The epidemic duration increased from 38 to 114 weeks, two years, two months and two weeks of the social pain. It worthed. The mayor would be proud of his people.

Mathematically speaking, a regime of chaos exists when the system behavior strongly depends on the initial conditions. Just a slight disturbance in the system's state, the final result changes radically, pointing to a completely different situation. In the case of epidemic evolution, Kermack & McKendrick had already noted such a chaotic characteristic when they state: "In this sense, the population density may be considered as a threshold density of the population for an epidemic with these characteristics. No epidemic can occur unless the population density exceeds this value, and if it does exceed..., to a first approximation, ... the size of the epidemic will be, ... equal to... twice the excess". That is to say, a slight change in the considered threshold and the epidemic gains twice as strong the former regime. It means a significant difference for a bit of variation. Just a chaotic regime. By clicking on the button [`Clean Seed`] or by checking out the box [`Recover Random Seed`], followed by clicking on [Recalc], you may observe significant variations of the infection curves. For the same parameters set, very different scenarios of the epidemic may present themselves. For me, this inconstancy points to a chaotic regime.

Now, coming back to our little village, let's say that it would be in a country like Germany or Japan. In countries like these, people can keep the same habits and maintain the social parameters constant. In countries like Brasil and so many others, people do not follow the authorities' instructions rigorously. They tend to release the social distancing policy when things seem to go fine and panic when the situation gets serious[2]. To deal with this condition, we have the option to check the box [`Soc Dist Var`], that is to say, "Variable Social Distancing". As seen previously, Eq (4) reflects the social panic described here. By placing the parameter "Civic Factor" at 3.0, we hope to correspond to what people do in a moderately developed civilization. Checking in the box [`Variable Mobility`] also set the same tendency concerned with social mobility. Let's redo the simulation (clicking on [`Recalc`]).

---

[2]This behavior cannot be credited simply to "ignorance", as some wish. In fact, most people need to work on a daily basis (not necessarily in formal jobs) to have the means to feed themselves and their kins.
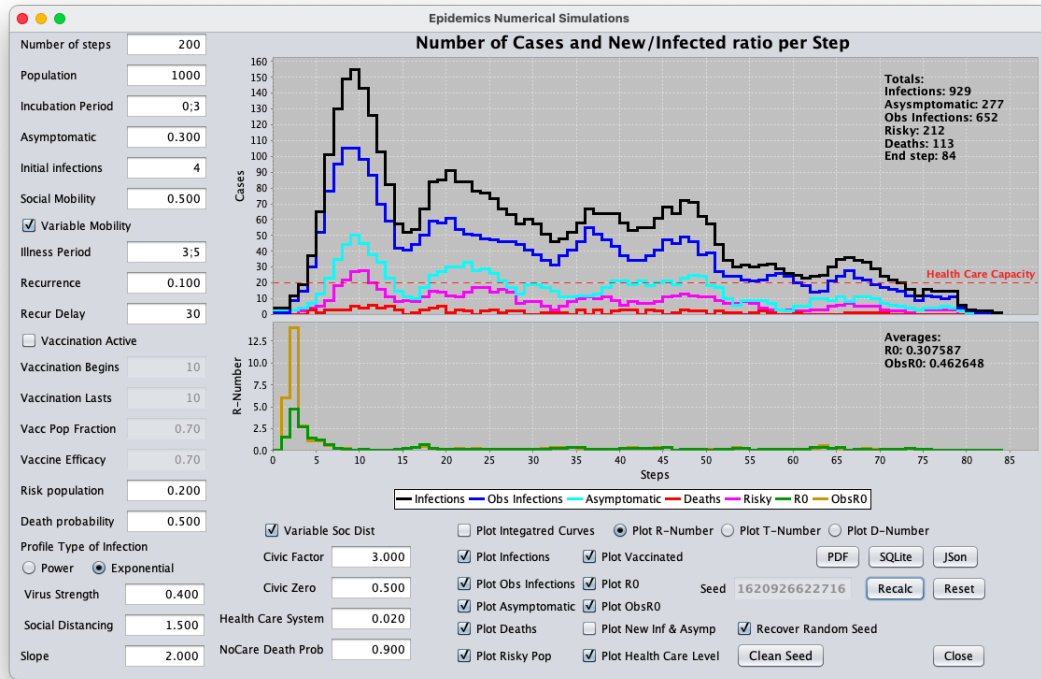
Figure 7: Variable social distancing and mobility: epidemic again.

As we see in Fig (7), the situation of a epidemic comes back. It means that waiting for an emergency to take the necessary protection measures does not work. It is practically the same as doing nothing. It takes longer with retaking outbreaks in 'waves'. Worse is the authorities to behave in contradiction, as is Brasil's case. The only hope is the vaccination.

Vaccination

Let's take the vaccination parameters the program suggests by default:

- Beginning at step ten;

- Lasting ten steps;

- Fraction of population to be vaccinated: 70%;

- Vaccine efficacy: 70%.

