IWOMB 2021 Workshop 1: Team Gillespie

Alastair Jamieson-Lane and Alejandra D. Herrera Reyes

June 2021

Introduction

Both Stochastic Differential Equations (SDEs) and Gillespies algorithm can be implemented using any scripting language that takes for fancy (Julia, Python, Matlab, R). For concreteness, we will be using the open source language Python, and encourage you to do the same. With that said, the goal of the workshop is to learn the numerical methods for stochastic models not to learn how to code with Python. If you and your team decide to work with another software that is completely fine with us. Perhaps you and your teammates will even decided to use multiple different programming languages, and will compare the results.

In part 2 of this worksheet, we describe the fictional disease *whalepox* which we made up for academic purposes.

PART 1: Simulating a birth-death process

This part will be attempted during our first meeting in the workshop. The instructor will be present and you'll work in teams to solve it.

One of the most studied stochastic process is a basic population model known as the birth-death process. For this model, we assume an individual dies at rate δ and reproduces at rate μ. Hence, if we have 100 individuals in our population the *total* birth rate is 100μ and the total death rate is 100δ. The birth-death process can be used to model classical animal populations (such as rabbits, trees or humans), but can also be used to model the early stages of an epidemic, with 'birth' and 'death' replaced by 'infection' and 'recovery'. Mathematically, we still have a rate of increase, and a rate of decrease, both proportional to the current population.

To get started on your simulations, we want you to simulate the birth-death process using the method we just taught you. For your first simulation, assume $I_0 = 1$. $\delta^{-1} = 1$ day, $\mu^{-1} = 0.8$ days. Discuss what you observe.

Extra questions:

1. What is the probability that the epidemic reaches $I(t) = 100$ before going extinct?
2. Explore different parameter values. For what parameter values does the epidemic grow? When does it shrink?
3. Talk to your teammates and see how they have implemented the simulation - have they done things exactly the same as you? Help out anyone who is stuck. Explaining things to others will help deepen both your understanding and theirs.
4. From a biology/evolution perspective, what is better for the infection: halving δ or doubling μ, or is it the same?
5. We told you in the introduction that the birth-death process is only an approximation for the early stage of the epidemic. When will this approximation break? why?

PART 2: Whalepox

Whalepox originated in Antarctica in 2026, and quickly devastated the Antarctic community. While previously assumed to be contained, the infection has been observed in Mexico, New Zealand, and on the campus of UP Mindanao in the early months of 2027. **You** *have been asked to model this infection, and provide uncertainty quantification.*

A few things you know in order to help get you started:

- *Whalepox is deadly in 8.6% of all cases cases. Those who recover are immune for life.*
- *For those infected, the average time to either death or recovery is twelve days ($\delta^{-1} = 12$ days).*
- *The total infection rate is proportional to an infection constant β, the number of susceptible individuals, S, and the* **fraction** *of the population which is still infectious I/N. Hence the total infection rate is $\beta S \frac{I}{N}$.*
- *During the early stages of the epidemic in Antarctic (when $S \approx N$), the daily number of whalepox cases was observed to double every nine days.*
- *The population of UP Mindanao in 2027 is 2300 students. The total population of the Philippines in 2027 is 118 million people.*

The following questions relate to modelling construction, and to policy decisions that might need to be made during an epidemic. Knowing about these questions, and the associated techniques in advance is useful, as often during epidemics, decisions need to be made quickly. Your task is to construct a basic SIR model using the information given, and then use this model to explore your simulations. The final goal should be that you and your team mates understand how to simulate, interpret and expand each part of the model.

Don't feel like you need to rush through all the questions; you are free to divide questions between your group members, ignore questions that don't seem interesting, or create questions of your own. If you want to dig into detail understanding one or another question this is fine. If you do divide questions between your group members, remember to come together at some point and discuss what you have observed.

(a) Given the initial doubling time observed in Antarctica, what is a realistic estimate of β? For this question, use the differential equation $\dot{I} \approx \beta I \frac{S}{N} - \delta I$ along with $I(9) = 2I(0)$.

(b) Construct a SIR type epidemic model using the epidemic parameters you have calculated. Use Gillespie's algorithm to model an epidemic on the UP Mindanao campus. Assume the initial number of infections is 3.

(c) From your simulation, how many people get infected? How long is the epidemic wave if you don't modify the parameters? Maybe, calculate the average trend with M=10 simulations, M=100 simulations, M=1000 simulations. Give an estimate of the mean length of the epidemic wave. How often does the epidemic go extinct before reaching 100 people?

(d) Which parameter in your model would you change in order to represent Lockdown/public distancing measures? What sort of information would you need in order to determine how much to change this parameter by?

(e) What would you add to your model in order to represent vaccination efforts? You will need to make modelling decisions in order to decide what the *rate* will be for vaccination, and how it will move people from one compartment to another in your model.

(f) Imagine that all infected people are placed into quarantine (or told to quarantine at home) when their infection is discovered. If it takes (on average) two days for the infection to be discovered, how would you model this change? HINT: One way includes more compartments, another way involves modifying your original parameters.

(g) Adapt your Gillespie simulation to modelling an epidemic across the entire country. Assume your initial number of infections is 60. What happens? What do you notice?

(h) How would you go about modelling an epidemic that was spreading between four different university campuses? What data would you want to access to build your model? How would splitting your system into multiple 'neighbourhoods' change your results?

(i) If one campus decides to have lockdowns, but the others don't, what does this do to your model? What happens to your total epidemic size if all campuses cut contact with one another, but internal contacts remain the same?

(j) The University can implement strict lockdowns, reducing contact by 60%, and costing two million dollars per day, or they can implement moderate lockdowns, reducing contact by 40%, but only costing one million dollars per day. Which is the better policy? How do you decide?

(k) What other epidemic response policies can you imagine? How would you "mathematically" implement them?

(l) In dealing with the fictitious whalepox, we have given you all the data that you need. In the real world, such data is **not** provided, and you must search for it yourself. Imagine you were modeling SARS, H1N1 or COVID, and only had access to data and research that was available in the early stages of the epidemic. What data would you use to calculate your infection, recovery and death rates? What information would you need to look up in order to determine the effect of closing schools or universities? How might this differ from country to country?

(m) Suppose the total infection rate changes slowly over time (for example, between summer and winter). How would you represent this in your simulation? What problems might arise?

(n) Different sorts of models are good for different problems. Talk to the other team. When is Gillespie the best algorithm to use? When are SDEs more useful? Is there a way to get the benefits of both approaches?