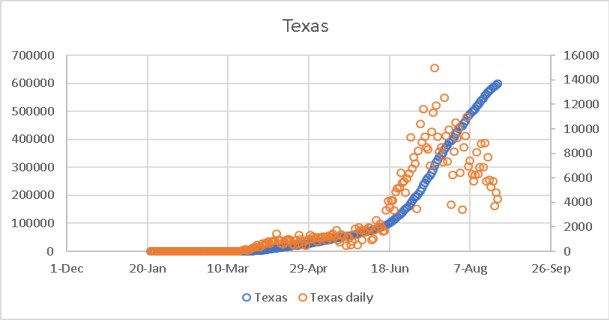
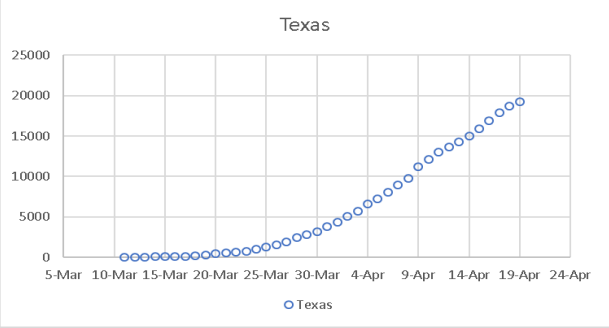
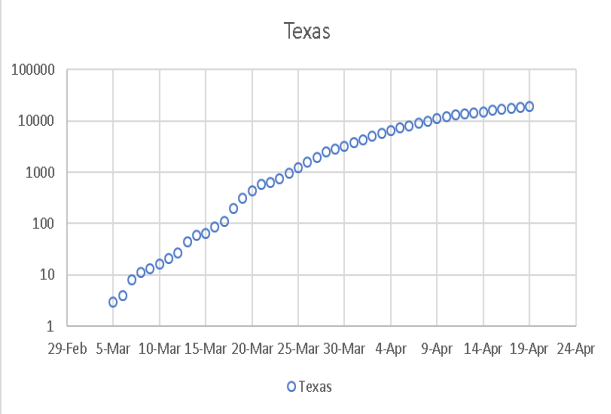
Explanation on R_o: If one can find the exponential part of the cumulative growth curve (a short section near the beginning of the curve that loses validity as N approaches zero), one finds the growth rate constant for an exponential, which is $K = d \ln(N) / dt$. [1]

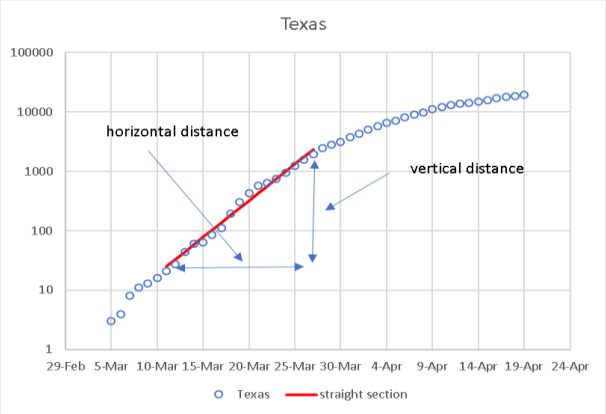
For example, here is some data on Texas' CV19 growth this year (the blue is cumulative), where the second graph zeros in on the very first days of this misery:





If we plot this semi-log (logarithmic vertical axis vs. arithmetic horizontal axis), this is what it looks like. The second graph highlights a straight line section, after which it begins to curve:





The straight section has a slope (ratio of rise over run, vertical over horizontal distance) that is 0.284 (done with a least squares regression). Afterwards, it is no longer exponential growth, since it curves off. This is called K, $K = 0.284$ for Texas. Notice this is a bit arbitrary, and avoids the values less than 40, at the extreme early stage.

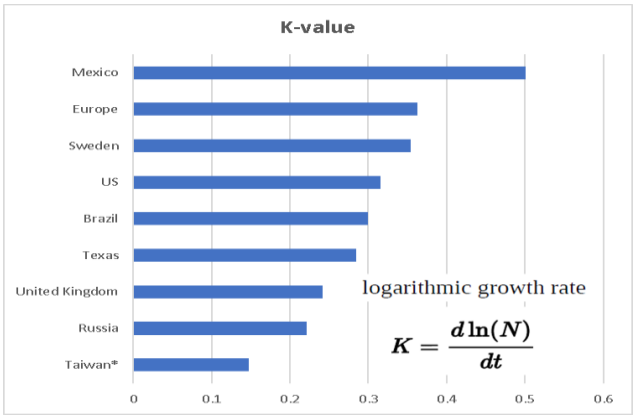
You can use this to find out infectious rate and contact rate, since R_0 is defined as the ratio of the contact rate to the infectious rate. β is often used for contact rate; τ is often used for infectious period, while γ is used for infectious rate, its inverse. So

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

R_0 is also found graphically as $R_0 = \exp(K\tau)$, where K is the slope of the exponential part of the cumulative curve just described.

