

Analysis of the COVID-19 Epidemic

This Analysis is Done Using Mathcad 14 Math Modeling Software

Our basic goal is to investigate and determine the properties of the COVID-19 Epidemic in the USA.

We will do an Analysis of the COVID-19 Epidemic for Four Countries: China, Italy, Spain, USA, and the World. We will plot the number of cases, find what models describe the outbreak, and then use them to predict the epidemic.

We will examine the Exponential Growth and the SIRD models.

We will develop techniques to abstract the Exponential and SIRD (α , β , γ , δ) model parameters.

COVID-19 is an exponential threat. Every day counts.

Table of Contents (Page Description)

- * Analysis of the COVID-19 Epidemic
- * General Description of the SARS-CoV-2 Virus and Conclusions
- * OUTLINE - Analysis of COVID-19 Epidemic
- * Susceptible, Infected, and Recovery Infectious Disease Epidemic Models
- * Mathematical modeling of infectious disease
- * Develop Discrete SIRD Model: Analyze Various Infection Outbreaks
- * SIRD Parameters for Typical Flu Season
- * SIR Model Normalized: Flu Season 2002-2007 and 2010-2013
- * COVID-19 SIRD Data, Model & Predictions Hubei Province, China
- * Methodology to Extract SIRD Parameters for Hubei
- * Extracting R_0 from Log Plots of Hubei Infected
- * USA Data Directly From Johns Hopkins
- * Methodology to Extract SIRD Parameters for USA
- * CDC Data on Deaths in USA
- * Assessment of Efficacy of US Government Intervention to Halt Spread
- * Italy COVID19 Cases Feb 15 - March 14
- * Spain COVID19 Cases Feb 1 - March 14
- * World COVID19 Cases Jan 21 - March 12
- * COVID-19 Epidemiology
- * Coronavirus: Why You Must Act Now - Stats

Tom Kotowski

VXPhysics.com

March 20, 2020

General Description of Virus and Conclusions

Some Characteristics of the COVID-19 Disease

The SARS-CoV-2 Virus is commonly referred to by the name of the disease, which is COVID-19. The later name is used for PR purposes because it was thought that the name of the virus SARS (Severe Acute Respiratory Syndrome caused by the corona virus) would cause some people to panic. This virus was discovered by observing that hospital patients were showing a very virulent type of pneumonia. Currently, according to the CDC, the incubation period for the novel coronavirus is somewhere between **2 to 14 days after exposure**. More than 97 percent of people who contract SARS-CoV-2 show symptoms within 11.5 days of exposure. The **average incubation period seems to be around 5 days**. For many people, COVID-19 symptoms start as mild symptoms and gradually get worse over a few days. **Transmission occurs primarily:** via respiratory droplets from coughs and sneezes within a range of about 6 feet (1.8 m).[23][24] Indirect contact via contaminated surfaces is another possible cause of infection.[25] Preliminary research indicates that the virus may remain **viable on plastic and steel** for up to **three days**, but does not survive on **cardboard for more than one day** or on copper for more than four hours.

Communicating best guesses on impact of COVID19 in US healthcare system

2020 Webinar of the American Hospital Association

USA Population: 350 million

- * $R_0 = 2.5$; Doubling time 7-10 days
- * Community attack rate = 30-40%
- * Cases requiring hospitalization = 5%
- * Cases requiring ICU care = 1-2%
- * Cases requiring ventilatory support = 1%
- * CFR = 0.5%

Estimate 27% of Population will be Infected

Community epi wave 2 months

US: 96 million cases (27% Population)

US: 4.8 million admissions

US: 1.9 million ICU

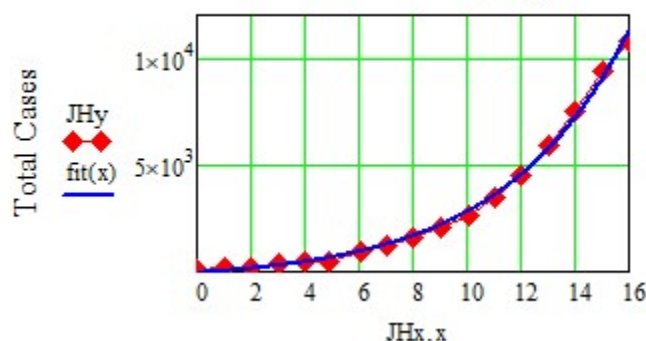
US: 1 PPV

US: 480,000 deaths

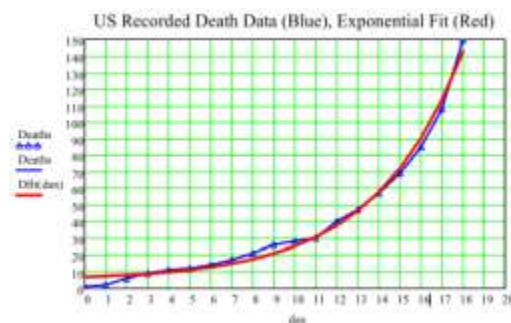
By applying our SIRD model to the Wuhan Virus, we found that the days to double was 3.62, time to recover was 41 days, and the Basic Reproductive Ratio, R_0 , was 2.74. See two pages after this to see the definitions of these parameters. Analysis of the USA data revealed that the number of days for infections to double was 2.309 and the Basic Reproductive Ratio is 2.74.

Exponential curve, see below is an excellent match to Both Total Number of Infected & Deaths vs. days.

USA- Total Cases, Exp Fit



Initial Total US Deaths, at the Present, is an Exponential Curve
It is Doubling every 4 days



With the Assumption that the Initial Rise of Epidemic can be Simply Modeled by an Exponential,
Below is a Rough Projection of Number Confirmed COVID-19 Cases vs Dates to One Million Cases

3/20/2020	11220	3/25/2020	33325	3/30/2020	97290	4/4/2020	282386	4/24/2020	818000
3/21/2020	13984	3/26/2020	41323	3/31/2020	120433	4/20/2020	349355	4/25/2020	1011790
3/22/2020	17403	3/27/2020	51214	4/1/2020	149056	4/21/2020	432181		
3/23/2020	21630	3/28/2020	63448	4/2/2020	184456	4/22/2020	534619		
3/24/2020	26859	3/29/2020	78578	4/3/2020	228238	4/23/2020	661311		

OUTLINE - Analysis of COVID-19 Epidemic

- Obtain Outbreak Data from China, Italy, Spain, USA, and the World
- Plot the Outbreak Data from January to March 17.
- Use the Exponential and SIRD Models to Model the Outbreaks in these Four Countries
- Begin with the System of Nonlinear Differential Equations for the SIRD Model
- Develop a Discrete Algorithm to Implement the SIRD Model Differential Equations
- For Reference, Apply the SIRD Model with Given Parameters for Typical US Flu Season
- Then apply Exp and SIRD Models to Existing Parameters for Countries up to March 15
- Develop Strategies to Extract Exponential and SIRD Parameters for the Four Countries
- Use the Levenberg-Marquardt Method: Minimize Residual Least Squares Error
- Develop a Strategy to Extract the α , β , γ Model Parameters: Use Levenberg-Marquardt
- Model China, does the SIRD Model with Abstracted α , β , γ Match the Data? No!
- Must Develop a New Strategy to Successfully Model the Data
- Instead of Just the Residual of the Infected, Combine Infected & Recovered Vectors
- Model Governmental Intervention with a Parameter δ Modification from some Time, T
- Revised Strategy: Combined Infected & Recovered, Extract α , β , and also δ Parameters
- Does Combined Vectors α , β , & δ SIRD Model Match the Data? Answer: **Yes.**
- Extracted Parameters: Determine Days for Doubling, Recover, & Reproductive Ratio
- Estimate R_0 from the Slope of the Linear Semi-Log Curve to Log(Infected) vs Time
- CDC has Limited Data for USA Cases - Jan 23 to March 17 - Plot Data
- Examine the number of deaths from COVID-19 in USA. Shows Exponential Rise.
- Show Data for Italy: Fit Exponential Function for Initial Phase of Epidemic
- Show Data for Spain: Fit Exponential Function for Initial Phase of Epidemic
- Data Represents Only the Initial Phase of the Outbreak in USA - Future?
- Use Data to Develop SIRD Model for other Countries. Find # Days to Double.

Susceptible, Infected, and Recovery Infectious Disease Epidemic Models

Data-Based Analysis, Modeling and Forecasting of the COVID-19 outbreak
<https://www.medrxiv.org/content/10.1101/2020.02.11.20022186v4.full.pdf> - March 5, 2020

The oldest and most common **Epidemiological Model** is the **SIRD model**, consisting of a set of four **differential equations**, which assigns every person in a population to be in one of 4 conditions or categories:

S = Susceptible to becoming infected. S_0 , Initial population (initial # of people who are susceptible),

I = Infected through contact with someone already infected. I_0 , Initial number of infected people

R = Recovered or Removal Group, either no longer sick/infected or dead.

D = Dead, no longer susceptible

Through time a person may **move from being susceptible to infected to recovered or dead**, so that the number of people in each category changes, but the total of $S + I + R + D$ remains some constant, N .

This is a compartmental model, with S, I, R, and D being compartments. Every person starts off in a given compartment and may then, in time, move to another. Graphically the compartment model looks like the plots below with the rates of movement between the compartments designated by the parameters: α , β , and γ .

In this model once someone recovers they are immune and can't be infected again. The model also assumes that a disease is passed from person to person. The SIRD model can't be used for diseases that spread by other modalities, such as eating exotic animals or being bitten by insects.

Epidemiological Parameters

Infection rate, α : Transmission rate, rate (number per day) that susceptible people become infected

Recovery rate, β : Recovery rate (number per day) that infected people recover. Also a Removal Rate

Death rate, γ : Fatality Rate from virus. Globally 3.7%, but varies with age.

CDC - **Flu in USA:** 36 million flu illnesses, 370,000 hospitalizations & 22,000 deaths.

Step Decrease in α , δ : Government Intervention Reduces Infection rate, α , by factor δ , $T = 21$ days later.

