

Methods in Modelling Covid 19 for both the UK and Brazil

Abstract

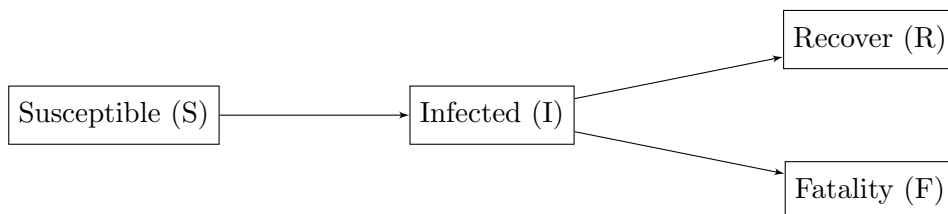
We will look at different methods to model the Covid 19 pandemic in both Brazil and the UK. Initially using a recurrence relation illustrating how different groups within the population as a whole evolve over time. We can then add complexity to our model removing assumptions previously made. Our model can then be calibrated to each of the countries in question separately based on their observed data. This will allow us to observe and compare the difference in key statistics such as the R value defined as the average number of additional people an infected person will infect between our countries in question. The UK imposed strict lockdown rules, in contrast Brazil implemented little to no intervention providing us with a clear comparison.

1 Introduction

Covid 19 was a worldwide pandemic beginning in 2020 encompassing the spread of the disease known as Coronavirus causing severe acute respiratory syndrome[1]. Countries around the world faced an enormous challenge aiming to curb the spread of the disease preventing it running rampant through the population owing to the disease's high infection rate and estimated fatality rate of 1%[2]. Designing a model to forecast cases and fatalities over time is paramount to allow governments to make informed decisions about the welfare of their citizens.

2 Simple SIR model

During the pandemic the evolution of the population can be distilled down into four different states. susceptible: those who have not been infected yet; infectious: those who are ill and infectious and then recovered or fatalities i.e. those who have recovered or those who haven't.



Initially we can make a couple assumptions omitting specific factors affecting infection rate such as duration and conditions of interactions between people and how easily the virus can be transmitted. By looking at each day i separately we can create a recurrence relationship for S , I , R and F values on day i dependent on the days before. ΔI_i , the number of people infected on any given day i can be described as proportional to the number of susceptible people times the probability a given person is infectious. (I_i/N) . d can be described as the number of days a person is infectious before they either recover or die.

$$\begin{aligned}\Delta I_i &= K S_i I_i / N \\ S_{i+1} &= S_i - \Delta I_i\end{aligned}\tag{1}$$

The Constant K in (1) can be described by the following equation:

$$K = R/d\tag{2}$$

As mentioned above R is the average number of additional people an infected person will infect and is different from country to country. The fatality rate can be written as K_f allowing us to express the number of people who have been infected, recovered and died on day $i+1$.

$$\begin{aligned} I_{i+1} &= I_i + \Delta I_i - \Delta I_{i-d} \\ R_{i+1} &= R_i + (1 - K_f)\Delta I_{i-d} \\ F_{i+1} &= F_i + K_f\Delta I_{i-d}. \end{aligned} \quad (3)$$

The population N can be seen to remain constant as shown below:

Proof.

$$\begin{aligned} N_i &= S_i + I_i + R_i + F_i \\ &= (S_{i+1} + \Delta I_i) + (I_{i+1} - \Delta I_i + \Delta I_{i-d}) + (R_{i+1} - (1 - K_f)\Delta I_{i-d}) + (F_{i+1} - K_f\Delta I_{i-d}) \\ &= S_{i+1} + I_{i+1} + R_{i+1} + F_{i+1} \end{aligned} \quad (4)$$

Therefore N_i must be constant for all i □

Using our new recurrence relationships we can now model infections for a general country's recoveries and fatalities against time i for different R values. We will use a log linear graph as it makes it easier to read the data.

