

Information Generation Rate and its Relationship with the Entropy of Non-Linear Models: Covid-19 Case, Peru 2020

Danny Villegas Rivas*, Manuel Milla Pino, Salli Villegas Rivas, Erick Delgado Bazan, Yary Pérez Pérez, Zadith Garrido Campaña, Martín Grados Vasquez, Cesar Osorio Carrera, Luis Ramírez Calderón, José Paredes Carranza, Ricardo Shimabuku Ysa

Received: 27 July 2020 / Received in revised form: 12 October 2020, Accepted: 19 November 2020, Published online: 24 November 2020
© Biochemical Technology Society 2014-2020
© Sevas Educational Society 2008

Abstract

In this paper, entropy was studied in non-linear models including exponential, Gompertz, and logistic, to estimate epidemiological parameters of interest in data from confirmed cases of infection by COVID-19 in Peru. The data related to the spread of COVID-19 in Peru comes from the information available on the INS-Peru institutional portal (2020). The Akaike information criterion (AIC) and the residual standard error (ERR) were considered to evaluate the entropy of the models. The estimation of the parameters of the models was carried out using maximum likelihood and by the Bootstrap method. The results showed that the entropy of the models is related to the information generation rate, associated with the differential in the number of tests applied. Entropy severely affected maximum likelihood estimators. The Bootstrap estimators showed better performance against EMV with the estimated peak of confirmed cases. Bootstrap estimators were significantly affected by sample size, especially when $n \leq 10$. The results of this research suggest considering the entropy and the information generation rate (differential in the application of tests for the diagnosis of COVID-19 in Peru), as well as the use of

Bootstrap estimators as an alternative to estimate parameters of epidemiological models.

Keywords: Nonlinear models, Information generation rate, Bootstrap.

Introduction

After the most recent H5N1 avian influenza epidemics and during the 2009 H1N1 influenza pandemic, the international scientific community in the public health area has made efforts within the framework of the imperative need to develop standardized research and collect data that will serve as support to face eventual pandemics (Sundus, et al., 2018; Shakeri, et al., 2018; Alzahrani, et al., 2019; Ren-Zhang, et al., 2020). On December 31, 2019, 27 cases of pneumonia of unknown etiology were identified in Wuhan City, Hubei Province in China. Wuhan is the most populous city in central China with a population of over 11 million. These patients presented most notably with clinical symptoms of dry cough, dyspnea, fever, and bilateral pulmonary infiltrates on imaging. All of the cases were related to the Wuhan Huanan Seafood Wholesale Market, which marketed fish and a variety of live animal species, such as poultry, bats, marmots, and snakes (Lu et al. 2020). The causative agent was identified from throat swab samples conducted by the Chinese Center for Disease Control and Prevention (CCDC) on January 7, 2020, and was subsequently named Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). The disease was named COVID-19 by the World Health Organization, known by its Spanish acronym WHO (World Health Organization, 2020).

To apply epidemiological models it is essential to understand the phenomena of complexity and chaos since chaos theory has been considered as a possible underlying explanatory model. The parameters associated with chaos are dimension measurements and information generation rates (entropy), understanding entropy as a measure of disorder. Since these analyses require large series of data that frequently make their calculation very difficult if not impossible in practical terms, theories and methods were devised to make the statistical study of regularity feasible, relating the information generation index with entropy, applied to small series of clinical data originated from complex “noisy” systems to demonstrate the existence or non-existence of chaos and non-linearity (Cuestas, 2013). Beyond the criteria that involve the evaluation of models, from the adjustment coefficient (R^2), the entropy (Information criteria), the number of parameters (Mallows' Cp), the residual standard error (ERR), which although they are

Danny Villegas Rivas*

Faculty of Forestry and Environmental Engineering, National University of Jaén, Cajamarca, Perú.

Manuel Milla Pino, Erick Delgado Bazan, Zadith Garrido Campaña,

Faculty of Civil Engineering, National University of Jaén, Cajamarca, Perú.