Time increment, T : Delay time until Governmental Intervention

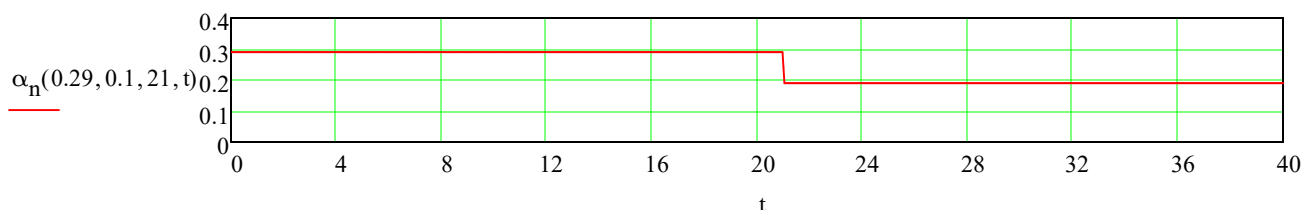
On March 14, 2020 President Trump decreed that starting on entry from Europe will be stopped for 30 days. Assume that α decreases by an increment of 0.1, 30 days from the beginning of this plot

The Infection rate α is now a step function $\Phi(t - T)$ of the initial value of α , with delay time, T , and decrease by a decrement, δ . **The decrement, δ , can have a huge effect.**

Government Intervention Model: Modified Infection Rate as a result:

$$\alpha_n(\alpha, \delta, T, t) := \alpha - \Phi(t - T) \cdot \delta$$

Decreased Infection rate, alpha, by decrement 0.1, delayed by 21 days



Discrete SIRD Compartmental Model - Epidemic Outbreak Equations

$$S(t) = S(t-1) - \frac{\alpha}{N} S(t-1) I(t-1) \quad \frac{S}{N} \text{ Fraction of Population}$$

$$I(t) = I(t-1) + \frac{\alpha}{N} S(t-1) I(t-1) - \beta I(t-1) - \gamma I(t-1)$$

$$R(t) = R(t-1) + \beta I(t-1)$$

$$D(t) = D(t-1) + \gamma I(t-1)$$

Mathematical modeling of infectious disease

https://en.wikipedia.org/wiki/Mathematical_modeling_of_infectious_disease

Compartmental SIRD Models in Epidemiology

Portions of this Page are a Description from Wikipedia:

Compartmental models are a technique used to simplify the mathematical modeling of infectious disease. The population is divided into compartments, with the assumption that every individual in the same compartment has the same characteristics. Its origin is in the early 20th century, with an important early work being that of Kermack and McKendrick in 1927.

Bio-mathematical deterministic treatment of the SIR model

The dynamics of an epidemic, for example the flu, are often much faster than the dynamics of birth and death, therefore, birth and death are often omitted in simple compartmental models. The SIR system without so-called vital dynamics (birth and death, sometimes called demography) described above can be expressed by the following set of ordinary differential equations:

This system of ordinary differential equations is **non-linear**, and **does not admit a generic analytic solution**. Nevertheless, significant results can be derived analytically, and with Monte Carlo methods, such as the Gillespie algorithm.

$$\frac{dS}{dt} = -\frac{\beta IS}{N},$$

$$\frac{dI}{dt} = \frac{\beta IS}{N} - \gamma I,$$

$$\frac{dR}{dt} = \gamma I.$$

- The terms dS/dt , dI/dt , dR/dt in the differential equations indicate the rate of change of the susceptible population size, the infected population size and the recovered population size, respectively.
- The terms β and α indicate the attack rate (number of susceptible persons get infected per day) and the recovery rate of the flu (inverse of the number of days a person remains infected), respectively.
- High value of α means a person will be infected by the flu for less number of days and high value of β means that the epidemic will spread quickly.
- Also, as can be seen from below, from the differential equations it can be shown that the population ($S + I + R$) is assumed to be constant.

Note: Authors may assign different meanings to the symbols α , β , γ , ρ .

Susceptible, Infected Recovered: The SIR Model of an Epidemic



Discrete SIRD Model - Infection Outbreak Equations

The number of people at any day, n, who are:

<u># Susceptible-S</u>	<u>Number (#) Infected - I</u>	<u># Recovered - R</u>	<u># Dead - D</u>
$S_t = S_{t-1} - \frac{\alpha}{N} \cdot S_{t-1} \cdot I_{t-1}$	$I_t = I_{t-1} + \frac{\alpha}{N} S_{t-1} \cdot I_{t-1} - \beta \cdot I_{t-1} - \gamma \cdot I_{t-1}$	$R_t = R_{t-1} + \beta \cdot I_{t-1}$	$D_t = D_{t-1} + \gamma \cdot I_{t-1}$

TWK's Algorithm to Calculate Values of S, I, R, D from Parameters

Set the delay time, T, for Intervention, δ , to 21 days

$$\text{SIRD}(S_0, I_0, \alpha, \beta, \gamma, \delta, N) := \begin{array}{l} S_0 \leftarrow S_0 \\ I_0 \leftarrow I_0 \\ D_0 \leftarrow R_0 \leftarrow 0 \\ (M_{0,0} \ M_{0,1} \ M_{0,2} \ M_{0,3}) \leftarrow (S_0 \ I_0 \ 0 \ 0) \\ \text{for } n \in 0..N \\ \quad \begin{array}{l} S_{n+1} \leftarrow S_n - \frac{S_n}{S_0} \cdot \alpha_n(\alpha, \delta, 21, n) \cdot I_n \\ I_{n+1} \leftarrow I_n + \frac{S_n}{S_0} \cdot \alpha_n(\alpha, \delta, 21, n) \cdot I_n - \beta \cdot I_n - \gamma \cdot I_n \\ R_{n+1} \leftarrow R_n + I_n \cdot \beta \\ D_{n+1} \leftarrow D_n + I_n \cdot \gamma \\ M_{n+1,0} \leftarrow S_{n+1} \\ M_{n+1,1} \leftarrow I_{n+1} \\ M_{n+1,2} \leftarrow R_{n+1} \\ M_{n+1,3} \leftarrow D_{n+1} \end{array} \end{array}$$

M

Let's look at the example of a typical flu season virus outbreak, $\alpha, \beta, \gamma, R_0$

Typical Flu Outbreak Parameters: $\alpha := 0.2$ $\beta := 0.14$ $\gamma := 0.01$ $R_0 := \frac{\alpha}{\beta + \gamma}$ $R_0 = 1.333$

A typical flu season lasts about 35 weeks (245 days). Week 1 generally is the first week of October and lasts until week 35 or the end of May. See plots of a typical flu season on the following page.

The basic reproductive ratio, R_0 is defined by epidemiologists as " R_0 " represents the average number of secondary cases that result from the introduction of a single infectious case in a totally susceptible population during the infectiousness period".

As such R_0 tells us about the initial rate of increase of the disease over a generation.

Note: Some models use proportions, rather than absolute numbers. Then S_0 is 0.9, $I_0 = 0.002$, $R_0 = 0.1$ and parameters $\alpha = 2$ and $\beta = 1.4$. For a flu season, S_{final} is about 0.52 and Infected peaks at 0.024.

If $S_0 < \beta/\alpha$ then Infections decrease monotonically to zero - this designates the season as nonepidemic.

The number of days it takes infections to double, t_{2x}	$t_{2x} := \frac{1}{\alpha}$	$t_{2x} = 5$
The number of days it takes to recover, t_{rec}	$t_{\text{rec}} := \frac{1}{\beta}$	$t_{\text{rec}} = 7.143$

The above shows that the number of infections quickly outnumber the number of recoveries

SIRD Parameters for Typical Flu Virus

$$\text{SIRD}(S_0, I_0, \alpha, \beta, \gamma, \delta, N) \quad \text{FluSIRD} := \text{SIRD}(999, 1, 0.29, 0.15, 0.01, 0.0001, 130)$$

$$\begin{array}{llll} \text{Susceptible - Sus.} & \text{Infected - Inf.} & \text{Recovered - Rec} & \text{Dead - D} \\ \text{Sus} := \text{FluSIRD}^{(0)} & \text{Inf} := \text{FluSIRD}^{(1)} & \text{Rec} := \text{FluSIRD}^{(2)} & \text{Ded} := \text{FluSIRD}^{(3)} \end{array}$$

The Sum of Each Row of the 4 Components = S0

$$j := 1..120 \quad \text{Total}_0 := 1000 \quad \text{Total}_j := \text{Sus}_j + \text{Inf}_j + \text{Rec}_j + \text{Ded}_j \quad \text{mode}(\text{Total}) = 1000$$

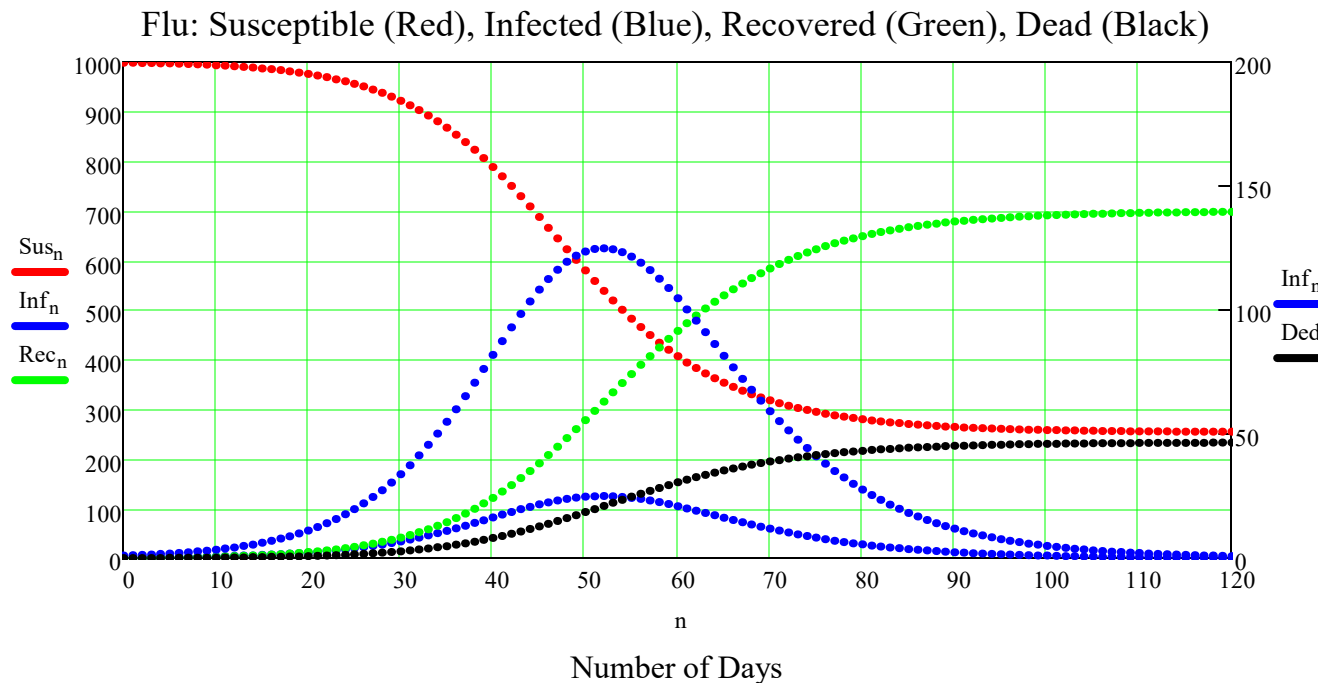
Below is a model for a **typical flu virus**. Because of **mutation, new strains** of influenza make **most people susceptible** ($S_n, \alpha = 29\%$) at the beginning of an outbreak. Interestingly, it shows that the number of infected people (blue curves) has reached a **peak after about 54 days (7 1/2 weeks)** and then falls to after 100-120 days (3 - 4 months).

This demonstrates that the SIRD model is a good representation for a flu season.

Note that there two different sized scales, the scale shown at the left is larger than the one at the right.

