

Estimación de la tasa de letalidad de COVID-19 en tres ámbitos mediante un análisis de regresión lineal simple

Inés Angélica de León Sigg¹, Oscar Pérez Veyna²

¹ Universidad Autónoma de Zacatecas, Unidad Académica de Contaduría y Administración Comercio y Administración S/N, Col. Progreso, Zacatecas, Zac. C.P. 98066

² Universidad Autónoma de Zacatecas, Unidad Académica en Estudios del Desarrollo Loma Dorada No. 303, Fracc. La Loma, Zacatecas, Zac. C.P.98068

Inesangelica.deleonsigg@unizacatecas.edu.mx

Resumen: Se presenta un análisis de regresión lineal simple para estimar la tasa de letalidad de la pandemia COVID-19 en tres ámbitos (México, Estado de Zacatecas, Zacatecas Capital), comprendiendo el periodo del 16 de marzo del 2020, fecha en que se detectó al primer paciente contagiado, al 25 de noviembre de 2020. Así mismo se calculan las Razones de Letalidad de los Casos y se propone la complementariedad de indicadores para la mejor comprensión del fenómeno. Los datos utilizados han sido tomados de la base de datos que CONACYT proporciona como información abierta al público.

Palabras clave: COVID-19, Regresión Lineal Simple, Letalidad, México.

Abstract: A simple linear regression analysis is presented to estimate the fatality rate of the COVID-19 pandemic in three areas (Mexico, State of Zacatecas, Zacatecas Capital), comprising the period of March 16, 2020, the date on which the first infected patient was detected, as of November 25, 2020. The Reasons for Case Lethality are also calculated and the complementarity of indicators is proposed for the better understanding of the phenomenon. The data used has been taken from the database that CONACYT provides as information open to the public.

Keywords: COVID-19, Simple Linear Analysis, Letality, Mexico.

1. Introducción

La pandemia provocada por el virus SARS-Cov2 (COVID-19 en adelante), ha sorprendido al mundo, ha mostrado la vulnerabilidad de la especie no obstante los importantes desarrollos científicos y tecnológicos que incluso han permitido al hombre incursionar en la conquista del espacio, en un esfuerzo continuo de casi 50 años. Sin embargo, ante una circunstancia de salud como la que vivimos, toda la capacidad científica se ha visto en desventaja respecto a la pausada pero consistente evolución de la infección. Frente al fracaso de los más avanzados sistemas de salud, que arrastran sin embargo las características que entraña un modelo económico que ha profundizado la desigualdad, la pobreza y la apropiación de los bienes comunes, resta preguntar si finalmente esta mortal experiencia permitirá entender que es necesario un ejercicio urgente de autoevaluación sobre el quehacer investigativo que converja en la reorientación de las líneas de investigación y las prioridades en salud, más allá de seguir modas en líneas de investigación que dictan laboratorios y colectivos desde las grandes y prestigiadas universidades en países desarrollados y a las cuales, por desgracia, nos hemos unido, dejando de lado la importancia de la salud comunitaria. Es difícil desarrollar con apertura este ejercicio necesario de autocrítica pues el trazo de la ruta de los investigadores tiene mucho que ver con la formación misma, sobre todo cuando ésta viene de visiones unitarias, espectros estrechos y muy ligados al interés que

defienden los investigadores basándose en argumentos propios de la disciplina, pero que suelen dejar de lado, como ha quedado claro frente a la pandemia, las verdaderas prioridades no solo en salud. Es un imperativo que la comunidad científica ubique en su agenda la necesidad de replantear más allá de la holgura que deja la autonomía de nuestras instituciones de educación e investigación, una que corresponda más con la contribución para abatir el pasivo que la academia y la investigación nacional tiene con la sociedad y sus sistemas e instituciones.

