Should we use the Gompertz function to model COVID-19 cases?

Shivam Patel

Imperial College London

Introduction

Recently, I saw a webinar by Michael Levitt[1], where he suggested that Gompertz function models COVID-19 data better than the Sigmoid function which is more commonly used. He used the examples of South Korea and New Zealand. I decided to test his hypothesis with more countries. This poster is the results of my test and answering the questions they have raised.

What is the Gompertz Function?

The Gompertz Function was shown by Benjamin Gompertz in 1825. It was initially used to calculate age related mortality rates. It is now used for many modeling purposes including the growth of tumors. Both the Gompertz function and the Sigmoid function are derived from the generalised logistic function[2],

$$\frac{d}{dt}I(t) = r \cdot I(t) \cdot s \cdot \left(1 - \left(\frac{I(t)}{N}\right)^{\frac{1}{s}}\right),$$

where $I(t)$ is the total number of cases at time t, r is the growth rate, s determines the location of maximum growth relative to the lower and upper asymptotes and N is the population size. Letting $s = 1$ will give the Logistic differential equation,

$$\frac{d}{dt}I(t) = r \cdot I(t) \cdot \left(1 - \left(\frac{I(t)}{N}\right)\right). \tag{1}$$

Futhermore taking the limit $s \to \infty$ gives rise to the Gompertz differential equation,

$$\frac{d}{dt}I(t) = r \cdot I(t) \cdot \log\left(\frac{N}{I(t)}\right). \tag{2}$$

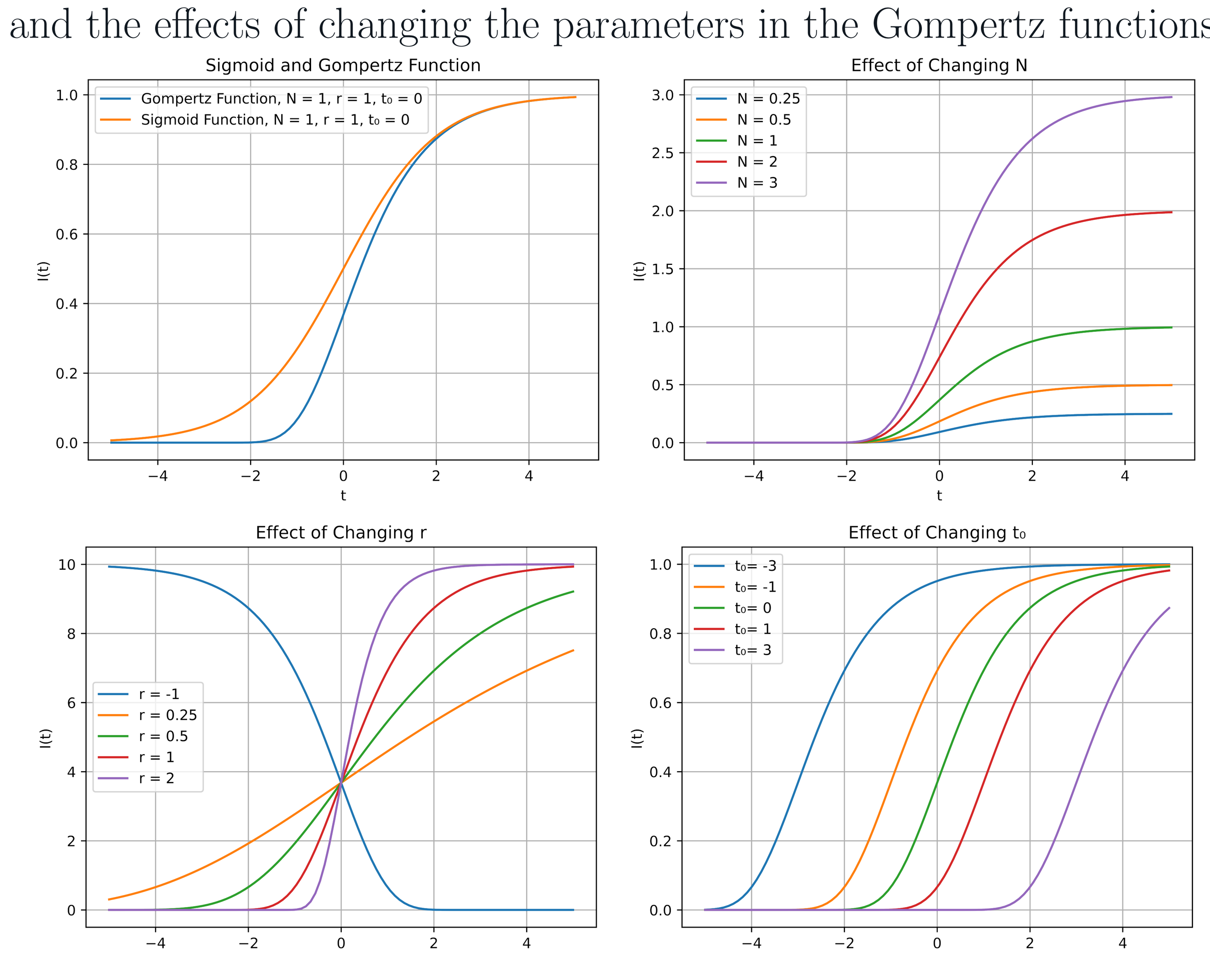
Solving (1) and (2) gives the Sigmoid function,