The blue curve for infected is shown on both scales, the top curve is used with the scale on the right.

$$\text{The black curve for the } R_{0_Flu} := \frac{\alpha - \delta}{\beta + \gamma} \text{ s use: } R_{0_Flu} := \frac{0.29 - 0.0001}{0.15 + 0.01} \text{ nly. } R_{0_Flu} = 1.812$$



**From Flu: Notice the number of infections drops off with the flu season.
The number of those that are susceptible drops below new infections**

SIR Model Normalized: Flu Season 2002-2007 and 2010-2013

Parameters and graphs are from the paper: "Forecasting seasonal influenza with a state-space SIR model", 2017

$$\text{SIRD}(S_0, I_0, \alpha, \beta, \gamma, \delta, N) \quad \text{FluSIRD} := \text{SIRD}(0.9, 0.0002, 0.2, 0.14, 0.01, 0.0001, 240)$$

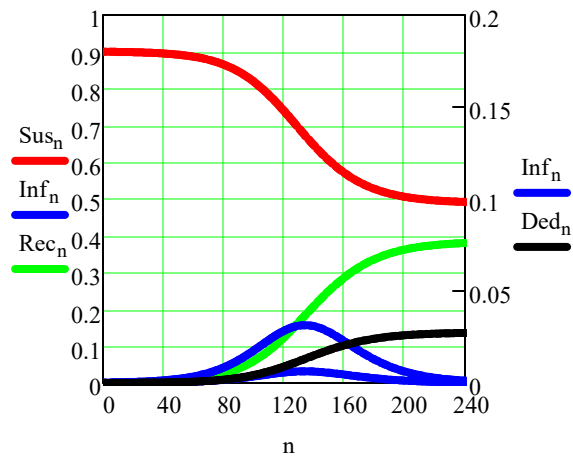
$$\begin{array}{llll} \text{Susceptible-Sus.} & \text{Infected - Inf.} & \text{Recovered - Rec} & \text{Dead - D} \\ \text{Sus} := \text{FluSIRD}^{\langle 0 \rangle} & \text{Inf} := \text{FluSIRD}^{\langle 1 \rangle} & \text{Rec} := \text{FluSIRD}^{\langle 2 \rangle} & \text{Ded} := \text{FluSIRD}^{\langle 3 \rangle} \end{array}$$

The Sum of Each Row of the 4 Components = S0

$$j := 1..120 \quad \text{Total}_0 := 1 \quad \text{Total}_j := \text{Sus}_j + \text{Inf}_j + \text{Rec}_j + \text{Ded}_j \quad \text{mode}(\text{Total}) = 0.9$$

Sus (Red), Infected (Blue), Recovered (Green), Dead (Black)

Flu: Susceptible



Number of Days Season: 35(Wks) x 7 = 245 days

Plot of Fraction of Infections/S0 for
Flu seasons 2002-2007 and 2010-2013

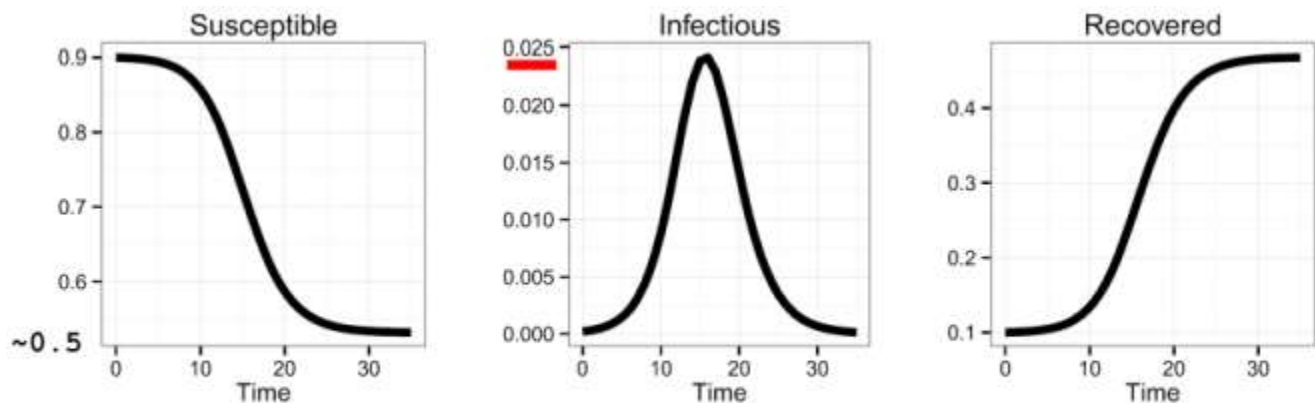
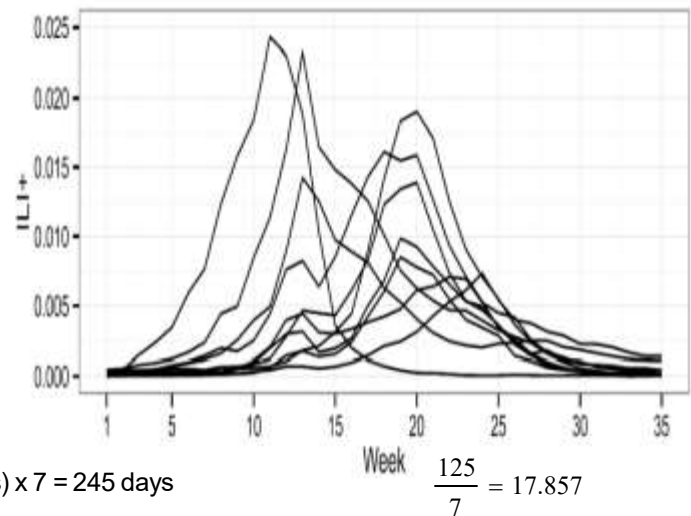


Fig. 2.

Simulated SIR curve with $S_0 = 0.9$, $I_0 = 0.0002$, $R_0 = 0.0998$, $\beta = 2$, and $\gamma = 1.4$.

COVID-19 SIRD Data, Model & Predictions Hubei Province, China

An outbreak of “pneumonia of unknown etiology” in Wuhan, Hubei Province, China in early December 2019 has spiraled into an epidemic. The virus is **SARS-CoV-2**, a coronavirus.

Simulations until the 29th of February of the cumulative number of recovered as obtained using the SIRD model. Dots correspond to the number of confirmed cases from the 16th of January to the 10th of February. The initial date of the simulations was the 16th of November with one infected, zero recovered and zero deaths. Solid lines correspond to the dynamics obtained using the estimated expected values of the epidemiological parameters

$$S_0 = 59 \times 10^6, \alpha = 0.193, \beta = 0.063/\text{day}, \gamma = 0.01;$$

Data from Hubei Province China from 1-22-2020 to 3-12-2020 - Population 58.5 Million

Hubei := READPRN("K-COVID19-Hubei-Complete.txt") HPop := $58.5 \cdot 10^6$ rows(Hubei) = 51

Data from Kaggle <https://www.kaggle.com/kimjihoo/coronavirusdataset>

COVID Data: Infected Dead Recovered
 HInf := Hubei^{<3>} HDed := Hubei^{<4>} HRec := Hubei^{<5>} **HInf₅₀ = 67781**

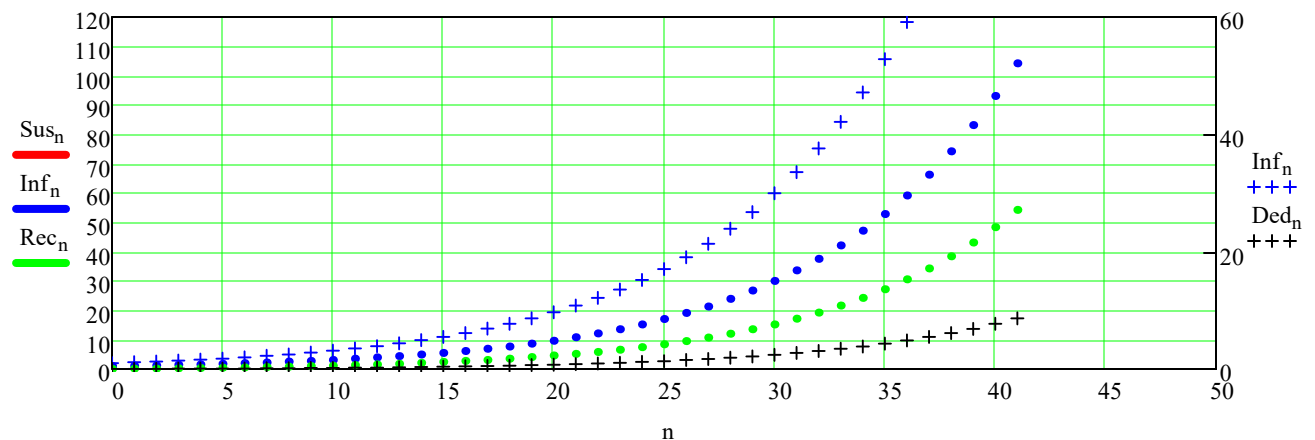
Calculate SIRD Model Values for Wuhan of COVID-19 to Feb 10, 2020

SIRD($S_0, I_0, \alpha, \beta, \gamma, \delta, N$) W19 := SIRD($59 \cdot 10^6, 1, 0.193, 0.063, 0.01, 0, 40$)

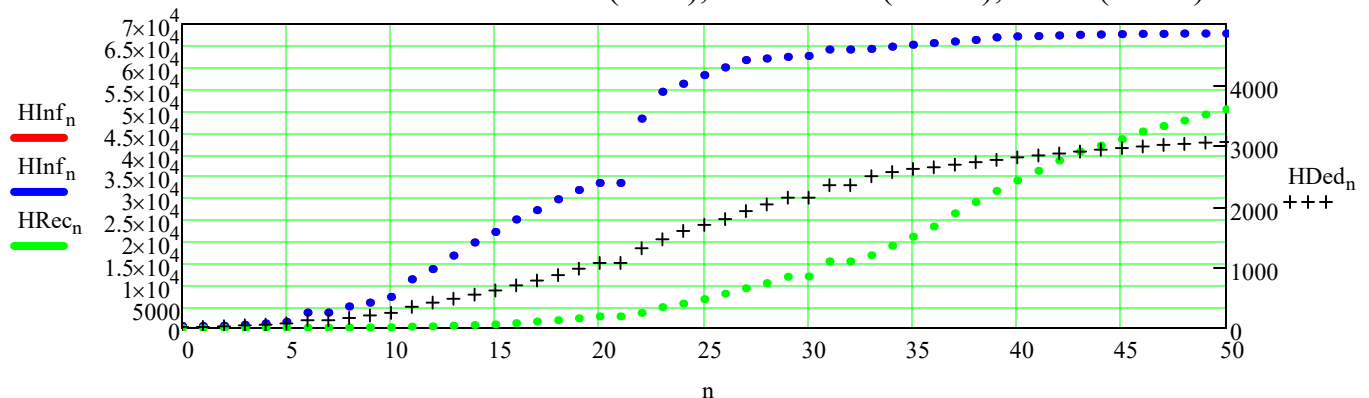
Model Results: Susceptible-Sus. Infected - Inf. Recovered - Rec Dead - D
 Sus := W19^{<0>} Inf := W19^{<1>} Rec := W19^{<2>} Ded := W19^{<3>}

Inf₄₀ = 93.051 $\alpha := 0.193$ $\beta := 0.063$ $\delta := 0.01$ $R_0 := \frac{\alpha}{\beta + \gamma}$ **R₀ = 2.644**

SIR Model: Sus (Red), Infected (Blue), Recovered (Green), Dead (Black)



Hubei COVID-19: Infected (Blue), Recovered (Green), Dead (Black)



Number of Days

Methodology to Extract SIRD Parameters for Hubei

Our methodology is similar to most papers, the major difference being we use Mathcad's error minimization function Minerr, and not MATLAB's fmin() error minimization tool.

