

Estimación

Jorge Esteban Mendoza Ortiz (418002863)*
Enrique Moctezuma González (315019344)**
Ana Victoria López Rodríguez (314614300)***

Resumen

sub_abstract

1. Materiales y métodos

Fuente de datos

El diseño del estudio fue ecológico. Los datos se obtuvieron del reporte titulado «Situación actual de la epidemia» publicado por la Secretaría de Salud de México el 20 de agosto de 2009 [8]. Se definió como unidad de análisis los casos de influenza A (H1N1) confirmados por el laboratorio del Instituto de Diagnóstico y Referencia Epidemiológicos (InDRE), citados en el reporte antes mencionado. El InDRE es la autoridad que recopiló los casos reportados de esta enfermedad a nivel nacional. Se definió como unidad de análisis de tiempo los días naturales.

Los cálculos presentados se realizaron utilizando el paquete estadístico R versión 3.5.3 y Matlab R2019a Update 5.

Descripción del modelo SIR

El modelo SIR fue desarrollado por Kermack y McKendrick en 1927 [1] y ha sido aplicado en diversos escenarios de epidemias. Este modelo estima el número de personas susceptibles de enfermar (susceptibles), el número de enfermos (infectados) y el número de personas que ya no pueden transmitir la enfermedad ya sea porque se han recuperado y han adquirido inmunidad o porque han fallecido (removidos).

Los supuestos del modelo SIR son: a) la población es homogénea y de tamaño fijo; b) en un momento dado, cada individuo sólo puede pertenecer a uno de los siguientes conjuntos: infectados, susceptibles o removidos; c) la interacción entre los individuos es aleatoria y d) no hay intervención externa que cambie la tasa de contacto

*esteban.mendoza@ciencias.unam.mx
**enriquemg@ciencias.unam.mx
***anaviclopez@ciencias.unam.mx

de la población. Se asumió que la población del país es constante y que el número de individuos susceptibles $S(t)$, infectados $I(t)$ y removidos $R(t)$ son variables que dependen del tiempo t .

El modelo SIR consta de un sistema de tres ecuaciones diferenciales:

$$\frac{dS}{dt} = -\kappa SI, \quad \frac{dI}{dt} = \kappa SI - \mu I, \quad \frac{dR}{dt} = \mu I,$$

donde κ es la tasa de contagio por unidad de tiempo (día) y μ la tasa de recuperación per cápita. Es la fracción de infectados recuperados en un día dado, que se calcula como el recíproco del período infeccioso estimado de la enfermedad y que en este estudio se estableció en tres días [5], *i.e.*, $\mu = 1/3$.

En términos informales podemos interpretar el modelo SIR como sigue: el cambio del número de susceptibles respecto al tiempo es inversamente proporcional a los contactos entre susceptibles e infectados; el cambio del número de infectados respecto al tiempo es directamente proporcional a los contactos entre susceptibles e infectados además de que el número se ve afectado por la tasa de recuperación; por último, el cambio del número de removidos respecto al tiempo es directamente proporcional al número de infectados, y estos se recuperan a una tasa fija μ .

Aplicación del modelo

Dado que uno de los supuestos del modelo es que no hay intervención externa que cambie la tasa de contacto de la población se estimó dos veces el parámetro κ para cada etapa de la primera ola pandémica que se presentó en México. Consideramos la primera etapa del inicio de la epidemia hasta la aplicación de medidas de aislamiento social, como la suspensión de actividades escolares (27 de marzo al 26 de abril de 2009). La segunda etapa se consideró después de la suspensión de actividades a nivel nacional y la aplicación de otras medidas para disminuir la tasa de contacto (26 de abril al 14 de mayo de 2009). Se añadió un retraso de tres días a los períodos, suponiendo un período de incubación de la influenza de tres días [5].