Así las cosas, después de un año de haberse detectado el primer caso de COVID-19 en China [1] y haber comenzado su propagación a nivel global, ahora sabemos que ha generado 61,869,330 de casos confirmados y ha dejado la impresionante cantidad de 1,448,896 muertes [2] al 29 de noviembre; los números continúan en ascenso. En México para el mismo período de tiempo se han registrado 1,285,959 de casos positivos y 121,436 defunciones [3]; la circunstancia sanitaria ha motivado la escritura de trabajos de naturaleza varia, derivados de diversos estudios respecto al COVID-19 con diferentes enfoques. Desde la perspectiva oficial, la difusión de información ha comprendido una rutina constante ante la cual, se percibe el escepticismo, en virtud de contar reiteradamente con la misma estructura completamente descriptiva de la información día con día. Si bien, al inicio de la pandemia se hizo del conocimiento del país el grupo de científicos que estarían detrás de los responsables, el hecho es que parece ser, con el tiempo, las visiones cambian y los equipos

de trabajo se disocian, seguramente por los criterios sobre los cuales se genera la información que se hace oficial. Es así que a diferencia de otros países entre los cuales se puede mencionar justo al país de origen de la pandemia: China, desde donde con extrema frecuencia es posible encontrar una sucesión de publicaciones basadas fuertemente en procedimientos estadísticos que buscan dejar claro a la comunidad científica, la estrategia que ese país ha seguido en la contención del virus causante de la COVID-19. Lo mismo es posible encontrar en grupos de investigación del Reino Unido, Australia, Nueva Zelanda, Nigeria entre otros. En tanto que la ausencia de publicaciones propias de nuestro contexto, están ausentes en revistas de alto impacto. Esta ausencia deja ver cierta debilidad en los grupos de investigación y sobre todo la ausencia de grupos, colectivos de investigadores y estudiantes de posgrado, que se inscriban en la línea que hoy ha quedado claro, estaba olvidada o minimizada. Una característica que vale la pena señalar es el tamaño de los grupos de investigación que generan los artículos que han estado mostrando los avances en la comprensión de los mecanismos de acción, la composición del virus, la estructura de su RNA para descifrar los mecanismos de infección así como también los avances sobre la elaboración de lo que se ha constituido en la esperanza de la humanidad para sobrellevar la vida con el virus, las vacunas.

En este esfuerzo, es importante colocar la atención en el alto grado de dependencia de nuestro país al grado de estar haciendo fila para poder recibir una dotación de las vacunas que han sido autorizadas. Tampoco México ha mostrado capacidad en el desarrollo de capacidades científicas y tecnológicas para el desarrollo de ese tipo de vacunas que hoy se ve, habrán de marcar el derrotero para unos y otros, los que sí puedan contar con ellas y los que no puedan contar con ellas.

Por el lado de la investigación en instrumentación y equipo, también México quedó expuesto frente a la alta demanda de ventiladores mecánicos que permitieran en algunos casos prologar la vida de las personas infectadas. Se debió hacer trabajo diplomático y un gran esfuerzo económico y aéreo para que nuestro país pudiera recibir no solo ventiladores sino también equipo personal para el cuerpo médico. Son estas las señales que nunca debieron aparecer en un país previsor y con orientación de sus recursos a la atención de lo que hoy se entiende es una ineludible prioridad, pues sin la salud de una sociedad, la economía no puede funcionar. Son aspectos que han detonado el interés por realizar una contribución aún desde un área administrativa que pretende incursionar desde los aspectos que son posible entender con la información disponible, pero con un serio empeño en posicionar el tema de la administración de la salud en las líneas de investigación de los programas de posgrado. El trabajo es motivador y se espera señale el interés por abrir la ciencia administrativa a un espectro amplio y de mayor compromiso con la sociedad en general.

En este trabajo se ha considerado la necesidad de intentar entender a través de un procedimiento estadístico, como es la relación entre los casos confirmados y la mortalidad provocada por COVID-19, para tres ámbitos: México, estado de Zacatecas y Zacatecas capital, en la intención de construir conocimiento que porte al diseño de acciones que prevengan situaciones futuras de

esta naturaleza, pues se tiene claridad que la pandemia ha modificado en un período relativamente corto, innumerables facetas de nuestra vida cotidiana. Muy probablemente, frente a la complejidad que entraña el monitoreo de la pandemia no solo en Zacatecas, se observa que algunos de los cambios serán de corto y mediano plazo, mientras que otros, dejarán su marca permanente. El trabajo consiste en un análisis para la estimación de la tasa de letalidad por causa de la pandemia COVID-19, a nivel nacional y con mayor énfasis en el Estado de Zacatecas. El tiempo comprendido en el análisis corresponde al periodo del 16 de marzo al 25 de noviembre de 2020.

