Introduction

P and N functions. The cumulative (or total) number of Covid-19 confirmed cases before day t is denoted by $P(t)$. We allow $t = 0$ to correspond to any given date. Think of P as "population" to remember the meaning of P.

The daily new cases N on day t is given by

$$N(t) = P(t+1) - P(t).$$

We sometimes also call $N(t)$ the...

- daily increase in cases on day t.
- daily cases on day t.
- new cases.

Note that with these definitions

$$\begin{aligned} P(t+1) &= P(t) + N(t) \\ &= (\text{Total confirmed cases before day } t) + (\text{New cases reported on day } t) \end{aligned}$$

Data sources. All data is sourced from the *COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University* accessible at the following link

- https://github.com/CSSEGISandData/COVID-19.

We reference this source as JHU in short.

Getting the data for your own explorations, or for students to explore. Raw data for countries and US states, together with their 7-day smoothing (see below) can be downloaded using a couple of websites that Isaac Flath built for the Consortium.

- https://isaac-flath.shinyapps.io/coronavirus2/
- https://isaac-flath.shinyapps.io/Covid-19/

These two sites also generate plots for any given time range for all the quantities we discuss in the problems we have written (P, N, relative change of P, log plots of P, etc).

Data smoothing. Covid-19 data (specially N-data) is very noisy, so in all the problems we have written we use the 7-day moving average of the data which we compute as follows ($\widetilde{P}(t)$ is the raw time series obtained JHU):

$$\begin{aligned} P(t) &= (\widetilde{P}(t) + \widetilde{P}(t-1) + \ldots + \widetilde{P}(t-6))/7 \\ N(t) &= P(t+1) - P(t) \end{aligned}$$

so, the new cases time series $N(t)$ gets smoothed through the smoothing of $P(t)$ and does not get smoothed separately. This makes the relation $N(t) = P(t+1) - P(t)$ continue to hold after the smoothing. We also make sure that we pick up JHU data well before $t = -7$, so the data is smoothed using the above formula even for $t = 0$ using the previous 7 data points in the JHU time series.

Figures 1 and 2 show the effect of this smoothing for the cumulative cases P and new cases N for Romania, with t in days since March 1, 2020 (thick blue is the smoothed data, thin black is the raw data). A side product of this smoothing is that data gets shifted forward (to the left) by 3.5 days.

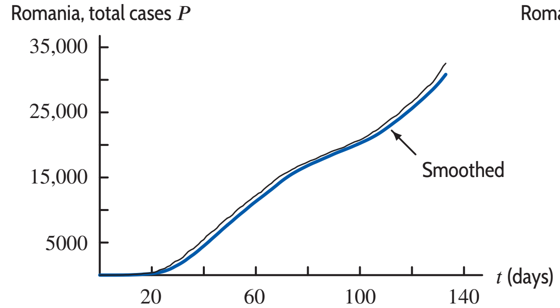

Figure 1

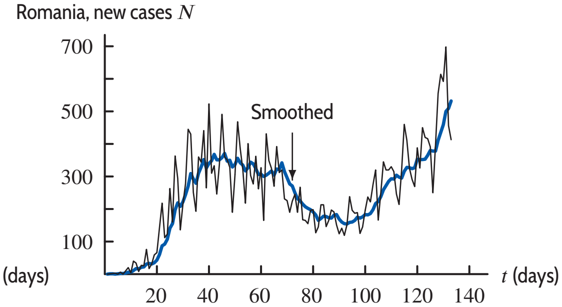

Figure 2

We will not use raw data $\widetilde{P}(t)$ (or $\widetilde{N}(t) = \widetilde{P}(t+1) - \widetilde{P}(t)$) in what follows unless explicitly stated otherwise.

Expected shape of P and N. The Covid-19 outbreaks have occurred in waves in many countries. Each wave has an N-plot which is bell-shaped and that will correspond to a sigmoid shaped P-curve.

