An epidemiological study : simulation and data analysis

Fouquet Damien, Rodriguez Louis, Saue Sigurd

March 2022

Abstract

The evolution of pandemics, namely the COVID-19 pandemic, has raised a number of questions and studies. Many epidemiologists have sought to understand, modelled and predict the pandemic behaviour in order to alert and to protect people. More generally, numerous researchers have tried to collect data about the pandemics. That is why our present study will aim to create a dynamic model to represent, to a certain extent, the evolution of a pandemic ; and pay a particular attention to the population dynamics in terms of births, deaths and mutations of the diseases. Then, we perform a number of simulations using the most relevant set of parameters to describe an hypothetical reality. Our main results show that the pandemic causes severe damages to the population during the first months but then flattens as an immunity is created and cements by the end of the year.

1 Model

In this section, we introduce the main components of our simulation, namely individuals, major functions (reproduction, transmission) and their associated parameters.

1.1 Hosts

We create the crucial element of our model : the hosts, who are set to receive diseases, to spread them, to reproduce, to die and so on. We launch each simulations with 100 hosts, 99 are healthy ones and 1 carries a disease. We characterise them with the following properties :

- A genetic code given by a triplet (α, β, γ) where α, β, γ are randomly chosen in $]0,1[$.

- A position given by a triplet (x, y, a) where $0 < x, y < EnvSize$ ($EnvSize$ being set to 500 for the simulations) and a is the angle.

- A level of immunity given by a random real between 0.5 and 1.5. The higher it is, the less likely an host will catch the disease. Plus, the immunity declines by a small amount at every step and is multiplied when the host has been cured.

- A level of infection status which is bound to be modified at each step of the simulation by diseases that the host may face : it increases when the host is infected and it decreases when the host is on the healing path and when it reaches 0, the host become sane again. It is set to 0 at the beginning except for the single infected person. When it comes to 1, the host dies.

- A level of fecundity or desire to father children *love* given by a random real between 0 and 1.

1.2 Diseases

Another major component of the model is the class disease. We characterise them with the following properties :

- A genetic code given by a triplet (α, β, γ) where α, β, γ are randomly chosen in $]0,1[$.

- A duration given by d where $d \in [\![1; 10]\!]$ and corresponds to the number of steps (ponderated) that the disease will last.

- A virulence ν given by a a random real between $[1, 10]$. It will decrease the level of health of an individual at each step by this amount (divided by a certain parameter so that the individual may live for a few steps even with a really harmful disease).

A disease may also be modified by a **mutation** : at each step, there is a probability of 10 % for each of the five components of the disease to mutate, meaning that their value will change by a small amount.

1.3   Major functions

Transmission function

When hosts are in a certain area (denoted *infradius*) around an infected host and their genetic is "close" enough then they become infected. A close genetic is defined by the parameter *closegenetic* which is a linear combination of the genetic code of an individual. The more immune the host is, the less likely the host will become infected.

Evolution function

During the simulation, hosts and diseases' attributes are going to be modified ; that will explain why hosts die, heal and endure different diseases. The latter is enabled by the evolution functions explained as follows. Disease hits harder infected individuals during a part of the simulation. The level of infection status defined for the hosts is associated with how infected individuals become "sicker" at each step of the simulation meaning that their infection status increases by ν divided by *immunity* (with another factor of ponderation) where ν is the virulence of the encountered disease.

    Hosts can heal as well. If he survives enough days, the level of infection the health status decreases by $Healing \times$ *immunity* (ponderated) $Healing$ is the coefficient associated with how infected individuals heal at each step of the simulation meaning that their health status increases by $infection \times healing$ of the disease.

Reproduction function

The reproduction function allows the population to grow (or grow anew) so that people who have died from their affliction do no lead the whole population to a premature extinction. It is defined using the parameter *love* as follows.

    Considering two individuals, if the sum of their *love* is greater than a certain value using *cs* and *ci* (*cs* or *ci* [Condition for Sane hosts or Condition for Infected hosts], with $cs \leq ci$ since an infected person should have more trouble to reproduce) multiplied by a "genetic" distance, then one individual is created.The new born will have randomly the genetic code of her parents (using *alpha*, *beta* and *gamma*) and her *love* is set to zero. Then their parameter *love* falls down to zero but can be restored by a certain amount at each step. We add the parameter *Maxchild* for the purpose of realism which is the maximum number of children by one individual. Also, there is a threshold of population, and people can not reproduce if the population is equal to this threshold.

