Proyecciones

México | Covid-19 semana 16, proyecciones SIR b(t), ARIMA y Comparativo Internacional

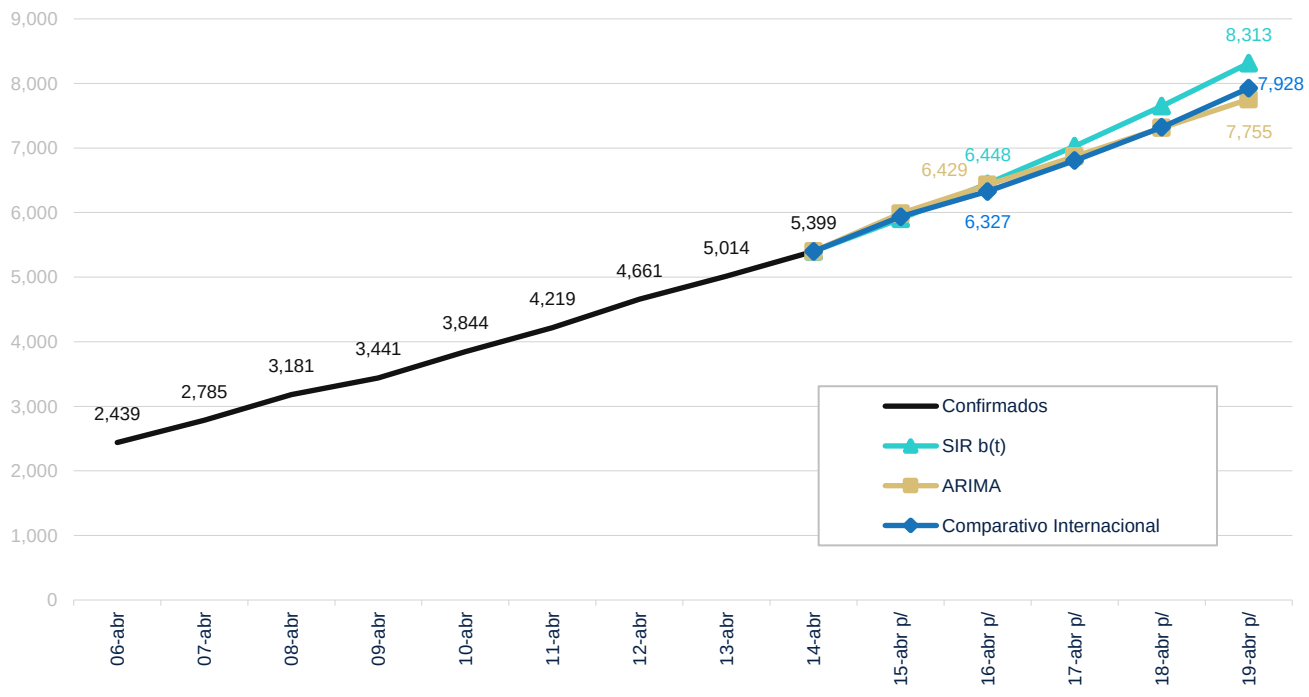
Juan José Li Ng / Carlos Serrano

15 abril 2020

- **Se presentan las proyecciones sobre el comportamiento del número total de casos de Covid-19 en México para la semana 16 de 2020, del 15 al 19 de abril.**
- **Para el término de la semana, los casos nuevos diarios podrían superar los 600.**
- **El jueves 16 de abril se proyecta que México podría tener entre 6,327 y 6,448 casos confirmados.**
- **El domingo 19 de abril se proyecta que México podría tener entre 7,755 y 8,313 casos confirmados.**

Se presentan las proyecciones sobre el comportamiento del número total de casos de Covid-19 en México para la semana 16 de 2020, de los días comprendidos del 15 al 19 de abril. Las proyecciones simplemente modelan el escenario posible a partir del comportamiento de datos anteriores observados.