Salli Villegas Rivas

Social Sciences Program, National Experimental University of the Western Plains Ezequiel Zamora, Guanare, Venezuela.

Yary Pérez Pérez

National Agrifood Training Program, Territorial Polytechnic University of Portuguesa "JJ Montilla", Guanare, Venezuela.

Martín Grados Vasquez, Cesar Osorio Carrera, Luis Ramírez Calderón, José Paredes Carranza

Graduate School, César Vallejo University, Trujillo, Perú.

Ricardo Shimabuku Ysa

Faculty of Mechanical and Electrical Engineering, National University of Jaén, Cajamarca, Perú.

criteria that allow calibrating the predictive capacity of the model, they are not a sufficient condition for these models to be used as instruments for decision-making. It is the circumstances surrounding the environment of the phenomenon that determine the quality of the information, that is, of the sample, and consequently the levels of entropy of the models.

In the case of COVID-19 in Peru, where on March 6, 2020, the first case of contagion was registered, marking the beginning of a public health problem that has led the Peruvian Government to take measures ranging from social distancing, mandatory social isolation, until mandatory social immobilization, and that to date (April 15, 2020) despite these mediations, the figure stands at more than 11,000 confirmed cases of contagion by COVID-19 in much of the national territory. In this sense, in Peru, the limitations in the acquisition of tests as a consequence of a global phenomenon, about which there is little knowledge given its recent appearance, has woven a series of situations that have impacted not only the daily life of people but also the possibility of having models that allow defining the behavior of COVID-19, in terms of fairly precise estimates concerning the peak of contagion, with which action lines can be established and the class and duration of measures to decrease the contagion rate. In this sense, the increase in the number of scientific investigations, and the proliferation of long and complex data sets, in recent years have expanded the scope in the applications of statistical methods (González-Díaz, 2016). That is why, in the face of the problems that this virus has generated, especially about models frequently used in epidemiology, which despite exhibiting a good fit, allow only a partial description of the behavior of this pandemic, but make estimation impossible of parameters that allow designing public policy strategies, the central object of this research is to study entropy in non-linear models, especially, exponential, Gompertz and logistic, to more accurately estimate epidemiological parameters of interest, namely the number and peak of COVID-19 infections.

Materials and Methods

The data related to the spread of COVID-19 in Peru comes from information available on the INS-Peru institutional portal (2020), for the period from March 6 to April 15, 2020. For modeling the estimation of the number of infected for COVID-19 in Peru growth the models that were considered include exponential, Gompertz, and logistic, that unlike the models frequently used in epidemiology, such as the SIR and SEIR, based on differential equations and that tend to make unrealistic estimates in the case of these epidemics, the growth models allow, in addition to modeling the behavior of the epidemic up to the phase where it would reach the peak of contagion, they would also be able to make estimates of the mentioned peak, which, as far as possible, would be consistent and asymptotically unbiased.

The growth models considered in this research are briefly detailed below:

According to Seber and Wild (1989), the Gompertz model is defined as follows:

$$f(x) = c + (d - c) \times \exp(-\exp(b(x - e)))$$

It is a response/growth curve across the true axis, that is not limited to non-negative values even though this is the range for most response and growth data.

If $b < 0$ the mean function increases, while it decreases for $b > 0$.

In practice, several reparametrizations of the model have been carried out.

According to Bruce and Versteeg (1992), the logistic model is defined as follows:

$$f(x) = c + \frac{d - c}{(1 + \exp(b(x - e)))}$$

Selection Criteria based on Information Measures

In this research, in addition to the widely known criteria to evaluate the goodness of fit of the models, such as the coefficient of determination (R^2) and the residual standard error (ERR), there are the information criteria or entropy indices.