SIR Model

$$\text{Model}(\alpha, \beta, \delta) := \text{SIR}(59 \cdot 10^6, 444, \alpha, \beta, \delta, 49)$$

Initial Guess (SIRD Params):

$$(\alpha, \beta, \delta) := (0.3, 0.063, 0.1)$$

Use Quasi-Newton Method: Minimize Least Squares Error

$$\text{LSQFit} = \sqrt{\sum (H_{\text{res}} - \text{Mod}(\alpha, \beta, \delta))^2}$$

Use Levenberg-Marquardt Method: Minimize Least Squares Error to Residual

Define a Residual to be the difference between the current data points HInf, HRec, HDed and SIRDModel

$$H_{\text{res}} := \text{stack}(H_{\text{Inf}}, H_{\text{Rec}}) \quad \text{Mod}(\alpha, \beta, \delta) := \text{stack}(\text{Model}(\alpha, \beta, \delta)^{\langle 1 \rangle}, \text{Model}(\alpha, \beta, \delta)^{\langle 2 \rangle})$$

$$\text{Residual}(\alpha, \beta, \delta) := H_{\text{res}} - \text{Mod}(\alpha, \beta, \delta) \quad R := \text{rows}(H_{\text{Inf}}) \quad R = 51$$

Condition to Minimize the Residual Least Squares Fit Error Using L-V Minerr Method

$$\text{Given} \quad 0 = \text{Residual}(\alpha, \beta, \delta) \quad 0.1 \leq \alpha \leq 0.4 \quad 0.01 \leq \beta \leq 0.1 \quad 0.05 \leq \delta \leq 0.25$$

Use Minerr to Extract Optimal Parameters:

$$(a, b, d) := \text{Minerr}(\alpha, \beta, \delta)^T$$

$$(a, b, d) = (0.276, 0.024, 0.243)$$

$$\text{Infected} \\ H_{\text{Inf}} := \text{Model}(a, b, d)^{\langle 1 \rangle}$$

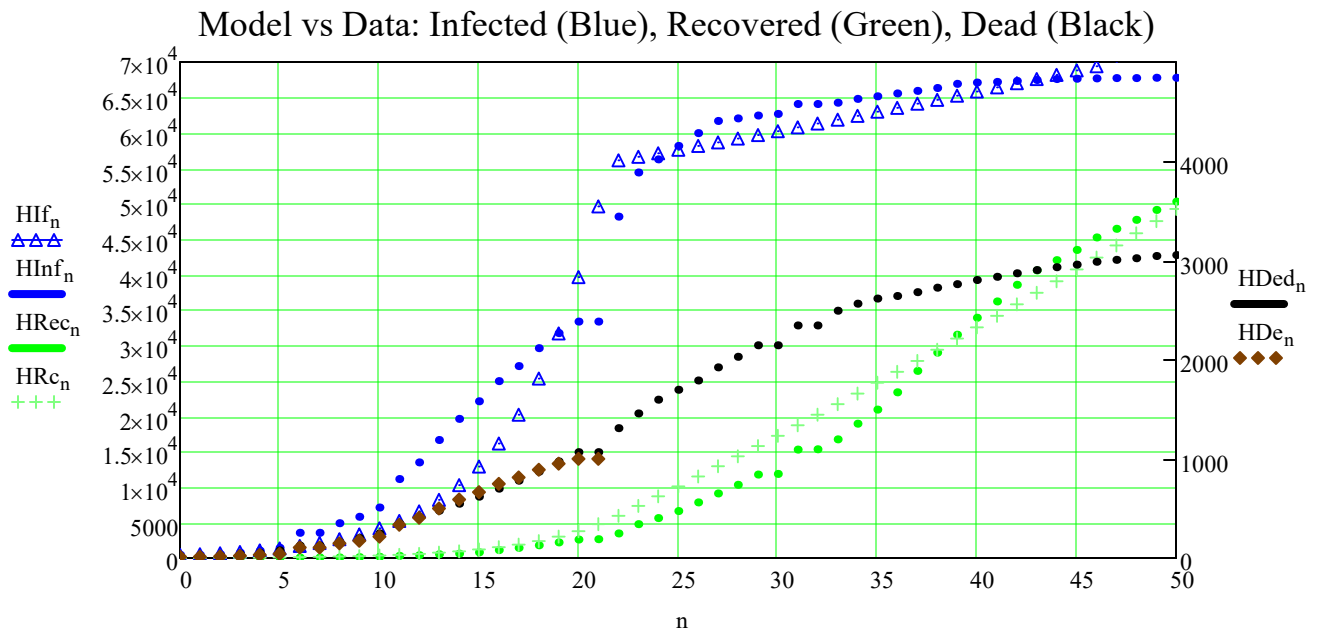
$$\text{Recovered} \\ H_{\text{Rec}} := \text{Model}(a, b, d)^{\langle 2 \rangle}$$

$$\text{Dead} \\ HD_{\text{d}} := H_{\text{Inf}} \cdot 0.03$$

$$\frac{\text{ERR}}{R} = 706.204$$

$$H_{\text{De}} := \text{submatrix}(HD_{\text{d}}, 0, 21, 0, 0)$$

Reasonably Good Match of SIRD Model (Symbols) to Data (Dots)



Extracted Parameters of Hubei/Wuhan COVID-19 Epidemic

The number of days it takes infections to double, t_{2x}

$$t_{2x} := \frac{1}{a}$$

$$t_{2x} = 3.62$$

The number of days it takes to recover, t_{rec}

$$t_{\text{rec}} := \frac{1}{b}$$

$$t_{\text{rec}} = 41.008$$

Basic Reproductive Ratio
before and after Intervention:

$$R_{0\delta} := \frac{a}{b + \gamma}$$

$$R_{0\delta} := \frac{\alpha - \delta}{\beta + \gamma}$$

$$R_0 = 8.035$$

$$R_{0\delta} = 2.74$$

Extracting R0 from Log Plots of Hubei Infected

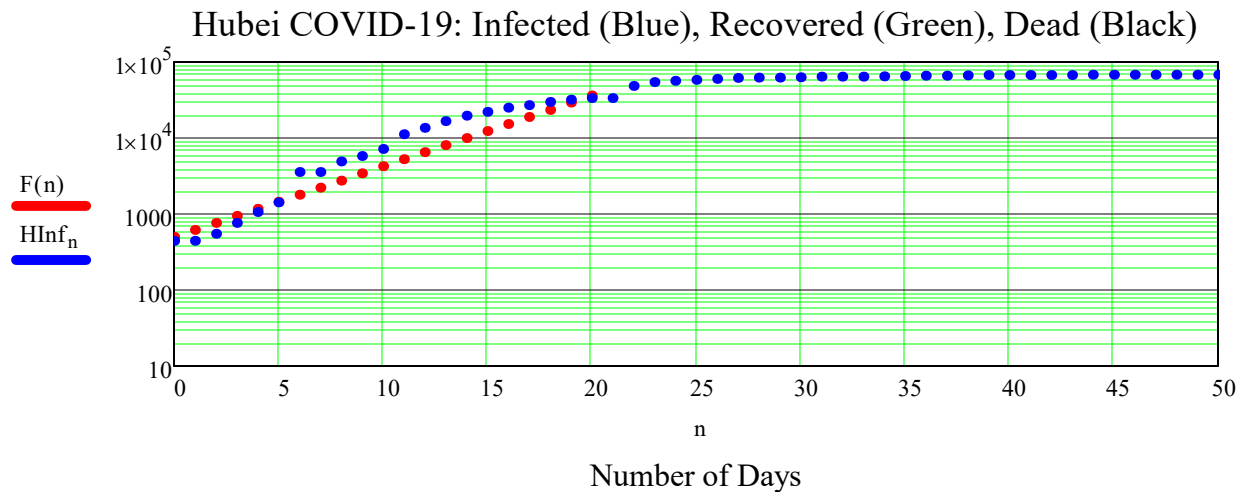
$$\text{Inf}(t) = \text{Inf}_0 \cdot e^{(R-1) \cdot \left(\frac{t}{\text{Infectious_Period}} \right)}$$

$$\log(\text{Inf}(t)) = \log(I_0) + (R-1) \cdot \frac{t}{\text{Infectious_Period}}$$

$$\text{HInf20} := \text{submatrix}(\text{HInf}, 0, 20, 0, 0)$$

$$k := 0..20 \quad X_k := k + 1 \quad F(k) := 400 \exp \left[(4-1) \cdot \frac{X_k}{14} \right]$$

Currently, according to the CDC, the incubation period for the novel coronavirus is somewhere between **2 to 14 days after exposure**. More than 97 percent of people who contract SARS-CoV-2 show symptoms within 11.5 days of exposure. The **average incubation period seems to be around 5 days**. For many people, COVID-19 symptoms start as mild symptoms and gradually get worse over a few days.