Gráfica 1. **Síntesis de las proyecciones de casos confirmados totales de Covid-19 para México, del 15 al 19 de abril de 2020**



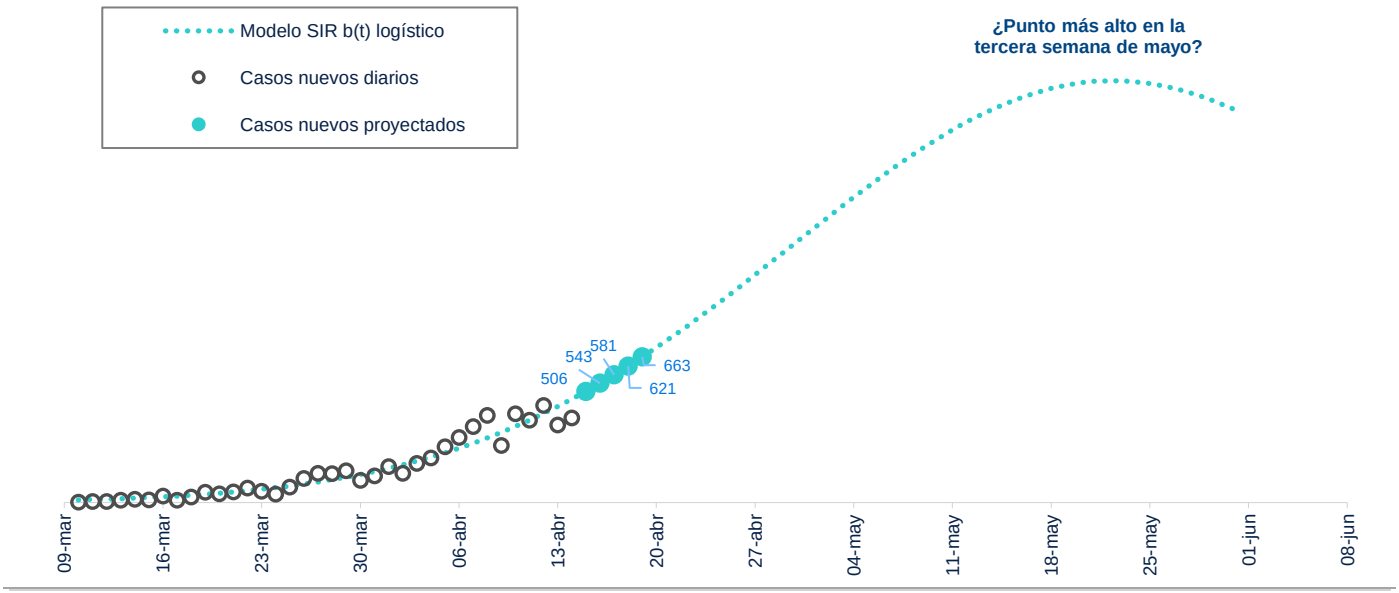
Fuente: BBVA Research. p/ proyección

1. Resultados modelo SIR b(t) logístico

SIR es la abreviatura de la familia de modelos con la estructura Susceptibles – Infectados – Removidos. A continuación, se muestran los resultados del modelo SIR b(t) logístico en donde los parámetros fueron ajustados considerando la evolución de los casos totales de Covid-19 en Alemania, Francia y Reino Unido y de otros países, principalmente asiáticos y europeos. Se observa que el punto más agudo del padecimiento en estos países se llega entre 80 a 100 días después de los primeros 5 casos reportados. El modelo se calibra considerando la temporalidad del punto más alto de casos nuevos o punto de inflexión, no considera el número de casos observado en estos países. Cabe señalar que el modelo fue ajustado para cuantificar y proyectar casos confirmados, si bien sabemos, e incluso las autoridades han reconocido, que el número total de individuos infectados es mucho mayor.

Así, con los datos actuales y considerando la evolución en otros países de este padecimiento, el modelo SIR ajustado para México indica que se podría **alcanzar el punto más alto de casos nuevos por Covi-19 durante la tercera semana de mayo**. Sin embargo, se observa que el crecimiento de los casos nuevos en México es más lento que en otros países que se encuentran en una fase más avanzada del ciclo de este padecimiento, por lo que el tiempo para llegar al punto más alto podría prolongarse.

Gráfica 2. **Proyecciones de casos nuevos confirmados de Covid-19 del modelo SIR b(t) logístico para México, del 15 al 19 de abril de 2020**

