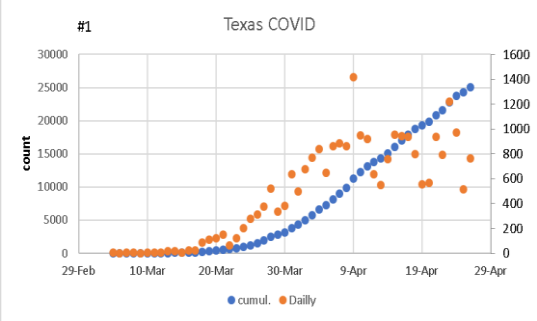
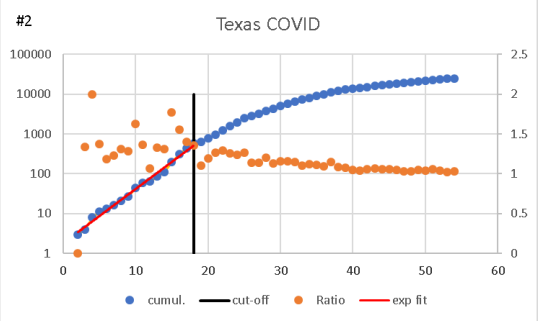
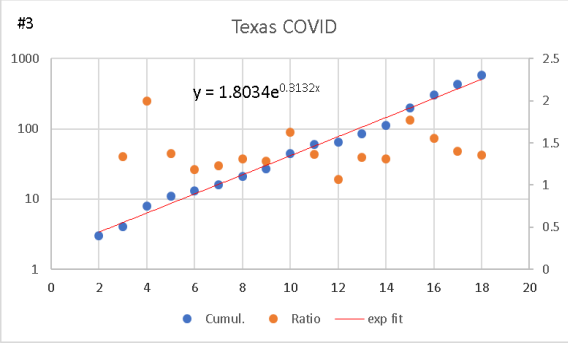
Explanation on R_o: If one can find the exponential part of the cumulative growth curve (a short section near the beginning of the curve (see graph #1) that loses validity as *count* gets close to zero), one finds the growth rate constant for an exponential. This will lead to the basic reproduction number, R_o.[1]

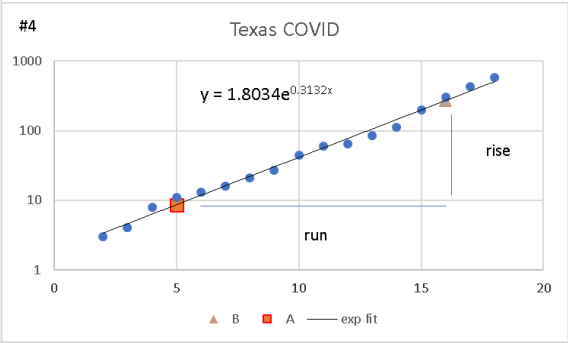
For example, here is some data on Texas' CV19 growth this year (the blue is cumulative), where #1 is the early days of the infestation. #2 is the same data transformed to a logarithmic scale. A log scale shows ratios and multiplicative effects better than a standard arithmetic scale. The orange data in #2 show ratios of successive confirmed case counts. Notice it is nearly constant (around 1.3) to the left of the vertical black line, before it starts dropping off. The red line is an exponential curve fit on the data prior to where the straight line starts to diverge. This is the true exponential part of the growth curve, before the number of people susceptible to infection drops off appreciably, and before any public health measures are taken. It is a basic characterization of the disease' behavior for its entire existence.



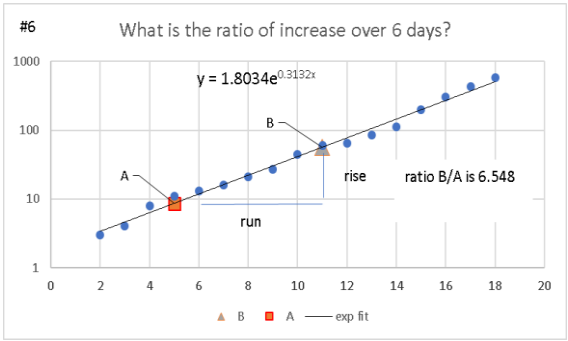


#3 is the same data, just zoomed in. It also shows the fitted curve equation, an exponential function. Notice the time constant is 0.313 (the value that *e* is raised to and multiplied by the *x* value, which in this case is the day number.) #4 shows the slope of this curve (with log vertical axis). The slope is constant throughout this stage, and is 0.313.





