

# Modelando la Covid-19 con solo una ecuación

## Modeling Covid-19 with a single equation

Celia T. González González<sup>1</sup>, Mariano Rodríguez Ricard<sup>2\*</sup>

**Resumen** Nuestro interés en este trabajo es la presentación de un modelo de dinámica de epidemia que consta de una sola ecuación. Se muestra la equivalencia del modelo con los clásicos del tipo SIR. Se destaca la conveniencia del modelo para una epidemia de la que se conoce el período medio de recuperación. Se muestra un procedimiento para el ajuste de parámetros de acuerdo a datos dados, y en particular, los que se consideraron cuando modelamos la epidemia de la Covid-19 en Cuba.

**Abstract** Our concern in this paper is to present a dynamic model of an epidemic disease with a single equation. We show the equivalence of this model with the classic SIR-type. The convenience of this model for an epidemic for which it is known the recovery period, is highlighted. We show a method for the estimation of parameters for a given data, in particular, those considered when we modeled the Covid-19 in Cuba.

### Palabras Clave

modelo SIR, dinámica de epidemia

### Keywords

SIR model, epidemiological dynamics

<sup>1</sup>Departamento de Matemática, Universidad de La Habana, La Habana, Cuba, [celia@matcom.uh.cu](mailto:celia@matcom.uh.cu)

<sup>2</sup>Departamento de Matemática, Universidad de La Habana, La Habana, Cuba, [rricard@matcom.uh.cu](mailto:rricard@matcom.uh.cu)

\*Autor para Correspondencia, Corresponding Author

## Introducción

Son bien conocidos los modelos de la dinámica de epidemias basados en un contacto aleatorio entre la población de infecciosos y la población de susceptibles [1]-[3]. El ejemplo más simple [4] es el conocido como modelo SIS (susceptible-infeccioso-susceptible), y que es representado por el sistema

$$dS/dt = \gamma I - \beta SI ; dI/dt = -\gamma I + \beta SI \quad (1)$$

donde  $S + I = 1$ ,  $S$  es la fracción de susceptibles e  $I$  es la fracción de infecciosos respecto de una población cuya cantidad es conocida,  $\beta$  es la tasa de contagio (o de contacto), y  $\gamma$  es la tasa de recuperación. Se asume en este modelo que los recuperados se tornan susceptibles nuevamente.

Otro modelo muy empleado es el llamado SIR en donde se asume que los recuperados no se vuelven a infectar. El modelo típico es el de Kermack-McKendrick [2],

$$dS/dt = -\beta SI ; dI/dt = \beta SI - \gamma I ; dR/dt = \gamma I \quad (2)$$

en la que se asume nuevamente que  $S + I + R = 1$ , y se refieren las fracciones a un tamaño de población constante. En ocasiones, el sistema ec.2 se complementa con una ecuación asociada a la fracción de fallecidos (removidos)  $\tilde{R}$  a partir de una tasa de letalidad  $\eta$ , resultando,

$$d\tilde{R}/dt = \eta dI/dt , \quad (3)$$

de modo que  $\gamma + \eta = 1$ , y así los no recuperados son los que fallecen. El sistema ecs.2-3 se refiere ahora a las fracciones de población representadas en la suma  $S + I + R + \tilde{R} = 1$ .

Como que en estos modelos se asume que la población total es constante y nos referimos a fracciones de población, resulta cómodo reducir variables en el sistema, por ejemplo eliminando la fracción de los susceptibles. Esto puede hacerse en cualquiera de los sistemas mencionados anteriormente.

El modelo que consideramos en este trabajo resulta de asumir conocido el período de recuperación, lo que permite eliminar también la tercera ecuación del sistema ec.2. Las ecuaciones son todas nildimensionales, y para ello hemos considerado fracciones de población respecto de una población total fija conocida, y una escala temporal de un día. Así, los parámetros en las ecuaciones tampoco tienen dimensiones físicas. Estos detalles los veremos en la Sección 1. En la Sección 2 se muestra la función objetivo y la manera en que se pueden estimar los parámetros de la epidemia. En la Sección 3 se muestran algunas figuras a partir de datos oficiales recopilados. Al final, en la Sección 4, damos unas breves conclusiones del trabajo.