Este trabajo sobre la letalidad de la enfermedad en el país y específicamente en Zacatecas, está basado en los estudios de [4] y [5], que realizaron en los países de Filipinas y Nigeria respectivamente.

El objetivo es estimar la proporción de letalidad en tres ámbitos (México, Estado de Zacatecas y Zacatecas capital), a partir únicamente de los casos confirmados con la enfermedad y los fallecimientos por esa misma razón, a través del uso de la técnica de regresión lineal simple. De la misma manera a partir de los mismos datos, calcular según la OMS, la Razón de Letalidad de los casos (CFR), con el fin de mostrar que la disposición de unos y otros indicadores son complementarios y aportan para la mejor comprensión del fenómeno y por tanto en la configuración de acciones para su pronta erradicación.

Primeramente, se hace un análisis a los datos contenidos en la base de datos pública del CONACYT correspondientes a los casos confirmados y al número de defunciones a nivel nacional. Después, el análisis se realiza para los casos específicos del Estado de Zacatecas, información obtenida del sitio web de CONACYT [3].

2. Marco teórico

2.1 Antecedentes

El 31 de diciembre de 2019 se reportó un clúster de neumonía causado por un patógeno desconocido, en Wuhan, una ciudad de aproximadamente 11 millones de habitantes en el centro de China. Un nuevo coronavirus [6], [7] fue identificado como el agente etiológico [8].

Para prevenir la dispersión, fue prohibido el transporte desde y hacia Wuhan, seguido por el cierre total de la provincia de Hubei un día después. En términos de la población protegida, este hecho constituye el mayor cerco sanitario en la historia de la humanidad [6].

2.2 Estimación de la mortalidad

Las tasas de letalidad ayudan a entender la gravedad de la enfermedad, a identificar las poblaciones en riesgo y a evaluar la calidad de la atención sanitaria.

Para evaluar la proporción de personas infectadas que tienen un desenlace mortal se utilizan dos medidas:

La razón de letalidad de la infección (IFR, por su acrónimo en inglés) que estima la proporción de muertes entre todas las personas infectadas. La razón de letalidad de los casos (CFR, por

su acrónimo en inglés) que estima la proporción de muertes entre los casos confirmados [9].

2.3 Métodos empleados en la estimación de la mortalidad.

El estudio realizado por Medina [4], se hizo al inicio de la pandemia con el objetivo de estimar la tasa de mortalidad en Filipinas y con eso proporcionar información para que se pudieran tomar medidas al respecto. De igual forma, Abubakar et al., [5], hacen un estudio respecto a la tasa de mortandad los primeros 44 días de la pandemia en Nigeria, con la intención de proporcionar información para una adecuada toma de decisiones. En ambos casos emplean la técnica de regresión lineal y no lineal para la estimación de la tasa de mortalidad.

3. Material y método

Considerando que el análisis consiste en el estudio de dos variables, se ha considerado a la variable de respuesta el número de casos de muerte por COVID-19, y la variable predictorora el número de casos confirmados como positivos de COVID-19; se ha seleccionado el análisis de regresión lineal simple [7].

En el caso de México, esta información se está analizando nueve meses después de haber iniciado el confinamiento, dentro de una situación en que los contagios no han disminuido pese a las medidas de contingencia que las autoridades han llevado a cabo; por los recientes repuntes y la vuelta a las restricciones, los casos no disminuyen y han puesto en seria dificultad al sistema de salud, Zacatecas no ha sido la excepción.

Para este estudio se ha considerado la base de datos que CONACYT pone a disposición como datos abiertos en su sitio web oficial [3], a partir de ahí, se toman los datos a nivel nacional y los datos correspondientes a los 57 municipios del Estado de Zacatecas.