The numbers above are supposed to be the "normal" ones for a generic vaccination process. The beginning at step ten is the 'reasonable' time that the authorities realize the epidemics and set up a consistent campaign. Lasting ten steps are also expected for the organized logistic to furnish and cast vaccines overall the territory. The result is seen in Figure 8.

As we see, it does help a little. In this particular case, from a level of 99%, we reach 62% for the best. Many retakes with different random seeds show this result is consistent. And so? What is the vaccination for? We see, retaking the social protection measures, that the immunization has efficacy only when:

1. It is applied at the very beginning of the epidemics and with a high degree of population coverage and drug efficacy, or;

2. The social protection measures are not abandoned because of the vaccination campaign.

It is clear that it is not possible to follow these actions. The only thing left to be sad: try to stay at home the longest and keep observing the social distancing measures. See Figure 8, below.
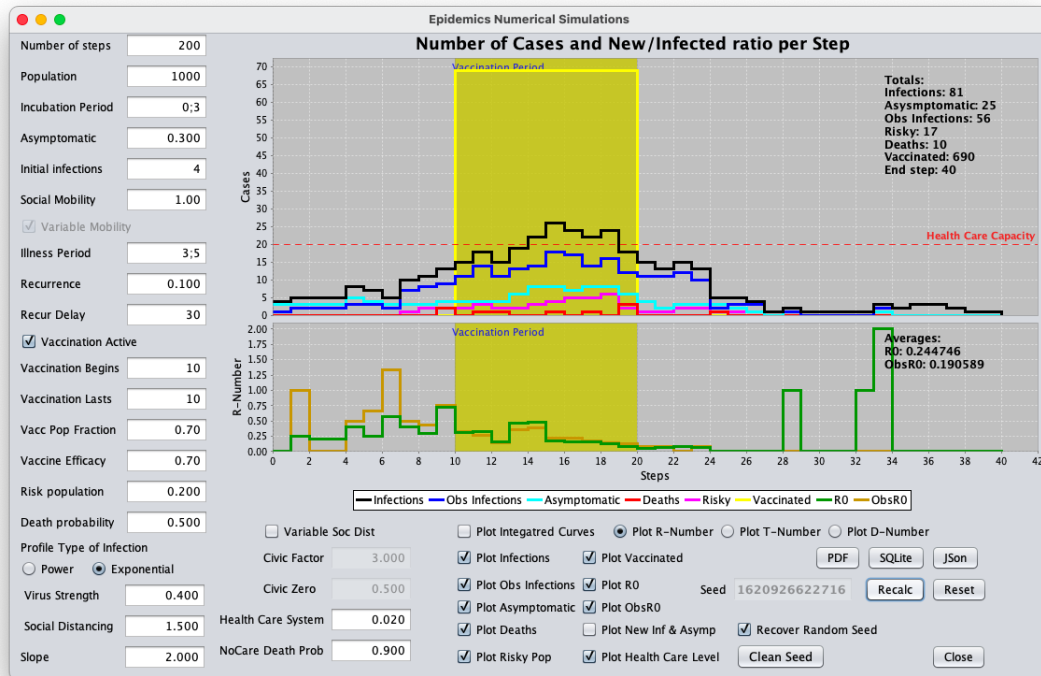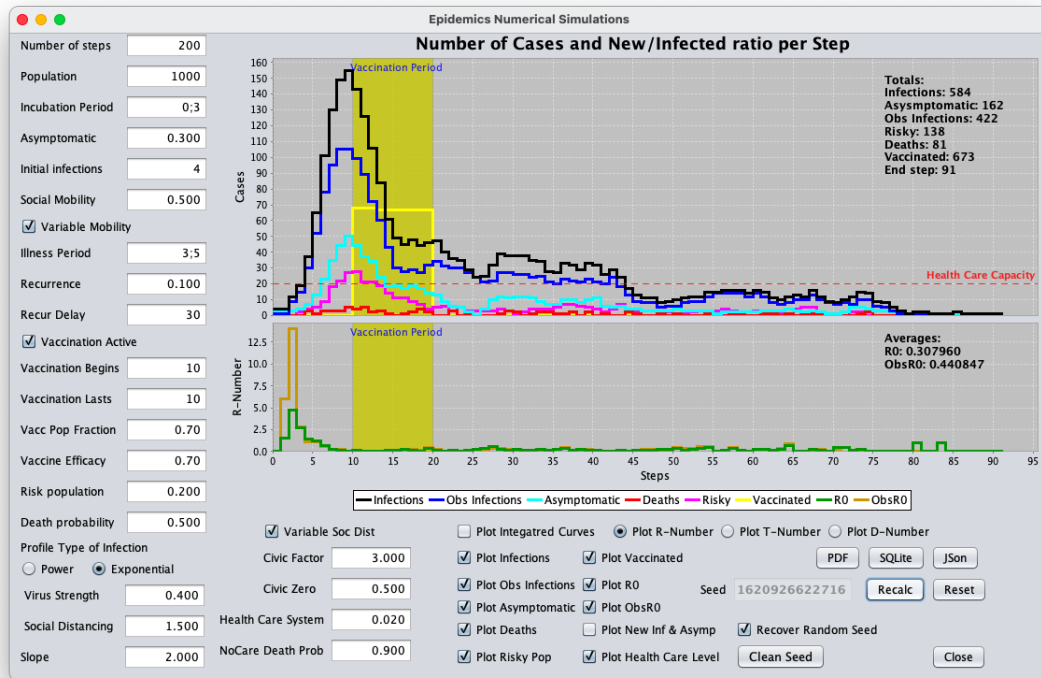
Figure 8: Epidemic with Vaccination for different social conditions.