Akaike Information Criteria (AIC)

This criterion is detailed in González and Landro (2018), who points out that if the problem consists of selecting the coefficients β that are as close as possible to the vector β^* , the distance between the distributions $f(Y/\beta^*)$ and $f(Y/\beta)$ can be characterized by an entropy measure of the form (see Akaike, H. (1978b)):

$$D(\beta^*, \beta) = \int_{-\infty}^{\infty} f(y/\beta^*) \ln[f(y/\beta)] dy - \int_{-\infty}^{\infty} f(y/\beta^*) \ln[f(y/\beta^*)] dy$$

(where the first addend of the second member represents the ability to fit of $f(Y/\beta)$ for $f(Y/\beta^*)$ and the second addend, for a given function $f(Y/\beta^*)$, is a constant). The minimization of the entropy measure implies the minimization of the information criterion (see Kullback, 1959):

$$KL(\beta^*, \beta) = -D(\beta^*, \beta) = \int_{-\infty}^{\infty} \{ \ln[f(y/\beta^*)] - \ln[f(y/\beta)] \} f(y/\beta^*) dy$$

Assuming that $\beta = \beta^* + \Delta\beta$ (where $\Delta\beta = [\Delta\beta_1 \ \Delta\beta_2 \ \dots \ \Delta\beta_k]^T$ is an arbitrary norm vector small), then the criterion $KL(\beta^*, \beta)$ admits a Taylor series expansion of the form:

$$KL(\beta^*, \beta^* + \Delta\beta) \approx \int_{-\infty}^{\infty} \left\{ \sum_i (\Delta\beta)_i \frac{\partial \log[f(y/\beta^*)]}{\partial \beta_i^*} - \frac{1}{2} \sum_i \sum_j (\Delta\beta)_i (\Delta\beta)_j \frac{\partial^2 \log[f(y/\beta^*)]}{\partial \beta_i^* \partial \beta_j^*} \right\} \cdot f(y/\beta^*) dy$$

If $f(y/\beta^*)$ is a regular function, the first term of the second member of this expression vanishes and, consequently, it follows that $KL(\beta^*, \beta^* + \Delta\beta) \approx \frac{1}{2} \|\Delta\beta\|_I^2$ (where $\|\Delta\beta\|_I^2 = \Delta\beta^T I(\beta^*) \Delta\beta$, where $\|\bullet\|_I^2$ is the Euclidean norm and $I(\bullet)$ is the information matrix of Fisher). Suppose that β is included in an s -dimensional space $\theta_s(1, 2, \dots, k-1)$, while the vector of the true values of the coefficients, β^* , is included in a k -dimensional space ($k > s$). Denoting by β_s^* the projection of β^* on θ_s in the sense of the Euclidean norm; it is shown that $2KL(\beta^*, \beta_s) \approx \|\beta_s^* - \beta^*\|_I^2 + \|\beta_s - \beta_s^*\|_I^2$ (where $\beta_s \in \theta_s$ and it is verified that $\beta_s \approx \beta_s^*$).

Replacing β_s by the vector of random variables $\hat{\beta}_s$ formed by the restricted maximum-likelihood estimators of β^* in θ_s and, taking into account that, for values of n that are sufficiently large, $n\|\beta_s^* - \hat{\beta}_s\|_I^2 \xrightarrow{d} \chi_s^2$, it is verified that $2E[KL(\beta^*, \hat{\beta}_s)] \approx \|\beta_s^* - \beta^*\|_I^2 + \frac{s}{n}$. This expression constitutes a measure of the deviations of $\hat{\beta}_s$ to the vector β^* and allows us to conclude that the expected value of this deviation includes a component that represents the error related to the selection of a coefficient space approximated by β_s^* and another which represents the error due to the estimation of the vector of the coefficients. Akaike showed that, under certain conditions of regularity, the likelihood ratio is:

$$LR(Y) = -2 \sum_{j=1}^n \log \left[\frac{f(y_i/\hat{\beta}_s)}{f(y_i/\hat{\beta}^{(MV)})} \right] \xrightarrow{d} \chi_{NC(k-s)}^2 (\|\beta_s^* - \beta^*\|_I^2)$$