Subsequent waves stack on top of one another and result in N-plots that look like a sum of bell-shaped curves (which may or may not drop all the way down to zero between the bells), and this results in a P-plot that will look like a sequence of stacked sigmoid s-curves. Figures 3 and 4 are the P and N plots for Australia between March 1 and August 16, 2020.

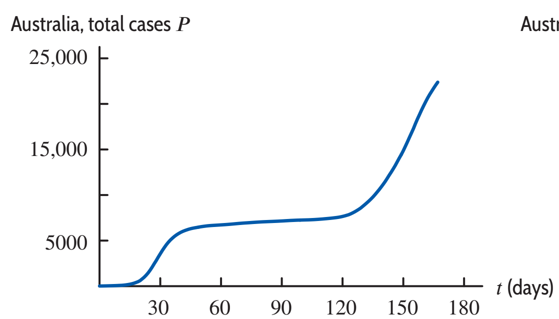

Figure 3

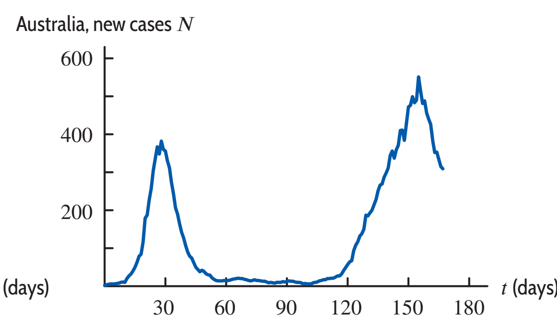

Figure 4

Of course, the outbreaks in many countries do not look so distinctly as a sum of waves. Besides possible data collecting issues or changes in data collecting strategies, drastic measures implemented to control the outbreak have noticeable effects on the way the outbreak developed. Figures 1 and 2 are good examples of expected shapes P and N plots when the waves are not separate.

Many more problems and problem ideas in the Instructor notes package. Visit `https://mcwg.github.io/covid` for many more problems and problem ideas.

Sample problems - Calculus - see the Instructor notes package for more

1. Figure 5 and Figure 6 are made from data[1] corresponding to the Covid-19 outbreak in Switzerland from March 20, 2020 to May 15, 2020. Determine which graph is the graph of the function $P(t)$, the total reported confirmed cases, and which is the graph of the function $N(t)$, the daily increase in confirmed cases.

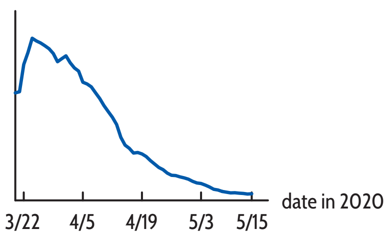

Figure 5

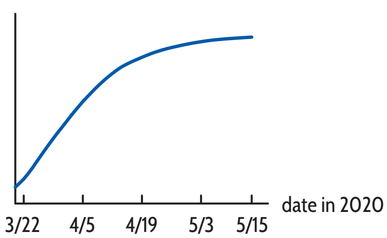

Figure 6

2. Between the dates 4/8/20 and 6/1/20 Brazil quickly became one of the countries with the highest number of confirmed Covid-19 cases.[2] Figure 7 shows graphs for the cumulative confirmed number of cases, P, and its daily relative increase $(1/P)dP/dt$. Note that P is increasing and concave up. Was P growing exponentially between 4/8/20 and 6/1/20? How does the plot of $(1/P)dP/dt$ help you decide?

