Covid Data Analysis

Emma Hughes 19335855

October 2021

1 Introduction

The aim of this exercise is to provide a numerical description of the number of Covid-19 cases in Ireland throughout the pandemic. It is assumed that the number of cases increase and decrease exponentially through time and the natural logarithm of cases per day is plotted against time to confirm this. This indeed is found to be the case, shown by the approximately linear increases and decreases of the data when plotted on a logarithmic scale. The logarithmic data is fitted linearly and the results of this fit are used to provide an exponential fit for the data on a linear scale.

2 Linear Regression

The exponentially rising part of a wave can be described as

$$n(t) = n_0 \exp(\lambda(t - t_0)) \tag{1}$$

where n(t) is the number of cases reported per day, t_0 is the day that marks the onset of the wave, n_0 and λ are treated as fit parameters that can be determined from a least squares fit. An increasing wave will have a positive λ and a decreasing wave will have a negative λ.

To linearly fit the data we take the natural logarithm of both sides of Equation (1) obtaining

$$\ln n(t) = \ln n_0 + \lambda(t - t_0) \tag{2}$$

In order to get an equation which we can use as a fitting function for the data we substitute $a = \ln n_0 - \lambda t_0$ and $b = \lambda$ into Equation (2), resulting in

$$\ln n(t) = a + bt \tag{3}$$

This equation is used as the fitting function in the curve fit function in the scipy library which preforms a least square fit of the logarithmic Covid data. The resulting fits are plotted on top of the data on a logarithmic scale as shown in Figure 1 and the fitting parameters obtained from the linear regression rounded to 3 significant figures are included in the table below.

Fitting Parameters	$\boldsymbol{n_0}$	$\boldsymbol{\lambda}$
First Wave Increase	1.19	0.223
First Wave Decrease	686	-0.0707
Second Wave Increase	9.07	0.0419
Second Wave Decrease	851	-0.0320
Third Wave Increase	208	0.123
Third Wave Decrease	3080	-0.0366
Fourth Wave Increase	291	0.0369

These results are as expected with positive decay constants for increasing waves and negative for decreasing. The initial number of cases at the start of each wave, n_0 increases as the pandemic goes on and the same pattern is observed in the peaks of the waves (n_0 for each wave decrease). This indicates the overall continuous rise of covid cases since the beginning of the pandemic.

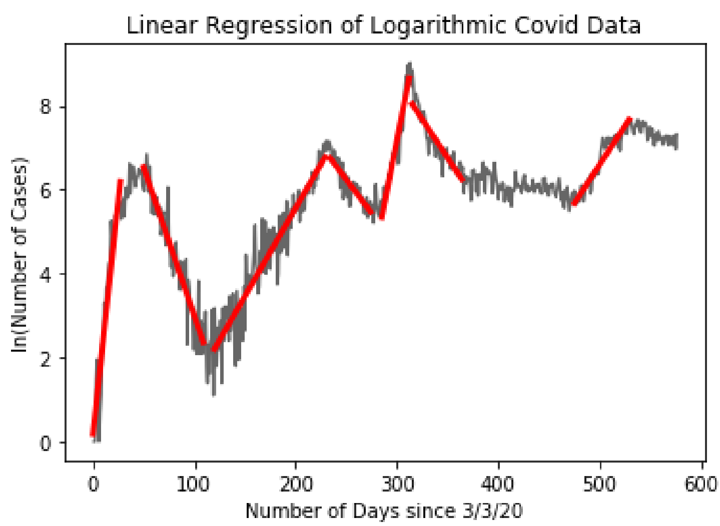

Figure 1: Data fitted linearly on a logarithmic scale

The values of n_0 and λ for the increase and decrease of each wave, along with the corresponding t_0 values are substituted in to Equation (1) and the resulting exponential functions for n(t) are plotted on the linear scale of the data as shown in Figure (2).

This fit worked reasonably well. At the base and throughout the body of the waves, the exponential fits are a very good approximation for the data. However at the fits tend to underestimate the height of the peaks. This is because in choosing the t_0 values I tended to underestimate the peak to ensure I did not include part of the decreasing wave in the increasing fit and vice versa.

I excluded the 4th wave decrease from the fit as in this data it has not reached the base of its wave form so it would be inaccurate to analyse the decrease as we

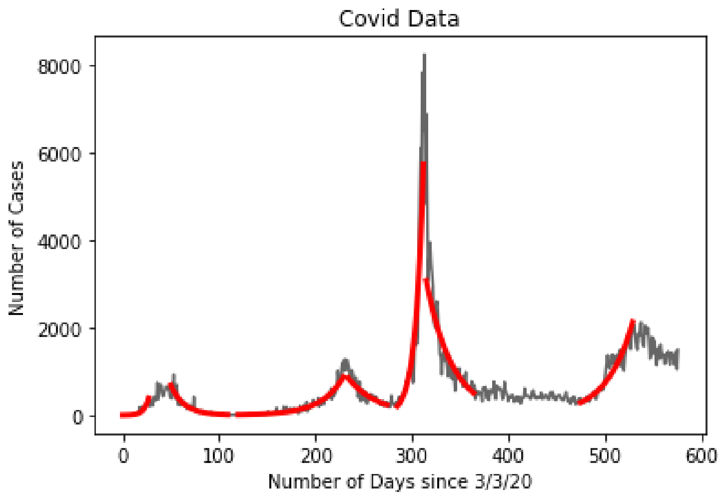

Figure 2: Data fitted exponentially on a linear scale

do not have all of the information here. I also excluded a section of very noisy data at the top of the first wave as it considerably skewed my fit so that the exponential fit did not lie on top of the data. This noise could perhaps indicate the lack of testing available at the beginning of the pandemic, with the highs in the noise indicating days when Ireland recieved a shipment of tests and lows when they had ran out.

The shape of the 4th wave looks different from the rest. The exponential curve fit does not work as well for it as it appears that the rate of increase of cases has slowed considerably. This is likely due to vaccinations decreasing the likelihood of catching the virus and hence decreasing the number of cases and rate of growth. Using the polynomial function $n(t) = At^2 + Bt + C$ to approximate the 4th wave increase we obtain a much better fit as shown in Figure 3.

The better fit of the polynomial function rather than the sign function for the 4th wave confirms that the rate of increase of Covid cases for the 4th wave has slowed due to vaccinations.

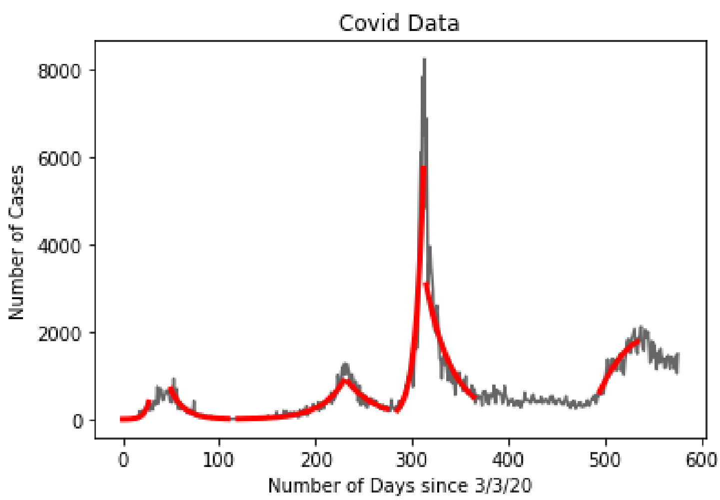

Figure 3: 4th wave fitted with a polynomial function