And therefore, that $\frac{1}{n}[LR(Y) + 2s - k]$ is an unbiased estimator of the measure $E[KL(\beta^* - \hat{\beta}_s)]$. The Akaike information criterion (AIC) consists of minimizing the logarithm of the likelihood function $-2L_n(Y, \hat{\beta}_s) + 2s$ ($s = 1, 2, \dots, k-1$) in which the first term represents the measure of the error due to the lack of capacity to adapt to the approximation and the second term defines the penalty factor. Under the assumption of normality of the assumed true model, its density function assumes the form:

$$f(\hat{Y}^*) = \frac{1}{(\sigma_{p\varepsilon}^* \sqrt{2\pi})^{n-p}} \exp \left\{ -\frac{1}{2\sigma_{p\varepsilon}^{*2}} \sum_{t=p+1}^n [Y_t - Y_t^*]^2 \right\}$$

and the likelihood function of the candidate model (\hat{Y}_t^p) will be of the form. Therefore, the Kullback-Leibler distance will assume the form:

$$f(\hat{Y}^p) = \frac{1}{(\sigma_{p\varepsilon} \sqrt{2\pi})^{n-p}} \exp \left[-\frac{1}{2\sigma_{p\varepsilon}^2} \sum_{t=p+1}^n (Y_t - \phi_1 Y_{t-1} - \dots - \phi_p Y_{t-p})^2 \right]$$

Hence, the Kullback-Leibler distance will assume the form:

$$KL = \frac{2}{n-p} E \left[\ln \left(\frac{f(\hat{Y}^*)}{f(\hat{Y}^p)} \right) / f(\hat{Y}^*) \right] = \ln \left(\frac{\sigma_{p\varepsilon}^{*2}}{\sigma_{p\varepsilon}^2} \right) + \frac{\sigma_{p\varepsilon}^{*2}}{\sigma_{p\varepsilon}^2} \frac{1}{n-p} \sum_{t=p+1}^n [m(\hat{Y}_t^*) - \phi_1 Y_{t-1} - \dots - \phi_p Y_{t-p}]^2 - 1.$$

Thus, substituting in this expression the coefficients ϕ_j , $\sigma_{p\varepsilon}^{*2}$ and $\sigma_{p\varepsilon}^2$ by their maximum-likelihood estimators, we obtain:

$$KL = \ln \left(\frac{\hat{\sigma}_{p\varepsilon}^{(MV)2}}{\hat{\sigma}_{p\varepsilon}^{*(MV)2}} \right) + \frac{\hat{\sigma}_{p\varepsilon}^{(MV)2}}{\hat{\sigma}_{p\varepsilon}^{*(MV)2}} + \frac{L_2}{\hat{\sigma}_{p\varepsilon}^{(MV)2}} - 1.$$

From this definition the following selection criteria results:

$$AIC_{(p)} = \ln(\hat{\sigma}_{p\varepsilon}^{(MV)2}) + \frac{2(p+1)}{n-p},$$

which allows obtaining an asymptotically efficient estimator $\hat{p} = \min_p AIC(p)$.

Bootstrapping Estimation

In addition to the maximum likelihood estimators of the parameters of the nonlinear models considered in this investigation, the estimation was performed using the Bootstrap method proposed by Efron (1979), which is one of the simplest methods used to obtain an estimator of a parameter $\beta = \beta(P)$ where P is the postulated statistical model. Alonso (2001) presents the Bootstrap method in a general situation:

Let be $Z = (Z_1, Z_2, \dots, Z_n)$ a data set generated by the statistical model P , and let be $T(Z)$ the statistic whose distribution $L(T; P)$ we wish to estimate. The Bootstrap method proposes as an estimator of $L(T; P)$ the distribution $L^*(T^*; n, P^*)$ of the statistic $T^* = T(Z^*)$, where Z^* is a data set generated by the estimated model \hat{P}_n . Note that if $\hat{P}_n = P$, then the distributions $L(T; P)$ and $L^*(T^*; \hat{P}_n)$ coincide. Then if we have a good estimator of P , it is logical to suppose that $L^*(T^*; \hat{P}_n)$ it will approach $L(T; P)$.

