XLI CILAMCE
XLI Ibero-Latin American Congress on Computational Methods in Engineering
16-19 November, 2020 FOZ DO IGUAÇU PARANÁ - BRAZIL

Gompertz model in the study of confirmed cases of COVID-19 in the city of Porto Alegre

Eduarda de C. Castro[1], Eliete B. Hauser[1]

[1]*Polytechnic School, Pontifical Catholic University of Rio Grande do Sul*
Av. Ipiranga 6681, Porto Alegre, 90619-900, Rio Grande do Sul, Brazil
eduarda.castro@edu.pucrs.br, eliete@pucrs.br

Abstract. In the present study, the behavior of the cumulative number of confirmed cases by COVID-19 in the city of Porto Alegre (RS) was analyzed based on official data from March 8th to July 31st, 2020, contained in the COVID-19 Bulletin and published daily by the City Hall. The analytical solution of the Gompertz differential equation was deduced and its exact solution was expressed by the Gompertz sigmoidal model, P(t). Nonlinear regression techniques were used to determine the parameters of the P(t) function. Simulations were performed, ranging the maximum number of contaminants from 100% to 20% of the total population of Porto Alegre. The results obtained were adequate with a Pearson correlation coefficient of at least 0.99. The models were used to predict the data for August 7th, 2020 and the largest relative percentage deviation was approximately 13% and 21% for August 19th, 2020. The inflection point of P(t) was calculated, which shows a change in the growth rates of the number of accumulated cases (large at the beginning of the process and changing to a slower growth).

Keywords: Nonlinear regression, Gompertz differential equation, Inflection point.

1 Introduction

COVID-19 has been a hot topic worldwide since the beginning of the year. The name refers to a communicable disease caused by the coronaviruses (a family of viruses common in many different species of animals), which has spread out of control and identified for the first time in Wuhan, China. The confirmed cases in Porto Alegre faced an increasing since the first case which occurred in March 8th, 2020. Until July 31st, 8627 cases were registered in the bulletin released by the City Hall, COVID-19 Bulletin number 131 [1]. For this reason, a large part of the population is working and studying at home, maintaining social isolation and mandatory usage of facial masks on the streets has been implemented by the government. With the appearance of the new coronaviruses, the disease was spread and transmitted from person to person causing deaths around the world: first across the Asian continent, and then, others. The COVID-19 is caused by the coronaviruses, called SARS-CoV-2, which has a clinical spectrum ranging from asymptomatic infections to several conditions, as shown in Ministry of Health [2]. According to the World Health Organization, the majority (about 80%) of patients contaminated can be asymptomatic or oligosymptomatic (few symptoms) and approximately 20% of the detected cases require hospital care because of breathing difficulty, of which approximately 5% may need ventilatory support.

Basically, the severity of the disease depends on the risk group affected. Young and healthy people may have mild and moderate symptoms that resemble a cold while older people or people who already have some type of disease can be severely harmed. The coronaviruses typically cause symptoms that are similar to others viral illnesses as fever, dry cough, headache, chills, loss of taste or smell, among others. In Porto Alegre, the disease caused 352 deaths until July 31st. However, the numbers continue to grow and almost half the number of deaths from COVID-19 occurred in July. This month brought together as the worst marks of the disease evolution to date.

The present study aims to deduce from the Gompertz sigmoidal model as an analytical solution of the Gompertz differential equation and to use the least squares criterion to determine the parameters of the linearized model [3]. In addition, this research work aims to estimate the inflection points, when the growth rate reaches its maximum value. Currently, the population of Porto Alegre is approximately one million and five hundred thousand inhabitants. Simulations were performed assuming the maximum number of contaminants from 100% to 20% of the total city population.

2 Methodology and Results

The Gompertz model is expressed by the first order ordinary differential equation, [3].

$$P'(t) = -b\ln\left(\frac{K}{P(t)}\right)P(t) \quad (1)$$

Where *P(t)* is the number of infected people, *b* is the growth rate and *K* is the asymptotic value of *P(t)*,

$$\left(K = \lim_{n\to\infty} P(t)\right) \quad (2)$$

Equation (1) is written in the differential form as:

$$\frac{dP}{dt} = -b\ln\left(\frac{K}{P}\right)P \quad (3)$$

Separating the variables and integrating:

$$\int \frac{dP}{P\ln(\frac{K}{P})} = \int -bdt \quad (4)$$

Performin the variable change on the left side of the equation Equation (3):

$$u = \ln\left(\frac{K}{P}\right) \quad (5)$$

Then:

$$du = \frac{1}{\frac{K}{P}}(-KP^{-2}dP) = \frac{P}{K}(-KP^{-2}dP) = \frac{P}{K}(-\frac{K}{P^2}dp) = -\frac{dP}{P} \quad (6)$$