What if you know for this particular disease that the infectious period is 6 days? You might want to know the multiplier starting with one infected person over that same period. The ratio is the same for any pair of data on this section of the curve that are separated by 6 days, it should be apparent. This is the basic reproduction number.



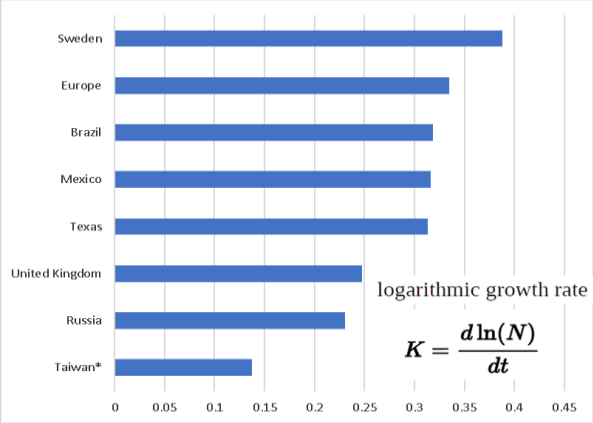
[1] Basic Reproduction Number. R_o is a measure of transmissibility: $R_o < 1$, disease disappears; $R_o = 1$, it's endemic; $R_o > 1$, epidemic. R_o is mentioned a lot during this epidemic, along with flattening of curves, with not a lot of understanding or relevance.

Mathematically, the above is just $count_n$ / $count_{n-6} = e^{K(1/\gamma)}$, where $1/\gamma$ is the infectious period, equal to six in this example, and K is the constant equal to the slope found earlier. There are lots of ways to find R_o numbers, and it depends on the disease and how probablistic you want to make it, but this is the fundamental concept.

You can use this to find out infectious rate and contact rate, since R_o is also defined as the ratio of the contact rate to the infectious rate. β is often used for contact rate; τ is often used for infectious period, while γ (or κ) is used for infectious rate, its inverse. So,

$$R_o = \frac{\beta}{\gamma} = \beta\tau$$

$R_o = e^{K\tau}$ where τ is the average infectious period for each person. Doubling time is also often mentioned. Reduce K by reducing contacts, reduce τ by isolation of infected individuals, for example. Here are some K values taken from Johns Hopkins data:



Example ($R_0 = 6.548$, $K_{TX} = 0.313$ / day)

$$R_o = e^{K\cdot\tau}$$

$$\tau = \frac{ln(R_o)}{K} \qquad \frac{ln(6.548)}{0.313} = 6 \quad \text{days}$$

Example, doubling time (2/1) :

$$2 = 1\cdot e^{K\cdot\tau}$$

$$\tau = \frac{ln(2)}{K} \qquad \frac{ln(2)}{0.313} = 2.2 \quad \text{days}$$

Interesting comment in open letter from Belgian doctors (Sep 2020):

https://docs4opendebate.be/en/open-letter/

Confirmed cases are supposed to be proven through testing and symptoms, but it may not always be true. The actual number of infected people can be several times what is measured with confirmed cases, though. So there is a difference between case fatality rate (CFR) and infection fatality rate (IFR). CFR is the ratio of the number of deaths divided by the number of confirmed (preferably through testing) cases of disease. IFR is the ratio of deaths divided by the number of actual infections.

Modeling

Interestingly, one way these epidemics are visualized and modeled is based on susceptibles contacting infected people at some average rate. The number of susceptibles is based on entire population[2] (which implies that some fraction of the population *isn't* susceptible; think vaccine or herd immunity). So, over time, these susceptible people are catching the disease at some rate. You could put it in pseudo-algebraic terms like this:

rate of change, susceptibles = contact rate, times number that are susceptible, times number of infectious [3]

[2] Another way to look at this is the ratio of susceptibles to the whole population is the basis for assigning a *probability* of infection of anyone in the whole population, not just susceptibles: contact rate X probability X infectious. The program of compartmentalizing by susceptibles, infectious, and recovered is known as the SIR model.

[3] Contact rate is average number of people a person has a *successful* infective contact with, per unit time. Hence contact tracing, that not only tells how many people a certain infected person has had contact with and infected, but hopefully who they were.