The models described above, their estimators (EMV & Bootstrap), and the model selection criteria (AIC & ERR) were determined in the R environment, using the "drc" package and the "boot" package (R Core Team 2020). For details see Appendixes 1 and 2.

Results

Table 1 shows the results of the evaluation of entropy in three non-linear models (exponential, logistic & Gompertz) adjusted to data from confirmed cases of contagion by COVID-19 in Peru in the period March 6 - April 15, 2020, and related statistics. There it is observed that, for each of the models considered, the entropy index (*AIC*) and the residual standard error (*ERR*) increase as the sample size increases (days considered in the study). In the same way, the variances $\hat{\sigma}_{\beta_i}^2$ of the maximum likelihood estimators (*EMV*) of the peak of the curve (*c* in the case of the exponential model and *d* in the Gompertz & logistic models) increase when *n* grows, which results in unstable estimators, even when the data shows a good fit, especially in the case of the logistic model (see Figure 1). On the other hand, Figure 2 shows a comparison of the estimates by maximum likelihood and *Bootstrap* of the peak of confirmed cases of contagion by COVID-19. There it is observed that the *Bootstrap* estimators show a better performance than the *EMV*, as well as a considerable increase in their value on day 35. However, for values

of $n \leq 10$ the *EMV* show a better performance than the *Bootstrap* estimators.

Discussion

About the results of the entropy measurement of the models, Cuestas, (2013) points out that entropy is related to the information generation rate, hence the increase in the *AIC* and *EER* values associated with the three models as *n* grows, it may be related to the rate of generation of official information expressed in the differential of the number of rapid and molecular tests applied. Regarding the *Bootstrap* estimators, Quintana (2003) points out that the error of the *Bootstrap* approximation to the distribution of the pivotal T_n is of order n^{-1} in probability, so the *Bootstrap* can not only allow approximating the probabilistic distribution statistics of interest when obtaining it is complex, but also allows to improve the normal approximation of the classical estimators, among them the *EMV*. In this sense, this may explain the performance of the *Bootstrap* estimators against the *EMV* when *n* grows, and in turn, the behavior of the *EMV* against the *Bootstrap* estimators when $n \leq 10$.

Table 1. Evaluation of the Entropy of Non-linear Models Adjusted to the Data of Confirmed Cases of Contagion by COVID-19 in Peru between March 6 - April 15, 2020.

Model	Day (n)	AIC	EER	Standard Error of the Estimator ($\hat{\sigma}_{\beta_i}$)			
				b	c	d	e
Exponential	5	16,759	0,918	-	3,916	3,663	1,688
	10	81,086	11,175	-	1979,617	7,758	300,304
	15	159,928	42,816	-	4414,170	23,613	261,318
	20	226,444	61,771	-	5523,214	29,102	280,223
	25	360,903	299,779	-	7324,569	59,657	240,469
	30	413,530	219,771	-	9990,748	82,813	197,884
	35	567,516	749,163	-	19743,778	259,320	440,470
	41	740,754	1910,435	-	24706,220	610,530	600,690
Gompertz	5	19,061	1,338	0,456	117,441	5,046	8,714
	10	61,233	4,047	0,029	2,520	1424,007	5,051
	15	115,949	9,658	0,014	4,509	1935,455	3,745
	20	159,989	11,500	0,023	4,868	67,288	0,580
	25	241,163	26,883	0,004	13,015	1781,100	3,380
	30	305,814	35,976	0,001	13,889	2887,740	1,577
	35	476,137	200,505	0,001	63,644	200460,000	1,761
	41	571,927	241,088	0,001	52,256	26707,000	1,458
Logistic	5	19,235	1,362	0,510	116,640	4,793	7,736
	10	60,086	3,822	0,084	3,179	7338,922	12,575
	15	114,991	9,354	0,804	6,384	371,240	2,560
	20	157,362	10,769	0,031	5,368	24,473	0,289
	25	242,296	27,499	0,025	28,825	3767,651	9,667
	30	298,883	32,051	0,011	21,401	8464,700	7,301
	35	459,827	158,831	0,010	43,261	589060,000	6,873
	41	561,218	211,572	0,015	54,610	3842,900	1,349