## 1. La ecuación del modelo

Por el momento, consideremos el sistema ec.2, en el que vamos a asumir que la fracción de recuperados se incluye en la de los susceptibles, resultando el sistema ec.1. Así, básicamente hay dos poblaciones, la de los no infecciosos (susceptibles

+ recuperados) representada por  $S(t)$  y la de los infecciosos (sintomáticos + asintomáticos) representada por  $I(t)$ . Se destaca que parte de la población de infecciosos pasa, al cabo de un cierto período de tiempo, al grupo de los recuperados. De esta manera,  $S(t) + I(t) = 1$ , y podemos eliminar la variable  $S$ , quedando la ecuación para los infecciosos (= infectados) en la forma

$$dI/dt = I(\sigma - \beta I) \quad (4)$$

en la que hemos considerado el parámetro

$$\sigma = \beta - \gamma \quad (5)$$

el que debe ser positivo para que exista una población de infecciosos. Esta ecuación tiene solo dos puntos estacionarios,  $I = 0$  e  $I = \sigma/\beta$ , el primero es inestable mientras que el segundo es asintóticamente estable. Recordando que  $I$  representa una fracción de población, se concluye que el segundo es el valor máximo de la fracción de población de infecciosos. La interpretación de este parámetro  $\sigma$  se aclara a continuación. La solución de la ec.4 resulta fácil de obtener, y se puede representar en la forma siguiente:

$$I(t) = L\kappa \frac{\exp(\sigma t)}{\kappa \exp(\sigma t) - \kappa + 1} \quad (6)$$

Aquí  $L = \lim_{t \rightarrow \infty} I(t)$  es el valor máximo de la fracción de infecciosos, y  $\kappa$  representa la fracción inicial de infecciosos relativa al total de infecciosos posible, esto es,  $I(0) = L\kappa$ . Por esta razón,  $\kappa$  resulta ser también un parámetro desconocido aunque se conozca el número de infectados en el momento inicial. De esta forma, para describir la dinámica de la epidemia tenemos que identificar tres parámetros:  $L$ ,  $\sigma$ ,  $\kappa$ . En este momento es conveniente aclarar la relación entre el número  $L$  que representa el máximo de la fracción infecciosos, y el bien conocido número reproductivo básico  $R_0$  de la dinámica de poblaciones. De la ec.5 se sigue

$$L = \frac{\sigma}{\beta} = 1 - R_0^{-1} \quad (7)$$

con lo cuál la condición evidente para que persista la propagación de la epidemia, es decir  $L > 0$ , resulta ser equivalente a que  $R_0 > 1$ .

Resta por ver la cuantificación de los recuperados a partir de la evolución en la ec.6, o de manera equivalente, la determinación de la cantidad de individuos infectados que se mantienen *activos* en un momento dado. Un elemento adicional que es brindado por los especialistas de salud es el llamado *período medio  $T$  de recuperación* de la enfermedad. Conociendo este período medio de recuperación, podemos introducir la fracción de infectados activos  $Y(t)$  en un instante de tiempo  $t > T$ , la que estará dada por

$$Y(t) = I(t) - I(t - T). \quad (8)$$

Precisamente es la curva dada por el gráfico de la función  $Y(t)$  la que interesa en la predicción, de modo que en ella se representa la cantidad de infecciosos activos en cada momento. Se

remarca el hecho de que con los tres parámetros mencionados en la ec.6 y el conocimiento del período medio de recuperación, ya tenemos la *curva* de la epidemia. Los tres primeros parámetros hay que determinarlos de los datos recopilados, mientras que  $T$  debe determinarse por los especialistas a partir del conocimiento de la enfermedad.

Complementariamente, la letalidad de la epidemia solo puede determinarse a partir de los datos que se tengan en cada lugar de infectación, y depende de los recursos hospitalarios que estén a disposición de los enfermos, y de las condiciones físicas de cada infectado. Eso hace que solo estimados estadísticos pudieran dar algún valor razonable de la tasa de letalidad en la ec.3. De ese modo, la tasa de letalidad no tiene que incluirse dentro de los parámetros a determinar en el sistema, pues se determina de manera independiente.

## 2. Estimación de parámetros

Tenemos que determinar la función  $I(t)$  dada en la ec.6 y, para ello contamos con el conocimiento de los datos recopilados para intentar hacer el ajuste de los parámetros que permita que el gráfico de la función  $I(t)$  brinde la mejor aproximación posible a los datos oficiales sobre la cantidad de individuos que se han contagiado en el transcurso de la epidemia. Esto es lo mismo que decir que los puntos sobre el gráfico *fracción de población infectada vs tiempo* que corresponden a los datos oficiales diarios disten lo menos posible de la curva de la función  $I(t)$ .