$$I(t) = \frac{N}{1 + \exp(-r \cdot (t - t_0))},$$

and the Gompertz function,

$$I(t) = N \cdot \exp(-\exp(-r \cdot (t - t_0))).$$

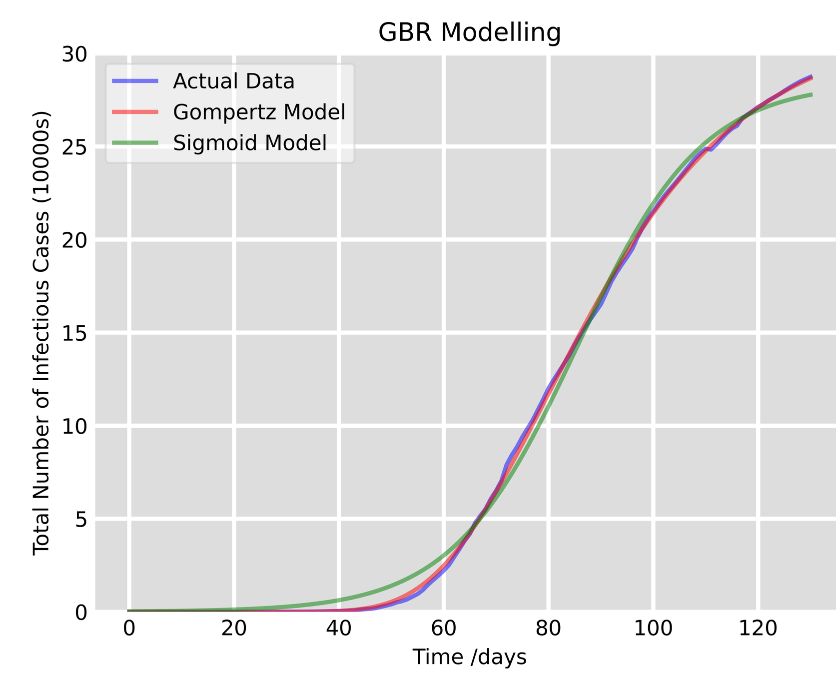
The following plots show the difference between the Gompertz and Sigmoid functions, and the effects of changing the parameters in the Gompertz functions.