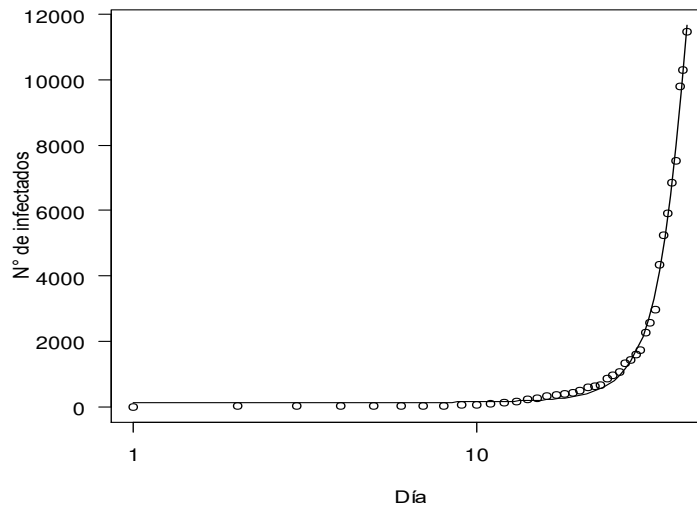


Figure 1. Adjusted Logistic Model on the Data from Confirmed Cases of COVID-19 Infection in Peru (March 6- April 15, 2020).

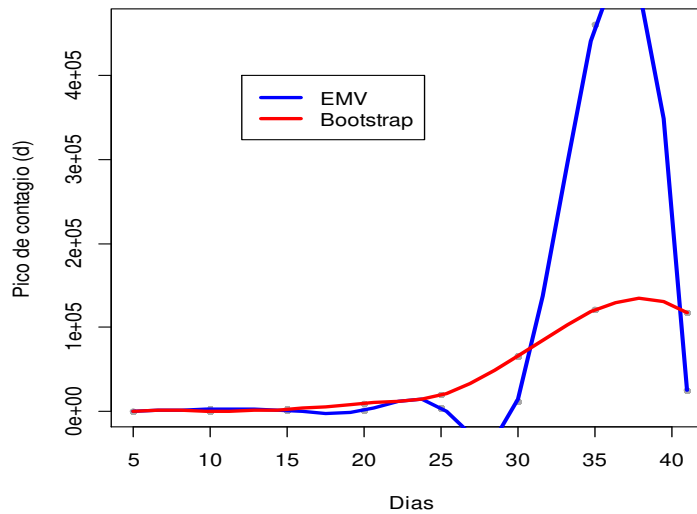


Figure 2. Maximum Likelihood Estimate vs. *Bootstrap* of the Peak of Contagion (d) using a Logistic Model on Data from Confirmed Cases of Contagion by COVID-19 in Peru (March 6-April 15, 2020).

Conclusions

The findings of this research fundamentally gravitate around the following aspects: first, it was evidenced that the entropy of the non-linear models considered in this work (exponential, logistic & Gompertz) is related to the information generation rate, which is associated with the differential in the number of tests applied. Likewise, entropy severely affected the maximum likelihood estimators. On the other hand, despite the effects of entropy, the *Bootstrap* estimators showed a better performance compared to the *EMV* with the estimated peak of confirmed cases, which showed greater consistency and stability of these estimators, in addition to

being less sensitive. The entropy associated with the rate of generation of information related to confirmed cases of contagion by COVID-19 in Peru. However, *Bootstrap* estimators were significantly affected by sample size, especially when $n \leq 10$. It is suggested to consider the entropy and the information generation rate (differential in the application of tests for the diagnosis of COVID-19 in Peru), as well as the use of *Bootstrap* estimators as an alternative, to estimate parameters of epidemiological models. Finally, the results of this research indicate that there is solid evidence to affirm that in the epidemiological area, especially in the case of COVID-19 in Peru, the criterion based on the entropy of the model allows the description of this phenomenon from a

model of mathematical ability to make predictions that contribute to the implementation of control-oriented strategies as a fundamental action within the public policy; however, based on the increase in the rate of information generation and the fact that the *Bootstrap* and *EMV* estimators show a trend towards stability, it is presumed that with a slight increase in the number of sampling days ($n \approx 50$), more precise estimates could be made of the peak of confirmed cases of COVID-19 infection in Peru.