Para estimar el parámetro κ en cada uno de los períodos considerados hemos recurrido a métodos numéricos. Para obtener una primera subestimación del parámetro hemos obtenido una solución linealizada de la segunda ecuación del modelo. Para obtener la linealización observemos primero que la segunda ecuación del modelo SIR cumple la siguiente igualdad:

$$\frac{dI}{dt} = \kappa SI - \mu I = \mu \left(\frac{\kappa SI}{\mu} - I \right) = \mu \left(\frac{\kappa S}{\mu} - 1 \right) I = \mu(\mathcal{R}_0 - 1)I, \quad (1)$$

donde $\mathcal{R}_0 = \frac{\kappa S}{\mu}$ es el número reproductivo básico cuya interpretación explicamos más adelante. Si hacemos $\beta = \mu(\mathcal{R}_0 - 1)$, obtenemos la ecuación diferencial $\frac{dI}{dt} = \beta I$, cuya solución podemos obtener analíticamente, y está dada por: $I(t) = I_0 e^{\beta t}$, donde $I_0 = I(0)$, con lo que nuestro modelo resultante es:

$$I(t) = I_0 e^{\beta t} + \varepsilon(t).$$

Si ignoramos el término de error y tomamos el logaritmo natural a cada lado de esta última expresión obtenemos la linealización deseada, puesto que:

$$\log I(t) = \log(I_0 e^{\beta t}) = \log I_0 + \beta t.$$

Haciendo $Y = \log I(t)$, $\alpha = \log I_0$, obtenemos el modelo lineal: $Y = \alpha + \beta t + \varepsilon(t)$, que es fácil de calcular aunque subestima los parámetros, puesto que hemos sustituido los términos de error originales por otros términos de error del modelo lineal.

Una vez que obtenemos dichas subestimaciones utilizamos el método de descenso de gradiente para hallar una aproximación numérica a los parámetros I_0 y β que minimizan la función de costo:

$$J(I_0, \beta) = \frac{1}{2n} \sum_{t=0}^n (I(t) - I_0 e^{\beta t})^2,$$

donde n es el número de días considerados en cada período menos uno porque el primer día de cada período es el tiempo $t = 0$.

El método de descenso de gradiente es un método iterativo que actualiza los parámetros I_0 y β del siguiente modo:

$$(I_0^{(i+1)}, \beta^{(i+1)}) := (I_0^{(i)}, \beta^{(i)}) - \alpha \nabla J(I_0^{(i)}, \beta^{(i)}),$$

donde:

$$\begin{aligned} -\nabla J(I_0^{(i)}, \beta^{(i)}) &= -\left(\frac{\partial}{\partial I_0} J(I_0^{(i)}, \beta^{(i)}), \frac{\partial}{\partial \beta} J(I_0^{(i)}, \beta^{(i)}) \right) \\ &= -\left(\frac{1}{n} \sum_{t=0}^n (I(t) - I_0^{(i)} e^{\beta^{(i)} t}) (-e^{\beta^{(i)} t}), \frac{1}{n} \sum_{t=0}^n (I(t) - I_0^{(i)} e^{\beta^{(i)} t}) (-I_0^{(i)} t e^{\beta^{(i)} t}) \right) \end{aligned}$$

es la dirección de máximo descenso de la función de costo dados los parámetros (I_0, β) en la i -ésima iteración del método y α es una «constante de aprendizaje» que determina la magnitud del cambio de los parámetros a cada iteración. Para α suficientemente pequeña el método converge a un mínimo local si la función de costo que se trata de minimizar es convexa, como en este caso.

Para calcular el número de habitantes en México a inicios del brote epidémico hemos recurrido a los datos publicados por el Instituto Nacional de Estadística y Geografía (INEGI) en el II Censo de Población y Vivienda 2005 [4] y el Censo de Población y Vivienda 2010 [3], que finalizaron sus levantamientos el 29 de octubre de 2005 y el 25 de junio de 2010, respectivamente.