Find the Slope to the Log Plot of Infected to Get an Estimate of R0

$$m := 0..15 \quad \text{HIm}_m := \text{HInf}_m \quad \text{daz}_m := m \quad \text{Assume Infectious Period is 14 days}$$

$$\text{LM} := \log(\text{HIm}) \quad \text{LN} := \log(\text{HInf}) \quad y(x) = s \cdot x + \text{int} \quad F(x) := (x-1)^T$$

$$\text{linfit}(\text{daz}, \text{LM}, F) = \begin{pmatrix} 0.126 \\ 2.587 \end{pmatrix} \quad \text{slope}(\text{daz}, \text{LM}) = 0.126 \quad \text{slope} = \frac{R-1}{\text{Infectious_Period}}$$

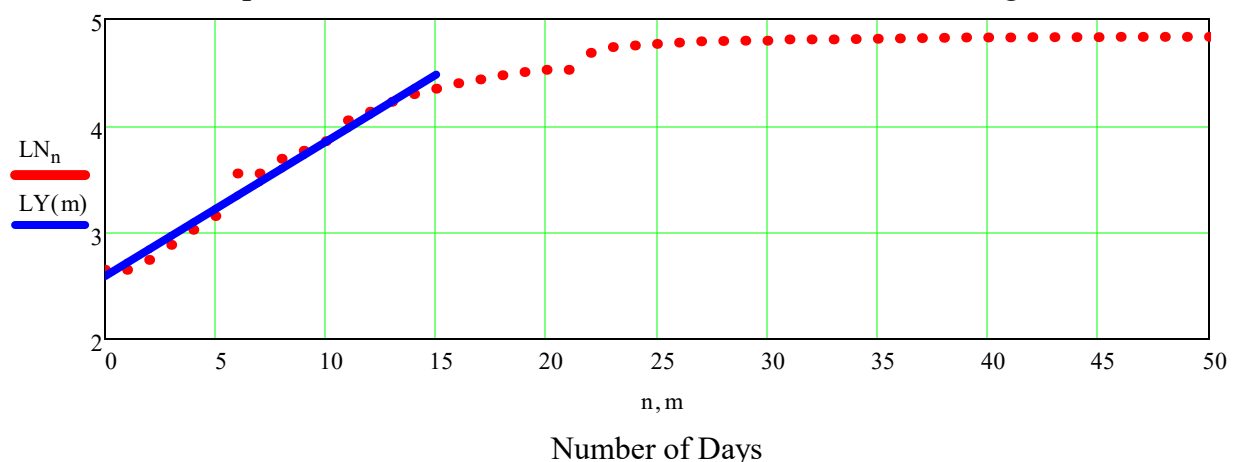
Linear Fit, LY, to Log of the Curve of Initial Infected:

$$\text{LY}(m) := 0.126 \cdot m + 2.587$$

$$\text{R0L}(\text{slope}, \text{Infectious_Period}) := \text{slope} \cdot (\text{Infectious_Period}) + 1$$

$$\text{R0L}(0.126, 14) = 2.764$$

Hubei Epidemic: Estimation of R0 from Linear Fit to the Log of Infected



USA Data Directly From Johns Hopkins



Vox Analysis of Center for Systems Science and Engineering - Johns Hopkins University Data

<https://www.vox.com/policy-and-politics/2020/3/13/21178289/confirmed-coronavirus-cases-us-countries-italy-iran-singapore-hong-kong>

Data abstracted from the plot in the above website using PlotDigitizer 2.6.8 on Log-Linear Scales

Read Data File: JH_USA := READPRN("JHU USA Vox COVID -100 0 Day March 20.PRN")

JHx := JH_USA⁽⁰⁾ JHy := JH_USA⁽¹⁾ Cases := JHy rows(JHy) = 17

Rc := rows(Cases) Rc = 17 i := 1..Rc - 1 days_i := i rows(Cases) = 17 rows(days) = 17

Calculate the Rate of Growth of Cases, Find Average Rate, and Days to Double from Avg. Rate

rate_i := 100 * $\frac{\text{New_Cases}_i}{\text{Cases}_{i-1}}$ Cases_{Rc-1} = 10760 rate := rate + 1 rate_{avg} := mean(rate) = 33.985

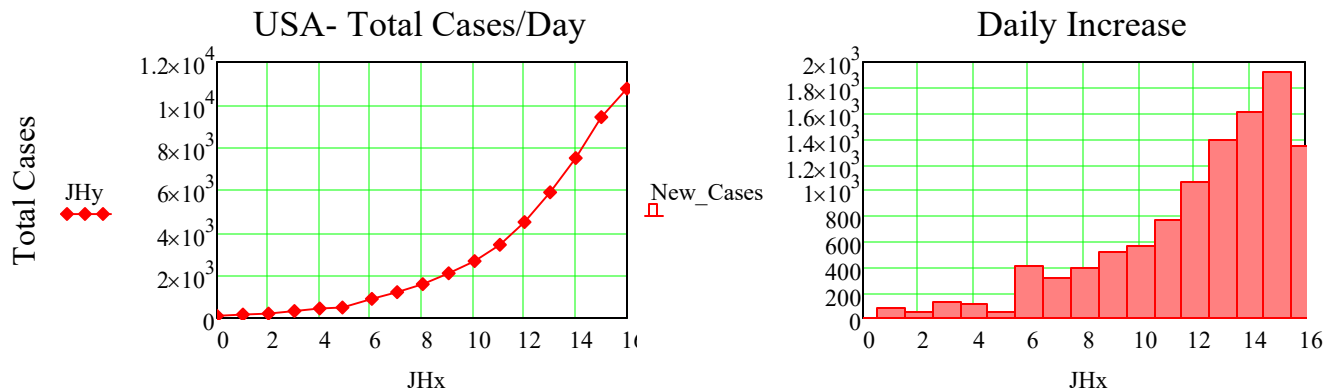
Calculate the Number of Days for Cases to Double - Dbl:

$$\text{Dbl} := \frac{\ln(2)}{\ln\left(1 + \frac{\text{rate}_{\text{avg}}}{100}\right)}$$

Days to Double: Dbl = 2.369

Dbl_Days(10755, 100, 16, 0) = 2.371

Exponential Growth --> US Epidemic



Guess for Parameters: a := 1 b := 0.01 c := 1 guess := (a b c)^T

Calculate Parameters for Exponential Fit, fit(x), to Total Number of Cases

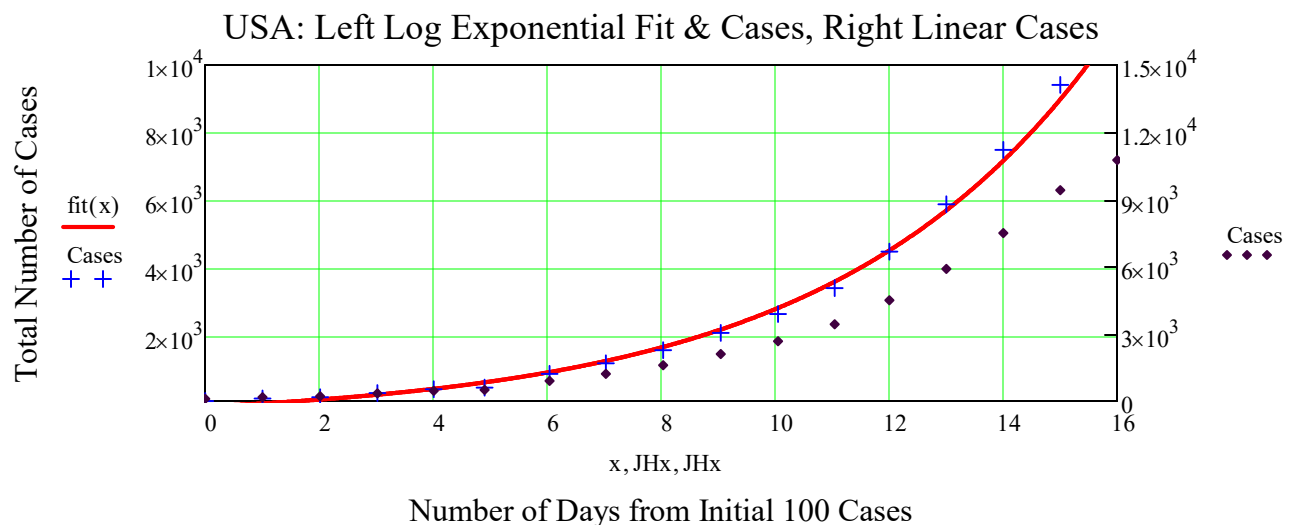
(a b c) := expfit(JHx, JHy, guess)^T (a b c) = (388.713 0.213 -450.912) fit(x) := a e^{b·x} + c

Population of US UProb := 331·10⁶ **Pop of FL** FProb := 21.0·10⁶ **Pop of Lee Co.** LCProb := 339·10³

US Reproductive Ratio R_{0L}:

LM := log(Cases) slope(JHx, LM) = 0.126

R_{0L}(0.137, 14) = 2.918



Methodology to Extract SIRD Parameters for USA

Our methodology is similar to most papers, the major difference being we use Mathcad's error minimization function Minerr, and not MATLAB's fmins() error minimization tool.

SIR Model

$$\text{ModU}(\alpha, \beta, \delta) := \text{SIR}(300 \cdot 10^6, 100, \alpha, \beta, \delta, R_c - 2) \quad \text{InfU}(\alpha, \beta, \delta) := \text{ModU}(\alpha, \beta, \delta)^{(1)}$$

Use Levenberg-Marquardt Method: Minimize Least Squares Error to Residual

Define a Residual to be the difference between the current data points JHy and SIRDModel

$$\text{ResidJH}(\alpha, \beta, \delta) := \text{JHy} - \text{InfU}(\alpha, \beta, \delta) \quad \text{Given} \quad 0 = \text{ResidJH}(\alpha, \beta, \delta)$$

Use Minerr to Extract Optimal Parameters:

$$(a \ b \ d) := \text{Minerr}(\alpha, \beta, \delta)^T \quad (a \ b \ d) = (47.394 \ 46.999 \ 0.1)^T$$

Infected

Recovered

Dead

$$\frac{\text{ERR}}{R} = 9.437$$

$$\text{UIf} := \text{ModU}(a, b, d)^{(1)}$$

$$\text{URc} := \text{ModU}(a, b, d)^{(2)}$$

$$\text{UDd} := \text{UIf} \cdot 0.03$$

$$n := 0..R_c - 1$$

$$\text{UDe} := \text{submatrix}(\text{UDd}, 0, 13, 0, 0)$$

Extracted Parameters of USA COVID-19 Epidemic

Calculate Days to Double, D2X and Fit Power Function to USA JH Data

$$\text{Guess} \quad \text{ex} := 1.4 \quad \text{FitPwr}(\text{ex}, t) := \text{JHy}_0 \cdot \text{ex}^t$$

$$\text{ResidJH}_2(\text{ex}) := \text{JHy} - \text{FitPwr}(\text{ex}, \text{JHx}) \quad \text{Given} \quad 0 = \text{ResidJH}_2(\text{ex}) \quad \text{Ex} := \text{Minerr}(\text{ex}) \quad \text{Ex} = 1.35$$