Los datos analizados a nivel nacional corresponden al periodo del 13 de enero, primer caso confirmado en la República Mexicana de acuerdo a [3], al 25 de noviembre; en el caso de Zacatecas, el periodo comprende del 16 de marzo, primer caso confirmado en el Estado según [3], al 25 de noviembre de 2020, datos que corresponden a los reportes diarios de los casos confirmados y de defunciones.

En el modelo de regresión lineal simple, la ecuación representativa del modelo general está dada por la Ec. (1).

$$E(y) = \beta_0 + \beta_1 x \quad (1)$$

es decir,

$$Defunciones = \beta_0 + \beta_1 * Confirmados \quad (2)$$

Donde β_0 es el punto de intercepción y para el caso opera como factor de ajuste y β_1 es la pendiente de la recta del modelo, es decir, la razón de cambio que sufren las defunciones por el cambio unitario de casos confirmados con COVID-19.

Para obtener dicha ecuación de predicción, se consideran los datos de la base de datos: 318 días para el cálculo nacional y 255 para el cálculo en Zacatecas, a partir del primer caso de contagio presentado en ambas partes.

4. Análisis de resultados

Oficialmente, de acuerdo a [8], el 28 de febrero se dio a conocer el primer caso positivo en el país, correspondiente a una persona que viajó a Italia y presentó síntomas de COVID-19 un día después de haber llegado a la Ciudad de México. Los datos presentados en la base de datos de CONACYT dan cuenta del primer contagio el día 13 de enero, dato similar que se presenta en la base de datos de European Centre for Disease Prevention and Control (ECDC) [9]. Para el presente estudio se tomó el dato proporcionado en la base de datos de CONACYT.

Actualmente, la cifra de contagios en el país asciende a 1,285,959 y una cifra de 121,436 defunciones por causa del COVID-19. Según INEGI [10], con información de la Encuesta Nacional de la Dinámica Demográfica (ENADID) 2018, la población total de México es de 125 millones de personas; la base de datos del ECDC presenta una cifra de 127,575,529 habitantes en la República Mexicana; por su parte, CONACYT, da una cifra de 127,792,286 habitantes en el país; para el presente estudio se ha considerado la población total que informa CONACYT en sus bases de datos.

4.1 Análisis nacional

La Figura 1 muestra el número de casos confirmados y el número de defunciones acumulados a causa de COVID-19, se puede observar claramente el aumento del número de casos.

La Figura 2 muestra la relación entre los casos confirmados con COVID-19 y las muertes a causa de la enfermedad a nivel nacional, se puede observar que conforme han aumentado los contagios, las muertes también han aumentado.

El resultado del análisis devuelve la Ec. (3), la cual indica que por cada caso confirmado de COVID-19 existe un aumento de 0.0685 muertes por dicha enfermedad.

$$Defunciones = 93.456 + 0.0685 * Confirmados \quad (3)$$

De igual forma, el análisis indica que existe una correlación del 80.89% entre el número de defunciones y los casos detectados de COVID-19. El coeficiente de determinación (R^2) explica el número de muertes por COVID-19 en función de los contagios COVID-19 que se dan, con un 65.44% de probabilidad. Hay un

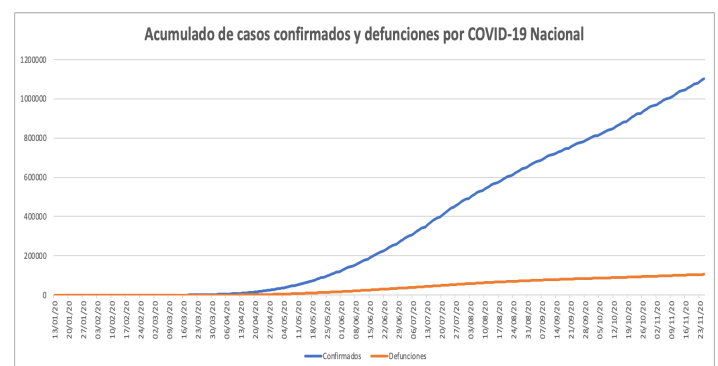


Fig. 1. Acumulado de casos confirmados y defunciones por COVID-19 a nivel nacional.