Outreaching

Three buttons are available to output the results of simulations in Epidemia. They line horizontally at the lower-left side of the panel: 'PDF', 'SQLite', and 'Json'. The first one, tagged as 'PDF', just produces a PDF version of the plot. Let's talk about the second, 'SQLite', following the third, 'Json'. The latter has a JSON file containing all information about the parameters used by the program to perform the simulations (the ones in the text fields, radio boxes, and checkboxes in the panel), all gathered at the header of the file.

Next, a sample of a JSON file as an output of a simulation.

```
{
Seed: 1616429969541
SimulationTime: "2021-03-22T13:19:48.558815"
NRounds: 200
NPopulation: 1000
Incubation:[
 0 3
 ]
Asymptomatic: 0.3
InitialInfected: 4
SocialMobility: 1.0
IllnessPeriod:[
 3 5
 ]
Susceptibility: 0.1
SusceptDelay: 30
Vaccination: FALSE
VaccinationBeginning:[
 10
 ]
VaccinationLasts:[
 10
 ]
VaccinationFraction:[
 0.7
 ]
VaccinationEfficacy:[
 0.7
 ]
RiskyPopulation: 0.2
DeathProbability: 0.5
ProbProfile: "Exponential"
VirusStrength: 0.8
SocialDistance: 1.0
ProfileSlope: 2.0
VarySocialDist: FALSE
VarySocialMobility: FALSE
CivicFactor: 1.0
CivicDisplacement: 0.5
HealthCareLevel: 0.02
DeathNoCareProb: 0.9
Counts:{
Infections:[
```

```
 4 9 11 31 52 88 123 164 232 297 337 381 413 427 426 414 405 369 343 305 281
257 235 203 174 150 108 89 70 60 43 32 25 22 11 4 3 2 0 0
 ]
Asymptomatic:[
 1 3 5 11 16 25 32 44 61 78 85 101 117 117 113 109 107 103 102 102 91 81 73
69 60 50 36 23 16 13 9 8 6 5 1 0 0 0 0 0
 ]
NewInfections:[
 4 5 2 20 21 38 36 45 74 77 66 66 71 65 58 49 38 42 37 37 38 33 20 11 15 6 3
8 6 6 2 1 0 0 0 0 0 0 0 0
 ]
NewAsymptomatic:[
 1 2 2 6 5 9 7 13 20 21 17 19 24 16 13 12 11 16 17 15 13 7 5 5 3 1 1 2 2 1 0 0
0 0 0 0 0 0 0 0
 ]
Risky:[
 1 2 3 6 11 15 23 31 41 53 62 75 76 86 92 90 90 79 75 67 61 56 57 50 41 37 27
25 18 13 10 9 7 7 5 2 1 1 0 0
 ]
Deaths:[
 0 0 0 0 0 0 0 1 2 2 3 5 0 5 12 2 10 6 14 5 9 2 5 8 3 6 3 5 4 0 1 1 0 1 2 1
0 0 0 0
 ]
Recovered:[
 0 0 0 0 2 1 4 5 10 24 19 34 51 54 49 45 68 57 61 57 48 40 38 36 27 39 24 20
12 19 11 6 3 10 5 0 1 2 0 0
 ]
AsympRecovered:[
 0 0 0 0 0 0 1 3 4 10 3 6 16 15 15 11 19 17 15 22 15 13 8 8 10 13 14 8 2 4 1 1
1 3 1 0 0 0 0 0
 ]
}
}
```

The fields in this file are easily recognizable with the fields in the Epidemia panel. What comes next are the values of the counts for the quantities plotted in the panel. We straightforwardly recognize them too: Infections, Asymptomatic, etc.

Finally, we find the button 'SQLite.' This task is by far the heaviest and hardest to execute. You must patiently wait for the end of the proceeding. It stores, besides the whole set of control parameters - as in the case of JSON file, all the state of each individual being in the simulation, step by step, or round by round, as I call the steps in this program. All of it in the format of an SQLite file, so you can use your favorite application to access and query data from it (I use "DB Browser for SQLite". It is free and friendly).

Each time you click on the button 'SQLite', after giving a name to your data set, it will store all these data into the file '~/EpidemiaWorkDir/Epidemia.sqlite' (if you are work in a Unix-like environment).

