COVID19 Pandemic evolution through Gompertz's equation

COVID19 Evolución de la pandemia mediante la ecuación de Gompertz

DOI: 10.46932/sfjdv2n3-074

Received in: May 1st, 2021

Accepted in: Jun 30th, 2021

Dr. Hector G. Riveros Rotgé

Instituto de Física UNAM

E-mail: riveros@fisica.unam.mx

Dr. David Riveros Rosas

Instituto de Geofísica

E-mail: driveros@igeofisica.unam.mx

ABSTRACT

The Gompertz equation describes the evolution of an epidemic. Applied to the data from France, Spain, Germany, and the United Kingdom, it is adjusted with slight deviations since each country implemented different restriction or distancing measures to prevent the spread. Some countries do not conform to Gompertz. With the data from Mexico, the graphs show that the measured and calculated data are overlapped. The deviations are not noticeable. We made adjustments with data up to recent dates. For the four countries, Gompertz predicts well the evolution of the epidemic. We tried to predict the growth of the epidemic using Gompertz. For Mexico, the maximum number of daily infections was predicted up to a month before it happened.

RESUMEN

La ecuación de Gompertz describe la evolución de una epidemia. Aplicada a los datos de Francia, España, Alemania y el Reino Unido, se ajusta con ligeras desviaciones, ya que cada país aplicó diferentes medidas de restricción o distanciamiento para evitar la propagación. Algunos países no se ajustan a Gompertz. Con los datos de México, los gráficos muestran que los datos medidos y los calculados se superponen. Las desviaciones no son perceptibles. Hemos realizado ajustes con datos hasta fechas recientes. Para los cuatro países, Gompertz predice bien la evolución de la epidemia. Intentamos predecir el crecimiento de la epidemia con Gompertz. Para México, se predijo el número máximo de contagios diarios hasta un mes antes de que se produjera.

1 INTRODUCTION

Since the middle XIX century, Benjamin Gompertz (English mathematician of a Dutch family) found an equation to describe the epidemic mortality. This equation has been used successfully to model the evolution of population growth at the cellular level [Frenzen & Murray 1986] [Selva-Castañeda et al. 2019], and animal species [Chevallier et al. 2020].

The coronavirus (COVID-19) is an infectious disease caused by a virus known as severe acute respiratory syndrome coronavirus 2 (SARS-COV2). The acronym COVID-19 represents the disease name

(COVI means Coronavirus, D means disease, and 19 represents the year of appearance). Diverses numerical models [references Gatell, Wong, Barrio y Miramontes] have been applied to predict the evolution of the disease. Only Barrio predicted the maximum contagion date with ten days of uncertainty. With a considerable number of parameters, these models have had a similar precision with the well-known Gompertz's equation with only three adjustable parameters.

This equation can be successfully used to model pandemics [Bijay & Balvinder 2020] [Martelloni & Martelloni 2020] with great precision (high correlation indexes) and with few adjustment parameters (only if the conditions for the spread of disease are holding). Although the Mexican government had implemented various provisions at different dates, like the program #quedateencasa (“#stayathome”), the healthy distancing, masks, and health traffic lights, there are no deviations from Gompertz's equation.

This work shows a good correlation from Gompertz's equation to model the behavior of the COVID19 pandemic. The modeling shows great precision despite the changing dispositions of the Mexican government to counter the pandemic. The equation analysis also made it possible to predict the maximum number of infections one month in advance. Gompertz's equation represents a simple and effective way to model the behavior of a pandemic like COVID 19.

2 METHODOLOGY

The Gompertz's equation is given as a function of time by equation (1) by:

$$N = A e^{-Be^{-Ct}} \quad (1)$$

Where A, B, and C are positive constants. The equation applied to a pandemic evolution can describe the number of confirmed cases (N). The constants depend on the pandemic spread conditions, like individual mobility, the number of risk contacts with other persons in the time, the initial number of infected people at the moment of the equation implementation, and the dispositions taken to counter the disease. In particular, the constant C represents a measure of the characteristic time, which depends on social conditions, restrictive rules, the number of persons in quarantine, and the number of tests carried out [Bijay & Balvinder 2020]. N represents the number of confirmed cases, which we can use to describe patients' growth in different countries. For this analysis, we use data from Spain, France, United Kingdom, and Germany since the beginning of the pandemic to July 18, 2020 .

