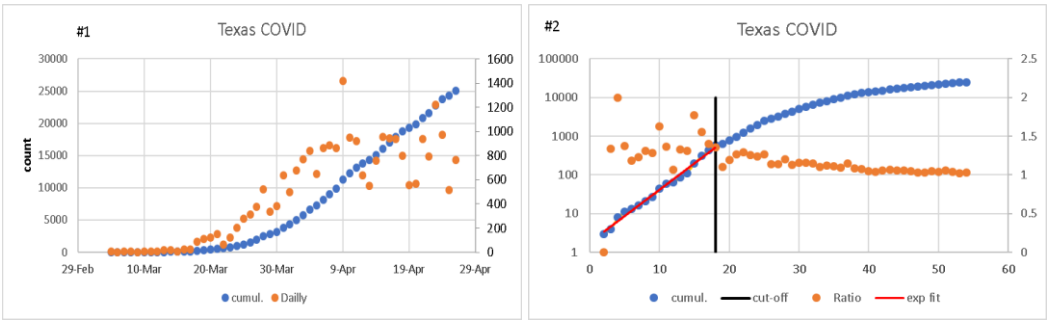
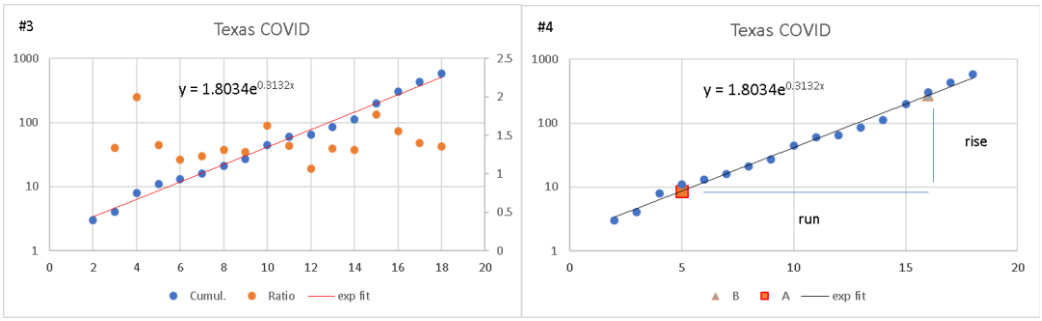
Explanation on R_o: If one can find the exponential part of the cumulative growth curve (a short section near the beginning of the curve (see graph #1) that loses validity as *count* gets closer to zero), one finds the growth rate constant for an exponential. This will lead to the basic reproduction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

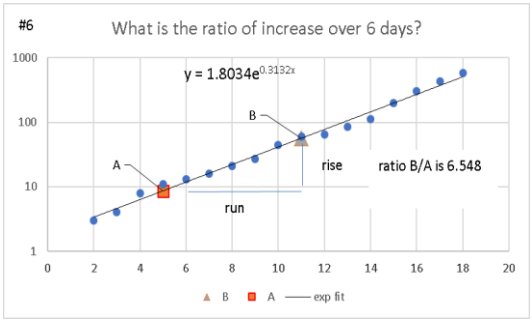
For example, here is some data on Texas' CV19 growth in 2020 (the blue is cumulative), where #1 is the early days of the infestation. #2 is the same data transformed to a logarithmic scale. A log scale shows ratios and multiplicative effects better than a standard arithmetic scale. The orange data in #2 show arithmetically *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 to one. 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, as will be shown.



#3 is the same data, just zoomed in. It also shows the fitted curve equation with its 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 its slope (with log vertical axis). The slope[2] is constant throughout this stage, and is 0.313.



What if you knew 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 is 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. Notice this is taken from *confirmed* cases, not *actual*, since actual is unknown. Assumes both are proportional.

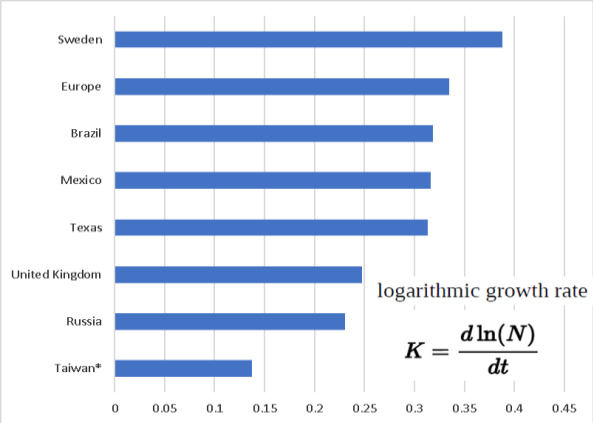
[2] Slope is *rise*/*run*, where *rise* is counted off logarithmically.

