Estimation of a state of Corona 19 epidemic in August 2020 by multistage logistic model: a case of EU, USA, and World

Milan Batista
University of Ljubljana, Slovenia
milan.batista@fpp.uni-lj.si

Abstract
The article provides an estimate of the size and duration of the Covid-19 epidemic in August 2020 for the European Union (EU), the United States (US), and the World using a multistage logistical epidemiological model.

Keywords: Covid-19, multistage logistic model, forecasting

Introduction
This article discusses the Corona 19 epidemic that outbreaks in Wuhan (China) in December 2019. At the outbreak, the epidemic final size and its duration are common questions. (Brauer, 2019a, 2019b; Fisman D, 2014; Hethcote, 2000; House, Ross, & Sirl, 2013). Different models are used to answer such a question. These include models based on logistic function or Richards function (Batista, 2020a; Pongkitivanichkul et al., 2020; Roberts, 2020; Zou Y et al., 2020), deterministic classical and enhanced SIR and SEIR models (Anastassopoulou, Russo, Tsakris, & Siettos, 2020; Giordano et al., 2020; S. B. He, Peng, & Sun, 2020; Loli Piccolomini & Zama, 2020; Lopez & Rodo, 2020; Maier & Brockmann, 2020; Ming, Huang, & Zhang, 2020; Nesteruk, 2020; Roda, Varughese, Han, & Li, 2020; Verity, Okell, & Dorigatti, 2020; W. Yang, Zhang, Peng, Zhuge, & Hong, 2020; Tang et al., 2020; Wu, Leung, & Leung, 2020; C. Y. Yang & Wang, 2020), statistical-based models (S. He, Tang, & Rong, 2020; Singhal, Singh, Lall, & Joshi, 2020; W. Yang, Zhang, Peng, Zhuge, & Hong, 2020; Zahiri, RafieeNasab, & Roohi, 2020; Zhan, Tse, Lai, Hao, & Su, 2020), time-series models (Agosto & Giudici, 2020; Ceylan, 2020), a new models (Nesterov, 2020; Nesteruk, 2020; Singhal, Singh, Lall, & Joshi, 2020).

There are at least two problems with the modeling of the epidemic. First, the question is whether a chosen model is an appropriate description of the epidemic, especially if the epidemic has several separate outbreaks or is dragging into a new wave. The second is that at the beginning of the outbreak or at a new wave, the parameters of the models are not known (Keeling & Rohani, 2008), or better they depend on the course of the epidemic. Therefore, prediction using such models are unlikely to be successful or should be used with caution, especially if used for long-term forecasting. However, when a model is a reasonable description of the epidemic, then the long term trend of an epidemic may be assessed by monitoring changes in the model parameters. It is clear that when the parameters of the model retain their values, a long-term prediction is possible because the epidemic curve is determined. We will call such an epidemic state stable; otherwise, the state is unstable. Here we stress that any new local outbreak or import of infected into the population can destabilize the situation; almost no model can predict this. The best that models can offer are solutions for selected scenarios that may or may not realize (Bettencourt & Ribeiro, 2008; Klepac, Kissler, & Gog, 2018). (Bettencourt & Ribeiro, 2008; Klepac, Kissler, & Gog, 2018).

In sequel will use a multistage logistic model to assess the state of the epidemic in EU, US and World. The model is not new. Two stages logistic model was used for modeling 2003 SARS outbreak in Toronto (Canada) (Hsieh & Cheng, 2006; Wang, Wu, & Yang, 2012), and the multistage logistic model was introduced by Chowell et (Chowell, Tariq, & Hyman, 2019) was used for modeling Spanish flu of 1918
in Genova, Switzerland (Chowell, Ammon, Hengartner, & Hyman, 2006). We note that a multistage model based on SEIR model was introduced by Abdulrahman (Abdulrahman, 2020).

The data used in this article are total confirmed cases up to 30 August 2020, as are daily reported by Worldometer¹. We do not enter into the question of how good and reliable these data are.

**The model**

The base of the multistage (or multi-wave) logistic model is the logistic model, which is also called a simple epidemic model (Bailey, 1975) or the SI (susceptible-infective) model (Frauenthal, 1980). The basic equation of the logistic model is (Daley & Gani, 2001; Frauenthal, 1980)

\[
\frac{dC}{dt} = rC \left(1 - \frac{C}{N}\right)
\]

where \(t\) is the time, \(C(t)\) is the total number of cases in time \(t\), \(r > 0\) is infection rate, and \(N\) is the size of the susceptible population. If \(C(0) = C_0 > 0\) is the initial number of cases, then the solution of (1) is

\[
C = f(t; N, r, C_0) = \frac{N}{1 + A \exp(-rt)},
\]

where \(A = \frac{N}{C_0} - 1\).