Here, the 'pragma' listing of the tables 'urbe' and 'individual', present in the Epidemia.sqlite file.
Urbe:

cid	name	type	notnull	dflt_value	pk	comments
0	nickname	text	0		0	A name you give
1	date_time	DATETIME	0		0	Given by the system

```
2      seed            int       0              0      For random series
3      rounds          int       0              0      Number of steps
4      npop            int       0              0      Number of individuals
5      incubation      text      0              0      Values of incubation
6      asymptomatics   double    0              0      Fraction of
7      initinfecs      int       0              0      Initial # of infected
8      socmobil        double    0              0      Social mobility
9      varmobil        boolean   0              0      If variable social mob
10     illnssprd       text      0              0      Values of illness time
11     recurrence      double    0              0      Fraction of recurrent
12     recurrencedelay int       0              0      Time for recurrence
13     vaccination     boolean   0              0      If vaccination active
14     vaccbegin       text      0              0      Starting times for vacc
15     vacclast        text      0              0      Lasting times for vacc
16     vaccfrac        text      0              0      Fraction to vaccinate
17     vacceffic       text      0              0      Vaccine Efficacy
18     riskypop        double    0              0      Fraction of risk pop
19     deathprob       double    0              0      Prob of death (regular)
20     infctprofil     text      0              0      Profile of infect curve
21     virstrngth      double    0              0      Virus strength
22     socdist         double    0              0      Social distancing
23     virslope        double    0              0      Slope of infec curve
24     varsocdist      boolean   0              0      If variable soc dist
25     civfac          double    0              0      Civic factor (Logistic)
26     civdist         double    0              0      Civic distancing (Log)
27     carelev         double    0              0      Level of health care sys
28     deathnocare     double    0              0      Death prob for no care
```

Individual:

cid	name	type	notnull	dflt_value	pk	comments
0	idindiv	int	1		1	Indiv ID
1	idurbe	int	1		2	Urbe ID
2	round	int	1		3	Step of simulation
3	infecround	int	1		0	Point when is infected
4	recurrence	boolean	0		0	If indiv may be recurr
5	illnssprd	int	0		0	Period it get sick
6	x	double	0		0	X-Position
7	y	double	0		0	Y-Position
8	maydie	boolean	0		0	If indiv is risky to die
9	asymptomatic	boolean	0		0	If indiv is asymptomatic
10	vaccinated	int	0		0	When it was vaccinated
11	dead	int	0		0	When it dies
12	incubation	int	0		0	Incubation time
13	immune	boolean	0		0	If it is immune
14	infected	boolean	0		0	If it is infected

For example, to query about the number of infections in a given simulation (urbe, in the Epidemia jargon), we could loosely query:

```
select round,count(*) as infections from individual
where infecround>0 and round>infecround+incubation
```
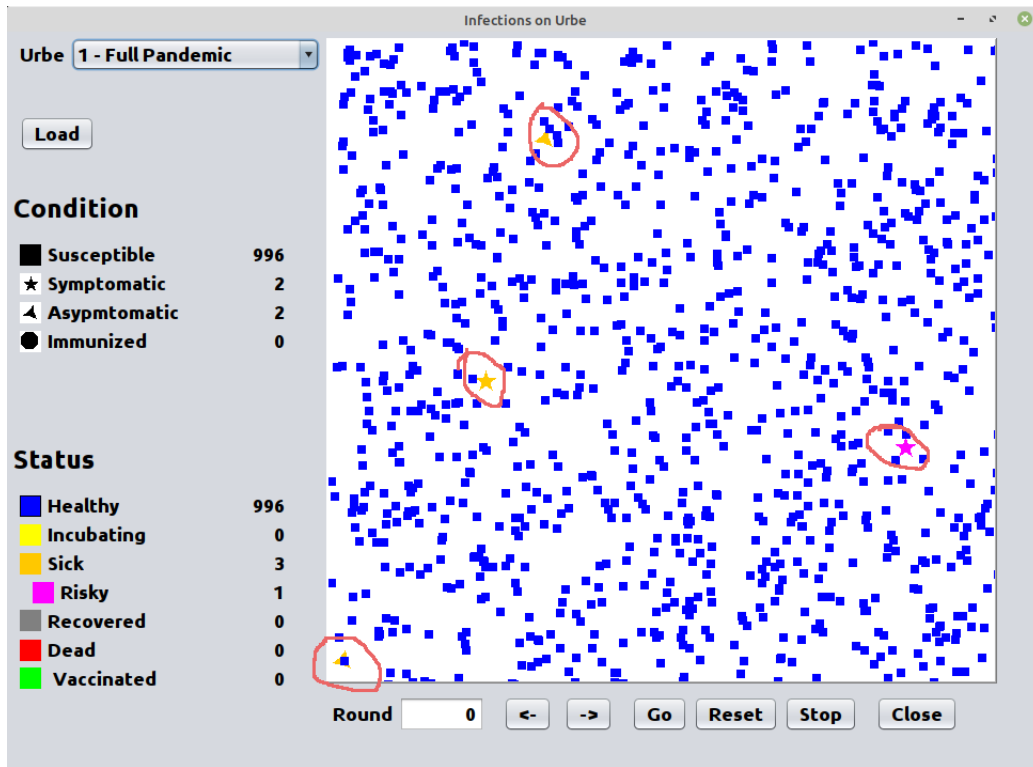
Figure 9: Window of Epidemic2D.

```
and round <=infecround+incubation+illnssprd
and idurbe=1
```