Fit JHy with a Power Function:

$$\text{Pwr}(t) := \text{JHy}_0 \cdot \text{Ex}^t$$

$$\text{D2X}(\text{Ex}) := \frac{\ln(2)}{\ln(\text{Ex})}$$

$$\text{D2X}(\text{Ex}) = 2.309$$

$$\text{Dbl_Days}(9415, 100, 15, 0) = 2.288$$

$$100 \cdot 2^{2.309} = 12188.753$$

The number of days it takes infections to double, t_{2x}

$$t_{2x} := \frac{1}{a}$$

$$t_{2x} = 0.021$$

The number of days it takes to recover, t_{rec}

$$t_{\text{rec}} := \frac{1}{b}$$

$$t_{\text{rec}} = 0.021$$

Basic Reproductive Ratio

$$\text{before and after Intervention:} \quad R_0 := \frac{a}{b + \gamma}$$

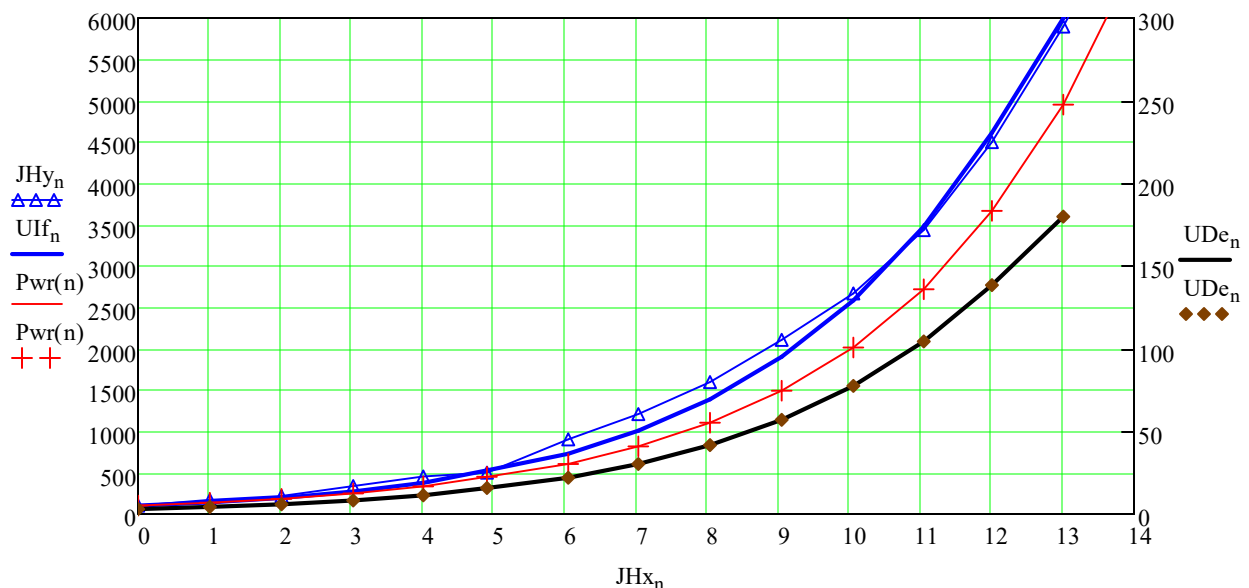
$$R_{0\delta} := \frac{\alpha - \delta}{\beta + \gamma}$$

$$R_0 = 1.008$$

$$R_{0\delta} = 2.74$$

Reasonably Good Match of US Data (JHy) by SIRD Model and Power Fit

Model vs Data: Infected (Blue), Recovered (Green), Dead (Black)



CDC: Data on USA Total Confirmed Deaths to March 19

Data Source For US Deaths to March 19:
Data From The European Center for Disease Control and Prevention (ECDC)
<https://ourworldindata.org/coronavirus>

```
Dd_US := READPRN("USA Total Confirmed Deaths to March 19 CDC.txt")
R := rows(Dd_US) = 19    Dd_US_{R-1,1} = 150    Deaths := Dd_US^{1}    z := 0..R-1    dax_z := z
Deaths_0 = 1
```

Fit Exponential Curve

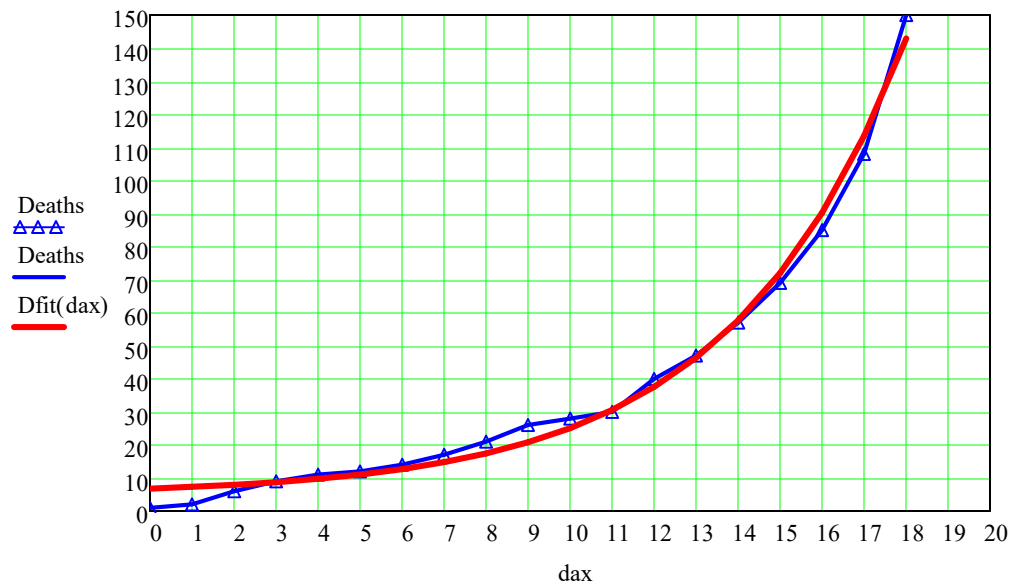
```
(a b c) := expfit(dax, Deaths, guess)^T    (a b c) = (1.805 0.241 5.015)    Dfit(x) := a e^{b·x} + c
```

Number of Days for Number of Total Deaths to Double

Dbf_Days(150, 1, 19, 0) = 2.628

Initial US Death Rate, at the Present Time, is an Exponential Curve It is Doubling every 4 days

US Recorded Death Data (Blue), Exponential Fit (Red)



Note the Difference Between the Number of Confirmed Infected and Actual Infected:

We can be certain that the number of deaths are accurate, because they are part of the public record.

We cannot get a realistic %deaths because the reported confirmed infected is not the same as the actual number of infected.

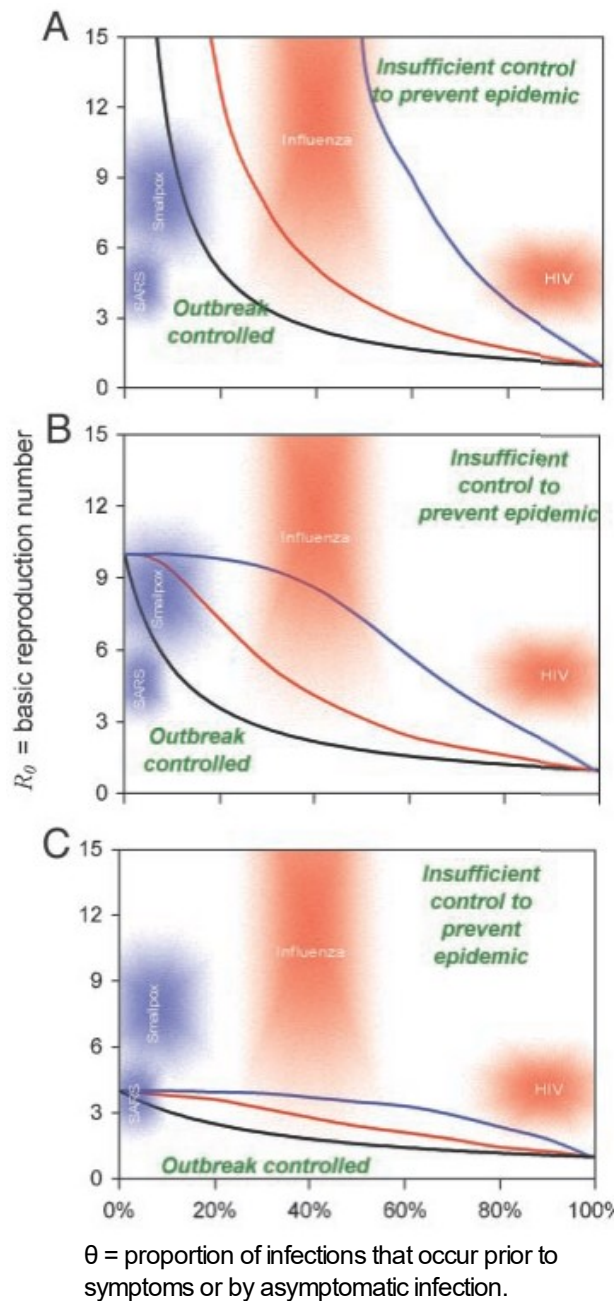
Not everyone goes to the hospital or is tested. These infections go uncounted. Because there may be a delay of 2 weeks after the actual infection before deaths are totaled. During those two weeks, every 2.2 days, the number of infected in the US doubles in number. See the above D2X (Days to Double or 2X) calculations. So the actual number infected at the time of death may be 10 to 27 times greater than the confirmed number of infected.

Assessment of Efficacy of US Government Intervention to Halt Spread

I am still working on this Section:

Factors that make an infectious disease outbreak controllable

Christophe Fraser* †, Steven Riley*, Roy M. Anderson, and Neil M. Ferguson



Criteria for outbreak control. Each curve represents a different scenario, consisting of a combination of interventions and a choice of parameters. For each scenario, if a given infectious agent is below the R_0 - θ curve, the outbreak is always controlled eventually. Above the curve, additional control measures (e.g., movement restrictions) would be required to control spread. Black lines correspond to isolating symptomatic individuals only. Colored lines correspond to the addition of immediate tracing and quarantining of all contacts of isolated symptomatic individuals. The black (isolation only) line is independent of distributional assumptions made (low or high variance), whereas the colored (isolation contact tracing) lines match the variance assumptions made in Fig. 1 (red high variance; blue low variance). The efficacy of isolation of symptomatic individuals is: 100% in A, 90% in B, and 75% in C. Contact tracing and isolation is always assumed 100% effective in the scenarios in which it is implemented (colored lines). Curves are calculated by using a mathematical model of outbreak spread incorporating quarantining and contact tracing

Italy COVID19 Cases Feb 15 - March 14

Growth Data and Curve Fit for WHO COVID-19: Total Cases in Italy

Source: CDC Data for Total Number of World Cases Reported to WHO (Virus is called (SARS-CoV-2))

https://en.wikipedia.org/wiki/Template:2019%E2%80%9320_coronavirus_outbreak_data/WHO_situation_reports

Read Data File: $WCases := \text{READPRN}(\text{"Italy COVID-19 Feb 15 to March 17 2020 Totals.txt"})$

$Rc := \text{rows}(WCases)$ $Rc = 32$ $i := 1..Rc - 1$ $days_i := i$ $\text{rows}(WCases) = 32$ $\text{rows}(days) = 32$

$WCases_{Rc-1} = 27980$

Calculate the Rate of Growth of Cases and Find Average

$WCases_0 = 3$

$NCs_i := WCases_i - WCases_{i-1}$

$\text{New_Cases} := NCs$

$\text{rate}_i := 100 \frac{NCs_i}{WCases_{i-1}}$

$WCases_{Rc-1} = 27980$

$\text{rate} := \text{rate} + 1$

$\text{rate}_{\text{avg}} := \text{mean}(\text{rate}) = 52.369$

Calculate the Number of Days for Cases to Double - Dbl:

$\text{Dbl} := \frac{\ln(2)}{\ln\left(1 + \frac{\text{rate}_{\text{avg}}}{100}\right)}$