2   Results

We aim to understand the evolution of hosts and diseases in this epidemiological model and pay a particular attention to the following indicators : the number of births, of deaths, of infected, sane and cured people and the mutations. Thus, in order to conduct a statistical analysis of the model we first gather the data and then, we decide to establish two different models, each of these aiming to represent a different aspect of reality : "classic" pandemic and mass extinction. For each model we perform ten simulations (which are available in the annex) and select the most relevant or representative graphs . Moreover, for purposes of making real-life adapted simulations, the latter are set to last approximately 3 years (10,000 steps, as 1 day equals to 10 steps) and it usually enables trends to become visible during this period.

2.1 Model 1 - Classic pandemic

We establish this model as our reference. Indeed, we set the parameters handling virulence and duration of the disease at a medium level.
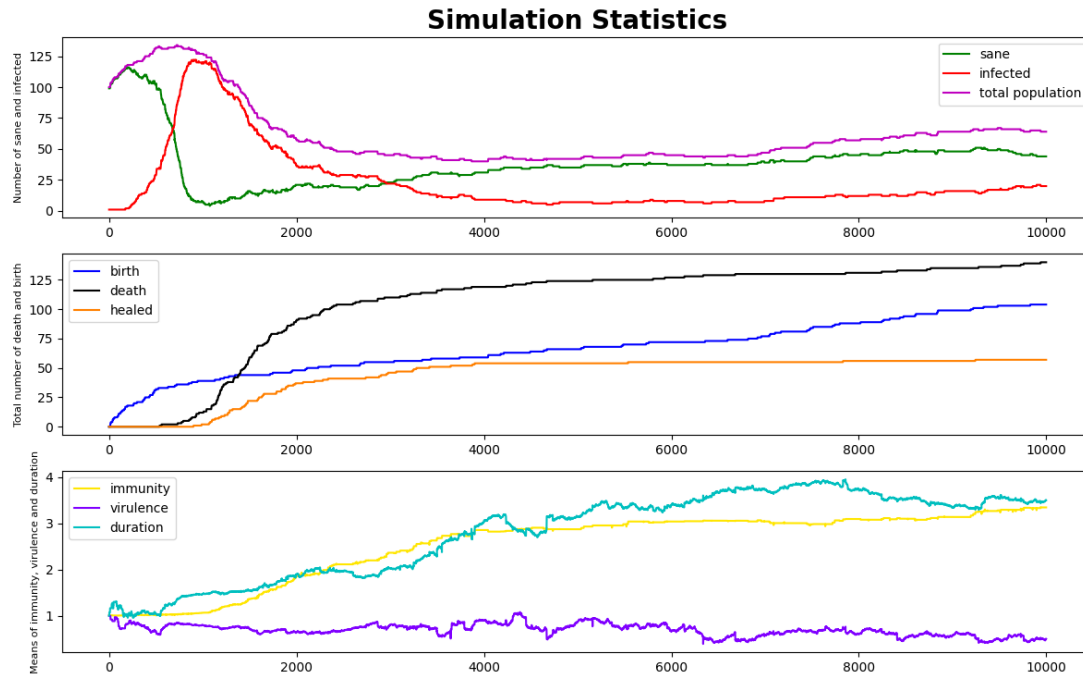

FIGURE 1: Population dynamics in a classic pandemic

In figure 1, we see an exponential increase of infected people but the population manages to grow as new individuals appear and no one dies. However,after 1000 steps infected people begin to succumb and the population declines up to 2000 steps before increasing again after 7000 steps, because there are new births. In the meantime, few people are able to heal from the diseases ; indeed, the number of cured people is half the one of the dead people at 2000 steps (100 against 50). In addition, after several months the number of deaths stabilizes while the number of healed people increases. Indeed, the immunity has a sharp increase at the same time. Furthermore, the births after a sharp increase during the first steps (reaching 50 new individuals), follows a more relaxed regime before increasing again. As the third graph shows, the yellow curve increases all the simulation long thereby we have reached a global immunity in this model and diseases have a much greater duration while the virulence never ceased to decrease (overall).

