During the initial growth phase of a pandemic, the number of cases shows exponential growth. As time progresses, the number of new cases each day starts to decrease, and eventually we reach a final number of cases. Thus, the cumulative number of cases form an approximately linear relationship with time in the log-linear scale in the initial phases, but then plateaus after a certain timepoint.

It is common in Epidemiology to model this initial exponential growth and subsequent slow-down using a logistic growth model. The cumulative incidences (the total number of cases by time ) can be approximated by

,

where is the exponential growth rate, , and

The new cases in a time period is thus

By taking the derivative of , we can estimate the change in number of total cases with respect to time. Hence, when looking between two subsequent time points, we can find the number of new cases.

The derivative of , the density of the logistic function, is , which looks a lot like a normal distribution! Hence, we can use a normal distribution to approximate the logistic function.

This led us to attempt to model number of new cases using a Gaussian curve.

We found this fit the number of new cases quite well for many countries, with some limitations of course. Since this density fit the number of new cases quite well in initial modeling, we decided to create a model that incorporates an term, and hence we included covariates for as well as .