However, round 0 would be missing. Round 0 has a particular condition because it gets infected beings coming from outside, which means they had been infected by other means than the normal process in the environment. To get the infected beings in round 0, we should query differently:

```
select round,count(*) from individual
where infecround=0
and round<=illnssprd
and idurbe=1
group by idurbe,round
```

A 'UNION' operation on these two last queries would give the complete set of infections over the simulation. Other results would be obtained in the same way. Recurrence would be obtained with more complex command, since its value comes from the table 'urbe'. As consequence, a join should be done:

```
select i.round,count(*) from individual i
inner join urbe u on (u.ROWID=i.idurbe)
where i.recurrence and i.infecround>-1
and i.round>i.infecround+i.incubation+i.illnssprd+u.recurrencedelay
and i.idurbe=2 group by round
```

Epidemic2D-app: Seeing what is going on

Complementing Epidemia, I've released a way of 'visualize' what happened in the previous simulations. The app 'Epidemic2D' does it. The Figure 9 presents a screenshot of it. Although the picture is (almost) self-explained, let me add some comments. This simulation, whose name, Full
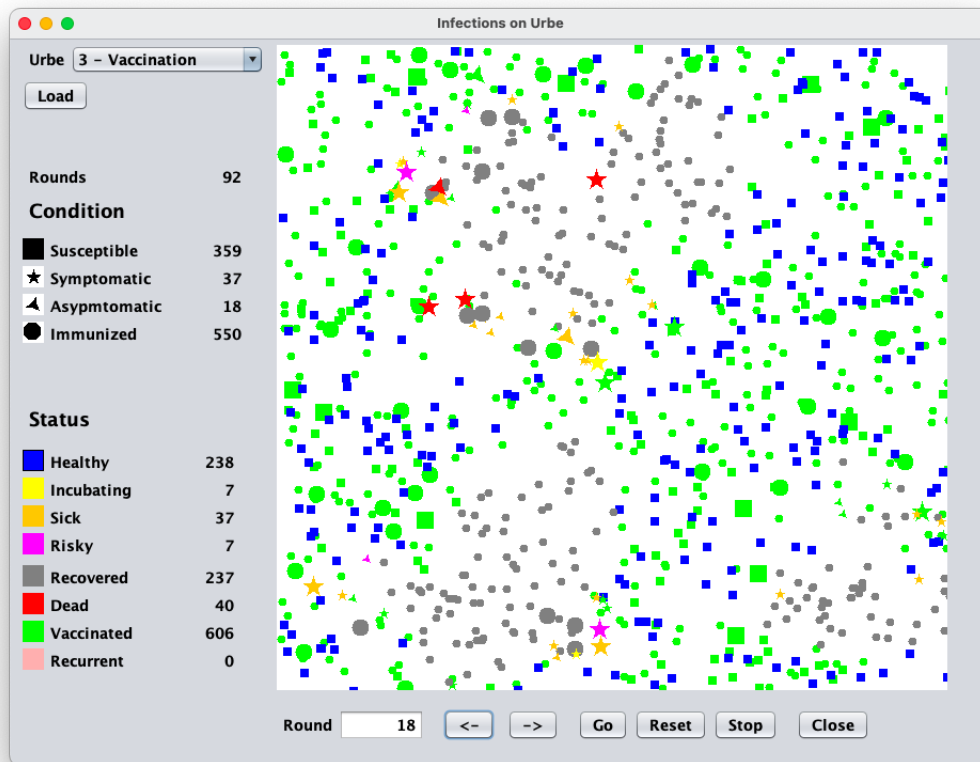
Figure 10: Step 18 of simulation "3 - Vaccination".

Epidemic, seen at the upper-left side of the panel, in the combo list, was done to help us retrieve what we have done, begins with four infected individuals. They are tagged with handmade circles (red). Because they are sick, they are painted orange (not risky) and magenta (risky to die). Two of them are asymptomatic, and this is why they are marked with a 3-vertex-star. The other two are symptomatic, and then they are 5-vertex-star. So, we have two sets of two asymptomatic and two symptomatic individuals, one of each set being at risk of dying. When the counts go on, we may see all kinds of possibilities happening and "sparkling" on the screen, like these four initial infected beings. Sparkling means get a bigger size. All the other beings are healthy (blue) and susceptible (squared).

Clicking on the button "Go," you may watch the evolution of the epidemic whose simulation you generated on the Epidemia panel and gave a nickname. In the mid-time, you see a myriad of multiple geometric forms and colors just indicating the state of each individual at that step. Click on the button "Stop" and spend some time closely observing what is happening. It may be very instructive. Figure 10 shows the simulation "Vaccination" at step 13.