2.2 Model 2 - Mass extinction

In the previously described model , an overall pattern begins to emerge : an immunity seems to appear and to strengthen as the simulation continues. Thereby, we wonder whether we can create a simulation where diseases are sufficiently harmful and long to beget a mass extinction. Setting a high virulence rate and duration level for the diseases, along an average fertility, we obtain the following graph.
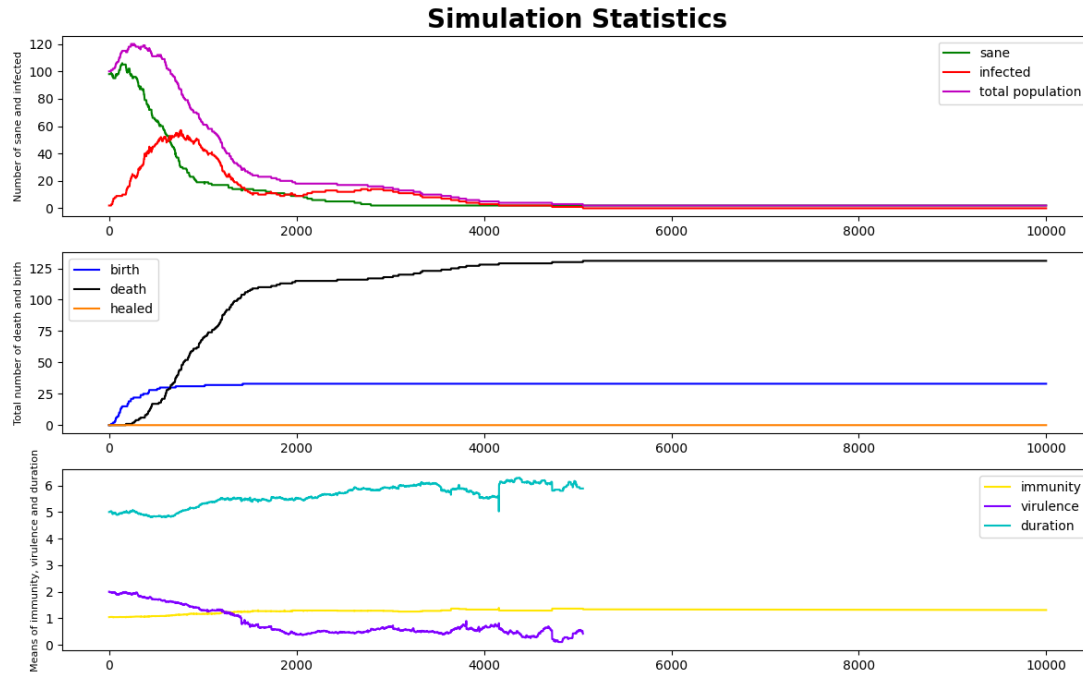
FIGURE 2: Population dynamics with a high duration and a lethal illness

The key element of this model is the fact that the duration and virulence graphs stop abruptly at around 6000 steps because there are no more infected and no more disease and therefore no more mutations. The disease and mutations are too strong for the hosts to develop immunity, so the immunity graph is flat.

In this case, the disease is extremely violent and every infected individual eventually dies. Moreover, only 1 to 5 % of the individuals survive after almost a year and even if we had kept the simulation going, they would have died of loneliness, without reproducing. However, there are other cases where enough people survive (above 5%) and it allows the population to grow anew. This singular case can be explained by the fact that setting a too high virulence causes the infected hosts to die before everybody become infected and then the disease disappears and the sane ones are able to survive and reproduce. This divergence can be seen on the annex.

Finally, although the virulence declines, diseases continue to affect people and that explains why individuals continue to die. Besides, at the end the number of deaths converges to almost 125 (same amount than in the previous model) but here, individuals die much faster. Thereby, this model predicts the decimation of the population.

3   Conclusion

To conclude, we notice that modelled pandemics are devastating during the first steps of the simulations but once a global immunity is reached, diseases become more benign and no longer kill individuals. In addition, we observe in the second model results differ a lot from the classic case by simply modifying few parameters. That being said, we must bear in mind that the simulation has its limits since only one hundred individuals are handled by the model and the chosen parameters may not be relevant in certain economies. Furthermore, it is certain that a more scientific study and conducted by epidemiologists or biologists notably, would convey more credible and robust results.

Annex

Here we put at your disposal a number of other simulations of all the models already commented and two other models (a long-term pandemic and a cyclic model) with a detail of all the global parameters used. `https://drive.google.com/drive/folders/11cu0lcHH2aniy1JN1xQgN8Bzpl_PaDLq?usp=sharing`