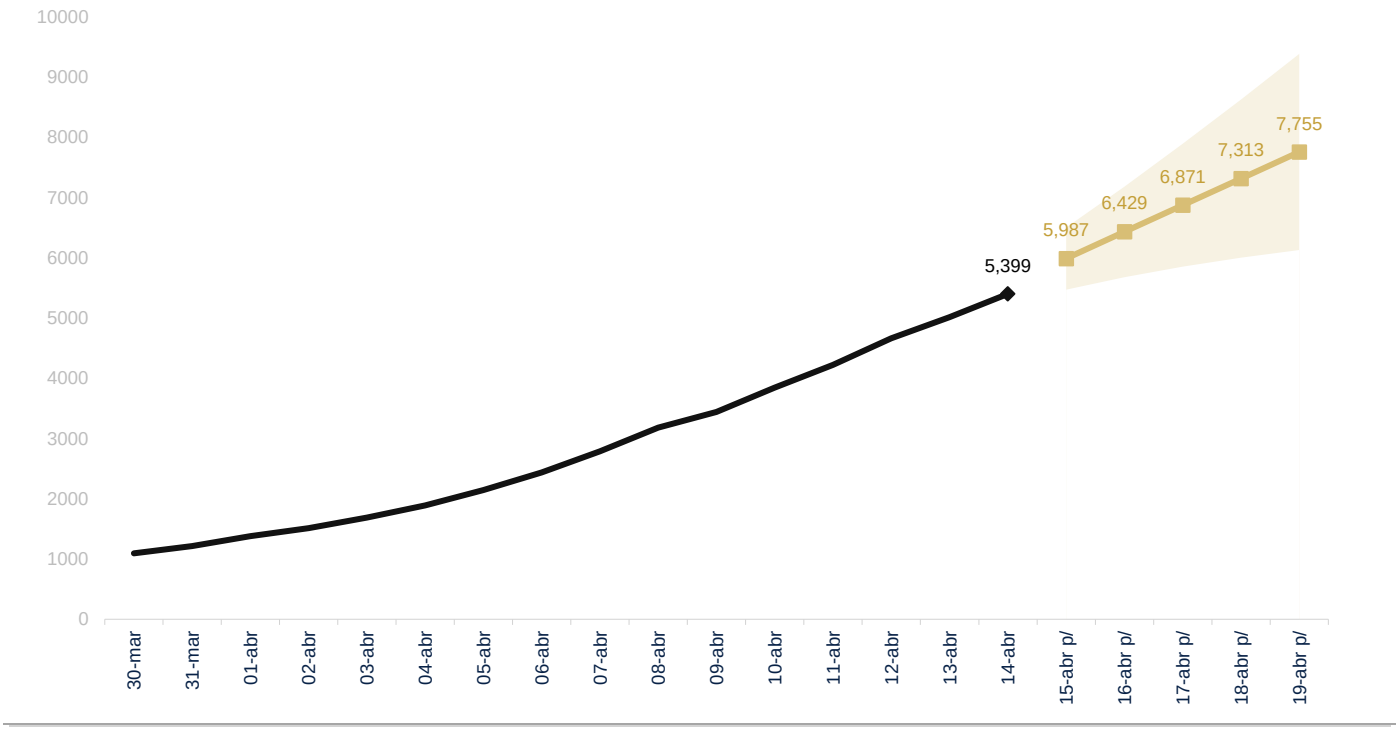


Fuente: BBVA Research. p/ proyección

2. Resultados modelo ARIMA

ARIMA son las siglas de los modelos AutoRegresivo Intregrado de Media Móvil, que es una técnica de series de tiempo en econometría. De continuar la tendencia actual observada, estos serían los resultados de la media de la mejor estimación ARIMA, así como su intervalo de confianza al 95%.

Gráfica 3. **Proyecciones de casos confirmados totales de Covid-19 del modelo ARIMA para México, del 15 al 19 de abril de 2020**

