Exploring COVID-19 Cases

Kevin Dang

April 15, 2020

Introduction

COVID-19, also known as the coronavirus was declared as a global pandemic by the World Health Organization on March 11, 2020. The first known case occurred in mid-November and the first confirmed death occured mid-January. The covid-19 dataset is likely to be inaccurate and incomplete due to numerous cases being labelled as influenza or pneumonia as a direct result of lack of testing and censorship by the Government of China. Unfortunately, COVID-19 has spread throughout the world with 1.7 million confirmed cases and 100,000 confirmed deaths at the time of writing. While there is little we can do about the past, we can look towards the future. Our analysis will aim to determine when the rate of increase in COVID-19 cases in Canada will start its decline.

Despite the fact that the coronvirus outbreak originated in Wuhan, China, the increase in the number of cases have declined for the country as a whole. There is a lot of uncertainty about how vastly underreported their numbers are, especially as some experts estimate there could be 40 times their current reported cases. However, this is the data we have and we can use China's cases to try and estimate when the increase in the number of cases in Canada will begin to slow if we follow proper social distancing and self-isolation protocols.

The methods contain an overview of the two models used for the analysis: logistic and exponential. In the results section, the models are plotted and the coefficients are determined. Various calculations are made regarding the maximum growth rate and upper limit of COVID-19 cases in Canada. The discussion section ties everything together in a concise summary.

Methods

Historically, epidemics and pandemics tend to follow a logistic growth pattern. The number of cases starts off small, then we see an exponential increase which eventually starts to slow down and reach a plateau. This is the pattern of a sigmoid. The formula for logistic growth is:

$$y(t) = \frac{c}{1 + a \cdot e^{-bt}} \tag{1}$$

Where y(t) is the number of cases at a time t, c is the maximum number of cases , a and b are constants that can be modified (b > 0). Also note that the maximum growth rate occurs at t = ln(a)/b, y(t) = c/2.

With this information we can figure out the number of days it took China to reach the maximum growth rate inflection point. We also know that this point represents half of the maximum number of cases. We can then compare this logistic model to Canada's cases, and estimate the time lag between Canada and China. The next step is to use an exponential model to extrapolate the data in Canada until we reach the maximum growth rate. Then we can estimate the maximum number of cases, or at least have a number where we think we will reach a plateau.
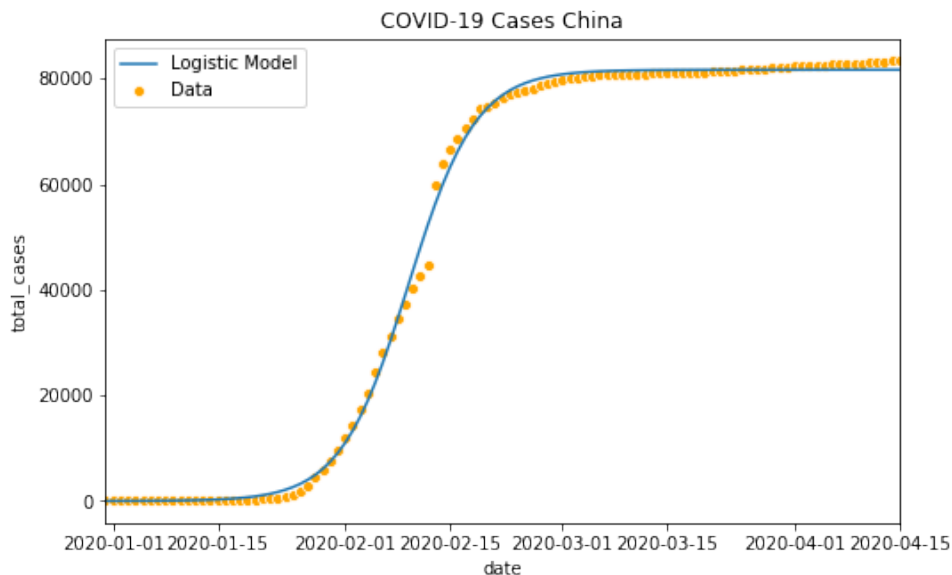
Exponential growth takes the form of:

$$y(t) = y_0 \cdot b^t \tag{2}$$

Where y(t) is the number of cases at a time t, y_0 is the initial value, b is the growth factor (number of people infected per infected individual).

Results

Figure 1: COVID-19 Cases in China, with a logistic model

The logistic model in Figure 1 has the following coefficients from equation 1: $a = 7467, b = 0.2207, c = 81684$. Plugging them into the equation, we get:

$$y(t) = \frac{81684}{1 + 7467 \cdot e^{-0.2207t}} \tag{3}$$

From equation 3, we know that the max growth rate of cases in China occurs at $t = \ln(a)/b = \ln(7467)/0.2207 \approx 40$ days, or on February 9, 2020. On this date, the model interpolates the number of cases as 40,842.