Fig. 1.- The graph shows the Gompertz fitting for Spain, France, UK, and Germany. Despite the different strategies to reduce contagions, the deviations from Gompertz are slight. In all cases, the Gompertz lines go through the data measured in each country.

It must be noted that the data deviations for the different countries are slight, although the changes in control strategies. These changes had a small effect on its evolution. Each country applied several provisions at different times. Countries that deviate from the expected curve, in particular conditions, can estimate the effectiveness of the provisions. For example, United States relaxed the mobility restrictions, and the number of contagions was incremented.

Another good example is the case of Mexico City. A sanitary traffic light was implemented to evaluate the contagion conditions and hospital occupancy in each state to return the population to its activities. The red light represented the continuity of the stay-at-home program and the closure of non-essential economic activities. The following colors (orange and yellow) allow a gradual reopening of activities until the green color allows the reopening of 100% public spaces, schools, and shops. When Mexico City passes to orange sanitary traffic light, the daily contagions leave the tendency of Gompertz's curve (figure 2). This change means infections number increased due to the reopening of some businesses and the increment of tests to detect the disease.

Fig 2. Daily cases in Mexico City, when the orange sanitary traffic light was implemented.

A MATLAB® was implemented to find the fitting parameter of Gompertz's equation. Gompertz is an excellent tool to describe and predict the pandemic evolution while the mobility condition remains constants. This work aims to find a criterion to determine the maximum of Gompertz's curve. This criterion was tested in different countries with similar results.

Every day the number of contagious was actualized by the health authorities of each country. Gompertz's equation was fitted with this data to determine the pandemic evolution as the days went by. The fitted equation permits calculating daily contagious (ΔN) and predicts the date for the curve's maximum (equation 2). $N(t)$ represents the number of infectious at the day (t), and when the derivative of $N(t)$ with respect to time is zero, the maximum occurs (M). At the beginning of the pandemic, the expected maximum date changes in a wide range of dates, but the maximum reaches a stable value as the days went by. It allows finding a criterion to determine how many days in advance are necessary to predict the maximum of the pandemic.

$$\Delta N = N(t) - N(t-1) \quad (2)$$

3 RESULTS

Figure 1 shows a good agreement between the Gompertz equation and the number of cases in different countries. The correlation coefficients are above 0.999. The problem to resolve is to determine the number of cases to predict the rest of the curve.

The data from Italy are enough to test the change of the curve fitting parameters while the first pandemic wave is in progress.

Fig.3.- Gompertz curve fitting data for different days in Italy

Gompertz's predictions for daily contagions at 28/02 (Feb, 28) and 6/03 (March, 6) were made with few cases because of the pandemic's beginning, and the corresponding curves are the lowest in figure 3. The prediction at 13/03 is the highest, and the curves for 20/03 and 27/03 are in the middle. After the maximum, for the dates 3/04 to 24/04, the predictions are accurate, and the curves are overlapping.

The confirmed contagions Italy of Gompertz's curve. But Mexico adjusts better, and these permits determine the maximum curve days before it happens. The federal sanitary authorities publish the contagious data every day, and it can be downloaded from the web page coronavirus.gob.mx. The main issue is that the last month reported doesn't have all the real confirmed cases because of a delay time between a confirmed case and the case is reported to the database.

For the case of Mexico, there are two data sources. First, the data released by the federal health authorities on their internet web page (coronavirus.gob.mx). Secondly, the data released at the evening press conference by the undersecretary of prevention and health promotion, Dr. López Gatell. The federal data delays until one month to reach the database from the contagious occurrence. In the evening conference press, are shown the cumulative cases of the day after at current day. When Gompertz's equation is applied, the curves change slowly every day as available data increases.