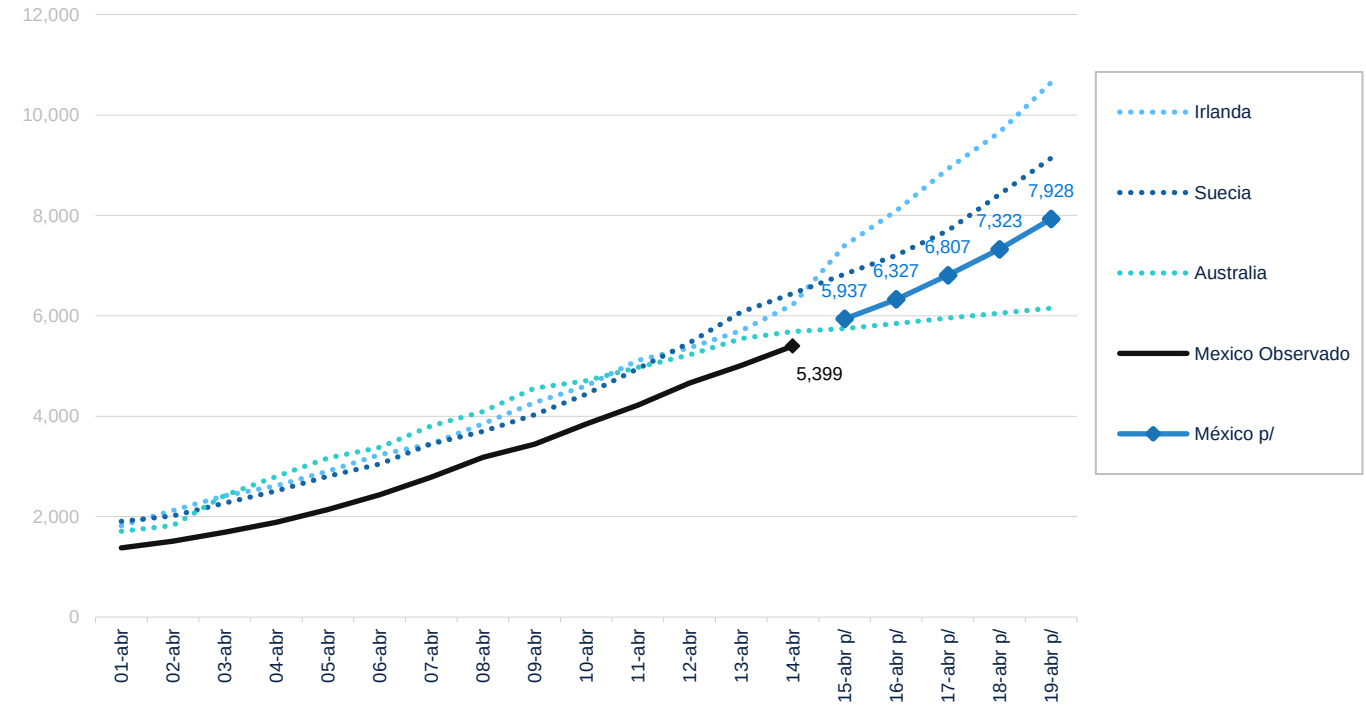


Fuente: BBVA Research. p/ proyección

3. Resultados Comparativo Internacional

Al comparar más de 180 países o regiones, se encuentra similitud entre los datos de casos confirmados totales de Covid-19 en México entre el 1º. y 14 de abril, con la experiencia observada previamente en la evolución de este padecimiento en Irlanda, Suecia y Australia. Después de hacer el empalme temporal con cada uno de estos países para este periodo de 14 días, se presentan los datos observados que tuvieron estos países los siete días posteriores y se promedian los datos para proyectar cómo podría evolucionar este padecimiento en México.

Gráfica 4. **Proyecciones de casos confirmados totales de Covid-19 por Comparativo Internacional para México, del 15 al 19 de abril de 2020**



Fuente: BBVA Research. p/ proyección

Anexo Metodología

1. Sobre los datos

Los datos para los casos diarios de Covid-19 se obtienen del Centro Europeo para la Prevención y Control de Enfermedades (ECDC, por sus siglas en inglés). Desde el inicio de la pandemia, el ECDC ha recolectado información sobre el número de casos y muertes por Covid-19 a partir de datos oficiales de las autoridades de salud alrededor del mundo.

Diariamente, entre las 6:00 y 10:00 horas del Centro de Europa (CET) un equipo del ECDC analiza fuentes de información de todo el mundo y las valida y documenta en una base de datos. Posteriormente una consulta de esta información con los casos y defunciones por día se publica diariamente en su página de internet. Esta base contiene registros de más de 180 países o regiones de todo el mundo (ECDC, 2020).

Debido a la diferencia horaria con México, los datos correspondientes a este país tienen algunas horas de retraso. Las modelaciones para México se realizan con la base de datos del ECDC, pero se hace un ajuste en caso de que se tenga información más actualizada publicada en el reporte y/o rueda de prensa diarios del Covid-19 que presenta la Secretaría de Salud de México (Hernández, 2020; SSA, 2020).

La detección de los primeros casos confirmados en cada país puede concentrarse en pocos días o tener una dispersión amplia. Esto puede afectar la modelación, por lo que se decidió crear una serie acotada y tomar los datos a partir de *d* número de casos para los modelos SIR y Comparativo Internacional.