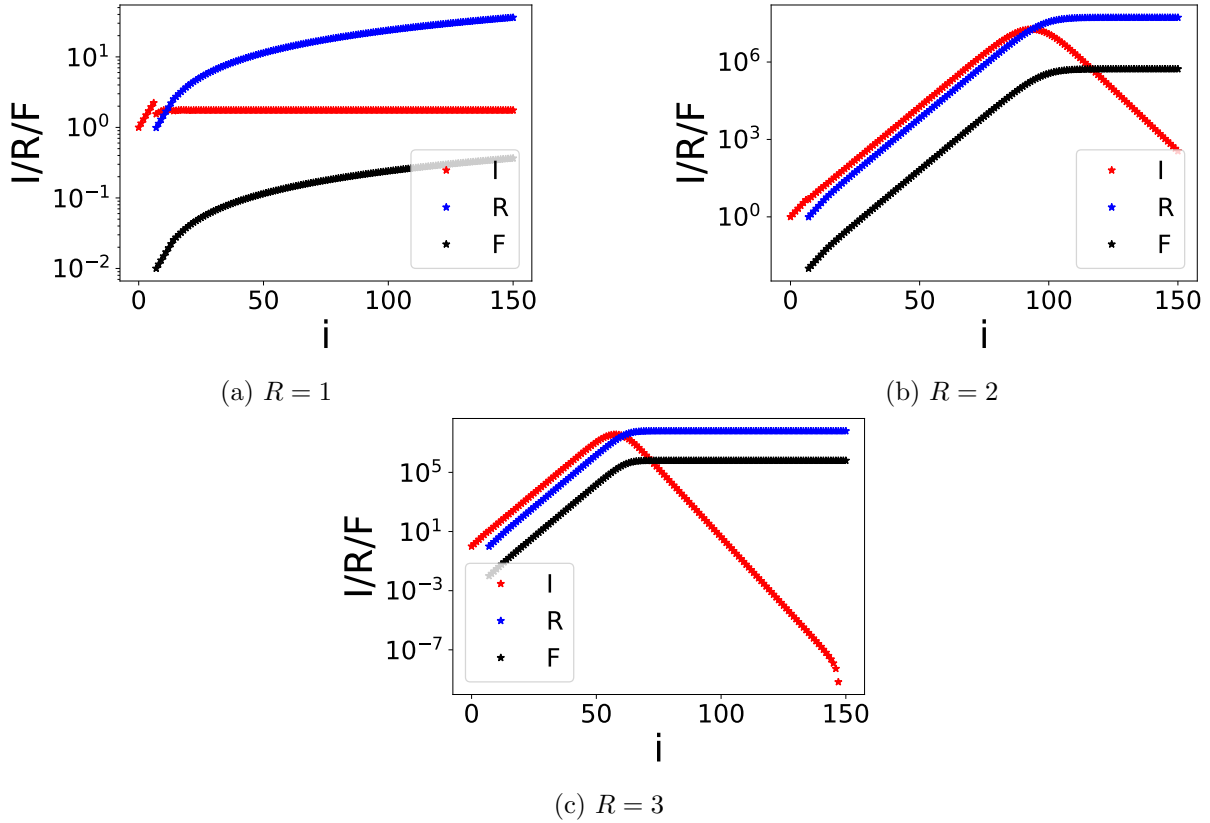


Figure 1: Modelling different R values

From the graphs above we can infer that the higher the R value the greater the number of total fatalities and concurrent infections. Looking at figure 1a $R=1$, we can observe that

infected persons remains constant over time due to straight flat red line. This makes sense as each person is only infecting one other on average so it is difficult for disease to spread through the population. For greater values of R such as 2 and 3 the number of infections increases exponentially until the proportion of people left to infect is too small no matter the R rate. This can be described by the term herd immunity where the a significant proportion of the remaining population has recovered and is no longer susceptible to the disease causing the rate of new infections to fall.