Fig. 4.- Gatell's data fitted with Gompertz equation using reported cases for the corresponding date.

Figure 4 shows calculated daily cases from the cumulative contagious predicted by Gompertz's equation. Each curve represents the prediction using reported data until the day of the calculation. June 13 shows the lowest curve, and August 16 the highest. The dates for the maximum for each curve from the lowest to the highest are: 16/Jul, 20/Jul, 23/Jul, 29/Jul, 31/Jul, 4/Aug, 4/Aug, and 4/Aug, respectively. The sliding of the maximums gets slower, and from August 2, all maximums are overlapping. With the reported data until June 23, Gompertz's equation predicted the maximum on July 20. The actual maximum occurred on July 25 with an error of 5 days.

Fig. 5.- Fit of federal data to Gompertz daily equation.

Figure 5 shows the daily data obtained from Gompertz's equation. The maximum dates from the lowest to highest curves are 10/May, 20/Apr, 19/Jun, 30/May, 9/Jul, 18/Aug, and 29/Jul. The dates of maximums of the curves from the lowest to highest are 15/Jun, 1/Jul, 14/Jul, 17/Jul, 19/Jul, and 25/Jul. Except for two curves, the other five are closely grouped. Since June 19, the date for maximum was forecast within an interval of 19 ± 11 days. The measured maximum occurs on July 25. Both dates differ six days for maximum.

Recent records show accumulated contagions and a 5-day moving average, with data up to June 30, 2021. The figure shows the adjustments with Gompertz using data up to January 9, 2021. Due to new infections, it was necessary to adjust two more Gompertz curves on different dates and take the data from the sum of the three Gompertz equations as the prediction. Figure 5 shows the total cases and the Gompertz sum overlapped. The 5-day moving average <5 days> of the measured data and the daily value of the Gompertz sum

Fig. 6.- Total cases and the Gompertz sum overlapped, the 5-day moving average <5 days> of the measured data and the daily value of the Gompertz sum.

The prediction is very good, until May 24, 2021, when what appears to be a regrowth begins to grow. There are still too few cases to fit a fourth Gompertz curve. Vaccination covers 30% of the population, but its effect would be to lower infections below the Gompertz curve; its effects are not yet noticeable.

4 DISCUSSION

One of the primary goals to analyze the contagious and deaths in a pandemic evolution is to find and evaluate strategies to reducing the spread of the virus and its lethality. There are three

alternatives to minimize casualties. The first is to reduce mobility, which does not seem possible for economic reasons. The second is to reduce infections by detecting the infected to isolate them. And the third is to lower the lethality of hospitals. From the national data, 0.2% of the population are infected; for Pemex and Nuevo Leon, 0.8% of its members are infected. These data do not include asymptomatic patients. Only Volkswagen measured 5000 asymptomatic workers and found 2% infected. We can estimate the actual percentage of infections between 1 and 2%. Measuring 100 asymptomatic will detect between 1 and 2 cases. In Doctor Simi's offices (local network of doctors' offices), they charge 800 pesos per analysis. If we assume a wholesale cost of 700 pesos per analysis, in 100, it would be 70 thousand pesos. According to Gompertz, the total number of infected will reach 1.26 million; that is, 700 thousand infections are missing. Of 100 infected, 74 are cured at home, 26 in hospitals, and 5 require intensive care. The costs at home are approximately 50 thousand pesos, in hospitals 480 thousand and intensive care units one million. In other words, like 18 million pesos to treat 100 cases. The cost of detecting 100 infected is between 3.5 and 7 million pesos. It is cheaper to prevent than to remedy. It is not easy to calculate an infected asymptomatic detected today to accounts of people who would be infected until September 2021.

Local lethality could be different due to the various clinical strategies followed for patient care. Identifying hospitals with low lethality should be a priority in the plan to combat the disease. However, the lethality data per hospital has not been made public, making it challenging to identify more effective procedures against the disease. We only have the lethality of the Ticomán General Hospital of 50%, which exceeds 2.2% of global lethality.