Mathematically, the above is $B/A = e^{K(1/\gamma)}$, where $1/\gamma$ is the infectious period, equal to six in this example, and K is the constant equal to the slope found earlier.[3] 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 the initial contact rate β, since β is defined as the product of the infectious rate γ and R_o. τ is often used for infectious period, while γ (or κ) is used for infectious rate, its inverse.

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

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

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

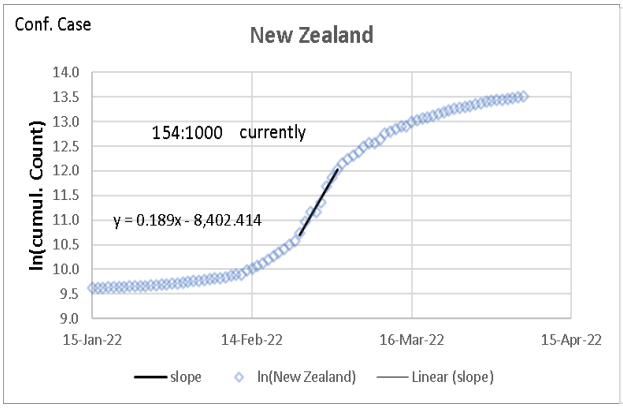
$$\tau = \frac{ln(R_o)}{K} \qquad \frac{ln(6.548)}{0.313} = 6 \text{ days}$$

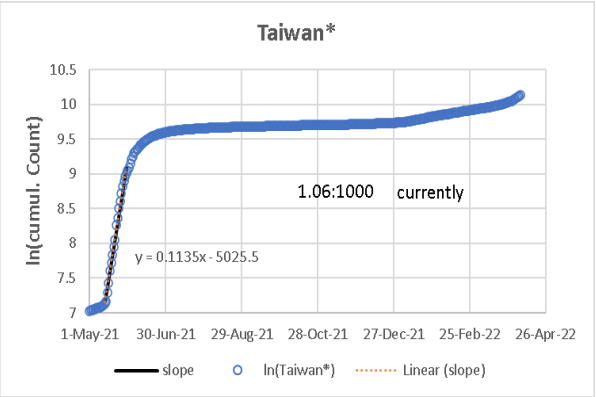
Example, doubling time (2/1) :

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

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

At this date, Taiwan, South Korea, China, and New Zealand have recently experienced a belated exponential growth period. Plotting this logarithmically shows it may be true, but the slopes are still very small:

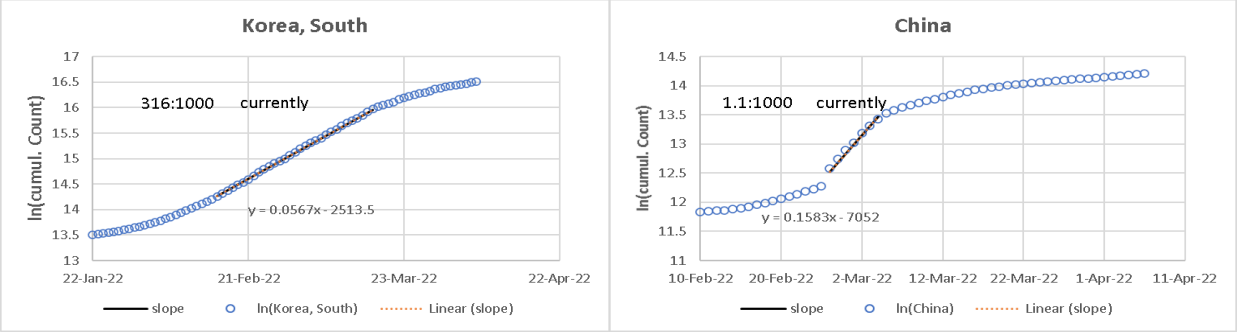




New Zealand and Taiwan clearly have "peaked" as far as the slope goes. New Zealand has the highest, but is still significantly less than the Texas K demonstrated above ($K = 0.313$). You can also see New Zealand is no longer "pristine" and is joining the rest of the world in this epidemic, exceeding 100:1000 confirmed cases.

[3] Exponent algebra: $B/A = \exp(Kt_n) / \exp(Kt_{n-6}) = \exp[K(t_n - t_{n-6})]$. $(t_n - t_{n-6})$ is the interval $1/\gamma = 6$.

