

Bachelor Dissertation

Kasper Engelen, Jonathan Meyer, Dawid Miroyan, and Igor Schittekat

University of Antwerp

Abstract. This document reports our findings regarding the final dissertation. The sections and subsections correspond to the assignments given to us. In this project we worked with a simulator called Stride, developed at the University of Antwerp. We explore various concepts within computational epidemiology through the use of this program.

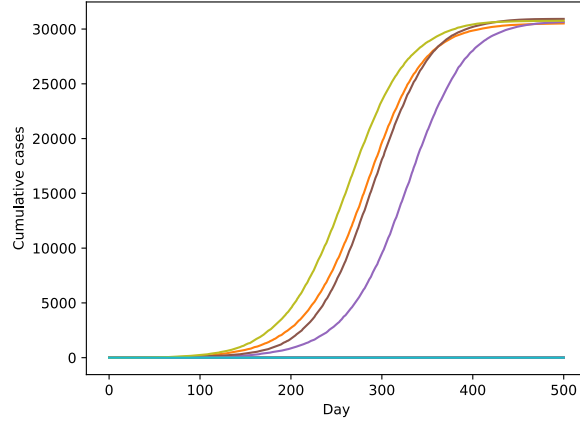
Keywords: Computational Epidemiology · Dissertation

1 Simulation

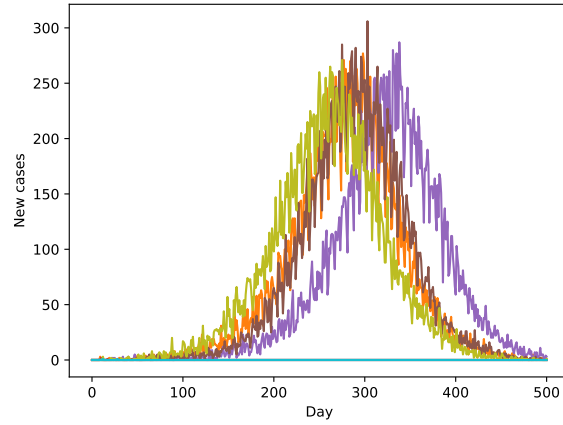
1.1 Stochastic Variation

The first topic we consider is stochastic variation. Since the simulation uses a pseudo-random number generator, it's useful to inspect the influence of this stochasticity on the results of the simulation. Using the Stan tool which is provided with Stride, we collected data on 100 simulations with an identical configuration file. The only difference between executions is the RNG seed. When plotting the results, it becomes apparent that the amount of new cases per day follows a normal distribution. For the data we collected, we calculated a mean of 33.33 new cases per day, and a variance of 1196.99. Looking at the plot for the cumulative cases per day, two 'categories' can be distinguished: 'outbreak' scenarios and 'extinction' scenarios. These cases are quite evenly spread, which explains the high variance for new cases per day. In the former, the curve has a sigmoid shape, indicating that the disease successfully spread among the population. The latter scenario corresponds to the curves that are almost constant and are bounded by a value well under the population size. In these cases, the disease did not manage to spread, resulting in only a few infected people at the end of the simulation. This is likely explained by the initial infection: if the first person to be infected is sufficiently isolated (either socially or by being surrounded by people who are immune), the disease doesn't have a chance to spread.

Figure ?? shows the different scenarios. In the cumulative cases plot the 'extinction' scenario is clearly visible at the bottom. The plot for new cases per day is very jagged, which can also be explained by the stochastic nature of the simulation.



(a) Cumulative cases per day.



(b) New cases per day.

Fig. 1: Plots using different seeds

1.2 Extinction Threshold

As discussed previously, it might be the case that only very few people become infected over the course of a simulation. This is referred to as extinction. There is a clear distinction between outbreaks and extinctions, so in this subsection we attempt to find an *extinction threshold*.

Figure 2 gives the frequencies of the amount of infected people at the end of the simulations. The distinction between outbreaks and extinctions is quite clear

from this histogram. There is one peak on the lower end of the x-axis, and there is a cluster on the higher-end.