It was a plenty agitated round. You find people getting infected, incubating, ill, risky ones dying, and many of them being vaccinated. Some are getting recovered. There are asymptomatic and not all over the terrain. All were happening simultaneously. At this point, the number of 'removed' beings from the set of infected (recovered or dead) is getting close to the new infected ones. Partly because of the social distancing measures, but mainly because the susceptible ones lack terrain.

The Costa's Semaphore

Portugal's Prime Minister, Mr. Antônio Costa, presented in a Press Interview a diagram for supporting decisions of the policy to face the Covid-19 epidemic disease in Portugal. The chart is now known by the designation Costa's Semaphore, regarding the three traffic's light colors, signing to the Public Health Officers the measures to be applied according to the "semaphore" colors, like in the
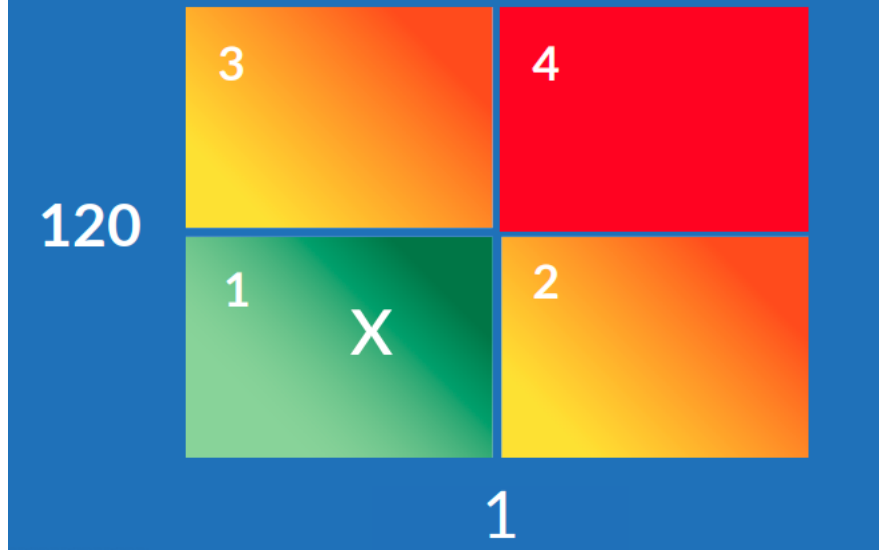
Figure 11: The Semaphore of Mr. Antônio Costa.

traffic conditions. The "Semaphore" is shown in Figure 11.

In the X-axis, we put the R-Number. As we have seen, the R-number is the number of sick people just notified (incidence) divided by the number of sick people in the round. The Y-axis goes the incidence: the number of new sick ones by a hundred thousand people in the population. So, there are four zones in the chart, called 'quadrants'. The first one, in the lower-left part, is the 'green' quadrant. As long as the conditions stay there, officials may relax the social restrictions. However, if the R-number or the incidence exceeds the first quadrant limits, R-number greater than one, or incidence greater than 120/100000 hab, we reach the yellow-to-orange zone, and the situation demands more tight measures. But, if the R-number and incidence both exceed their threshold levels, reaching the red zone, the social measures should be extreme.

To fit this chart into the Epidaemia scheme, we should first understand the meaning of the quantity in the ordinate axis. The adopted one by Mr. Costa must have a relationship with the sanitary and social conditions in Portugal. Consequently, it means that it could take other values for other places or countries. So, we should seek an "objective" parameter, not depending on the place's conditions, but universal quantities, as we see on the abscissa axis, which refers to the R-number that directly affects the dynamics of the disease evolution. Thus, I found the parameter

$$H = \frac{\# \text{ of hospitalized}}{\text{Absolute Health Care capacity}},$$

as adopted when discussing Social Distancing (page 9). Costa's Semaphore points to a 2-D phase space criterium to control the variable Social Distancing and Mobility in our Epidaemia scheme. Still, it would be too much work for too little benefit. I say it because this 2-D phase space is excessive. I mean, there would be no advantage in traversing a path in an arc of a circle centered at the origin of the coordinates in the panel. Only the distance to the center matters.

On the other hand, we may adapt the variable social distancing and mobility scheme to the one introduced by Portugal's Prime Minister, Mr. Costa. Now, instead of simply the variable H, we take:

$$\rho = \sqrt{R^2 + H^2},$$

to drive the Logistic function regulating the social distancing and mobility variation. To align this procedure to the Costa's, I have to add new parameters. I called it Emergency Factor $E_f = (R_0, H_0)$:

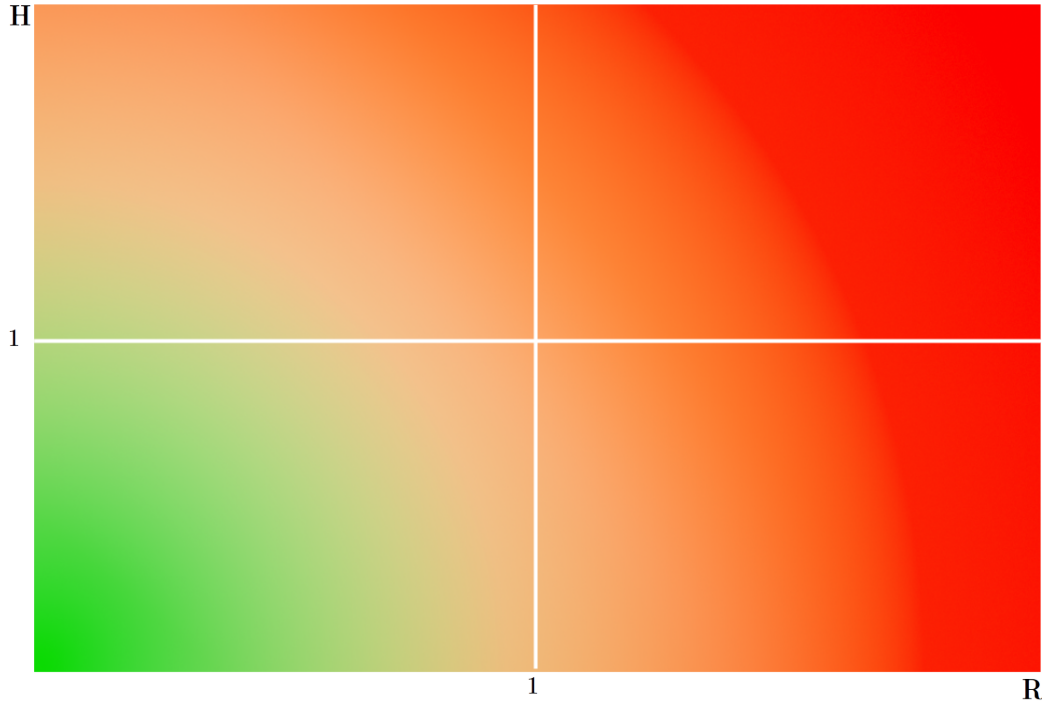$$\rho = \sqrt{\left(\frac{R}{R_0}\right)^2 + \left(\frac{H}{H_0}\right)^2}, \tag{8}$$

25

Figure 12: Costa's Semphore adapted to Epidaemia simulations.

to regulate the position of the Logistic function, so that we can rewrite it from the Eq (2):

$$C(\rho) = \Lambda(\rho) = \frac{(\Lambda_0 - \Gamma_0)}{1 + \exp\left(-a(\rho - 1)\right)} + \Gamma_0, \tag{9}$$

where 'C' stands for "Costa's Number". The role of the emergency factors (in both directions) is to displace the center of the logistic function. Adding values to one of the parameters (R or H) will strengthen the Costa Semaphore's calculation into tighter social measures. If we want to give incidence (H) greater sensibility to the threshold for a lockdown, we decrease the value of H_0. The same for the R-number.

The resulting panel, inspired in the Costa's Semaphore is shown in the Figure (12).

I adopted the following values for testing Costa's Semaphore strategy:

```
SocialMobility: 3.0
VarySocialMobility: TRUE
Vaccination: FALSE
VirusStrength: 0.4
SocialDistance: 2.0
VarySocialDist: TRUE
CivicFactor: 5.0
CivicDisplacement: 1.0
EmergencyFactor:[  1.0 0.5  ]
```

The other parameters remain with their default values. We may observe, changing the seed for the random numbers series, after many experiments, that if we maintain the threshold for H near the unit, the strategy will not work. The parameter H tends to increase more times than the parameter R. The idea is to make H more sensitive. If we put the additive parameter H_0 to near the unit (positive), Costa's strategy works much better. The Figure 13 shows that making $H_0 = 0.5$ flatten the infection curve with a significant dropping in the number of deads, giving the government time for a mass vaccination policy when possible. Otherwise saying, when the authorities realize that 50% of the hospital beds are occupied (assuming $R < 1$), they should turn on the yellow light.
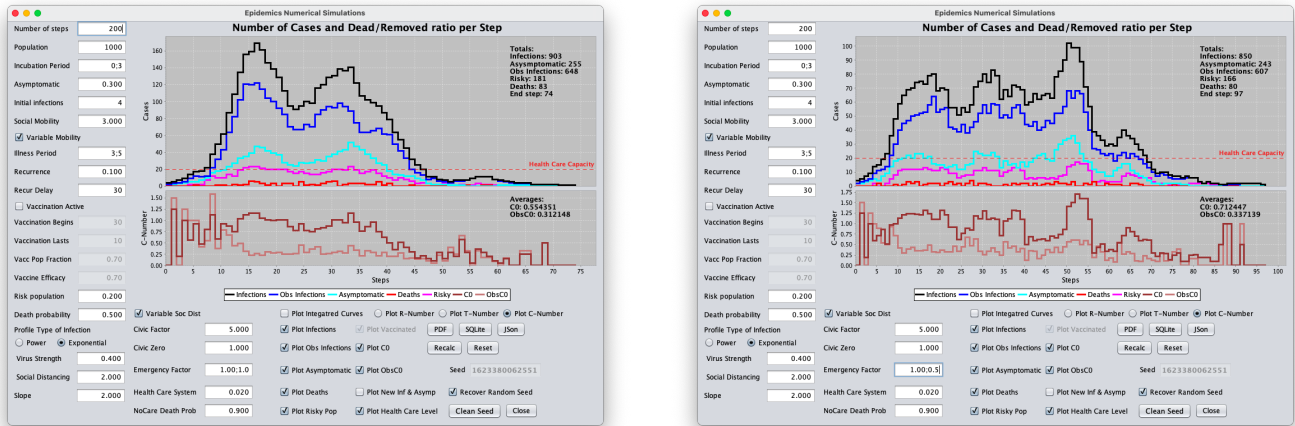
Figure 13: Simulation of Costa's Semaphore policy. Notice that the parameter "Emergency Factor" was included in the panel, below "Civic Dist" in form of a 2D-array: the first element refers to R, the second to H. We see that adopting $H_0 = 1$ will not suffice to prevent collapses in the health system (figure on the left). We should impose, at least, $H_0 = 0.5$ (50% of beds occupied, figure on the right). Also, notice that the third option for plotting the lower panel changed to 'C-number', that is to say, the Costa Semaphore threshold (parameter ρ described in the text).