Now, assuming that epidemic is composed of \(n_w\) mutually separated waves, then the model (2) can be generalized as follows

\[
C = \sum_{k=1}^{n_w} f_k(t | N_k, r_k, C_{0,k}, \theta_k) = \sum_{k=1}^{n_w} \frac{N_k}{1 + A_k \exp[-r_k(t - \theta_k)]},
\]

where \(A_k = \frac{N_k}{C_{0,k}} - 1\). For each wave, we, therefore, have four unknown parameters: \(N_k > 0, C_{0,k} > 0,\)

\(r_k > 0\), and \(\theta_k > 0\) is the time delay, where for the first wave \(\theta_1 = 0\). We did not incorporate parameter \(A_k\) into the exponential function because we want \(\theta_1\) to be in the data time range. The total number of unknown parameters is therefore

\[
p = 3 + 4(n_w - 1) = 4n_w - 1.
\]

If the number of data are \(n\) then \(n > p\) and therefore the number of waves that can be detected is limited by

\[
n_w < \frac{n + 1}{4}.
\]

From (3), we can calculate the total size of the susceptible population

\[
N = \lim_{t \to \infty} C = \sum_{k=1}^{n_w} N_k.
\]

¹ https://www.worldometers.info/coronavirus/
We follow Daley and Gani and define the end of the epidemic when the number of infectives is within 1 of its final size (Daley & Gani, 2001). Thus, to determine the end of the epidemic must be solved numerically (3) for \( t \) where we set 
\[ N - 1 = \sum_{k=1}^{\infty} f_k (t_{\text{end}}, N, r, C_{0,i}, \theta_k) \Rightarrow t_{\text{end}} \ . \quad (7) \]

The above model was implemented in the Matlab program fitVirusXX (Batista, 2020b). In the program, the parameters of the model are estimated by the ordinary least-squares method by minimizing the following expression 
\[ \sum_{i=1}^{n} \left[ C_i - \sum_{k=1}^{\infty} f_k (t_i, N, r, C_{0,i}, \theta_k) \right]^2 = \min , \quad (8) \]

where \( C_1, C_2, \ldots, C_n \) are the reported cases in times \( t_1, t_2, \ldots, t_n \). The above function has many possible local minimum values. Therefore, a heuristic approach was used with the brute force search method to determine a quasi-minimum of (8).

**Results**

*European Union.* Figure 1 shows that the course of the epidemic in the EU so far can be described in three waves. EU countries introduced strict quarantine in March so that the first wave peaked in early April and then weakened by mid-June. Already in April, a second smaller wave appeared, but it did not have a pronounced peak; its effect was only that the first wave dragged on into June. After the release of the measures, a new summer wave began to rise in early July.
From the graph in Figure 2, we can see that the epidemic in the EU has so far shown no signs of calming down. The estimate of the final number of infections, as well as the duration of the epidemic, has been steadily increasing since April. In August, the course of the epidemic passed into a markedly unstable phase, where, as can be seen from Figure 2, no clear trend can be observed. The current estimate shows a final 3.8 to 4.5 million infections and a duration of 600 to 700 days, i.e., until the winter of 2022.

**The United States.** In Figure 2, we can see that the epidemic in the U.S. has two waves, the first smaller reaching its peak in early May and the second larger at the end of August. In the graph in Figure 4, we can see that the trend in predicting the size of the epidemic and its duration was linear, then began to rise sharply at the end of June and reached its peak in mid-June with an estimate of 10 million final infections. This was followed by an unstable period of declining size estimates, and in the last two weeks of August, this estimate stabilized at about 6.3 million total infections. The estimate of the duration of the epidemic stabilized at 434 days, i.e., the epidemic is expected to last until May 2021.
Figure 3. Current state and forecast of Covid 19 in USA

World. Similar to the EU, we see in Figure 5 that the current course of the epidemic around the World can be divided into three waves. The first peaked in April, the second wave extended the epidemic in June, followed by a stronger third wave, which peaked in mid-August. From the graph in Figure 6, we can see that the course of the epidemic until the beginning of June was a steady increase in the estimation of its final size and duration. This was followed by a period of an indistinct but growing trend, which is not yet showing signs of calming down. The current estimate of the final size is 30 to 40 million infections and a duration of 450 to 700 days, i.e., the epidemic could drag on into winter 2022.
Figure 5. Situation and forecast of Corona 19 epidemic in the World

Figure 6. Daily-predicted epidemic size and duration for Covid 19 in the World.

Conclusions

We first note that each case under consideration has a different course of estimates of the final size and duration of the epidemic. The absence of a unique pattern makes the prediction of the final size and duration of the epidemic difficult. Namely, even if a convergence of parameters is achieved, we cannot be sure that this is the last phase of the epidemic; a new outbreak is possible at any time.
For now, only parameters for the US achieve convergence. It can be estimated that about 6.5 million peoples will be infected, and the epidemic will last about 450 days, i.e., until May 2021, if, of course, no third wave will emerge.

A new wave of epidemics is rising in the EU, which is expected to peak in October. However, the daily estimates of the model parameters for the EU are extremely unstable, so that every forecast so far is questionable and will certainly change. Similarly, we can conclude about the course of the epidemic around the World. The epidemic has so far crossed the top of the third wave, but the situation is not yet stable.

In the end, we stress that these predictions are not final but only reflect the current data.

References