3 Exponential Model

At the onset of the pandemic the number of people infected is extremely small in comparison to the total population. As a result we can assume for the first d days S is constant and equal to N, using this simplification (1) reduces down to:

$$I_{i+1} = I_i + \Delta I_i \quad (5)$$

Moreover, again similarly we can assume recoveries are negligible considering only the first d days of the pandemic resulting in the following equation for I_{i+1} :

$$I_{i+1} = I_i + R(I_{i-d} - I_{i-d-1}) \quad (6)$$

From figure: 1b we can see that initially the infections and fatalities can be seem to be exponentially increasing. As a result we can aim to model this using an exponential distribution of the following form.

$$\begin{aligned} I_i &= Aexp(\lambda i) \\ F_i &= Aexp(\lambda i) \end{aligned} \quad (7)$$

Substituting in 7 into 6 we can derive an expression for R in terms of λ and d.

$$\begin{aligned} Aexp(\lambda(i+1)) &= Aexp(\lambda i) + R(Aexp(\lambda(i-d)) - Aexp(\lambda(i-d-1))) \\ e^\lambda e^{\lambda i} &= e^{\lambda i} + R(e^{\lambda i} e^{-d\lambda} - e^{\lambda i} e^{\lambda(-d-1)}) \\ e^\lambda &= 1 + R(e^{-d\lambda} - e^{-\lambda(d+1)}) \\ R &= \frac{e^\lambda(e^\lambda - 1)}{e^{-d\lambda}(e^\lambda - 1)} \\ &= exp(\lambda(d+1)) \end{aligned} \quad (8)$$

Similarly, another parameter epidemiologists use is the doubling time defined as the time for the number of fatalities or cases to double. We can derive this value τ using our exponential model for the beginning of the pandemic.

$$\begin{aligned} F &= Aexp(\lambda t) \\ 2F &= Aexp(\lambda(t+\tau)) = 2Aexp(\lambda t) \\ \lambda(t+\tau) &= \log(2) + \lambda t \\ \tau &= \frac{\log 2}{\lambda} \end{aligned} \quad (9)$$

Using a separate program and the function 7 to model the data we can use our fatalities data and formula 9 to find λ and τ for both equations. We do this as the observed fatality data is more reliable than for cases. The fitting will be performed using the log of the data generating a log linear graph of fatalities against days for both countries assuming d=6.5.



Figure 2: Fitting exp model to fatalities data

Country	λ	Doubling time	R	A
UK	0.294	2.35	9.10	6.69e-5
Brazil	0.0793	8.74	1.81	71.3

It can be seen above that UK had a significantly larger R value suggesting the infection spread much more rapidly than in Brazil initially. In addition, the shorter doubling time for fatalities supports this conclusion as infections directly lead to fatalities.

4 More general SIR Model

A more general model still consists of the same four stages as mentioned above depending on time d when initially exposed although now once you are exposed to the virus you are not infectious for a few days. After they can then contaminate others and will ultimately recover or not. In our model we will not look at displaying symptoms as an informative state as many people with Covid 19 do not display symptoms but are still infectious. However, now we will introduce the probability distributions to further our model. The probability to be infected on day d will be proportional to the probability to meet someone who was infected on day $d - 1$ times the respective probability, P_{Inf} that a person is infectious after 1 day, plus the probability that they were infected on day $d - 2$ times the probability of infection after 2 days, plus ... This can be summarised but the following equation.

$$\Delta I_i = R S_i \sum_{d=1}^{N_i} \frac{\Delta I_{i-d}}{N} P_{Inf}(d) \quad (10)$$

where n_i is the number of days such that $P_{Inf}(d) = 0$ if $d > n_i$. The distribution of fatalities and recoveries are shown to be similar in [3] as a result define this probability distribution as $P_r(d)$ for both.

$$\begin{aligned} \Delta F_i &= K_f \sum_{d=1}^{n_r} \Delta I_{i-d} P_r(d) \\ \Delta R_i &= (1 - K_f) \sum_{d=1}^{n_r} \Delta I_{i-d} P_r(d) \end{aligned} \quad (11)$$

Again where n_r is the number of days such that $P_r(d) = 0$ if $d > n_r$. We can now substitute 11 into 3 to give our general equation. Our Probability distributions can be defined as follows argued in [2] and [3]