Sean $p_{05} = 103,263,388$ y $p_{10} = 112,336,538$ el número de habitantes en México el 29 de octubre de 2005 y el 25 de junio de 2010, respectivamente. Existe una diferencia de 1700 días entre ambas fechas. Estas cantidades cumplen la ecuación:

$$p_{05} \left(1 + \frac{r}{100} \right)^{1700} = p_{10},$$

donde r es la tasa de crecimiento diaria que obtenemos del siguiente modo:

$$r = 100 \left[\left(\frac{p_{10}}{p_{05}} \right)^{\frac{1}{1245}} - 1 \right].$$

Dadas estas ecuaciones, estimamos que el número de habitantes en México el 27 de marzo de 2009 es:

$$S_0 = p_{05} \left(1 + \frac{r}{100}\right)^{1245} = 109,832,769,$$

porque hay una diferencia de 1245 días entre el 27 de marzo de 2009 y el 29 de octubre de 2005.

El número de susceptibles al inicio del segundo período se ha calculado numéricamente a partir del modelo obtenido del primer período.

Por último, después de haber obtenido numéricamente los parámetros β , I_0 y S_0 , de (1) calculamos el número reproductivo básico \mathcal{R}_0 , que está dado por:

$$\mathcal{R}_0 = \frac{\beta}{\mu} + 1. \quad (2)$$

Si además por definición $\mathcal{R}_0 = \frac{\kappa S_0}{\mu}$, sustituyendo en (2) deducimos que la tasa de contagio por día κ está dada por:

$$\kappa = \frac{\beta + \mu}{S_0}.$$

El código utilizado junto con los datos originales y simulados, así como el código fuente de este reporte pueden ser consultados en [6].

2. Resultados

Para el primer período considerado —del 27 de marzo al 26 de abril de 2009— obtuvimos una tasa de contagio $\kappa_1 = 5.326540e-09$ y un número reproductivo básico $\mathcal{R}_1 = 1.755086$.

Para el período comprendido del 27 de abril al 14 de mayo de 2009, durante el cual se aplicaron en México medidas de aislamiento social, obtenemos una tasa de contagio $\kappa_2 = 2.470645e-09$ y un número reproductivo básico $\mathcal{R}_2 = 0.814052$.

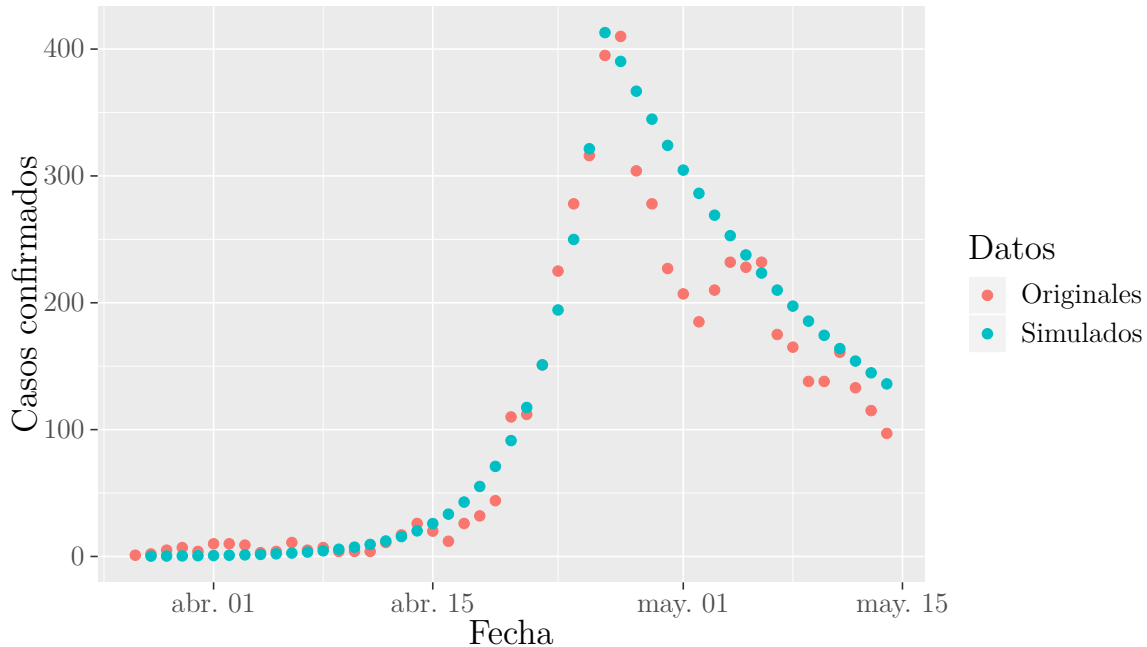
Para una comparación de la distribución original de casos confirmados y la distribución generada numéricamente por nuestro modelo, refiérase a la figura 1.