Days to Double: $\text{Dbl} = 1.646$

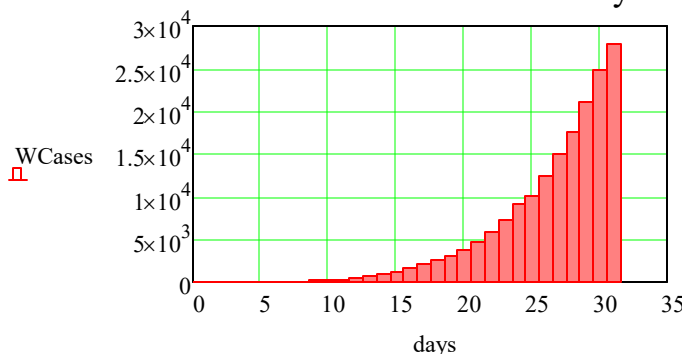
Using Exp Dble Fn Method:

$F_n = N_0 \cdot e^{t}$ $\text{Days}_{\text{double}} = 2.3$

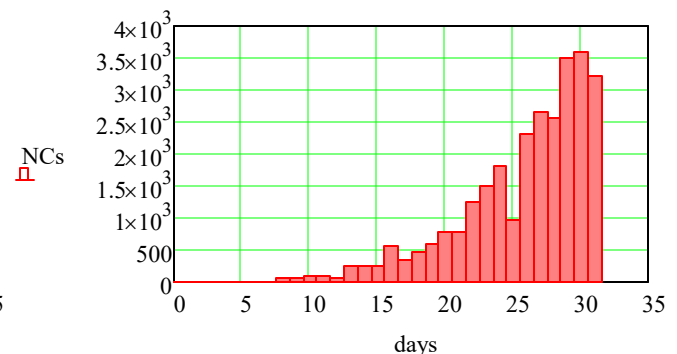
Exponential Growth --> Italy Epidemic

Note: the large increase day 25 is because of a change in reporting from laboratory confirmed to all confirmed.

Total Number of Cases/Day



New Cases

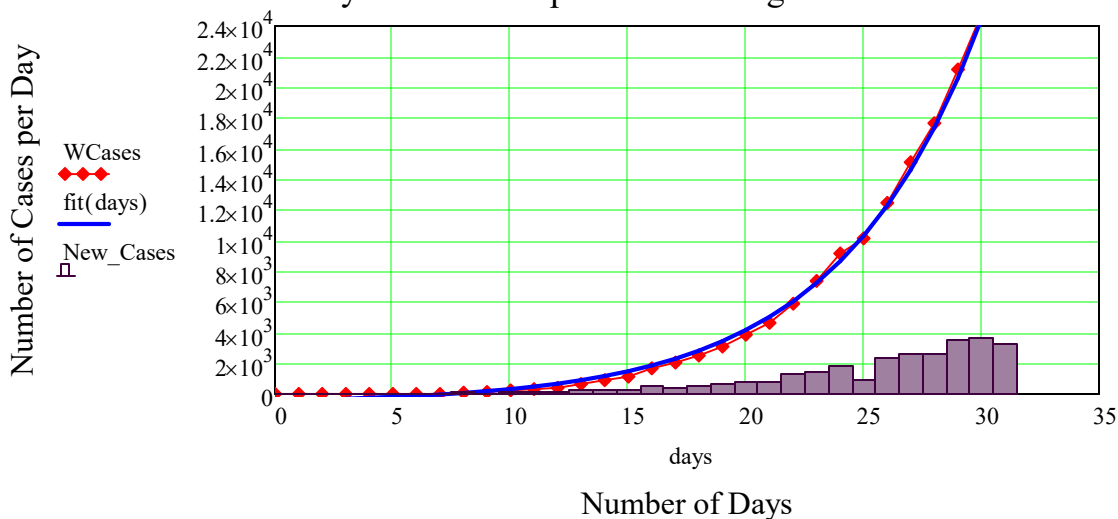


Guess for Parameters: $a := 1$ $b := 0.01$ $c := 1$ $\text{guess} := (a \ b \ c)^T$

Calculate Parameters for Exponential Fit, $\text{fit}(x)$, to Total Number of Cases

$(a \ b \ c) := \text{expfit}(\text{days}, WCases, \text{guess})^T$ $(a \ b \ c) = (171.664 \ 0.166 \ -585.889)$ $\text{fit}(x) := a e^{b \cdot x} + c$

Italy- Cases & Exp Fit - Lin&Log Scale & New Cases



Italy Reproductive Ratio R_{0L} :

$LW := \log(WCases)$

$\text{slope}(\text{days}, LW) = 0.147$

$R_{0L}(0.04, 14) = 1.56$

Spain COVID19 Cases Feb 1 - March 14

Growth Data and Curve Fit for WHO COVID-19: Total Cases in Spain

Source: CDC Data for Total Number of World Cases Reported to WHO (Virus is called (SARS-CoV-2))

https://en.wikipedia.org/wiki/Template:2019%E2%80%9320_coronavirus_outbreak_data/WHO_situation_reports

Read Data File: $WCases := READPRN("Spain COVID-19 Feb 1 to March 17 2020 Totals.txt")$

$Rc := rows(WCases)$ $Rc = 32$ $i := 1..Rc - 1$ $days_i := i$ $rows(WCases) = 32$ $rows(days) = 32$

$WCases_{Rc-1} = 9191$

Calculate the Rate of Growth of Cases and Find Average

$New_Cases_i := WCases_i - WCases_{i-1}$ $NCs := New_Cases$
 $rate_i := 100 \frac{New_Cases_i}{WCases_{i-1}}$ $WCases_{Rc-1} = 9191$ $rate := rate + 1$ $rate_{avg} := mean(rate) = 43.443$

Calculate the Number of Days for Cases to Double - Dbl:

$$Dbl := \frac{\ln(2)}{\ln\left(1 + \frac{rate_{avg}}{100}\right)}$$

Rate Days to Double: $Dbl = 1.921$

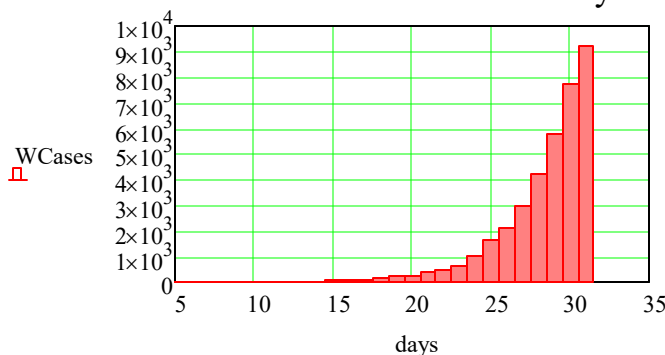
Using Exp Dble Fn Method:

$$Fn = N_0 \cdot e^{t} \quad Days_{double} = 2.3$$

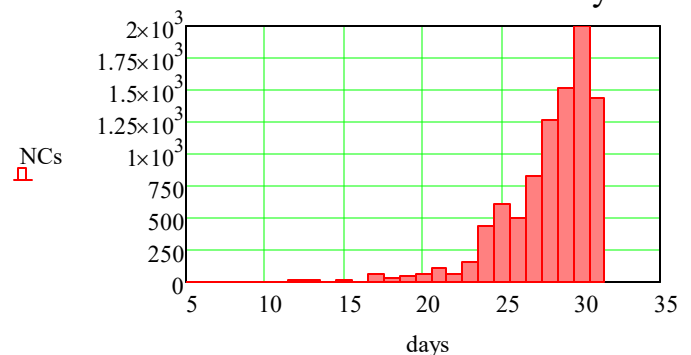
Exponential Growth --> Spain Epidemic

Note: the large increase day 25 is because of a change in reporting from laboratory confirmed to all confirmed.

Total Number of Cases/Day



Number New Cases Daily

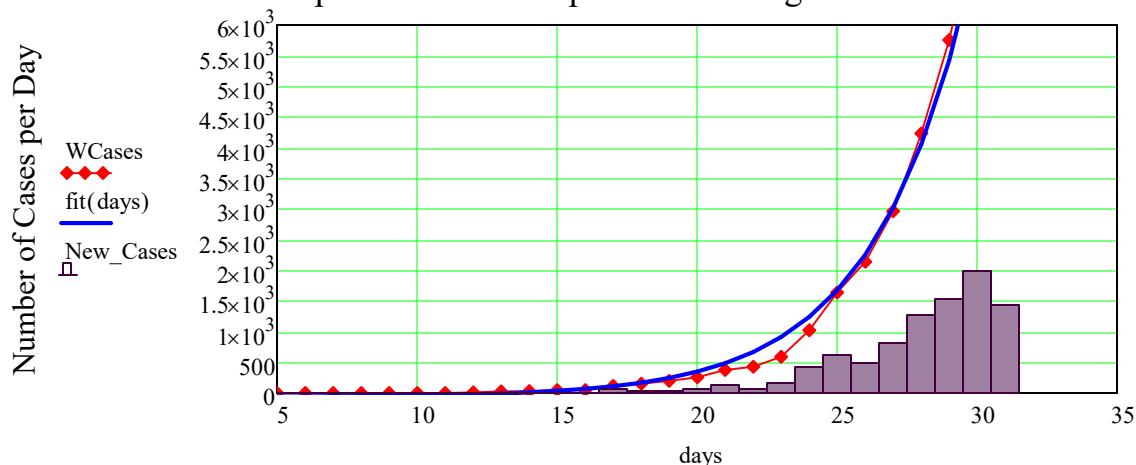


Guess for Parameters: $a := 1$ $b := 0.01$ $c := 1$ $guess := (a \ b \ c)^T$

Calculate Parameters for Exponential Fit, $fit(x)$, to Total Number of Cases

$(a \ b \ c) := expfit(days, WCases, guess)^T$ $(a \ b \ c) = (1.308 \ 0.288 \ -62.962)$ $fit(x) := a e^{b \cdot x} + c$

Spain- Cases & Exp Fit - Lin&Log Scale & New Cases



Spain Reproductive Ratio R_{0L} :

$LW := \log(WCases)$ $slope(days, LW) = 0.139$

$R_{0L}(0.04, 14) = 1.56$

World COVID19 Cases Jan 21 - March 12

Growth Data and Curve Fit for WHO COVID-19: Total Cases in World

Source: CDC Data for Total Number of World Cases Reported to WHO (Virus is called (SARS-CoV-2))

https://en.wikipedia.org/wiki/Template:2019%E2%80%9320_coronavirus_outbreak_data/WHO_situation_reports

Read Data File: $WCases := READPRN("World\COVID-19\Jan\21\to\March\15\2020\Totals.txt")$

$Rc := rows(WCases) \quad Rc = 53 \quad i := 1..Rc - 1 \quad days_i := i \quad rows(WCases) = 53 \quad rows(days) = 53$

Calculate the Rate of Growth of Cases and Find Average and Days to Double from Average Rate