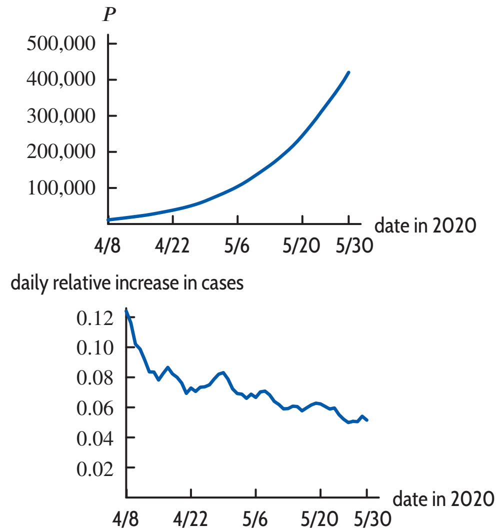

Figure 7

3. The graphs in Figure 8 show the daily increase in confirmed cases, N, around their initial peak of the 2020 Covid-19 pandemic for South Korea, Israel, Poland and Azerbaijan.[3] Match each graph with the corresponding cumulative confirmed cases graph in Figure 9. Provide explanations for your choices.

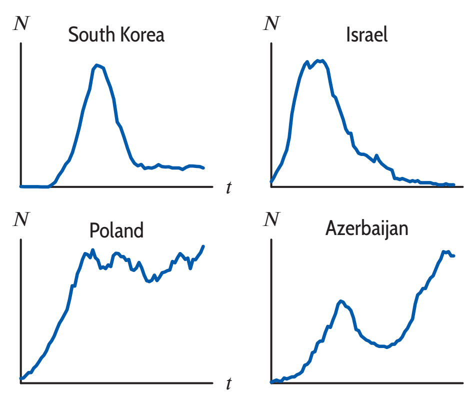

Figure 8

[1] https://www.anzics.com.au/wp-content/uploads/2020/02/2018-CCR-Report.pdf accessed on May 23, 2020.

[2] Data from Johns Hopkins University, downloaded from https://github.com/CSSEGISandData/COVID-19 accessed on June 1, 2020.

[3] Data from Johns Hopkins University, downloaded from https://github.com/CSSEGISandData/COVID-19 accessed on June 20, 2020.

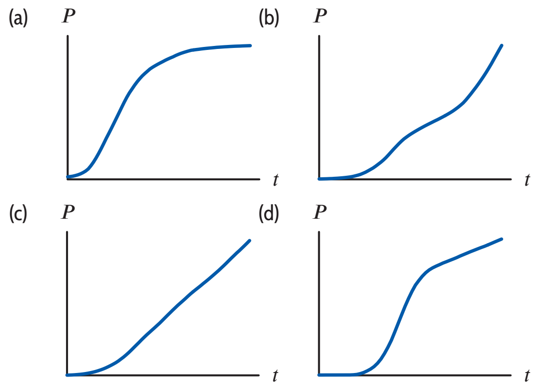

Figure 9

4. The total cumulative number of confirmed cases during the 2020 Covid-19 outbreak grew approximately exponentially in many countries during specific periods of time. For example, in Bangladesh[4], the total cumulative number of confirmed cases, P, grew about 6.2% per day during a two week period between 5/8/20 and 5/22/20.

Explain why the daily increase in cases, N, was also growing exponentially during this period of time. At what rate was it growing?

5. You plan to model the Covid-19 pandemic in your community. Your current data on total cases P is shown in Figure 11.

(a) Why might a logistic model seem appropriate?
(b) If a logistic model is correct, has your community now had fewer than half, about half, or more than half of the total number of cases it will eventually see?
(c) Using a logistic model, select the reasonable estimate of the ultimate number of cases by the end of the epidemic: 8,000; 10,000; 20,000; 30,0000.
(d) Another group of modelers has reported a prediction that is twice as high as your logistic prediction. How should you react?

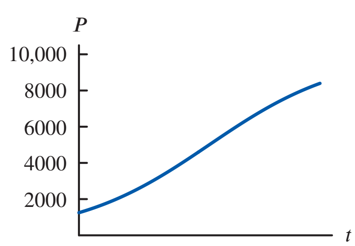

Figure 10

6. Figure 12 shows the daily percent change in the cumulative number of confirmed Covid-19 cases in Taiwan[5] during a 1-month period between 3/19/20 and 4/17/20.

