Introduction

P and N functions. The cumulative total number of Covid-19 cases confirmed up to and including day t is denoted by $P(t)$. We allow $t = 0$ to correspond to any given date.

The number of new cases confirmed on N on day t is given by

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

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

$$\begin{aligned} P(t) &= P(t-1) + N(t) \\ &= (\text{Total cases confirmed up to and including day } t) + (\text{New cases confirmed on day } t) \end{aligned}$$

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.

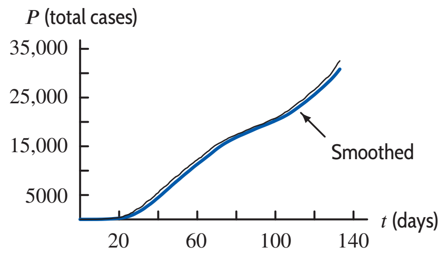

Figure 1

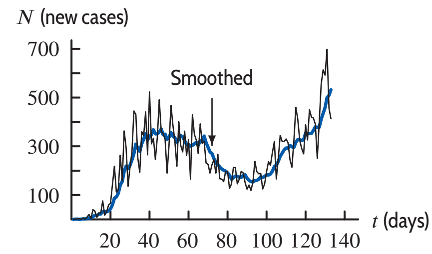

Figure 2

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.

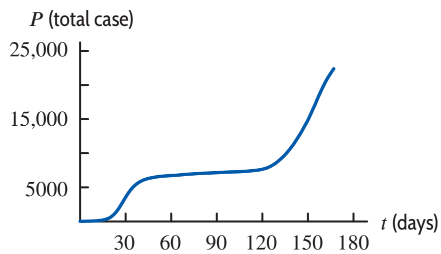

Figure 3

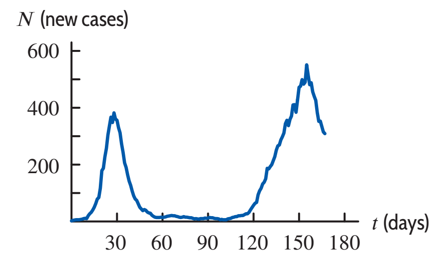

Figure 4

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.

More problems and problem ideas at the MCWG Covid resource page. Visit `https://mcwg.github.io/covid`.

Sample problems - Differential Equations

1. The total number, P, of Covid-19 cases in Iceland[1] confirmed up to and including day t, where $t = 0$ is March 4, 2020, was approximately logistic during the initial phase of the outbreak. Figure 5 shows dP/dt, approximated by N, the number of daily new cases, against P, together with a good fit parabola.
   - **(a)** Use Figure 5 to find a logistic model for P assuming 20 total confirmed cases up to March 4, 2020.
   - **(b)** Compare the model's predictions to the actual values of P in Table 1.

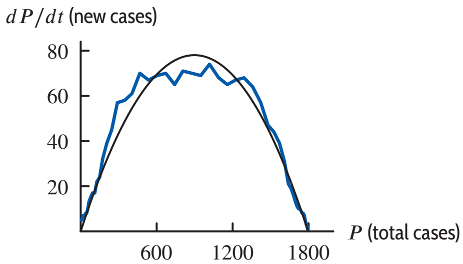

Figure 5

Table 1

Date	t	Total cases, P
March 18	14	173
April 1	28	1017
April 15	42	1696
April 29	56	1792
May 13	70	1801
May 27	84	1804

2. If $P(t)$ is growing logistically, then

$$\frac{dP}{dt} = kP\left(1 - \frac{P}{L}\right)$$

where k, and L are positive constants and $0 < P < L$.
   - **(a)** Is $\ln P$ an increasing, decreasing or constant function of t?
   - **(b)** Is the graph of $\ln P$ concave up, concave down, or straight?
   - **(c)** What does the graph of P vs t look like when plotted with a logarithmic scale on the vertical axis?