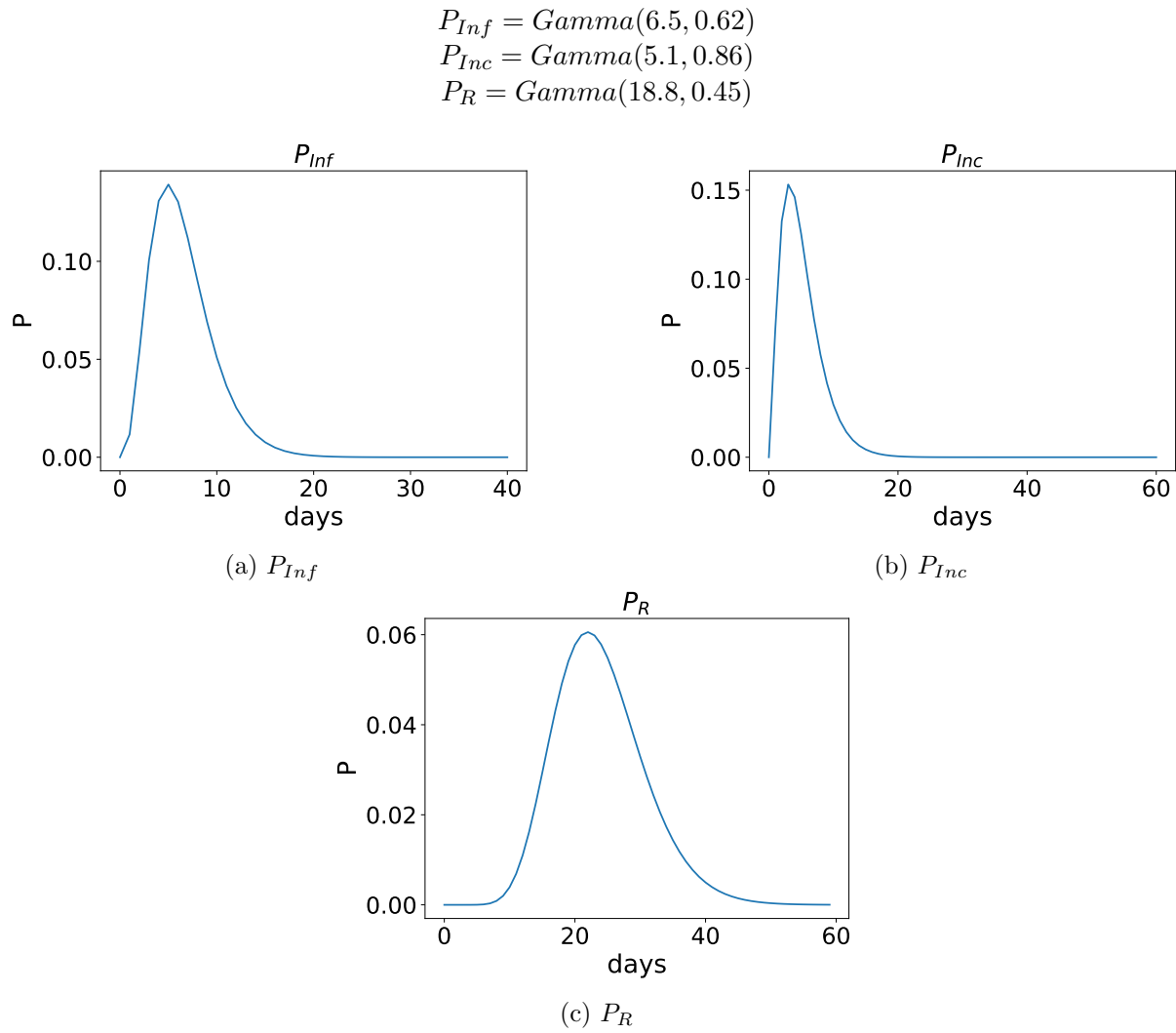


Figure 3: Modelling different R values

The probability of recovering $P_r(d)$ can be thought of as the sum of all possible combinations for n days then recovering $n-d$ days later. eg for $d=4$

$$P_r(d) = \sum_{n=0}^d P_{Inc}(n)P_R(d-n) \quad (12)$$

Days Incubating	Days till recovery after inc
4	0
3	1
2	2
1	3
0	4

By assuming that $P_{Inc}(d)$ and $P_R(d)$ are 0 outside the domain $d[0, N]$ for some N , it can be shown that $Pr(d)$ is 0 for $d < 0$.