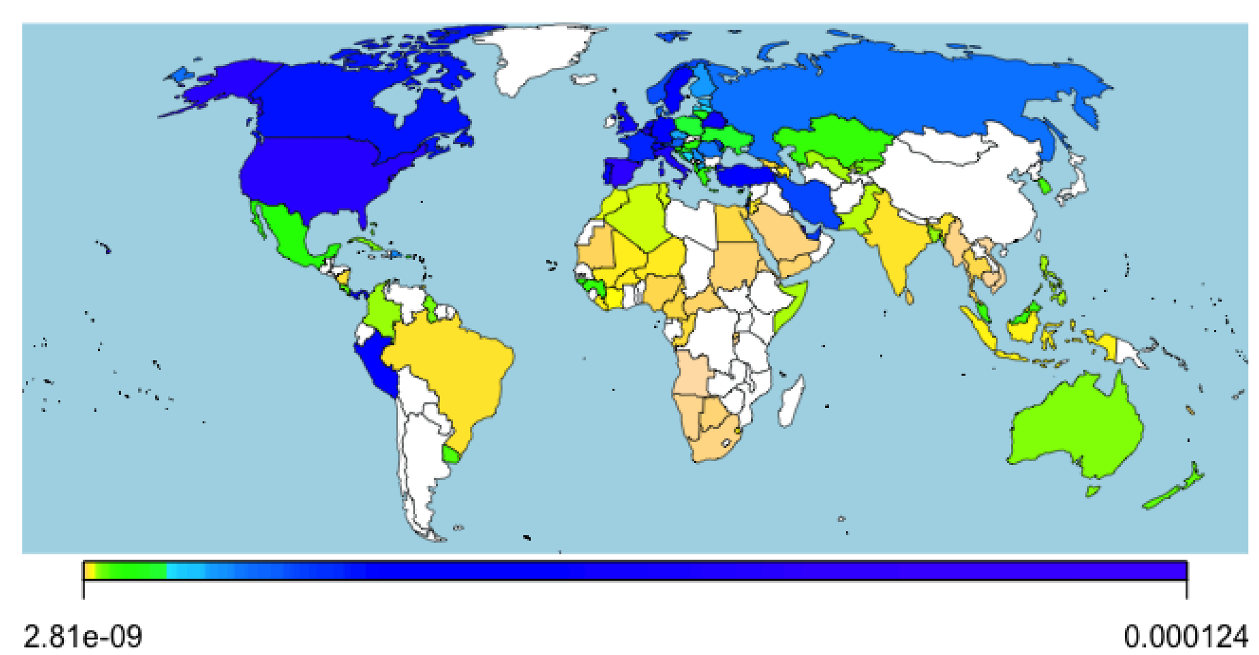
Is the Gompertz Model more accurate?

To fit the models, I used iterative gradient descent procedures such as the Levenberg Marquadt algorithm. From modelling 210 locations using both the Sigmoid function and the Gompertz function, and taking the root mean squared error of both models, I have found the Gompertz function produces a more accurate model for 136 locations, the greatest improvement being in USA. The link to my source code can be found in the References and Oral URL section.

Here is a graph of the total number of cases in Great Britain with both the Sigmoid and Gompertz models overlaid.



Here is a plot showing the countries where the Gompertz Model is more accurate than the Sigmoid Model and the normalised improvement by the shading.



Can the Gompertz Model be improved?

The Gompertz model is limited in many ways. It assumes a constant population, no chance of reinfection, no exposed period and no maternally-derived immunity. However it can be improved in a similar way to the SIR model being improved to the SIRS, SEIR, or MSEIRS models[3]. If one considers $\frac{\beta I}{N}$ in the SIR model to be equal to $r \cdot \log\left(\frac{N}{I}\right)$, then the SIR model now includes the Gompertz function. This is because the force of infection can be nonlinear.