Appendix 1. R Code to Adjust Non-linear Models on Data Related to the Number of Confirmed Cases of COVID-19 Infection in Peru until April 2020

```
library(drc)
library(stats)
library(kableExtra)
plot(dia, infectados, xlab = "Tiempo", ylab = "N")
tablamod <- data.frame(
  Modelo = c("Exponencial", "Gompertz", "Logístico", "Log-
logístico", "Weibull"),
  Parametro.fct = c("EXD.3()", "G.4()", "L.5()", "LL.5()", "W1.4()")
)
kable(tablamod, caption = "Tabla 1. Códigos para modelos en fct")
%>%
kable_styling(full_width = F) %>%
column_spec(1, bold = T, border_right = T) %>%
column_spec(2, width = "10em")
dataset1 <- data.frame(dia, infectados)
dataset1
result.G <- drm(infectados~dia, data = dataset1, fct = G.4())
summary(result.G)
plot(result.G, xlab = "Día", ylab = "N° de infectados")
AIC(result.G)
result.Exp <- drm(infectados~dia, data = dataset1, fct = EXD.3())
summary(result.Exp)
plot(result.Exp, xlab = "Día", ylab = "N° de infectados")
AIC(result.Exp)
result.Log <- drm(infectados~dia, data = dataset1, fct = L.5())
summary(result.Log)
plot(result.Log, xlab = "Día", ylab = "N° de infectados")
AIC(result.Log)
```

Appendix 2. R Code to Estimate Non-linear Model Parameters through Bootstrap on Data related to the Number of Confirmed Cases of COVID-19 Infection in Peru until April 2020

```
library(Bootstrap)
library(ISLR)
library(drc)
library(stats)
library(kableExtra)
plot(dia, infectados, xlab = "Tiempo", ylab = "N")
tablamod <- data.frame(
  Modelo = c("Exponencial", "Gompertz", "Logístico", "Log-
logístico", "Weibull"),
  Parametro.fct = c("EXD.3()", "G.4()", "L.4()", "LL.4()", "W1.4()")
)
```

```
kable(tablamod, caption = "Tabla 1. Códigos para modelos en fct")
%>%
kable_styling(full_width = F) %>%
column_spec(1, bold = T, border_right = T) %>%
column_spec(2, width = "10em")
dataset1 <- data.frame(dia, infectados)
dataset1
result.Log <- drm(infectados~dia, data = dataset1, fct = L.4())
summary(result.Log)
AIC(result.Log)
set.seed(1)
library(ISLR)
indices.train <- sample(x = nrow(dataset1), size =
0.5*(nrow(dataset1)), replace = FALSE)
datos.entrenamiento <- dataset1[indices.train,]
datos.test <- dataset1[-indices.train,]
n=nrow(dataset1)
head(result.Log)
result.Log$fit$par[3]
meanstar1=mean(dataset1$infectados)
sdstar1=sd(dataset1$infectados)
R = 1000
Fstar = numeric(R)
for (i in 1:R) {
siminfectados= mnorm(n, mean=meanstar1, sd=sdstar1)
simtiempo=dataset1$dia
simdata = data.frame(siminfectados,simtiempo)
result.Log <- drm(siminfectados~simtiempo, data = simdata, fct =
L.4())
Fstar[i]=result.Log$fit$par[3]
}
Fstar
mean(Fstar)
hist(Fstar,main="",xlab="Pico de
contagio",ylab="Densidad",breaks=90,freq=FALSE).
```