That needs to be refined, since the number in the population that is susceptible will decrease, as they are converted to infectious. That means the rate is negative. So, add a negative sign, and abbreviated a bit:

susceptibles change = −contact rate X susceptible X infectious

That's a bit cumbersome, so if you follow Leibniz' suggestion, and write a rate of change as the change divided by the corresponding change in time, where the symbol for change (difference) is a *d*, *S* for susceptibles, *t* for time:

$$\frac{dS}{dt} = -contact\ \ rate\ \ X\ \ susceptibl\ \ e\ \ X\ \ infectiou\ \ s$$

People have come up with less cumbersome algebraic symbols, like β for the contact rate, and *S* for susceptibles. Likewise, the average number of infectious would be well served with the letter I, so *I*:

$$\frac{dS}{dt} = -\beta SI$$

That's easier on the eye. It's important to remember all these quantities change over time, so they can be shown as functions of time (e.g. *S* written as *S*(*t*)), which clutters things up a bit. For example, deaths from an epidemic (not shown here) would reduce the value of *S*(*t*) over time also. *S* is ever decreasing (at least in this model), starting at the very beginning where just about everyone was susceptible and where R_o was found, keep in mind.

$$\frac{dS(t)}{dt} = -\beta S(t) I(t)$$

Furthermore, the number of infectious changes, too. So, another, new equation, for the compartment of infectious:

$$\frac{d\,I(t)}{dt} = \beta S(t) I(t) - \gamma\, I(t)$$

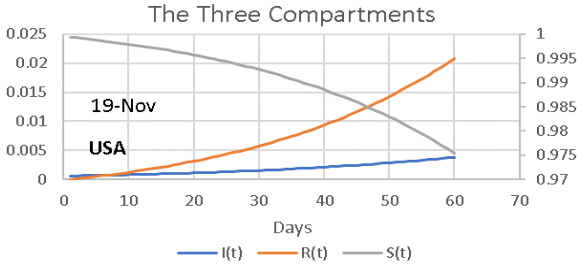
Notice the first term on the right side of the equation is just the negative of the change in susceptibles, since we're *adding* the same amount to infectious, that was *subtracted* from susceptibles before. So the change in infectious relates to the change in susceptibles, since members from one group are going over to the other. There is also the infectious period (not to be confused with contact rate). After a certain time, infectious either recover or die (this is kind of grim). So that has to be subtracted from the right side of the equation (the second term). γ is a common symbol for the frequency associated with the period of infection. Therefore γ^{-1} is the average period of infection. γ is the third letter in the Greek alphabet, gamma. $\gamma I(t)$ should be recognized as the rate of infectious by death or recovery, taken away from the overall rate.

The SIR model includes one more item in addition to the above, the recovered cases, *R*(*t*). The three compartments should add up to a constant *N*, the total population, throughout their evolution:

$$\frac{dS}{dt} = -\beta\ S(t) I(t)$$

$$\frac{dI}{dt} = \beta\ S(t) I(t) - \gamma\ I(t)$$

$$\frac{dR}{dt} = \gamma\ I(t)$$



where $S(t) + I(t) + R(t)$ = constant. That's true, because they are all feeding off each other. An approximation to these expressions can be made by adding increments, using Euler's method:

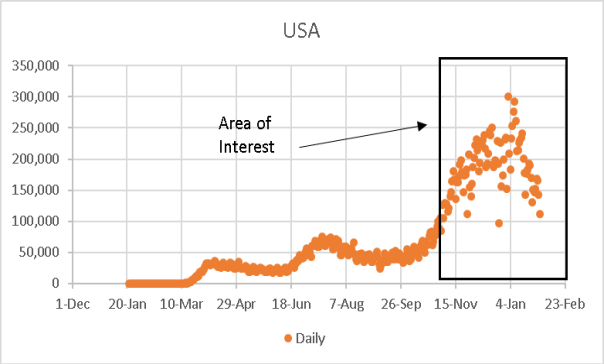
$$f(t + \Delta t) = f(t) + \Delta t\,[\Delta f(t)/\Delta t]\ , \text{ because } \frac{df}{dt} \approx \frac{\Delta f}{\Delta t}\ .$$

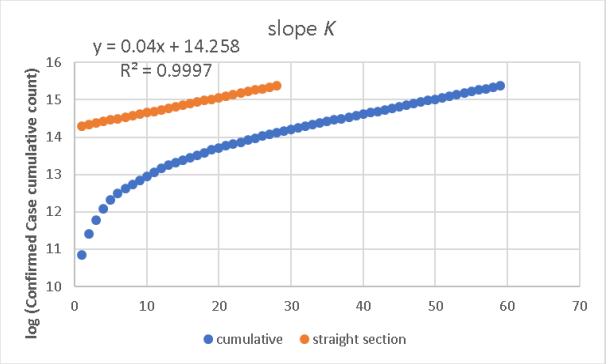
That way a table can be built for *I*(*t*), for example.