$WC := WCases \quad New_Cases_i := WCases_i - WCases_{i-1} \quad NewC := New_Cases$
 $rate_i := 100 \frac{New_Cases_i}{WC_{i-1}} \quad WC_{51} = 1.425 \times 10^5 \quad rate := rate + 1 \quad rate_{avg} := mean(rate) = 15.185$

Calculate the Number of Days for Cases to Double - Dbl:

$Dbl := \ln(2) \cdot \left(\ln \left(1 + \frac{rate_{avg}}{100} \right) \right)^{-1} \quad \text{Days to Double: } Dbl = 4.903$

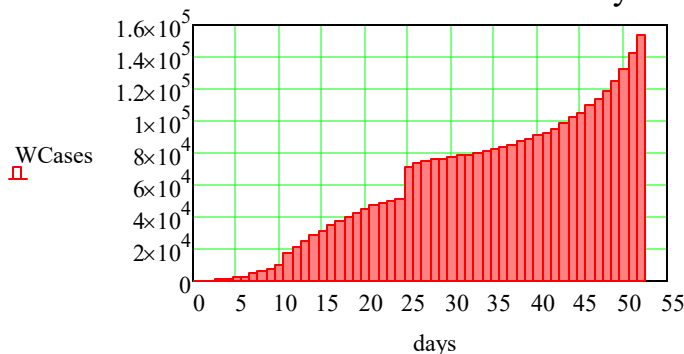
Calculate Number of Days to Double (D2X) and Fit Power Function to World Data

$FPr(ex, t) := WC_0 \cdot ex^t \quad ResW(ex) := WC - FPr(ex, days) \quad \text{Given } 0 = ResW(ex) \quad Ex := Minerr(ex) \quad D2X(Ex) = 5.475$

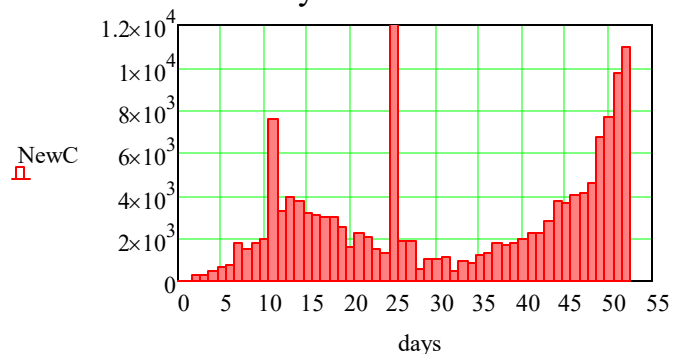
Exponential Growth --> World Epidemic

Note: the large increase day 25 is because of a change in reporting from laboratory confirmed to all confirmed.

Total Number of Cases/Day



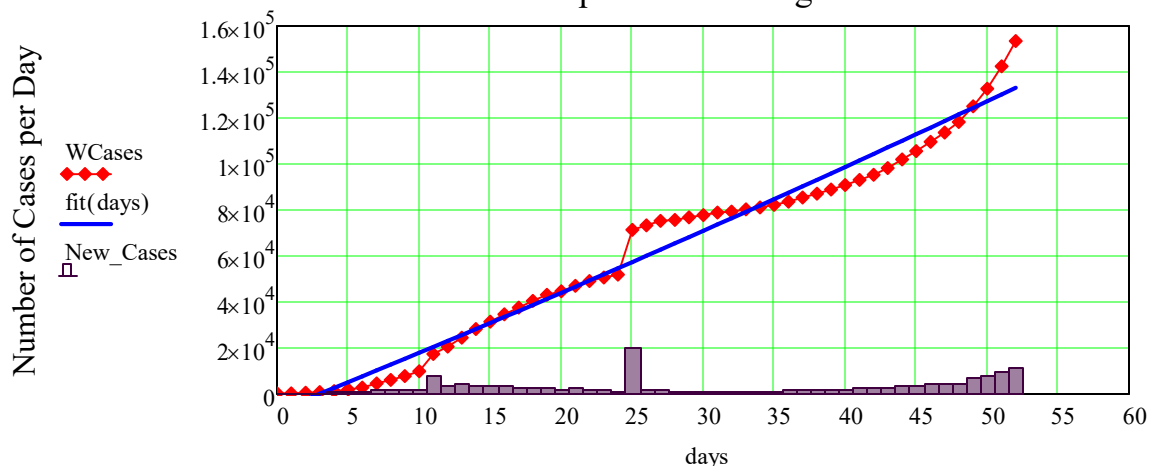
Daily % Rate of Growth



Calculate Parameters for Exponential Fit, fit(x), to Total Number of Cases

$(a \ b \ c) := expfit(days, WCases, guess)^T \quad (a \ b \ c) = (8.066 \times 10^5 \ 0.003 \ -8.14 \times 10^5) \quad fit(x) := a e^{b \cdot x} + c$

World- Cases & Exp Fit - Lin&Log Scale & New Cases



World Reproductive Ratio R_{0L} :

$LW := \log(WCases) \quad slope(days, LW) = 0.04 \quad R_{0L}(0.04, 14) = 1.56$

COVID-19 Epidemiology

AHA: COVID-19 Best Guess Epidemiology

2020 Webinar of the American Hospital Association

Communicating best guesses on impact of COVID19 in US healthcare system

USA Population: 350 million Estimate 27% of Population Infected

- | | |
|--|--|
| * $R_0 = 2.5$; Doubling time 7-10 days | Community epi wave 2 months |
| * Community attack rate = 30-40% | US: 96 million cases (27% Population) |
| * Cases requiring hospitalization = 5% | US: 4.8 million admissions |
| * Cases requiring ICU care = 1-2% | US: 1.9 million ICU |
| * Cases requiring ventilatory support = 1% | US: 1 PPV |
| * CFR = 0.5% | US: 480,000 deaths |

PREPARE FOR DISEASE BURDEN ROUGHLY 10X SEVERE FLU SEASON

Study: Nowcasting and Forecasting the International Spread of COVID-19

Nowcasting and forecasting the potential domestic and international spread of the

2019-nCoV outbreak originating in Wuhan, China: a modeling study, Wu, Leung, January 31, 2020

Findings:

In our baseline scenario, we estimated that the **basic reproductive number** for 2019-nCoV was **2.68 (95% CrI 2.47–2.86)** and that 75,815 individuals (95% CrI 37 304–130 330) have been infected in Wuhan as of Jan 25, 2020. The epidemic **doubling time** was **6.4 days** (95% CrI 5.8–7.1). We estimated that in the baseline scenario, Chongqing, Beijing, Shanghai, Guangzhou, and Shenzhen had imported 461 (95% CrI 227–805), 113 (57–193), 98 (49–168), 111 (56–191), and 80 (40–139) infections from Wuhan, respectively. If the transmissibility of 2019-nCoV were similar everywhere domestically and over time, we inferred that epidemics are **already growing exponentially in multiple major cities of China** with a **lag time behind the Wuhan outbreak** of about **1–2 weeks**.

Nowcasting and forecasting the potential domestic and international spread of the

2019-nCoV outbreak originating in Wuhan, China: a modeling study, Wu, Leung, January 31, 2020

Transmission occurs primarily via respiratory droplets from coughs and sneezes within a range of about 6 feet (1.8 m).[23][24] Indirect contact via contaminated surfaces is another possible cause of infection.[25] Preliminary research indicates that the virus may remain **viable on plastic and steel** for up to **three days**, but does not survive on **cardboard for more than one day** or on copper for more than four hours

China - Required Hospital Resources %:

Mild (Stay at Home) 80.9%, Severe (Hospital) 13.8%, Critical (ICU) 4.7%

Coronavirus: Why You Must Act Now - Stats - Tomas Pueyo

<https://medium.com/@tomaspueyo>

Crucially, these true cases weren't known at the time. We can only figure them out looking backwards: About a 9 day delay from the official count catching up to the actual cases. **11 days** after lockdown that reported cases started to decrease.

But the US is vastly undertesting because it doesn't have enough kits. **Real cases may be 6X larger** than reported. There are dozens of countries with exponential growth rates. As of today, most of them are Western.

Find true number

through deaths. If you have deaths in your region, you can use that to guess the number of true current cases. We know approximately how long it takes for that person to go **from catching the virus to dying** on average (**17.3 days**). That means the person who died on 2/29 in Washington State probably got infected around 2/12.

The **average doubling time** for the coronavirus (time it takes to double cases, on average). It's **6.2**. That means that, in the 17 days it took this person to die, the cases had to multiply by $\sim 8 (=2^{17/6})$. That means that, if you are not diagnosing all cases, one death today means 800 true cases today.

France claims 1,400 cases today and 30 deaths. Using the two methods above, you can have a range of cases: between 24,000 and 140,000. the number of true cases in France is likely to be **between one and two orders or magnitude higher** than it is officially reported.

Bar Chart Data: --If you stack up the orange bars until 1/22, you get **444 Total cases**. Now add up all the gray bars. They add up to **~12,000 Total cases**. So when Wuhan thought it had 444 cases, it had **27 times more**. If France thinks it has 1,400 cases, it might well have tens of thousands

Epidemic Spread - Remember there are two factors

Infectious Disease Dynamics, Derek Cummings

- The number of individuals infected by each infectious case. (R_0)
- The time it takes between when a case is infected and when that case infects other people.

Difference in the Serial Interval

The average length of time between when a case is infected and when s/he infects others, This serial interval is different for these two pathogens

- Influenza ~2.5 days
- Measles ~ 18 days

A Third Factor, θ

- Defined as the proportion of transmission occurring prior to symptoms
- Measure of how much shorter latent period is than incubation period
- Proposed by Fraser and colleagues

Reproduction Ratio, R_0 : Example - Sexually Transmitted Infection

$$R_0 = \beta \times c \times D$$

β is the proportion of contacts that become infected

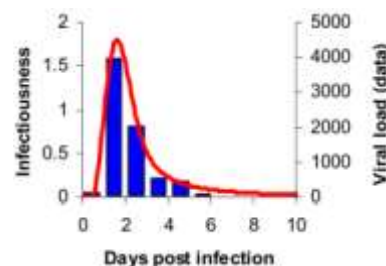
c is the number of contacts per day

D is the duration of infectiousness

Estimates of R_0 of several pathogens

Measles – 12	Scarlet Fever – 6	Smallpox – 6
Pertussis – 15	Mumps – 10	Influenza – 2
Chicken Pox – 9	Rubella – 8	HIV – 5
Diphtheria – 4	Polio – 6	Dengue – 4
Schistosoma japonicum – 3		

Could also use a proxy for infectiousness, viral load ---->
in infectiousness, viral load in oropharyngeal secretions, for example



Targeted interventions to stop transmission depend upon [being able to identify cases](#)

- Isolation, quarantine, screening travelers, prophylactic use drugs all depend on identifying people before transmit
- **Delays dramatically reduce effectiveness**
 - if your interventions don't identify people until after they've done the bulk of their transmission, they don't work
- The serial interval identifies the time-scale of response
- How quickly can we identify cases?