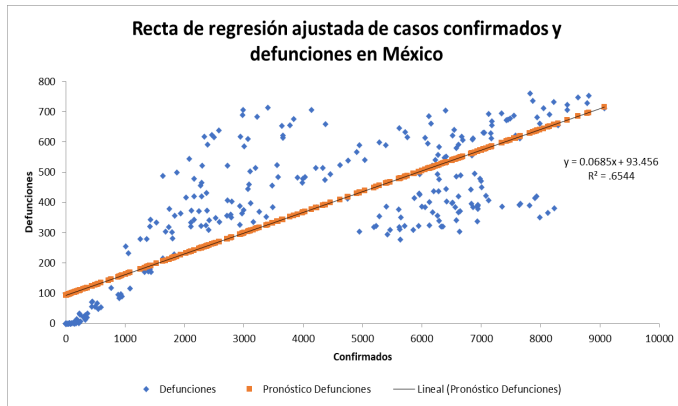


Fig. 2. Recta de regresión ajustada de casos confirmados de COVID-19 a nivel nacional.

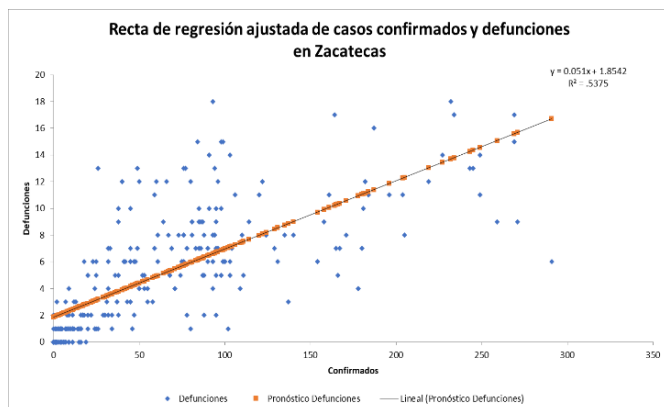


Fig. 3. Recta de regresión ajustada de casos confirmados de COVID-19 y defunciones en el Estado de Zacatecas.

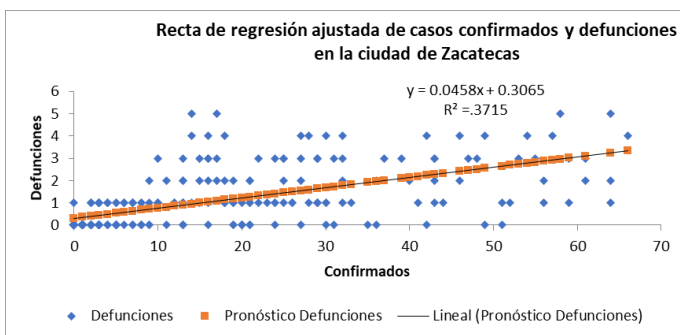


Fig. 4. Curva de regresión ajustada con línea de tendencia de casos confirmados de COVID-19 en la ciudad de Zacatecas.

34.66% de probabilidad de fallecimiento que debe ser explicado por otras causas diferentes del contagio.

4.2 Análisis para el Estado de Zacatecas

Con respecto al estado de Zacatecas, la Figura 3 demuestra la relación existente entre los casos confirmados y las muertes derivadas por COVID-19.

La ecuación representativa está dada por la Ec. (4), es decir, por cada caso confirmado de COVID-19 existe un aumento de 0.051 muertes por esa causa en Zacatecas.

$$\text{Defunciones} = 1.8542 + 0.051 * \text{Confirmados} \quad (4)$$

El resultado muestra una correlación entre el número de defunciones y los casos confirmados por COVID-19 del 73.31%; el modelo cuenta con un coeficiente de determinación (R^2) del 53.75% y explica en ese porcentaje, las defunciones por COVID-19 a partir de los casos confirmados. En este caso, hay un 46.25% de probabilidad de fallecer que requiere explicación distinta al contagio.