Q3

Proof. Using 12

$$\begin{aligned}
 P_r(d) &= \sum_{n=0}^d P_{Inc}(n)P_R(d-n) \\
 d &< 0 \\
 d &\notin [0, N] \\
 P_r(d) = P_R(d) &= P_{Inc} = 0
 \end{aligned} \tag{13}$$

The largest value that d can be such that $P_r(d)$ is non-zero is $d = 2N$. □

Proof. $P_r(d) = 0$ when either: 1) $P_R(d-n) = 0$ or 2) $P_{Inc}(n) = 0$

1) Zero when $d-n > N$

2) Zero when $n > N$

So boundary for both is when $n=N$, $d > 2N$

$$\begin{aligned}
 P_r(2N) &= \sum_{n=0}^{2N} P_{Inc}(n)P_R(2N-n) \\
 &= P_{Inc}(N)P_R(N) \\
 &> 0
 \end{aligned} \tag{14}$$

Since functions $P_{Inc}(n)$ or $P_R(2N-n)$ are zero for all other values of n. We can also look at $d = 2N + 1$ to check not larger and non-zero.

$$\begin{aligned}
 P_r(2N+1) &= \sum_{n=0}^{2N+1} P_{Inc}(n)P_R(2N-n+1) \\
 &= \sum_{n=0}^N P_{Inc}(n)P_R(2N+1-n) + \sum_{n=N+1}^{2N+1} P_{Inc}(n)P_R(2N-n+1) \\
 &= 0 \text{ for all } n
 \end{aligned} \tag{15}$$

Therefore $2N$ is largest value d can be such that $P_r(d)$ is non-zero.

$P_r(d)$ can also be shown to itself be a probability distribution with $\sum_{d=0}^{\infty} P_r(d) = 1$ and $P_r(d) > 0$ for all d. □

Proof.

$$\begin{aligned}
 \sum_{d=0}^{\infty} P_r(d) &= \sum_{d=0}^{\infty} \sum_{n=0}^d P_{Inc}(n)P_R(d-n) \\
 &= \sum_{d=0}^{2N} \sum_{n=0}^d P_{Inc}(n)P_R(d-n) + \sum_{d=2N+1}^{\infty} \sum_{n=0}^d P_{Inc}(n)P_R(d-n) \\
 &= \sum_{d=0}^{2N} \sum_{n=0}^d P_{Inc}(n)P_R(d-n) \\
 &= \sum_{d=0}^{2N} P_R(d-n) \sum_{n=0}^d P_{Inc}(n) \\
 &= 1
 \end{aligned} \tag{16}$$

□

5 Results

Using the model we have just constructed we can again look at the UK and Brazil to estimate R but also I_0 , the assumed initial infections within our model. We can start with the values estimated by our simpler model earlier, $R=9.10$ for the UK and $R=1.81$ for Brazil. For our starting I_0 values we can use $I_0 = 1$ for the UK and $I_0 = 10000$ for Brazil. Then we can fine tune these starting values to more accurately fit our data.