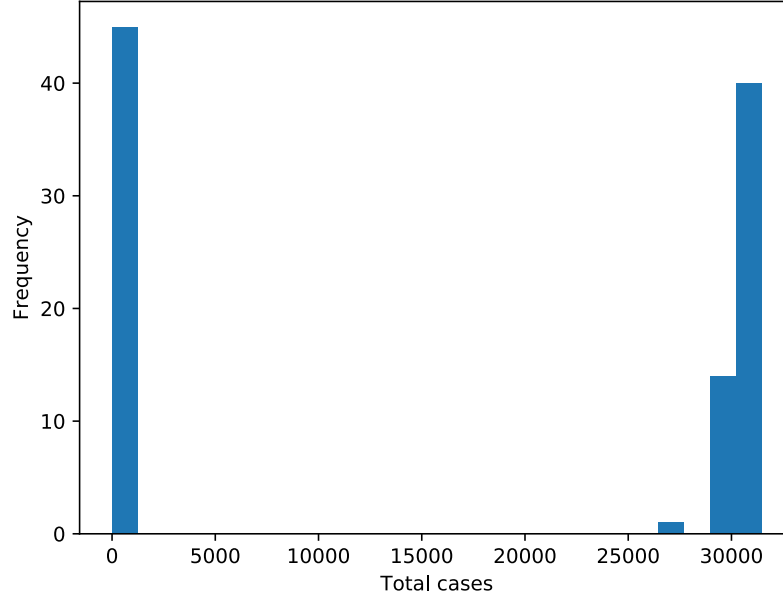


Fig. 2: Histogram showing the total amount of cases.

1.3 Immunity Level

In order to make assumptions about the population's immunity level, we look at the simulated results for different values. Upon experimentation, it becomes apparent the immunity level is approximately 70% of the population. Figure 3 shows the average new cases per day for different immunity levels. The reference curve is also included. This plot was generated by taking the average of new cases per day over 20 simulations per immunity level. Using PyStride to simulate outbreaks, we narrowed down the immunity I level to $70.5 \leq I \leq 71.75$.

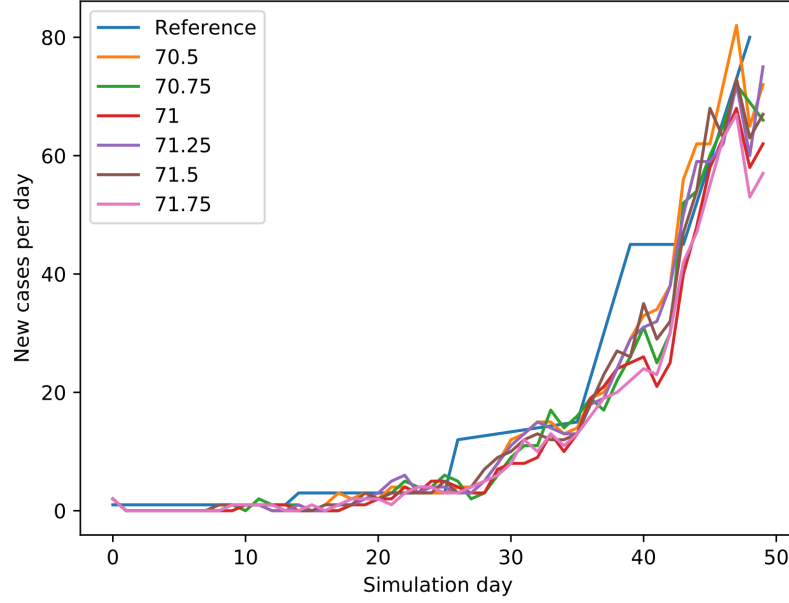


Fig. 3: New cases per day for different immunity levels.

1.4 Estimating R_0

Placeholder

Table 1: Table captions should be placed above the tables.

| Heading level | Example | Font size and style |
|-------------------|---|---------------------|
| Title (centered) | Lecture Notes | 14 point, bold |
| 1st-level heading | 1 Introduction | 12 point, bold |
| 2nd-level heading | 2.1 Printing Area | 10 point, bold |
| 3rd-level heading | Run-in Heading in Bold. Text follows | 10 point, bold |
| 4th-level heading | <i>Lowest Level Heading.</i> Text follows | 10 point, italic |

Displayed equations are centered and set on a separate line.

$$x + y = z \tag{1}$$

Theorem 1. *This is a sample theorem. The run-in heading is set in bold, while the following text appears in italics. Definitions, lemmas, propositions, and corollaries are styled the same way.*

Proof. Proofs, examples, and remarks have the initial word in italics, while the following text appears in normal font.

For citations of references, we prefer the use of square brackets and consecutive numbers. Citations using labels or the author/year convention are also acceptable. The following bibliography provides a sample reference list with entries for journal articles [?], an LNCS chapter [?], a book [?], proceedings without editors [?], and a homepage [?]. Multiple citations are grouped [?,?,?], [?,?,?,?].

References