South Korea of the four stayed on the log part of the curve the longest, but it can be seen its slope is also the smallest of all four:



Another question is why does the "true" exponential part of the curve disappear so quickly, after a few days? The reproduction number is the contact rate divided by the infectious rate, but if the contact rate decreases, while the infectious rate doesn't, the *effective* reproduction number is going to decrease, too. And that is what happens, since the pool of susceptibles is decreasing as they become infected, which reduces future "successful" contacts.

Assuming long lasting immunity[4], eventually the effective reproductive number (call it R_{eff}) is going to reach unity, and no more epidemic. This is a gradual process, and explains the usual bell-shaped curve the daily numbers assume. Let $s(t)$ be the fraction of the locale's population that are susceptible, then $R_{eff}(t) = R_o s(t)$. Where is $R_{eff}(t) \leq 1$?[5] It's where $R_o s \leq 1$, or $s \leq 1/R_o$. This the fraction of susceptibles remaining, so small it "starves" the epidemic out of existence. The more useful number is $1 - 1/R_o$, the percentage that is infected or recovered, which is *herd immunity*. So the original characterization of the epidemic by the basic reproduction number lasts throughout the epidemic, regardless of masks, lockdowns, and even vaccines, as the *effective* reproduction number changes.

The purpose of the vaccine is to supplement the natural progress to herd immunity. This can reduce sickness and death, and also reduce the time that the virus might mutate into resistant strains, which has the potential to start the whole thing over again. At one point, Dr Fauci recommended 90% vaccination of the population, which seemed to not take into account the already infected people. This suggests Fauci is taking the conservative path, and only counting what has been positively identified. (tests themselves aren't absolutely positive; vaccines aren't 100% effective, either) One can only speculate, but certainly can understand how the health authorities would want to get rid of this thing once and for all, and as soon as possible, so take the conservative path. However, there is a cost for any endeavor.

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 number of deaths divided by the number of *confirmed* (preferably through testing) cases of disease. IFR is the number of deaths divided by the number of *actual* infections. Ex.: if IFR = 0.34%, and CFR = 1.5%, 1.5%/0.34% = 4.4, *actuals* multiplier.

---

[4] The latest experience with the delta variant indicates it's a bit more complicated than just achieving herd immunity. Perhaps this is because new strains have resistance to the original immunity; the simple arithmetic of susceptibles becoming recovered with immunity is no longer completely valid. See below.

[5] See footnote number 1.

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 the entire population[6] (which implies that some fraction of the population *isn't* susceptible; think vaccine, natural immunity, or epidemic-acquired immunity, or even better, overall health helps resist infections). 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 [7]

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:

$$\frac{dS}{dt} = -contactrate\ X\ susceptible\ X\ infectious$$

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, starting at the very beginning where just about everyone was susceptible and where R_o was originally found, keep in mind.

$$\frac{dS(t)}{dt} = -\beta S(t)I(t)$$ (β assumed a constant in this example; may not always be true)

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

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

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, as shown earlier. 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, which is subtracted from the overall rate.

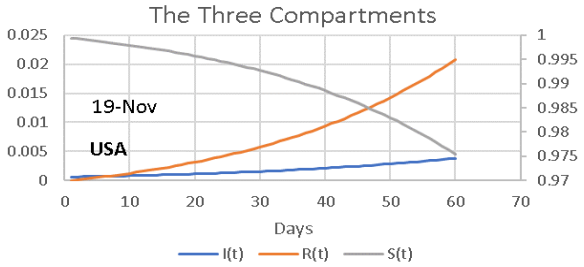
The SIR model includes one more compartment in addition to the above, the recovered cases, *R*(*t*). The three compartments should add up to a constant *N*, the total population, throughout their evolution:

---

[6] 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.

[7] 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.

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



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 using discrete time increments, using Euler's method, as might be in a spreadsheet:

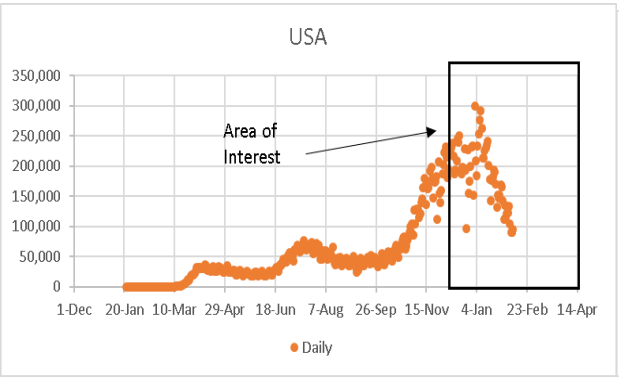
$f(t + \Delta t) = f(t) + \Delta t\, [\Delta f(t) / \Delta t]$, because $\frac{df}{dt} \approx \frac{\Delta f}{\Delta t}$. Most data is tracked daily, so Δt is usually one day.