3. Figure 6 shows the daily relative change in the total number of confirmed Covid-19 cases in Taiwan[2] from day t to day $t + 1$, where $t = 0$ is March 19, 2020.
   - **(a)** What was the daily relative change in cases on March 29, 2020?
   - **(b)** Was the total number of confirmed Covid-19 cases, $P(t)$, growing or falling during the 30 day period?
   - **(c)** Do you expect $P(t)$ to be growing exponentially during the 30 day period?
   - **(d)** Figure 7 shows the daily relative change in confirmed cases against P (not t). Use the fact that $(1/P)dP/dt \approx (P(t + 1) - P(t))/P(t)$ to write down a differential equation that you expect P to satisfy.
   - **(e)** 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.

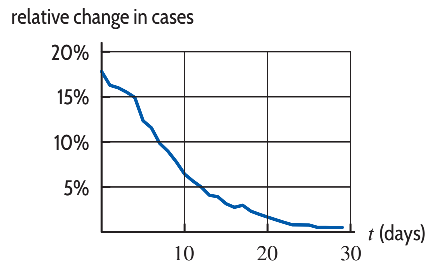

Figure 6

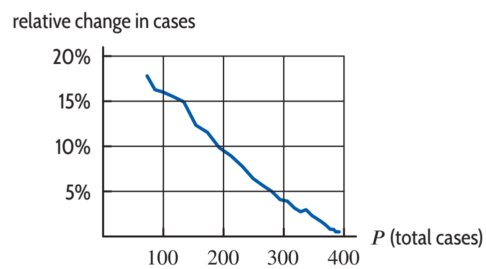

Figure 7

4. In some situations, there are some positive constants b and k so that for the total number $P(t)$ of Covid-19 cases up to an including day t

$$\frac{1}{P}\frac{dP}{dt} = be^{-kt}.$$

Show that any function that satisfies this equation can be written for some L and a as

$$P(t) = Le^{-e^{-k(t-a)}} = L\exp\left(-e^{-k(t-a)}\right).$$

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

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

5. Figure 8 shows the the cumulative total number of confirmed Covid-19 cases, $P(t)$, for the Covid-19 outbreak in Massachusetts t days since March 1, 2020, and Figure 9 shows the daily new confirmed cases, $N(t)$, computed as $N(t) = P(t+1) - P(t)$ (a discrete version of the derivative dP/dt). Note that that N/P is the relative change in cumulative cases from one day to the next. Figure 10 shows the plot of the log of this percent change, given by $\log_{10}(N/P)$. In this problem use the apparent straight line trend in Figure 10 to find a model for P.

(a) Explain why the straight line trend in Figure 10 for $0 \le t \le 110$ indicates that N/P is decaying exponentially during the first 3-4 months of the outbreak.

(b) Figure 11 shows that a good linear fit for the data in $0 \le t \le 110$ is given by $\log_{10}(N/P) \approx -0.1 - 0.0242t$. Use this, together with $P(70) \approx 77{,}000$, to find a Gompertz model for the data for $0 \le t \le 110$ of the form

$$P(t) = Le^{-e^{-k(t-a)}} = L\exp\left(-e^{-k(t-a)}\right).$$

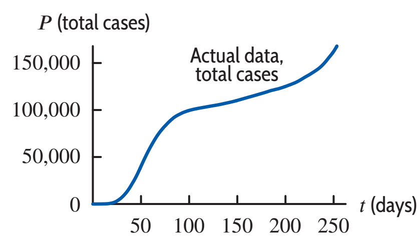

Figure 8

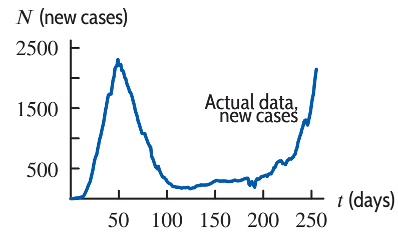

Figure 9

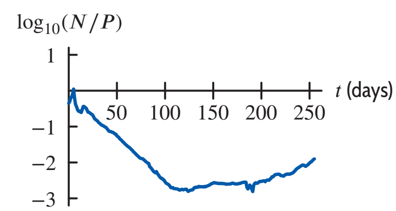

Figure 10

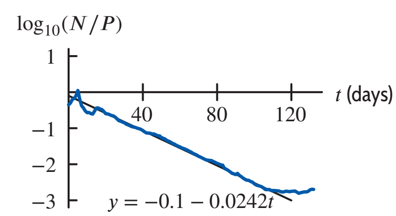

Figure 11