$$\frac{dI}{dt} \approx \frac{\Delta I}{\Delta t} = \beta\ S(t)I(t) - \gamma\ I(t)\text{, so } I(t+\Delta t) \approx I(t)\left[1 + \Delta t\left(\beta\ S(t) - \gamma\right)\right].$$

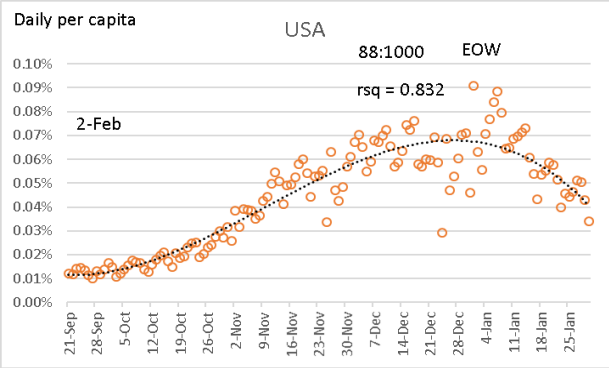
(Also, $S(t+\Delta t) \approx S(t)\left[1 - \Delta t\ \beta\ I(t)\right]$ and $R(t+\Delta t) \approx R(t) + \Delta t\ \gamma\ I(t)$.)

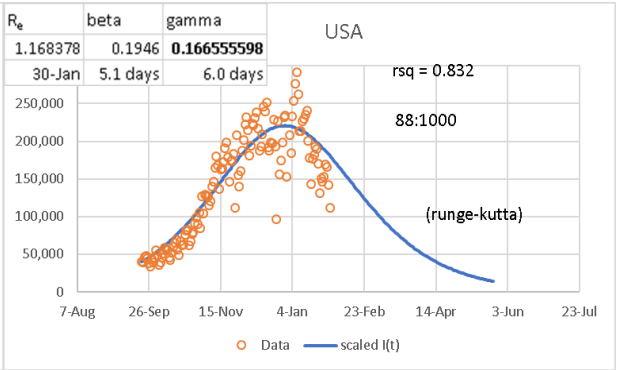
Taking the last few weeks of US numbers (just keep in mind, the US is comprised of thousands of similar sub-waves, at different magnitudes and starting points), and assuming that γ is (1/3)/day, taking daily increments (Δt = 1 day, in other words), a curve fit can be made, using R_o (use the definition from page 1) as a starting point as the two parameters to fit. The area of interest is shown below, with a logarithmic representation to the right, showing a slope of 0.041. Since R_o can be approximated as $e^{K/\gamma}$, R_o is initially estimated, and then further modified through a least squares fit. R_o is the value used to estimate herd immunity, which helps determine how much vaccine to administer. $R/N > 1 - 1/R_o$, R being Recovered.





One other thing interesting about the table so generated (it's called numerical integration, by the way) that you can see on the three compartments graph on the previous page is its arbitrary scale. That means the data to be fitted and the table of values has to be scaled to the actual counts.





It's not a great fit, but gives some idea of how it works. This reproduction number is very different than the original back in March, so probably ought be called R_{eff} or R_e instead. It is almost not an epidemic, but keep in mind this is a composite value of a very simplified demonstration. (The original R_o number from the beginning of March in the US could be 5.02, based on γ = 0.2 and K = 0.323.) Using the scaling factors, the original scale can be recovered, and the table filled out towards the end, as in the chart on the right.

In actual models, effective values are used, which are dynamic, meaning they change over time. For example, $\gamma(t)$ or $R_e(t)$, effective infectious rate and effective reproduction number. $R_e(t)$ is roughly equivalent to R_o times $s(t)$, where $s(t)$ is a fraction between 1 and 0 of the fraction S/N of susceptibles at time t. The above simplification describes three compartments, but more sophisticated models will add another compartment of *exposed* to susceptible, infected, and recovered, for example. Most of these models will show R_o from 2.5 to 6 or even more, but it's not directly comparable to these results. In fact, the value of R_o depends strongly on the model used, so are not interchangeable. But due to the uncertainties involved with mathematical modeling, one has to wonder what's the point of modeling beyond the simplistic level shown here.

It probably won't happen exactly as shown above, because of the uncertainty in any projection and also the lockdowns and extreme methods that will likely occur (already occuring), which will change parameters like β and R_e. The actual infection count is probably 5 to 10 times the confirmed cases. Another interesting thing to ponder is that once a country has achieved herd immunity the natural way, it won't need vaccines and won't need social distancing, masks, travel restrictions, etc. like the other countries that didn't will have to do for years. It will just need a lot of caskets.

P.S. See file "*Tri-County example*" for sensationlism example, too common with this epidemic.