2. Modelo SIR b(t) logístico

El modelo SIR es el más usado en epidemiología para modelar la dinámica de una nueva enfermedad dentro de una población determinada. Es un modelo clásico que se le atribuye al estudio publicado por Kermack y McKendrick (1927) sobre la epidemia de peste bubónica en Mumbai a principios del siglo XX (Perasso, 2018). Consiste básicamente en tres componentes (Jones, 2007; Kawata y Kyo, 2013):

- **Susceptibles (S),** número de individuos completamente susceptibles que pueden contraer el padecimiento;
- **Infectados (I),** número de individuos que están infectados y que pueden infectar a otros; y
- **Removidos (R),** número de individuos que dejan de pertenecer a la población infectada. Dependiendo de las especificaciones, esta categoría puede variar en su acepción y uso. En general, este grupo se integra por las personas que se encuentran en cuarentena y aislamiento voluntario, más los hospitalizados y más las defunciones.

El modelo toma su nombre por las iniciales de estos tres componentes. El modelo SIR supone que las personas en la categoría de removidos ya no pueden infectar a más personas, y que adquieren inmunidad a la enfermedad,

que puede ser de carácter temporal, por lo que no regresan a la población susceptible[1]. En esta versión del modelo SIR no se consideran las dinámicas vitales, es decir los nacimientos y las defunciones que no se atribuyen a la enfermedad (Hethcote, 1989). Así, cada persona de la población total (N) se encuentra, en cierto tiempo (t), únicamente dentro de uno de los tres grupos anteriormente mencionados; es decir:

$$N = S(t) + I(t) + R(t)$$

en donde $N, S(t), I(t), R(t) \in \mathbb{N}\ \forall t \geq 0$.

La dinámica del modelo inicia con la infección de la primer o las primeras persona(s) (I_0), que se infectaron de manera exógena al modelo (caso cero o casos importados). Cada persona infectada puede infectar a la población susceptible a razón de la tasa $\beta \cdot S/N$. Nótese que mientras mayor es el número de personas infectadas, mayor va ser la población susceptible que se puede infectar en el siguiente periodo. Por este efecto, la población susceptible disminuye en $\beta \cdot S/N \cdot I$, mientras que el total de personas infectadas aumenta en esa cantidad.

Cada persona infectada tiene una tasa de probabilidad de pasar al grupo de removidos igual a γ. Por esta dinámica, en cada periodo disminuye el total de personas infectadas en γI, las cuales pasan al grupo de personas removidas.

La dinámica del modelo se puede expresar a partir del siguiente sistema de ecuaciones diferenciales ordinarias (EDO)[2] de primer orden:

$$\frac{dS}{dt} = -\beta \frac{S}{N} I$$

$$\frac{dI}{dt} = \beta \frac{S}{N} I - \gamma I$$

$$\frac{dR}{dt} = -\gamma I$$

Así, se cumple la siguiente condición:

$$\frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = 0$$

La condición inicial del sistema está dada generalmente por $\{S = N - I_0,\ I = I_0,\ R = 0\}$.

[1] Los modelos SIS (Susceptibles – Infectados - Susceptibles) se usan cuando no se genera inmunidad en la población después de la infección (Kawata y Kyo, 2013).
[2] En Borrelli y Coleman (2005) se puede consultar en qué consisten y cómo se resuelvan las EDO.

2.1. Número básico de reproducción (R_0)

En el modelo SIR, el número básico de reproducción (basic reproduction number o R_0) está dado por:

$$R_0 = \frac{\beta}{\gamma}$$

R_0 se define como el número esperado de nuevos individuos infectados de la población susceptible producidos por una persona a lo largo de su periodo infeccioso. Diversas estimaciones indican que el COVID-19 tiene un R_0=2.2, aunque otros autores encuentran que puede ser de 3.49 (Riou, 2020; Sanche 2020; Wang et. al., 2020). Esto quiere decir que cada persona infectada puede transmitir el virus y enfermar en promedio a otras 2.2.

El número básico de reproducción es fundamental en los modelos epidemiológicos, al cual son muy sensibles. Como se señala en Perasso (2018), la búsqueda de una herramienta para medir el riesgo de una enfermedad y prevenir la aparición, o crecimiento, de una epidemia, llevó al concepto de R_0. En general:

- **Si $R_0 \leq 1$**, se puede tener un Equilibrio Libre de Enfermedad (Disease-Free Equilibrium o DFE) que es globalmente estable, lo que llevará a que la enfermedad eventualmente desaparezca.