Naturalmente, hay muchas formas de enfocar este problema y algunas de ellas muy simples. Sin embargo, podemos considerar el siguiente enfoque. Sea la desviación cuadrática media

$$F(L, \sigma, \kappa) = \frac{1}{M} \left( \sum_{i=1}^M d^2(a_i, I(t_i)) \right)^{1/2} \quad (9)$$

dada para un valor fijo de  $t$ . Ahora  $a_i$  representa la fracción de infecciosos en el  $i$ -ésimo día seleccionado ( $t_i$ ) obtenida de los datos oficiales, mientras que  $M$  representa el total de días considerados. Además,  $d(\cdot, \cdot)$  representa una distancia en  $\mathbb{R}$ . El problema consiste en encontrar un juego de valores de los parámetros para el cuál la función objetivo en la ec.9 alcance un valor mínimo local en el espacio de parámetros. Puede que exista más de un mínimo local para esta función objetivo, pero al menos en el ejemplo que nos ocupa, se pudo identificar uno solo. Pudiera también ocurrir que no se garantice la convergencia al mínimo por un método de aproximaciones sucesivas dado, pero no fue ese el caso en el ejemplo estudiado. Otro aspecto a destacar es que el juego de parámetros en donde se logra la mejor aproximación depende de la selección de los valores y días seleccionados. Por supuesto, debería ocurrir que, a medida que se consideren más puntos, los juegos de parámetros que representen la mejor aproximación se encuentren cerca de un cierto *conjunto de atracción* en el espacio de parámetros y el valor mínimo alcanzado por la función objetivo será menor, al precio de cálculos cada vez mayores. Nuestro interés es solo indicar la tendencia de la dinámica para que se puedan tomar decisiones apropiadas. Un papel

muy importante lo tiene el exponente de crecimiento  $\sigma$  y la determinación de la cantidad máxima de individuos  $R$  de los que resulten ser infectados a lo largo de la epidemia.

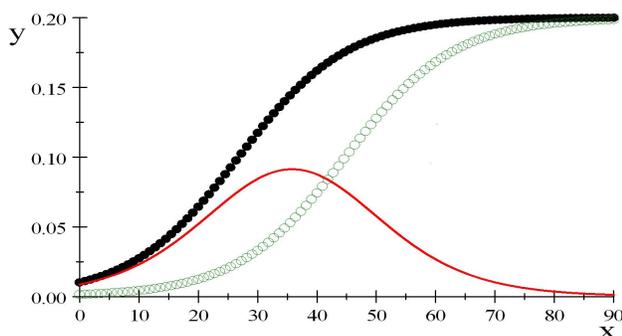
Se puede obtener la fórmula para encontrar el valor  $t = \tilde{t}$  en el que alcanza el máximo de la función  $Y(t)$  que describe la fracción de infecciosos activos. Esto se logra derivando respecto de la variable  $t$  en la ecuación ec.8 teniendo en cuenta la fórmula ec.6, e igualando a cero. Así

$$\tilde{t} = \frac{1}{2\sigma} \ln \left( \frac{1 - 2\kappa}{\kappa^2} e^{\sigma T} \right). \quad (10)$$

Se aclara que la tasa de letalidad en la ec.3 también puede ser otro parámetro más a determinar teniendo en cuenta las cifras oficiales de los fallecidos, a través de una función objetivo apropiada. Entendemos que los valores obtenidos en ese caso pueden ser muy diferentes de los que se hayan tenido en realidad.

### 3. Modelado de la Covid-19

En esta Sección mostraremos diferentes gráficos que fueron obtenidos en la predicción de la curva de infecciosos activos para la epidemia de la Covid-19 en Cuba, lo que a su vez servirá de ilustración de la metodología indicada.



**Figura 1.** Línea de puntos oscuros= $I(t)$ ; Línea de puntos claros= $I(t-T)$ ; Línea continua= $Y(t)$ . El eje X representa los días considerados, mientras que el eje Y representa la fracción de infectados. La fracción de infectados se consideró sobre una población de referencia de 10000 individuos.

Los cálculos para la obtención del mínimo local de la función objetivo se hicieron considerando que el día inicial fue el 03/04/2020 y se consideraron los datos de 30 días posteriores para la construcción de dicha función. Para determinar el mínimo local de la función objetivo se empleó una variante del Método de Newton para la obtención de las sucesivas aproximaciones.