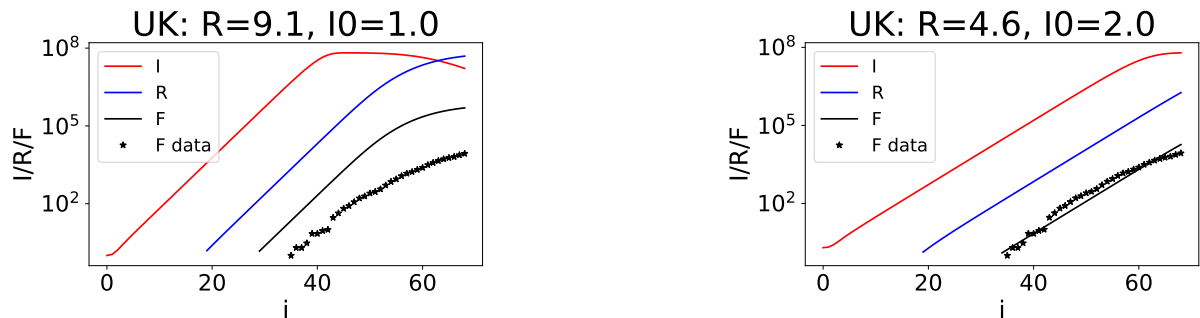


Figure 4: UK: Fine tune parameters to data

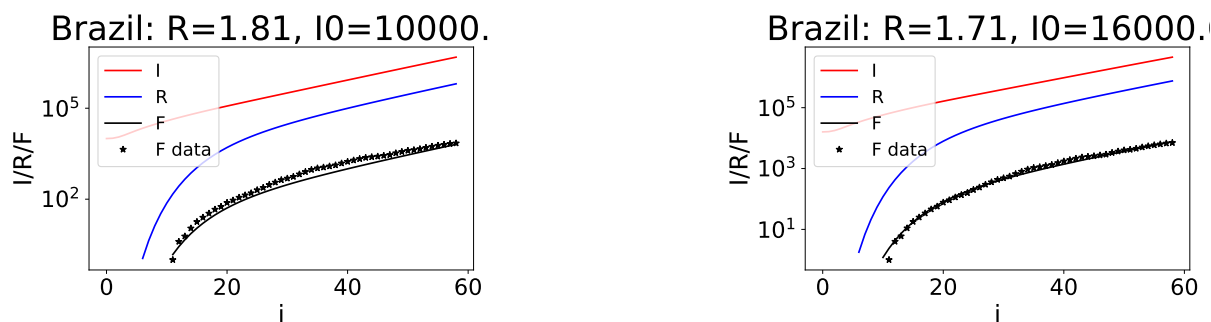


Figure 5: Brazil: Fine tune parameters to data

The profile of fatalities for the United Kingdom and Brazil are very different. On the logarithmic plot for the UK it is pretty much a straight line from the onset. This can be explained by the UK's implementations of various lockdowns and other curbing measures such as social distancing and mask wearing. As a result the straight line is produced suggesting the factors that such herd immunity that curb the exponential distribution never came into effect.

While for Brazil it is curving down owing to the lack of intervention meaning the virus eventually came to its natural conclusion for our model where either everyone has recovered or not.

6 Conclusion

In conclusion, our model can be seen to shown the differences and similarities for the Covid 19 pandemic in both the UK and Brazil. Each of our countries had extreme exponential growth in both cases and fatalities. During the early onset of the pandemic owing to the disease's high infection rate R values were high and greater than 1. However, after the initial period the evolution of the pandemic began to diverge between the countries with the UK continuing this exponential growth seen as a straight line on our log linear graphs. The major reason for this

Q5

Q6

would be the impact of lockdown rules disrupting the transmission of the disease. In comparison, in Brazil no counteracting measures were taken so the disease was left to come to its own conclusion whereby herd immunity had been achieved with enough people within the population having had the disease. This resulted in the fatality and cases curves to tail off and curve back down. As a result our model performed best at predicting the pandemic in Brazil since our recurrence relation could more accurately describe the evolution of the population overtime without unaccounted for events like lock downs occurring.

References

- [1] Wikipedia Covid-19 pandemic(2020) https://en.wikipedia.org/wiki/COVID-19_pandemic
- [2] Flaxman *et. al*, *Report 13: Estimating the number of infections and the impact of non-pharmaceutical interventions on COVID-19 in 11 European countries* <https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/covid-19/>
- [3] Stephen A. Lauer *et. al*, *The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application* Ann. Intern. Med. doi:10.7326/M20-0504