- **Si $R_0 > 1$**, el DFE es inestable y se llega a un Equilibrio Endémico (Endemic Equilibrium o EE) que es globalmente estable. Dependiendo de la magnitud de R_0, el EE puede implicar que una proporción importante de la población susceptible contraiga la enfermedad.

Cuando la proporción de la población susceptible es muy alta, en donde es casi nula la inmunidad, como para el caso del virus SARS-Cov-2 causante de la enfermedad Covid-19, las acciones de política pública y de la sociedad deben buscar reducir las posibilidades de contagio. En la actual pandemia por Covid-19, se han implementado medidas que buscan evitar conglomeraciones, realizar aislamientos y promover el distanciamiento social para disminuir el número de nuevos casos y de defunciones. Estas acciones serán efectivas si logran disminuir la magnitud de R_0. Cuando estas políticas y medidas aplicadas logran hacer que $R_0 \leq 1$, se estará muy próximo al pico máximo de casos nuevos y se observará una disminución gradual de las nuevas infecciones.

Cuando el número de infecciones llegue a un nivel bajo, los gobiernos y la sociedad tenderán a dejar o relajar las restricciones impuestas, con lo que aumentará R_0. Si esto se hace de manera anticipada y/o sin las políticas y medidas adecuadas para mantener $R_0 \leq 1$, se puede llegar a un nuevo brote del padecimiento.

A largo plazo, un equilibrio globalmente estable en una “vida cotidiana normal” solo se puede lograr con la inmunización de la población susceptible a través de la vacunación o de un tratamiento efectivo.

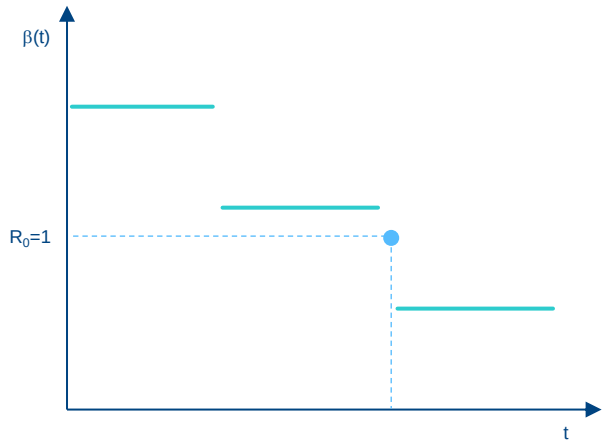
2.2. $R_0(t)$ y la función $\beta(t)$

Ante el brote de una enfermedad de gran trascendencia como lo es el COVID-19, los gobiernos y las sociedades no son impasibles, sino que reaccionan para eliminar o contener la transmisión del padecimiento, lo que hace que el número básico de reproducción (R_0) disminuya. Estas acciones no se hacen en una sola intervención discreta, sino que puede ser en diferentes etapas o en acciones y medidas incrementales a lo largo del tiempo.

Para una adecuada proyección de la evolución del padecimiento, es necesario modelar la magnitud y evolución de R_0. Para el modelo SIR, en el tiempo t el número básico de reproducción (R_0) está en función de los parámetros β y γ, es decir: $R_0(\mathrm{t}) = \beta(t)/\gamma(t)$. Por simplicidad mantendremos fijo $\gamma(t) = \gamma$, por lo que $R_0(\mathrm{t}) = \beta(t)/\gamma$, y así los

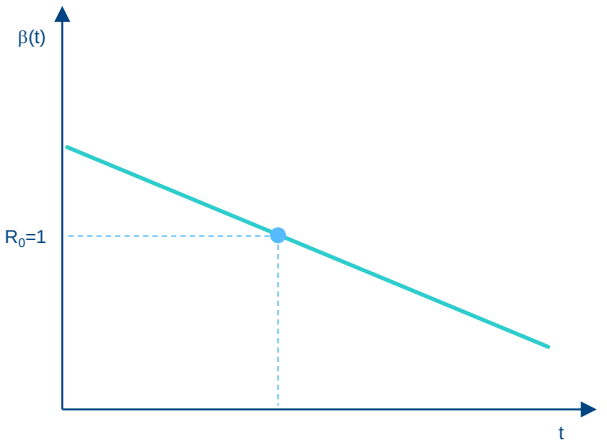
cambios se modelarán en una variable. En las siguientes gráficas se muestran algunos casos de formas funcionales que podría tener la función $\beta(t)$, y por tanto $R_0(t)$.