[1] Basic Reproduction Rate. 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. The real trick is figuring out the *effective* reproduction rate, R_e.

$R_o = e^{K\tau}$ 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 ($R_0 = 2.6$, $K_{TX} = 0.284$ / day)

$$R_0 = e^{K \cdot \tau}$$

$$\tau = \frac{ln(R_0)}{K} \qquad \frac{ln(2.6)}{0.284} = 3.4 \text{ days}$$

Example, doubling time (2/1) :

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

$$\tau = \frac{ln(2)}{K} \qquad \frac{ln(2)}{0.284} = 2.4 \text{ days}$$

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

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

Reminder from introductory statistics: A low incidence infection is very sensitive to false positives. For example, say 6% incidence, and 95% accurate test. Out of 1,000 people that means 60 people are infected. Therefore, out of 1,000 – 60 = 940, 5% are false positives (or 48 people). On the other hand, 95% of the 6% (or, 5.7%) that are infected are true, and so 5.7% of 1000 = 57. The number of false positives is almost equal to the number of true positives in this example! This is a pitfall of widespread testing with a test that is less than 99% accurate. In this example, it would almost double the estimated number of positive cases which could be used to set policy (lockdowns, masks, etc.). However, current practice is to test people with symptoms or prior exposure, so the sample incidence is no doubt much higher than the population incidence, which reduces this effect.[2]

Modeling

Interestingly, one way these epidemics are visualized and modeled is based on susceptibles contacting infected people at some average rate. These susceptibles are quantified as a percentage or fraction of the entire population[3] (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 of susceptibles = contact rate X fraction X infectious [4]

---

[2] Notice 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 with SARS-CoV-2. Much uncertainty on false positives, and therefore on the real IFR.

[3] 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.

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

That needs to be refined, since the fraction of the population that are susceptibles will decrease, as they are converted to infectious. That means the rate is negative. So, add a negative sign:

$$susceptibl\ \ es\ change\ \ = -contact\ \ \ rate\ \ X\ \ \ fraction\ \ \ X\ \ infectiou\ \ s$$

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*:

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

People have come up with less cumbersome algebraic symbols, like β for the contact rate, and a real fraction for the fraction, *S*/*N*, where *S* is susceptibles and *N* is the number of the total population. Likewise, the average number of infectious would be well served with the letter I, so *I*:

$$\frac{dS}{dt} = -\beta \frac{S}{N} I$$

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 may clutter things up a bit. For example, deaths from an epidemic (not shown here) would reduce the value of *N*(*t*) over time. *S*/*N* is almost always less than 1.0, keep in mind.

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

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

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

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.

Here's where an alternative explanation for R_o shows up. This is the basic reproduction "rate" (although it's not really a rate, more on that later). If you want to solve for the infectious rate function *I*(*t*) from above, you do the following:

$\frac{d\,I(t)}{dt} = \beta \frac{S(t)}{N(t)} I(t) - \gamma\, I(t)$ or rewritten as $d\,I(t) = \left[\beta \frac{S(t)}{N(t)} I(t) - \gamma\, I(t)\right] dt$. So, separating the *I* variable,

$\frac{dI(t)}{I(t)} = \left[\beta \frac{S(t)}{N(t)} - \gamma\right] dt$. In algebra, you do the same thing to both sides of an equation, so the next step is to integrate both sides:

$\int \frac{dI(t)}{I(t)} = \int_t^{t+\tau} \beta \frac{S(x)}{N(x)} - \gamma\, dx$, using *x* for dummy variable of integration. This integrates the right side of the equation from some time *t* to a little later time at $t + \tau$; in other words, over the interval τ, but at a specific place in time.

$$\int \frac{dI(t)}{I(t)} = \int_t^{t+\tau} \frac{dI(x)}{I(x)} dx = \ln I(t+\tau) - \ln I(t)\ \ , \text{ or } \ln \frac{I(t+\tau)}{I(t)}\ \ \ , \text{ left side simplification.}$$

$$\ln \frac{I(t+\tau)}{I(t)} = \int_t^{t+\tau} \beta \frac{S(x)}{N(x)} - \gamma \, dx$$ And since $$\exp\left[\ln \frac{I(t+\tau)}{I(t)}\right] = \frac{I(t+\tau)}{I(t)}$$, the following:

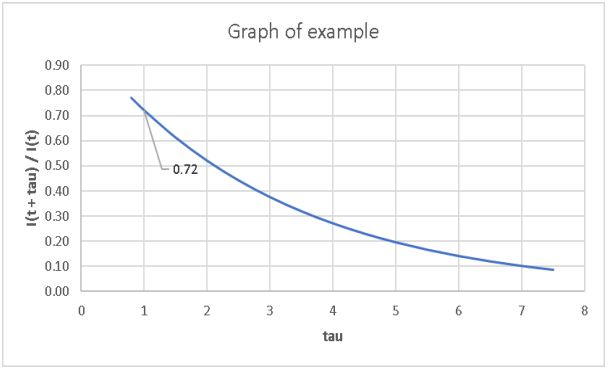
$$\frac{I(t+\tau)}{I(t)} = \exp\left[\int_t^{t+\tau} \beta \frac{S(x)}{N(x)} - \gamma \, dx\right]$$, applying the exponential function to the equation.

Finally, to R_o. It is expedient to combine β and γ into one parameter, which simplifies this expression.

$R_0 \equiv \frac{\beta}{\gamma}$, which makes the above expression $I(t+\tau) = I(t)\exp\left[\gamma \int_t^{t+\tau} R_0 \frac{S(x)}{N(x)} - 1 \, dx\right]$, rearranged slightly, too. The basic reproduction "rate" is a combination of the contact rate and the infection period in this case, in other words.

Such a thing is most likely to be numerically integrated, since it is rare to have an explicit algebraic form for $S(t)$ or $N(t)$ that can be integrated in a closed form. It should be clear that epidemics don't stay as pure exponential functions for long, or else everyone in the world would have gotten sick already, probably within a month of it starting! So this mysterious integral could lead to a gamma or gamma-like distribution, as so often is the case. Or, if over the interval from t to $t + \tau$, you have reason to believe S/N is relatively constant, the integral simplifies to $\tau\gamma$ $[R_0(S/N) - 1]$. Or

$$I(t+\tau) = I(t)\exp\left[\tau\gamma\left(R_0 \frac{S}{N} - 1\right)\right]$$ for interval τ. R_0 governs growth of this exponential function over a short period, and is dimensionless, since it is a ratio of contact rate and infectious rate with the same units, time^{-1}. That also explains why it is a bit of misnomer to call it a rate, when it has no unit dimensions, although it's common to hear it called a rate. This R_o is arrived at a little differently than the first example, as it's what *determines* this exponential function rather than being the *result* of the exponential function, but you can also see when this R_0 is close to 1, depending on what S/N is, the ratio of I's can also equal one[5]; that is no growth, it's endemic, as with the first definition. These two R_os are the same in this case, but definitions can vary with models. This modeling is helpful in trying to determine things like contact rate and infectious rate, parameters in these equations that ought to help inform the decision making on containment or eradication strategies. If the ratio S/N slowly decreases because there are fewer susceptibles (i.e. death, widespread immunity, vaccines), you can see the ratio $I(t + \tau)/I(t)$ will decrease, also. This is just a sketch of what it's about; you can imagine that there are probably a million different ways to try to model this.



What if the infectious period is 3 days, the contact period 4 days, and the ratio of susceptibles to the population is 3%, for the period of interest of 1 day? β = ¼, γ = 1/3, so R_o = (1/4)/(1/3) = 0.75. The growth multiplier is exp[1 X 1/3 X (¾ X 3% - 1)], or 0.72. This is negative growth, since it is less than 1. If you increase interval τ arithmetically, you will see the growth multiplier decrease exponentially in this particular example:

Cases can be reported over time as a cumulative value, or as a daily value. It's two ways of saying the same thing; the example at the beginning of this article concentrated on cumulative growth. The cumulative, or the total number of cases up to time t, $T(t)$, follows the equation $dT/dt = \beta S/N I$.

[5] Remember, $e^0 = 1$. One could also define R_{equiv} or R_{eff} as $R_o(S/N)$, which could also be dynamic as a function of time as $R_{eff}(t)$ = as $R_o(S(t)/N(t))$. There could be multiple equations besides the $I(t)$ = infectious described here, such as $S(t)$ = susceptibles or $R(t)$ = recovered that would help in their solution.