Querying Epidemia.sqlite for Costa's Semaphore output

I have put the file Queries.sql available in the GitHub Epidemia system of files for those who want to perform queries into the database Epidemia.SQLite. It contains a set of creating views instructions to make the database easier to use. I put the views in the most general form I could, so after loading it into the app DB Browser for SQLite, we can get views for, namely, generate data for Costa's Semaphore Panel.

Below we see a sample from the file Queries.sql for the creation of the the view sick_in_urbe:

```
-- Sick individuals
create view if not EXISTS sick_in_urbe as
  select * from individuals_in_urbe
    where immune=true and infected=true
  and infecround>-1 and round>=infecround+incubation
  and round<=infecround+incubation+illnssprd;
```

It takes the columns from an already created view named `individuals_in_urbe`, which is simply the join of the tables `urbe` with `individual`:

```
-- Individuals in urbe
create VIEW if not EXISTS individuals_in_urbe as
  select i.*,u.* from individual i
  inner join urbe u on (u.ROWID=i.idurbe);
```

The other views follow the same logic.

Now we are able to combine the different views in order to produce the number we are interest in. Here is an example producing a list of R-numbers:

```
-- Create temporary view named cnt_gettingsick
-- from the already created permanent view gettingsick_in_urbe
with cnt_gettingsick as (
-- Query defining the view
select idurbe,nickname,round,count(*) as gettingsick
```

```
    from gettingsick_in_urbe group by round),
-- Create temporary view named cnt_sick
cnt_sick as (
-- Query defining the view
select idurbe,nickname,round,count(*) as sick
 from sick_in_urbe group by round)
-- Query from the joining on the two temporary views created above
-- to produce the R-number
select g.round,cast (g.gettingsick as float)/(s.sick-g.gettingsick) as R
from cnt_gettingsick g
inner join cnt_sick s
  on (g.idurbe=s.idurbe and g.round=s.round)
-- Define the 'urbe' under interest
where g.nickname='Full Epidemic'
```

Comments:

- In the "tongue" of SQL, comments are signed with "$--$", such that everything after it in the line is ignored;

- Firstly, we use the statement "with [name] as" to create temporary views. They will get joined to produce the final query once we need the columns of the two temporary views just created;

- Names beginning with "`cnt_`" mean that if refers to counting rows;

- Columns lists have to contain `idurbe` or/and `nickname` that belong to the table urbe;

- To obtain R-number we divide the columns `gettingsick` by (`sick-gettingsick`) because this difference is the actual number of sick people (actives) in the previous round;

- The division simply produce only '0' as a result. The numerator should be cast to a float number for SQL to consider a float result.

Similarly, we generate a query to get the number of risky sick people divided by the health care capacity.

Conclusions

I introduce here an application aiming, in elementary conditions, to simulate a epidemic evolving in a social environment determined by three parameters: social mobility, social distancing, and virus strength. Additionally, one can choose the profile of the probability function of the contagious process: an exponential function, being possible to select the exponent (slope) of the power function inside the exponential function. One can quickly go to a Gaussian distribution. The other possibility is the power-law function, such that one can even put fractional exponents. Optionally, one can set on a variable social distancing and social mobility behavior. Canonical features described by the so-called 'logistic function' dictate how these parameters depend on the actual number of infected beings in a step.

The results are compatible with the popular theories on the theme. Series of simulations allow us to achieve certain conclusions about an epidemic. To illustrate the simulations, I released an app to plot step-by-step the evolution of the infections in a two-dimensional graphic panel.