El punto en el espacio de parámetros en donde se alcanzó el mínimo de la función objetivo fue:

$$L = 0,2005 ; \kappa = 0,0505 ; \sigma = 0,1095,$$

y el valor de esa función en el punto de mínimo fue

$$F = 5,0616 \times 10^{-4}.$$

Se consideró en los cálculos que el período medio de recuperación fue de 18 días, resultando de la ec.10 la igualdad

$$\tilde{t} = \frac{1}{2\sigma} \ln \left( \frac{1 - 2\kappa}{\kappa^2} \right) + 9. \quad (11)$$

El gráfico en la fig.1 corresponde a los valores anteriores. En ese gráfico aparecen tres curvas: la de puntos oscuros corresponde al gráfico de la curva que representa la fracción de individuos infectados, la de puntos claros corresponde a la fracción de infectados que dejan de serlo, bien por que se han recuperado o porque han fallecido. Aunque de la figura 1 se puede estimar el tiempo  $\tilde{t}$  en donde se alcanza el máximo de infecciosos activos, también se puede determinar de la ec.11, obteniéndose que  $\tilde{t} = 36$  días aproximadamente a partir del día inicial, es decir, del 3 de abril del 2020.

La fracción de fallecidos se estimó a partir de una tasa de letalidad que es estimada por vía estadística, y que en aquél momento fue  $\eta = 0,04$ . Digamos que la fracción de fallecidos (o removidos) fue obtenida de

$$\tilde{R}(t) = \eta I(t). \quad (12)$$

Notemos que a partir de la ec.6, la fórmula en la ec.12 permite una estimación del número de fallecidos en cada día.

### 4. Conclusiones

Las conclusiones que se obtienen del sistema de Kermack-McKendrick son equivalentes, para un mismo tamaño total de población, a las que se obtienen mediante la ecuación sencilla ec.4. La dinámica asociada a esta ecuación es bien conocida, se identifica usualmente como la *ecuación logística* y que se interpreta de maneras diferentes según sea el fenómeno que se esté modelando. Lo básico que hemos considerado es la existencia de dos clases: la de los susceptibles, que incluye a los recuperados; la de los infecciosos, a la que hay que restarle los que son recuperados a lo largo de la epidemia (estos pasan a la clase de susceptibles, al menos formalmente) y los removidos. La población de referencia que consideramos en la Sección 3 fue de 10000 individuos, a partir de las cantidades de infecciosos detectados en los primeros días de la epidemia.

Entre las ventajas que tiene el proceder que hemos presentado aquí se tienen las siguientes: la solución de la ecuación para la fracción de infecciosos se expresa analíticamente (ec.6) y eso permite que el ajuste de parámetros se haga de manera simple. Eso también determina la posibilidad de describir la función que representa la fracción de infecciosos activos de manera analítica, y también una fórmula para la obtención del tiempo donde se alcanza el máximo de esta población de infecciosos activos.

Es conveniente aclarar que estos modelos de la dinámica de epidemias reflejan mejor la realidad cuando los individuos, por alguna razón, están más ajenos a la forma de transmisión, aunque sepan que para contagiarse hay que tener algún contacto directo o indirecto con un individuo enfermo. Más precisamente, estos modelos no tienen en cuenta ni el temor

al contagio que puedan tener los individuos, ni el alcance de las medidas gubernamentales. De modo que es de esperar que los estimados de la cantidad de individuos enfermos mediante estos modelos, sean superiores a la real.

### Acknowledgments

Los autores queremos expresar nuestra felicitación al Comité Editorial de la Revista por la idea de conformar un número dedicado al modelado de la Covid-19 en Cuba. Agradecemos la aceptación de este trabajo en dicho número.

### Referencias

- [1] Edelstein-Keshet, L.: *Mathematical Models in Biology*, volumen 46. SIAM, Classics in Applied Mathematics, Philadelphia, 2005.
- [2] Kermack, W.O. y A.G. McKendrick: *Contribution to the mathematical theory of epidemics*. Roy. Stat. Soc. J., 115:700–721, 1927.
- [3] Murray, J.D.: *Mathematical Biology: I. An Introduction*, volumen 17 de *Interdisciplinary Applied Mathematics*. Springer, New York, third edición, 2001.
- [4] Ricard, M.R., C.T. González González y R.C. Bassanezi: *Epidemiological model with fast dispersion*. En Mondaini, R. y R. Dilon (editores): *BIOMAT 2005, Proceedings of the International Symposium on Mathematical and Computational Biology*, páginas 245–264. World Scientific, 2006.