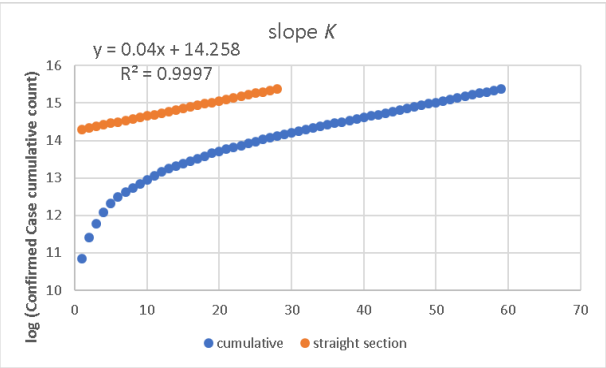
That way a time-based sequential table can be built for $S(t)$, $I(t)$, and $R(t)$.

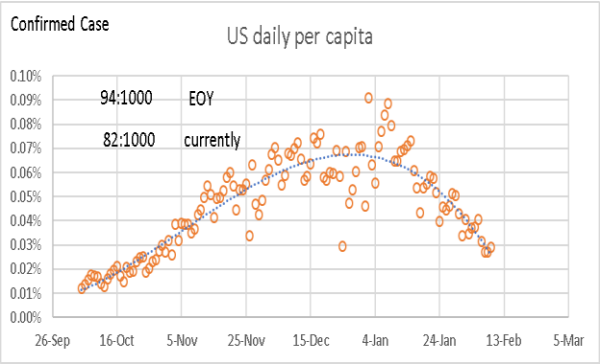
$\frac{dI}{dt} \approx \frac{\Delta I}{\Delta t} = \beta\, S(t)I(t) - \gamma\, I(t)$, so $I(t+\Delta t) \approx I(t)\left[1 + \Delta t\left(\beta\, S(t) - \gamma\right)\right]$. (If that isn't clear, do the algebra.)

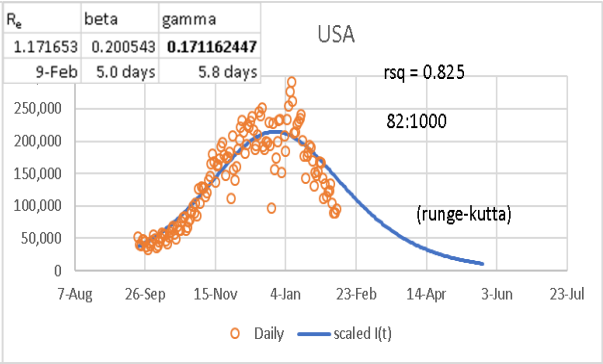
(Also, $S(t+\Delta t) \approx S(t)\left[1 - \Delta t\ \beta\, I(t)\right]$ and $R(t+\Delta t) \approx R(t) + \Delta t\ \gamma\, I(t)$. More algebra!)

Taking the last few weeks of US numbers (just keep in mind, the US is comprised of thousands of similar sub-waves from sub-locales, all at different magnitudes and starting points), and assuming that γ is (1/3)/day, taking daily increments (Δt = 1 day), 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.









This is probably the simplest way to do a numerical simulation, and can be enhanced by Runge-Kutta or more sophisticated numerical techniques.

It's not a great fit, but gives a very simplified idea of how it works. This reproduction number is very different than the original number back in March '20, so ought be called something different like 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. There are plenty of other factors not even mentioned here, such as change in infectious period due to treatments or natural mutations, the length of time of the epidemic, or regional, age, and ethnic variability, diet and exercise trends for a locale, changes in immunity with new strains and probably other unknown factors.

The above description describes three compartments, but more sophisticated models can add another compartment of *exposed* to susceptible, infected, and recovered, for example (SEIR model). 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 or R_{eff} depends strongly on the model used, so are not interchangeable between models. 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 certainly gets more complicated once waves of new variants start happening, with variability in immunity from prior strains and vaccines.

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 occurring), 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 might not need vaccines, nor 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.