The most effective way to reduce hospital deaths is to study the procedures of low-lethality hospitals to implement them in high-lethality ones.

Case lethality is defined as the ratio of deaths to the number of confirmed cases. The global lethality percentage is 2.2%, while in Mexico is 9.2%. This number is very imprecise as it depends on the tests that are applied to identify confirmed cases. However, within a country with similar conditions for identifying cases, comparisons can be made and provide helpful information. The government does not publish lethality data for each hospital but publishes it by type of institution. Data changes daily, and lethality drops as medical treatments improve. Recently, Dr. Mariano Sánchez from CIDE defines a hospital lethality as the ratio of deaths to hospitalized cases. He defines it this way considering that most confirmed patients are outpatients, they are cured at home, and only 26.4% of confirmed patients are hospitalized. Of 505,751 confirmed cases, 133,444 were hospitalized, and 48,972 died; this gives an average lethality of 36.7% for the hospital system. The data mention 6321 deaths outside of hospitals.

SEDENA hospitals (Secretary of National Defense) have the lowest lethality rate of 13.5%, Private hospitals of 15.7%, SEMAR (Marine Secretary) of 19.3%, SSA (Secretary of health and assistance) 31%, ISSSTE 32.2%, and IMSS (Mexican Social Security Institute) 44.6%. The government should study their medical treatments to spread them to other hospitals. If SEDENA and SEMAR are saturated, send them field hospitals. Within each hospital, the history of each patient makes it possible to detect those who have the best treatments. There are reports that different anticoagulants, anti-inflammatories, and/or antivirals reduce deaths.

5 CONCLUSIONS

The Gompertz equation accurately predicts the first part of confirmed cases using the data after passing the maximum for Italy, Spain, France, Germany, and the United Kingdom. In the case of Mexico, the maximum of the pandemic could be predicted a month in advance. The Gompertz equation also allowed modeling a second and third wave of infections by adjusting new parameters with the data from increased cases. A difference was observed in the lethality of hospitals belonging to different types of institutions. It would be convenient to know the lethality data for each hospital to identify the particular medical procedures most effective in reducing the lethality.

REFERENCES

- Bijay Kumar Sahoo and Balvinder Kaur Sapra. A data driven epidemic model to analyse the lockdown effect and predict the course of COVID-19 progress in India. *Chaos, Solitons and Fractals* 139 (2020) 110034
- Gabriele Martelloni, Gianluca Martelloni, Analysis of the evolution of the Sars-Cov-2 in Italy, the role of the asymptomatics and the success of Logistic model, *Chaos, Solitons and Fractals* (2020), doi: <https://doi.org/10.1016/j.chaos.2020.110150>
- O. Torrealba-Rodriguez, R.A. Conde-Gutiérrez, A.L. Hernández-Javier. Modeling and prediction of COVID-19 in Mexico applying mathematical and computational models. *Chaos, Solitons and Fractals* 138 (2020) 109946
- C. L. Frenzen and J. D. Murray. A Cell Kinetics Justification for Gompertz' Equation. *SIAM Journal on Applied Mathematics* 46(4), July 1986 DOI: 10.1137/0146042
- Chevallier Damien Baptiste and Mourrain Marc Girondot. Modelling leatherback biphasic indeterminate growth using a modified Gompertz equation. *Ecological Modelling* 426, April 2020 DOI: 10.1016/j.ecolmodel.2020.109037
- Antonio Rafael Selva Castañeda, Erick Ramírez Torres, Narciso Antonio Villar Goris, Maraelys Morales González, Juan Bory Reyes, Victoriano Gustavo Sierra González, María Schonbek, Juan Ignacio Montijano and Luis Enrique Bergues Cabrales. New formulation of the Gompertz equation to describe the kinetics of untreated tumors. *PLoS ONE* 14(11):e0224978, November 2019, DOI: 10.1371/journal.pone.0224978