References

- Akaike H (1978b) On the likelihood of a time series model. *The Statistician*. 27: 237-242.
- Alonso A (2001) Bootstrap confidence intervals for indicators in regression. *Marrow*, 10 (1), 11-20.
- Alzahrani, M. I., Alzahrani, M. I., Alsolami, E. A., Al Sulami, A. H., Abdulrahman, A., Alhejaili, M. A. A.,... & Ahmed, Z. (2019). Evaluative Role of Family Physician in Diagnosis and Management of Gout in Primary Health Care Centers: A Simple Literature Review. *Archives of Pharmacy Practice*, 10(3), 126-129.
- Bruce RD, & Versteeg DJ (1992) A statistical procedure for modeling continuous toxicity data. *Environ. Toxicol. Chem*, 11(1), 1485–1494.
- Cuesta E. (2013). Predictability of the spatial and temporal spread of the AH1N1 influenza epidemic in Argentina. Available at: http://lildbi.fcm.unc.edu.ar/lildbi/tesis/cuestas_eduardo_jose.pdf
- Efron B. (1979). Bootstrap methods: Another look at the jackknife. *Annals of Statistic*. 7(1): 1- 26.

- González C. (2016). Statistical significance and p-values. *Cuban Journal of Hygiene and Epidemiology*, 54(3), 60-62.
- González M, & Landro A. (2018). Information criteria and stochastic complexity. *Journal of Research in Financial Models*, 7(1), 21-40.
- Kullback S. (1959). *Information theory and statistics*. New York: Wiley.
- Lu H, Stratton CW, & Tang Y. (2020) Outbreak of pneumonia of unknown etiology in Wuhan China: the mystery and the miracle. *J. Med. Virol*, 92(4), 401-402.
- National Institute of Health [Internet]. Lima: INS; 2020 [cited 2020 Apr 9]. Available at: <https://web.ins.gob.pe/>
- Organization for Economic Cooperation and Development (2006) Current approaches in the statistical analysis of ecotoxicity data: A guidance to application – annexes. Paris: p. 80.
- Quintana M. (2003) Bootstrap estimation with sequential data. [Doctoral Thesis]. Spain: Department of Mathematics, University of Las Palmas de Gran Canaria; 2003. Available at: https://accedacris.ulpgc.es/bitstream/10553/19572/4/0287981_00000_0000.pdf
- R Core Team (2020) R Development Core Team. R: A language and environment for statistical computing. Disponible en: <http://www.r-project.org>, <http://www.bioassay.dk>
- Ren-Zhang, L., Chee-Lan, L., & Hui-Yin, Y. (2020). The awareness and perception on Antimicrobial Stewardship among healthcare professionals in a tertiary teaching hospital Malaysia. *Archives of Pharmacy Practice*, 11(2):50-9
- Seber GA, & Wild CJ. (1989) *Nonlinear Regression*, New York: Wiley & Sons.
- Shakeri, H., Rahmanian, V., Shakeri, M., & Mansoorian, E. (2018). Study Of Anti-Hbs Antibody Titer And Associated Factors Among Healthcare Staff Vaccinated Against Hepatitis B More Than Ten Years In Hospitals Of Jahrom In 2016. *Pharmacophore*, 9(1), 156-61.
- Sundus, A., Ismail, N. E., & Gnanasan, S. (2018). Exploration of healthcare practitioner's perception regarding pharmacist's role in cancer palliative care, Malaysia. *Pharmacophore*, 9(4), 1-7.
- World Health Organization (2019) Technical Focus: Early Clinical and Epidemiological Investigations on COVID-19 for a Public Health Response; 2019. [Internet] [cited 2020 Apr 9]; Available at: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/early-investigations>.
- World Health Organization, WHO Director-General's Remarks at the Media Briefing on 2019-nCoV. 2020. [Internet]. [citado 08 de abril de 2020]; Disponible en: <https://www.who.int/dg/speeches/detail/who-director-general-s-remarks-at-the-media-briefing-on-2019-ncov-on-11-february-2020>.