Figure 2: COVID-19 Cases in Canada, compared with China 63 days earlier
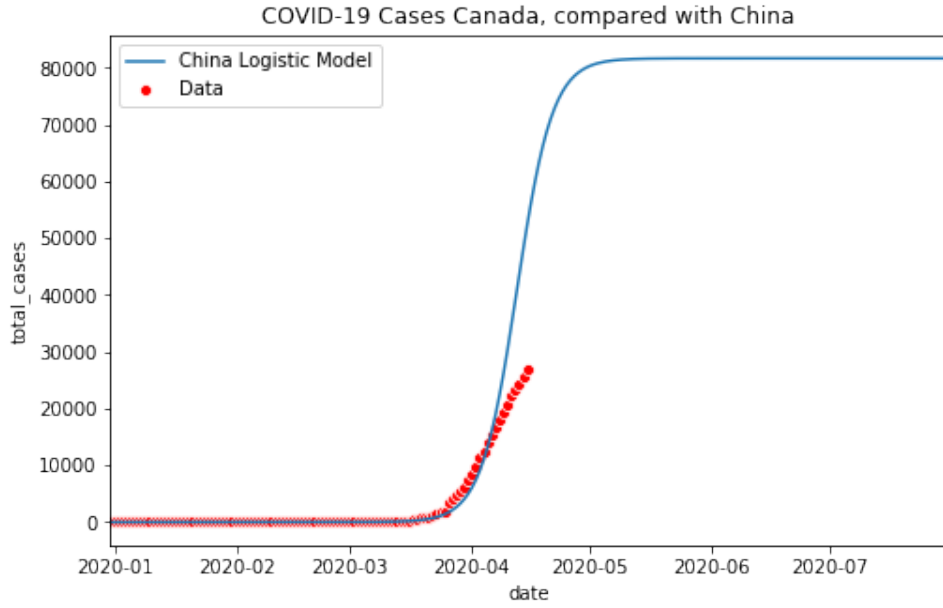
We note from Figure 2 that there is approximately a 63 day delay - Canada did not report their first case until 26 days after data was collected, and the acceleration began much later (about 37 days to yield a 63 day shift). This 63 day shift does not fit the Canada data points well, but it is a helpful guide to estimate the length of the delay.

Figure 3: COVID-19 Cases in Canada, with an exponential model
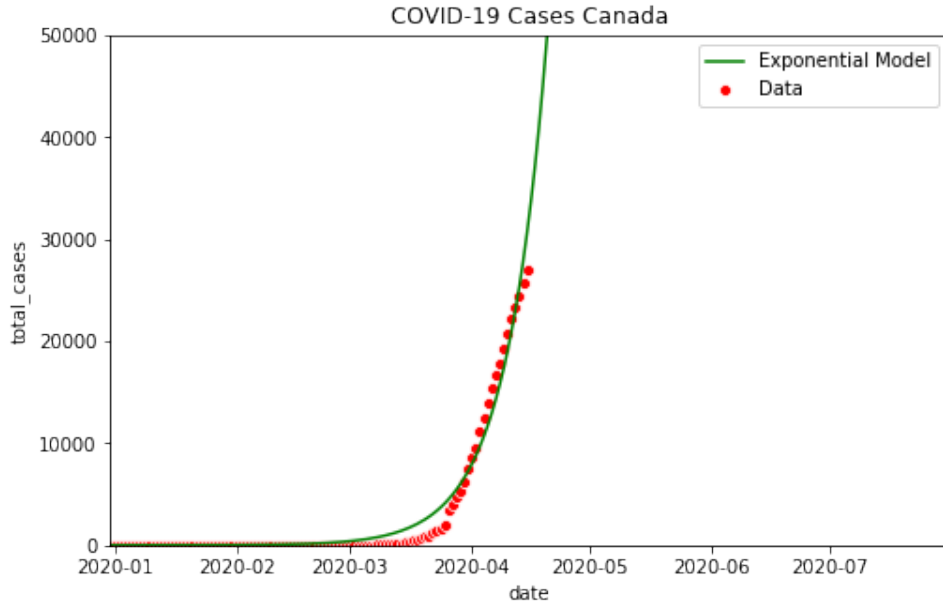
The exponential model in Figure 3 has the coefficients $y_0 = 1.177, b = 1.101$. Equation 2 can be filled in as:

$$y(t) = 1.177 \cdot 1.101^t \tag{4}$$

If Canada's cases were to reach the peak around the same time as China but with the 63 day lag, that means the maximum growth rate would occur at $t = 40 + 63 = 103$ days

3

(April 12). The exponential model gives the number of cases y(t) as 22,984. Now if we were to extrapolate the data points, using a logistic model then we expect the value of c from equation 1 to be twice that of 22,984, which is 45,968. This would represent the upper bound of cases in Canada given by the model. This is not a hard limit however, as we note from China's graph that some data points did go over, but it is a good guide for estimating when we will see a plateau.

Cases in China surpassed the theoretical upper bound of $c = 81684$ on March 26, which is 46 days after their maximum growth rate date of February 9. If Canada follows the same social distancing and self isolation protocols as China, then the theoretical upper limit of cases in Canada may be reached on May 28.

Discussion

Using a logistic model for China we determined the date on which the maximum growth rate in COVID-19 cases occurred. The total cases recorded on February 9 is about half of their current reported cases. Applying this principle to Canada, we estimated that the peak growth rate occured on April 12 (roughly 23,000 cases), so we expect that the upper bound of cases will be about 46,000 on May 28.

The main limitations of this analysis is that we cannot control for various scenarios. We do not know when the government will decide to reopen schools and businesses, or whether there are enough tests being administered. In addition, the Canadian government is not nearly as strict in enforcing social distancing rules as China. There is expected to be a large margin of error, but perhaps we can get a good look into the future about 6 weeks from the time of writing.

References

"Coronavirus Source Data", Hannah Ritchie, https://ourworldindata.org/coronavirus-source-data

"Who Declares Covid-19 a Pandemic." Domenico Cucinotta, Maurizio Vanelli, 19 March 2020, https://www.ncbi.nlm.nih.gov/pubmed/32191675

"1st Known Case Of Coronavirus Traced Back To November in China", Jeanna Bryner, 14 March 2020, https://www.livescience.com/first-case-coronavirus-found.html