3. Discusión

El número reproductivo básico \mathcal{R}_0 que hemos estimado es un parámetro positivo del modelo que tiene una relevante interpretación [2]. Si $0 < \mathcal{R}_0 < 1$, no habrá un brote epidémico porque las personas están recuperándose más rápido de lo que ocurren nuevos contagios, pero si $\mathcal{R}_0 > 1$ hay un brote epidémico dado que ocurre el escenario inverso: nuevos contagios ocurren más rápido de lo que la gente se recupera.

Las medidas tomadas por el gobierno de México redujeron el número reproductivo básico de $\mathcal{R}_1 = 1.755086$ a más de la mitad y dentro del umbral en el que el brote epidémico retrocede $\mathcal{R}_2 = 0.814052$. Utilizando nuestro modelo estimamos que si el gobierno de México no hubiera tomado medidas para reducir la tasa de contagio y se

Figura 1: Distribución de casos confirmados de influenza A (H1N1) según la fecha de inicio de síntomas, México, 27 de marzo al 14 de mayo de 2009.



hubiera permitido a la epidemia seguir su curso inicial, el número de removidos habría alcanzado cerca de 78,500,000 después de 120 días.

Cabe mencionar que si en la estimación de parámetros utilizamos únicamente el modelo linealizado, los parámetros obteniendo son $\mathcal{R}_1 = 1.477143$ y $\mathcal{R}_2 = 0.813669$, resultados que son más cercanos a la literatura sobre el tema. En [7] los autores utilizan la misma técnica de estimación y obtienen estimaciones para cada estado de la República Mexicana de \mathcal{R}_0 , con una media aritmética nacional de 1.26. En el período posterior a las medidas profilácticas los investigadores estiman que la mediana nacional del número reproductivo básico desciende a 1.01.

Referencias

- [1] Nicolas Bacaër. *A Short History of Mathematical Population Dynamics*. Springer, 2011.
- [2] Pauline van den Driessche. «Reproduction numbers of infections disease models». En: *Infectious Disease Modelling* 2.3 (ago. de 2017), págs. 288-303. DOI: 10.1016/j.idm.2017.06.002.
- [3] Instituto Nacional de Estadística y Geografía. *Censo de Población y Vivienda 2010*. Español. Jun. de 2010. URL: <https://www.inegi.org.mx/programas/ccpv/2010/default.html>.

- [4] Instituto Nacional de Estadística y Geografía. *II Conteo de Población y Vivienda 2005*. Español. Oct. de 2005. URL: <https://www.inegi.org.mx/programas/ccpv/2005/default.html>.
- [5] James E. Fielding y col. «Systematic review of influenza A(H1N1)pdm09 virus shedding: duration is affected by severity, but not age». En: *Influenza and other respiratory viruses* 8.2 (mar. de 2014), págs. 142-150.
- [6] Jorge Esteban Mendoza Ortiz, Enrique Moctezuma González y Ana Victoria López Rodríguez. *InfluenzaSIRModel*. Español. Ago. de 2019. URL: <https://github.com/esteban-mendoza/InfluenzaSIRModel>.
- [7] Estela Navarro-Robles y col. «Modelo para estimación del comportamiento epidémico de la influenza A (H1N1) en México». En: *Revista PanamERICANA de Salud Pública* 31.4 (abr. de 2012), págs. 269-274.
- [8] Secretaría de Salud. *Situación actual de la epidemia*. Ago. de 2009. URL: http://web.archive.org/web/20090823104214/http://portal.salud.gob.mx/sites/salud/descargas/pdf/influenza/situacion_actual_epidemia_200809.pdf.