4.3 Análisis para la ciudad de Zacatecas

A diferencia de los casos anteriores, y siguiendo la misma metodología, para la capital zacatecana se obtuvo un coeficiente de determinación del 37.15%, un modelo poco consistente; sin embargo, el P-value en el análisis de varianza generó un resultado de $1.608e^{-26}$, es decir, existe una relación estadísticamente significativa entre las defunciones y los casos contagiados de COVID-19. La figura 4 muestra la gráfica de regresión lineal para el análisis de la ciudad de Zacatecas, donde también se puede observar la ecuación resultante que indica que por cada caso confirmado de COVID-19 existe un aumento de 0.045 muertes.

La ecuación representativa de la recta de regresión lineal está dada por la Ec. (5).

$$\text{Defunciones} = 0.3065 + 0.0458 * \text{Confirmados} \quad (5)$$

En todos los casos, el análisis de varianza resultó con un P-value menor que 0.05, por lo que existe una relación estadísticamente significativa entre las defunciones y los casos confirmados de COVID-19, con un nivel de confianza del 95%.

Los resultados del Análisis de Varianza para los modelos ajustados, se presentan en la tabla 1; es importante señalar que conforme aumenta el ámbito para el cual han sido estimados los modelos, aumenta el coeficiente de regresión (B_1) en tanto que el error típico (S_{B_1}) disminuye conforme aumenta el ámbito. Es claro que en este último, hay una importante influencia del tamaño de la muestra.

En consecuencia, el valor de la estadística t-Student es mayor cuanto más grande es el ámbito y con ella, también el nivel de significancia estadística que alcanza cada modelo. Como se señaló, los modelos estimados para México y estado de Zacatecas, arrojaron coeficientes de determinación superiores al 0.5 mientras que el modelo para la ciudad de Zacatecas solo alcanzó un 0.37.

La tabla 2, presenta los resultados del cálculo de la Razón de Letalidad de los Casos (CFR) para los tres ámbitos.

Es evidente que, para fines de divulgación, resulta más sencillo de comprender la interpretación de la Razón de Letalidad de los Casos, pues expresa el porcentaje de fallecimientos por la enfermedad de COVID-19 respecto a la cantidad de casos confirmados. En tanto que el coeficiente de regresión expresa la razón de cambio en la cantidad de fallecimientos, conforme cambia unitariamente la cantidad de casos confirmados, con un

Tabla 1. Coeficiente de regresión, error típico, prueba t y significancia estadística para cada modelo.

Ámbito	B1	S _{B1}	t	p-value
México	.0685	0.00279	24.463	6.729E-75
Edo Zac	.051	0.00297	17.149	2.902E-44
Zacatecas	.0458	0.00380	12.035	1.608E-26

Tabla 2. Resultados para Razón de Letalidad según ámbito

Ámbito	CFR (%)
México	9.54
Edo Zac	8.00
Zacatecas	6.40

adicional, que es el poder predictivo de los modelos de regresión. Ambos resultados son importantes y es deseable que las instituciones responsables de conducir el esfuerzo por la contención de la pandemia, echen mano de las herramientas analíticas que mejor les permitan comprender la complejidad del fenómeno.

Los indicadores propuestos por la OMS tienen la gran ventaja de su aceptación por los miembros de la organización, en tanto que los indicadores que arrojan los modelos, son la propuesta de este trabajo.

5. Conclusiones y recomendaciones

Este análisis desde la perspectiva que permiten los modelos de regresión lineal simple ha permitido visibilizar a partir de la información disponible en CONACYT, que la situación en el México y en el caso particular del estado de Zacatecas, ha continuado a la alta con un número cada vez mayor de personas contagiadas de COVID-19, esta conclusión se deriva de que los coeficientes de los tres modelos, muestran signo positivo; lo que es más preocupante es el incremento del número de defunciones por la misma pandemia, que, a pesar de las acciones que los gobiernos han implementado, no cede y por el contrario, los datos más recientes dan cuenta del evidente rebrote.