(a) What was the daily percent change in cases on 3/29/20?
(b) Was the total number of confirmed coronavirus cases, P, growing or falling during that period?
(c) From the graph, do you expect P to be growing exponentially during that period?
(d) Figure 13 shows the graph of the daily percent change in cases against P (not t). Thinking of daily percent change in cases as $(1/P)dP/dt$, use it to write down a differential equation that you expect P to satisfy.
(e) Use separation of variables or a computer/calculator to solve the differential equation and find a formula for P in terms of t. Use the fact that $P = 73$ when $t = 0$ to find the value of the constant of integration.

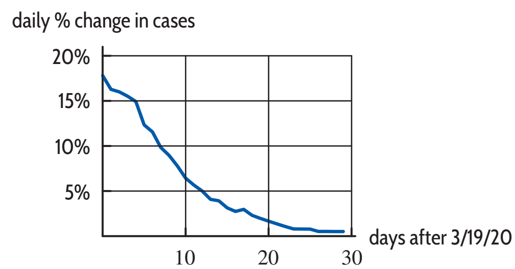

Figure 11

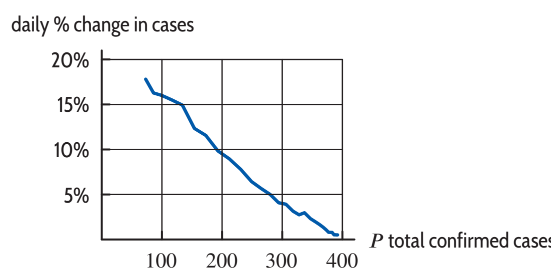

Figure 12

7. The cumulative total number of confirmed Covid-19 cases, P, in Belgium[6] during the Spring 2020 pandemic between the dates 3/15/20 and 5/5/20 is shown

[4]Data from Johns Hopkins University, downloaded from https://github.com/CSSEGISandData/COVID-19 accessed on June 1, 2020.

[5]Data from Johns Hopkins University, downloaded from https://github.com/CSSEGISandData/COVID-19 accessed on April 26, 2020.

[6]Data from Johns Hopkins University, downloaded from https://github.com/CSSEGISandData/COVID-19 accessed on June 15, 2020.

in Figure 14. The plot of dP/dt against P is shown in Figure 15, and the plot of $(1/P)dP/dt$ against P is shown Figure 16. (Here we used the estimates $dP/dt \approx P(t+1) - P(t)$, the daily increase in cases on day t, and $(1/P)dP/dt \approx (P(t+1)-P(t))/P(t)$, the percent daily increase in cases on day t.)

(a) If the trend in the data in Figure 15 continues, approximately how many people would you expect to get the virus during the course of the epidemic in Belgium? How about using Figure 16?

(b) Which of the 3 figures makes it easiest to decide visually whether a logistic model is a good fit for the data? Explain, and use it decide if a logistic function will fit the Belgium data well.

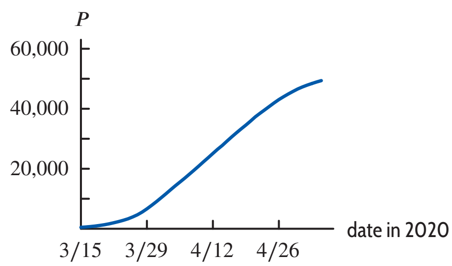

Figure 13

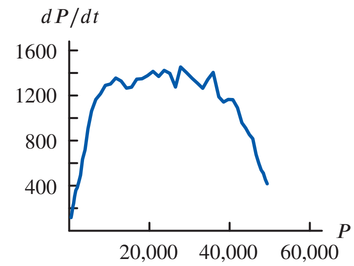

Figure 14

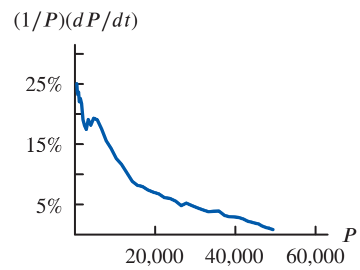

Figure 15