Gráfica 1. **Función $\beta(t)$ discreta**



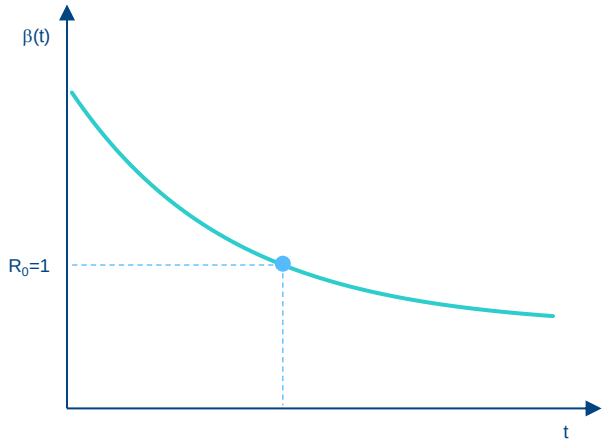
Fuente: BBVA Research.

Gráfica 2. **Función $\beta(t)$ lineal**



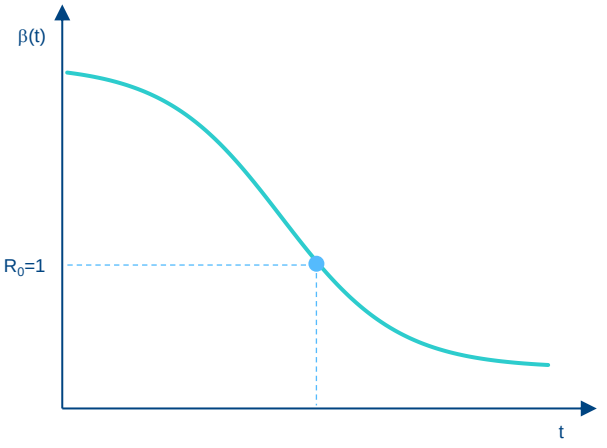
Fuente: BBVA Research.

Gráfica 3. **Función $\beta(t)$ exponencial**



Fuente: BBVA Research.

Gráfica 4. **Función $\beta(t)$ logística**



Fuente: BBVA Research.

La primera gráfica representa una función $\beta(t)$ con dos intervenciones discretas que disminuyen la transmisión de la enfermedad. Esta forma funcional podría ser adecuada para modelar intervenciones en fases o etapas, sin embargo, puede tener la desventaja de que la magnitud y los resultados de las intervenciones pueden variar significativamente entre países, además de que la evidencia en otros países en los cuales la epidemia broto antes, los resultados se observaron de forma gradual y no discreta. Las siguientes tres gráficas muestran formas funcionales de impactos graduales o continuos: forma lineal, exponencial y distribución logística. Se descartó el

uso de la forma lineal dado que en los algoritmos de optimización se obtenían algunas veces valores para $\beta(t)$ fuera de los parámetros aceptables. Se calibraron los modelos usando los últimos dos tipos de funciones y finalmente para este estudio se decidió elegir la forma funcional logística.

La función de densidad logística se define con dos parámetros: μ que define la localización o la media, y s que define la escala. La varianza es proporcional a la escala y está dada por: $\sigma^2 = \pi^2 s^2/3$. El modelo SIR (t) logístico se calibra con la información de estos parámetros observada en países en donde el padecimiento se encuentra más avanzado. Este modelo solo se usa para analizar las partes de crecimiento y descenso del padecimiento durante un periodo incremental de políticas y acciones de restricción social para contener el padecimiento. Cuando se empiezan a levantar o relajar estas medidas, este modelo ya no es el más adecuado.

3. Modelo ARIMA

Tras analizar los datos del padecimiento, se encontró que se pueden modelar de forma relativamente adecuada en el corto plazo a través de técnicas de series de tiempo. Para las variables univariadas analizadas (y_t) se optimizó con el mejor proceso AutoRegresivo Intregrado de Media Móvil de orden p, d y q, o ARIMA(p, d, q), con base en el Criterio de Información de Akaike (AIC). Una variable y_t es un proceso ARIMA(p, d, q) si:

$$\beta(L)\ (1-L)^d y_t = \theta(L)\epsilon_t$$

en donde L es el operador rezago, y

$$\beta(L) = 1 - \beta_1 L\ - \beta_2 L^2\ - \cdots - \beta_p L^p$$

$$\theta(L) = 1 - \theta_1 L\ - \theta_2 L^2\ - \cdots - \theta_q L^q$$

$$\epsilon_t \sim i.i.d.(0, \sigma^2)$$

Dado lo reciente del inicio de la serie de tiempo y de su comportamiento observado, no se consideraron los ajustes por estacionalidad.