Replacing Equation (4) and Equation (5) in Equation (3):

$$\int \frac{1}{u}du = \int -bdt \quad (7)$$

$$ln|u| = -bt + c \rightarrow e^{ln|u|} = e^{-bt+c} \quad (8)$$

So that $a = \pm e^c$,

$$u = ae^{-bt} \quad (9)$$

Replacing Equation (4) in Equation (8):

$$ln(\frac{K}{P}) = ae^{-bt} \quad (10)$$

Then:

$$e^{ln(\frac{K}{P})} = e^{ae^{-bt}} \rightarrow \frac{K}{P} = \square^{ae^{-bt}} \rightarrow P = \frac{K}{e^{ae^{-bt}}} \tag{11}$$

This way, the general solution for the Gompertz equation differential is obtained

$$P(t) = Ke^{-ae^{-bt}} \tag{12}$$

Denoting *a=exp(A)*, the Equation (11) is equivalent to:

$$P(t) = Kexp(-\exp(A - bt)), \tag{13}$$

The linearized expression of the Equation (12) is

$$ln\left(ln\ \left(\frac{K}{P}\right)\right) = ln\ \ A - b\ t \tag{14}$$

Applying the least squares criterion [4] [5] in this linearized form, the parameters *A* and *b* are estimated by solving the system of two linear equations, considering $(t_i, p_i), i = 0..n$,

$$\begin{cases} (n+1)\ lnA\ \ - b\sum_{i=0}^{n} t_i = \sum_{i=0}^{n} ln\ \ (ln(K/p_i)) \\ lnA\ \ \sum_{i=0}^{n} t_i - b\sum_{i=0}^{n} t_i^2 = \sum_{i=0}^{n} t_i\ \ ln\ \ (ln(K/p_i)) \end{cases} \tag{15}$$

Using the number of confirmed cases of COVID-19 in the city of Porto Alegre, from March 8^{th} to July 31^{st}, 2020 [1], simulations were performed for different values of the K parameter (maximum contamined population) ranging from 100% to 20% of the total city population. In Figure 1 are illustrated the Gompertz models for parameter *K* equal to 100%, 60% and 20% of the population of Porto Alegre.

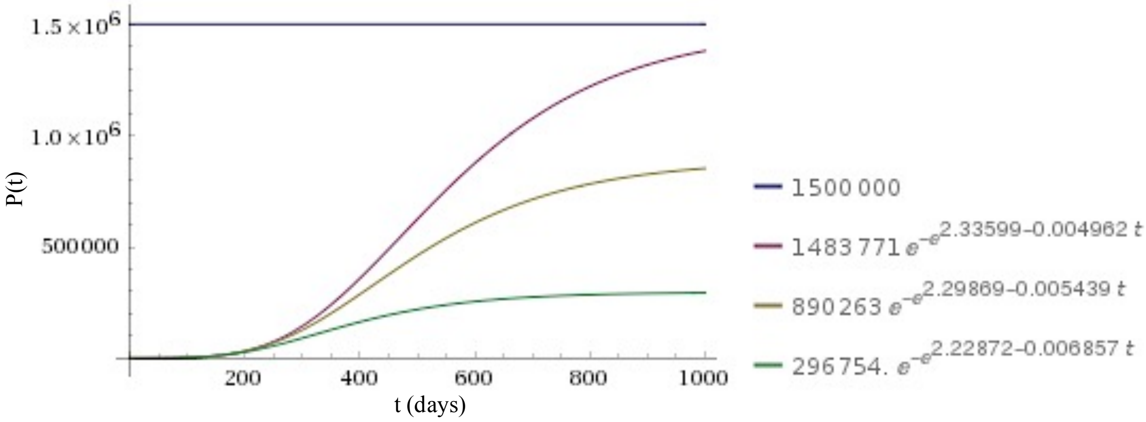

Figure 1. Gompertz model for parameter *K* equal to 100%, 60% and 20% of the population of Porto Alegre

Considering the maximum number of contaminants from 100% to 20% of the total population of Porto Alegre, the parameters *A* and *b* were calculated and they are shown in the Table 1.

Table 1. Parameters and Correlation Coefficient obtained from Parameter "K"

Parameter "K" (Percentage of the Population of Porto Alegre)	Parameter "*A*"	Parameter "b"	Relative Error on July 31st	Pearson Correlation Coefficient
100% = 1483771	2.335992	0.004962	0.146740748	0.996427
90% = 1335394	2.328185	0.005054	0.146667213	0.996457
80% = 1187017	2.319516	0.00516	0.145790729	0.996491
70% = 1038640	2.309781	0.005286	0.144976173	0.99653
60% = 890263	2.298692	0.005439	0.143709332	0.996575
50% = 741885	2.285835	0.005632	0.142213901	0.99663
40% = 593508.4	2.270588	0.005888	0.140363573	0.996697
30% = 445131	2.251995	0.006255	0.137864598	0.996785
20% = 296754.2	2.228717	0.006857	0.133143922	0.996908

Analyzing the data in Table 1, it follows that the results obtained are excellent for the period from March 8th to July 31st, 2020, with a lower Pearson correlation coefficient equal to 0.996457.