Original SIR Model

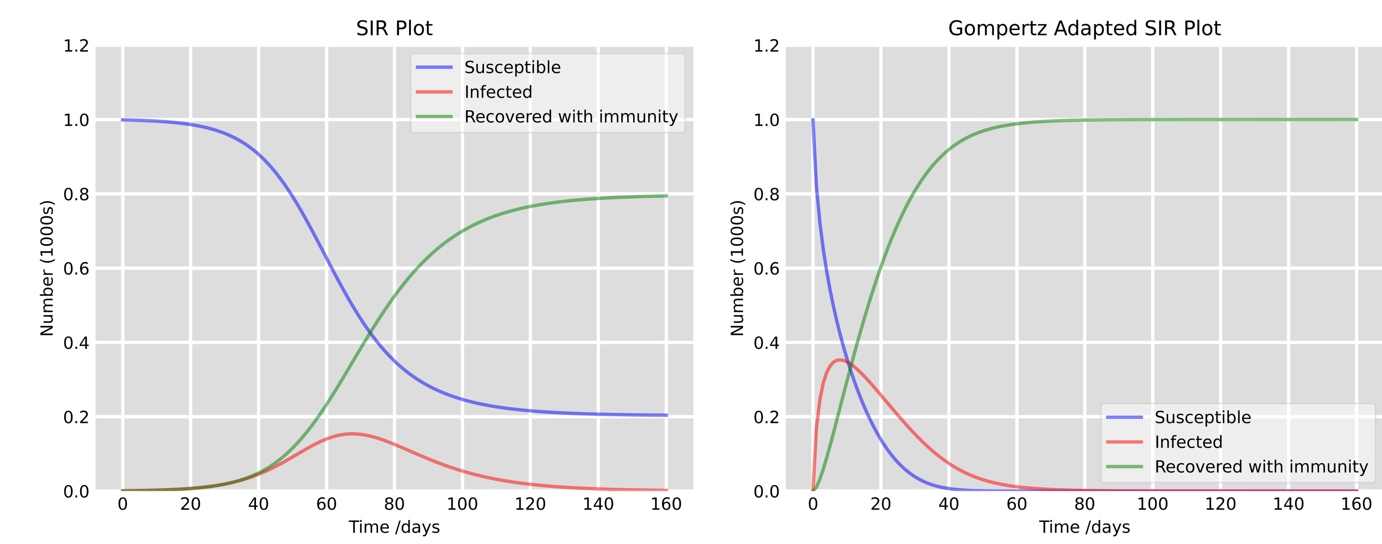
$$\begin{aligned}\frac{dS}{dt} &= -\frac{\beta IS}{N}\\ \frac{dI}{dt} &= \frac{\beta IS}{N} - \gamma I\\ \frac{dR}{dt} &= \gamma I\end{aligned}$$

$$\longrightarrow$$

Gompertz Adapted SIR Model

$$\begin{aligned}\frac{dS}{dt} &= -S \cdot r \cdot \log\left(\frac{N}{I}\right)\\ \frac{dI}{dt} &= S \cdot r \cdot \log\left(\frac{N}{I}\right) - \gamma I\\ \frac{dR}{dt} &= \gamma I\end{aligned}$$

What does this model look like?



Conclusion

My results show that the Gompertz model is more accurate in many locations but having a more accurate model does not necessarily constitute a better model. One could use a polynomial function and get an even closer fitting model but this would be useless as a polynomial function would then suggest the number of cases grows or falls to ± infinity, something which is impossible in the real world.

However the Gompertz function follows a similar S-shaped graph to the Sigmoid function however the difference is Gompertz function is not symmetric. The Gompertz functions show a higher rate of growth initially and do not tamper off as quickly as the Logistic function.
My results could support the following conclusions:

- Responses such as lockdowns and quarantines have been effective in slowing down the infection.
- The more people who are infected, the more immune individuals there are in the population and therefore a smaller force of infection.

This result could also highlight the pitfalls of the logistic function. One being the assumption that individuals are in contact equally within a population, which may not be the case.

To conclude I think we should consider using the Gompertz function to model COVID-19 data because it could be more accurate. But there is no clear answer to should we use the model or not. Both the Logistic Model and Gompertz Model should be computed and used to better inform the epidemiologists. A variety of epidemic models should be computed as the results from all of these models can help in making further conclusions on the nature of the epidemic.

References and URLs

Oral:https://imperial.cloud.panopto.eu/Panopto/Pages/Viewer.aspx?id=c45fde0a-0165-4172-9bed-abd60174bebe

Source Code:http://tiny.cc/ShivamPatelM1RSourceCode

[1] M. Levitt. [Online]. Available: https://www.youtube.com/watch?v=Uw2ZTaiN97k.

[2] C. Bauckhage, “The math of epidemic outbreaks and spread (part 3) least squares fitting of gompertz growth models,” 2020.

[3] H. W. Hethcote, “The mathematics of infectious diseases,” *SIAM review*, vol. 42, no. 4, pp. 599–653, 2000.