De lo expuesto, se desprende la importante necesidad de reponderar la medicina comunitaria y la epidemiología pero además es necesario dotar a los médicos de éstas y otras disciplinas, de herramientas que se sabe, son determinantes en todo empleado de la investigación: la formación estadística. Será determinante el rescate de la información que arrojan las visitas de pacientes a hospitales, diseñar sistemas informáticos modernos y de gran capacidad de análisis para que basados en algoritmos computacionales y poderosos sistemas informáticos, dispongan los sistemas de salud, de la información sobre aspectos de la propagación de enfermedades entre la comunidad. La modelación estadística ha mostrado sus bondades, sin embargo, la consistencia y calidad de la información, así como la periodicidad de la misma, son determinantes en la mejora significativa de la capacidad predictiva de los modelos.

La organización de la sociedad deberá operar de forma tal que, en perfecta sincronía con las autoridades sanitarias, diseñen y dispongan de estrategias para que los tiempos de respuesta ante brotes o amenazas, sean reducidos entre la detección y la acción.

Es clara de la experiencia multicitada de Wuhan, China que, detener la movilidad de las personas, imponer el distanciamiento social y la limpieza de manos fueron determinantes para detener la propagación.

La conducción de la estrategia contra este tipo de problemas sanitarios, deberá estar en manos de personal capacitado y de gran compromiso social, más allá de actores que lejos de aportar para la acción colectiva, se empeñan en inducir incertidumbre y con ella, la desobediencia y sus consecuencias.

A la pregunta de ¿Cómo será nuestra sociedad Post COVID-19?, seguramente deberemos avanzar en comprender mejor a la naturaleza y sus límites para la sustentabilidad humana así como mantener el hecho de que sin salud, pocas cosas funcionan, aún en las economías otrora más pujantes.

Referencias

- [1] J. Ma, «Exclusive | Coronavirus: China's first confirmed Covid-19 case traced back to November 17,» 13 March 2020. [En línea]. Available: <https://www.scmp.com/news/china/society/article/3074991/coronavirus-chinas-first-confirmed-covid-19-case-traced-back>. [Último acceso: 25 de noviembre 2020].
- [2] WHO, «WHO Coronavirus Disease (COVID-19) Dashboard,» World Health Organization, November 2020. [En línea]. Available: <https://covid19.who.int/>. [Último acceso: 28 de noviembre 2020].
- [3] CONACYT, «Covid-19 México,» Gobierno de México, noviembre 2020. [En línea]. Available: <https://datos.covid-19.conacyt.mx/#DownZCSV>. [Último acceso: 28 noviembre 2020].
- [4] M. Medina, «Preliminary Estimate of COVID-19 Case Fatality Rate in the Philippines using Linear Regression Analysis, » 7 Abril 2020.
- [5] A. Abubakar, A. Suleiman, U. Aliyu y S. Abubakar, «Estimation of the case fatality rate of COVID-19 epidemiological data in Nigeria using statistical regression analysis, » *Biosafety and Health*, 9 September 2020.
- [6] W. Mendenhall y T. Sincich, Probabilidad y estadística para ingeniería y ciencias, México: Prentice Hall, 1997.
- [7] Gobierno de México, «Se confirma en México caso importado de coronavirus COVID- 19,» 28 de febrero 2020. [En línea]. Available: <https://www.gob.mx/salud/prensa/077-se-confirma-en-mexico-caso-importado-de-coronavirus-covid-19>. [Último acceso: 28 de noviembre 2020].
- [8] ECDC, «Download the daily number of new reported cases of COVID-19 by country worldwide, » 28 November 2020. [En línea]. Available: <https://www.ecdc.europa.eu/en/publications-data/download-todays-data-geographic-distribution-covid-19-cases-worldwide>. [Último acceso: 28 de noviembre 2020].
- [9] INEGI, «nacionales, Estadísticas a propósito del día mundial de la población (11 de julio) datos,» 9 Julio 2020. [En línea]. Available: https://www.inegi.org.mx/contenidos/saladeprensa/aproposito/2020/Poblacion2020_Nal.pdf. [Último acceso: 25 de noviembre 2020].