4. Modelo Comparativo Internacional

Este modelo busca predecir el comportamiento del padecimiento de un país a partir de la serie de tiempo de países que se encuentran en una etapa más avanzada. Se tiene información diaria de más de 180 países o regiones del mundo de la base de datos de la ECDE. Para cada país y considerando diferentes rezagos, se elige el mejor ajuste a partir de minimizar la suma de los cuadrados residuales. Posteriormente, se hace un promedio de los países que presentan el mejor ajuste.

5. Referencias

Borrelli, Robert L. y Courtney S. Coleman (2005). *Ecuaciones diferenciales. Una perspectiva de modelación.* Oxford AlfaOmega. México.

Centro Europeo para la Prevención y Control de Enfermedades [ECDC] (2020). *Data on the geographic distribution of COVID-19 cases worldwide. Información diaria.* Recuperado de: https://www.ecdc.europa.eu/en/publications-data/download-todays-data-geographic-distribution-covid-19-cases-worldwide.

Hernández, Rodrigo (2020). *COVID-19 en México: casos por entidad y cronología del coronavirus.* La Razón. Recuperado de: https://www.razon.com.mx/mexico/coronavirus-en-mexico-casos-por-entidad-federativa-y-historico-de-casos-covid-19-muertes-secretaria-de-salud-en-que-fase-de-contingencia-esta-mexico/

Hethcote, Herbert W. (1989) *Three Basic Epidemiological Models*. En: Levin S. A., T. G. Hallam y L. J.Gross L.J. (editores) *Applied Mathematical Ecology*. Springer. Berlin.

Jones, James H. (2007). *Notes on R0.* Department of Anthropological Sciences. Stanford University.

Kawata, Yukichik and Koki Kyo (2013). *Analysis methods for infectious disease using mathematical models.* Munich Personal RePEc Archive Paper No. 68939.

Kermack W. O. y A. G. McKendrick (1927). *A contribution to the Mathematical Theory of Epidemics*. Proceedings of the Royal Society of London. Vol. 115, Núm. 772, pp. 700-721.

Perasso, Antoine (2018). *An introduction to the basic reproduction number in mathematical epidemiology.* ESAIM: Proceedings and surveys. Septiembre, Vol. 62, pp. 123-138.

Riou, Julien y Christian Althaus (2020). *Pattern of early human-to-human transmission of Wuhan 2019 novel coronavirus (2019-nCoV), December 2019 to January 2020.* Euro Surveillance. Vol. 25, Núm 4.

Sanche, Steven, Yen Tin Lin, Chonggang Xu, Ethan Romero-Severson, Nick Hengartner, y Ruian Ke (2020). *High contagiousness and rapid spread of severe acute respiratory syndrome coronavirus 2*. Emerging Infectious Diseases. Vol. 26, Núm. 7.

Secretaría de Salud [SSA] (2020). *Nuevo Coronavirus en el mundo COVID-19. Comunicado Técnico Diario.* Recuperado de: https://www.gob.mx/salud/es/archivo/prensa

Wang, Y., X. Y. You, Y. J. Wang, L. P. Peng, Z. C. Du, S. Gilmour, D. Yoneoka, J. Gu, C. Hao, Y. T. Hao y J. H. Li (2020). *Estimating the Basic Reproduction Number of COVID-19 in Wuhan, China.* Zhonghua liu xing bing xue za zhi = Zhonghua liuxingbingxue zazhi.

AVISO LEGAL

Este documento ha sido preparado por BBVA Research del Banco Bilbao Vizcaya Argentaria, S.A. (BBVA) y por BBVA Bancomer. S. A., Institución de Banca Múltiple, Grupo Financiero BBVA Bancomer, por su propia cuenta y se suministra sólo con fines informativos. Las opiniones, estimaciones, predicciones y recomendaciones que se expresan en este documento se refieren a la fecha que aparece en el mismo, por lo que pueden sufrir cambios como consecuencia de la fluctuación de los mercados. Las opiniones, estimaciones, predicciones y recomendaciones contenidas en este documento se basan en información que ha sido obtenida de fuentes estimadas como fidedignas pero ninguna garantía, expresa o implícita, se concede por BBVA sobre su exactitud, integridad o corrección. El presente documento no constituye una oferta ni una invitación o incitación para la suscripción o compra de valores.