The Gompertz models (second column of Table 2) were used to estimate the cumulative number of cases for August 7th and August 19th, 2020. For August 7th, the largest relative percentage deviation was approximately 13%. For August 19, the maximum percentage relative error was 21% when compared to 18,704 (the cumulative number of confirmed cases released by COVID-19 Bulletin on the same date).

Estimates to determine the inflection point require resolving the nonlinear equation $P''(t) = 0$. The solution is $t = A/b$. Also, $P(A/b) = K/e$, where $e \cong 2.718281$ (the Euler number). Then, $(A/b, P(A/b))$ is the inflenction point. The estimates were calculated for each percentage K of the population as shown in Table 2. The smallest inflection point occurs around $t=325$, representing February 5th, 2021 for K equal to 20% of the total population of Porto Alegre.

Table 2. Inflection Points and the Respective Estimates

Parameter "K" (Percentage of the Population of Porto Alegre)	P(t)	Inflection Points (A/b days)	Estimated P(A/b)
100% = 1483771	1483771exp(-exp(2.335992-0.004962t))	470.776	545849
90% = 1335394	1335394exp(-exp(2.328185-0.005054t))	460.662	491264
80% = 1187017	1187017exp(-exp(2.319516-0.00516t))	449.519	436679
70% = 1038640	1038640exp(-exp(2.309781-0.005286t))	436.962	382094
60% = 890263	890263exp(-exp(2.298692-0.005439t))	422.631	327509
50% = 741885	741885exp(-exp(2.285835-0.005632t))	405.866	272924
40% = 593508.4	593508.4exp(-exp(2.270588-0.005888t))	385.63	218340
30% = 445131	445131exp(-exp(2.251995-0.006255t))	360.031	163755
20% = 296754.2	296754.2exp(-exp(2.228717-0.006857t))	325.028	109170

For the parameter *K* equal to 20% of the population of Porto Alegre, Figure 2 shows the inflection point in *t=325*, representing the day February 5^{th},2021.

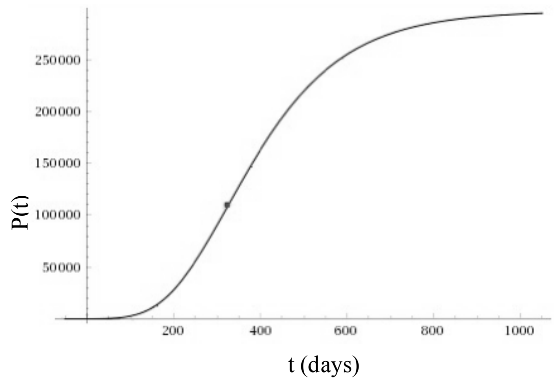

Figure 2. Inflection point of P(t), t=325, for parameter K equal to 20% of the population of Porto Alegre

3 Conclusions

The main objective of this study was achieved. Using techniques of separation and substitution of variables, the general solution of the Gompertz differential equation was constructed and expressed by the Gompertz sigmoidal model, P (t), linearizable. Nonlinear regression techniques were used to determine the parameters of the P(t) function. Simulations were performed, varying the maximum number of contaminants from 100% to 20% of the total population of Porto Alegre. The results obtained were excellent for the period from March 8th to July 31st, 2020 with Pearson correlation coefficient below 0.996457.

The models were used to forecast data for August 7th, 2020 and the largest relative percentage deviation was approximately 13%. In predictions made for August 19th, 2020 the largest relative percentage error was 21%. The inflection point of P(t) was calculated, which shows a change in the growth rates of the number of accumulated cases (large at the beginning of the process and changing to a slower growth).The smallest inflection point occurs around $t=325$, representing February 5th, 2021 for K equal to 20% of the total population of Porto Alegre.

It is concluded that the modeling used in this study is adequate for the present date. For future estimates other models must be incorporated and the analysis developed in this study must recalculate the parameters, considering the daily update of the published data (including for the previous months). Data quality and new models may be necessary due to the expected decrease in the number of new cases of COVID-19 contamination.

References

[1] City Hall of Porto Alegre. *COVID-19 Bulletin number 131*. Porto Alegre, Brazil, 2020. Available in: <http://lproweb.procempa.com.br/pmpa/prefpoa/sms/usu_doc/2020_07_31_boletim_covid_sms_131.pdf>. Accessed: August 20th, 2020.
[2] Ministry of Health. *About the disease*. Brazil, 2020. Available in: <https://coronavirus.saude.gov.br/sobre-a-doenca#interna>. Accessed: August 20th, 2020.
[3] E. B. Hauser. *Notas de Aula de Cálculo Numérico*. Porto Alegre, PUCRS, 2020.
[4] S. C. Chapra and R. P. Canale. *Numerical methods for engineering*. (5.ed.), São Paulo: McGraw-Hill, 2008.
[5] R. L. Burden and J. D. Faires. *Numerical analysis*. (8.ed.), São Paulo: Thompson, 200