

Historia de números en una pandemia

Santi García Cremades

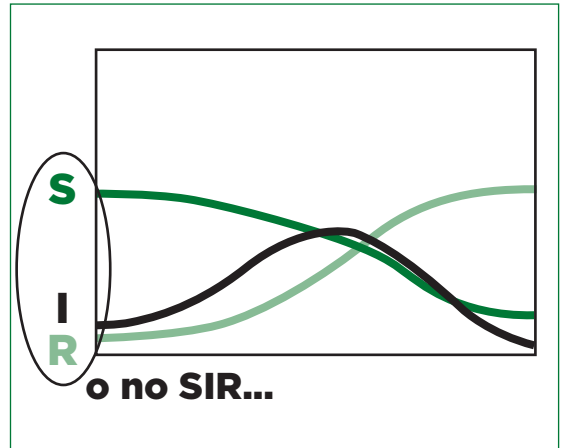
Matemático, profesor y divulgador científico

Modelos por doquier

Modelos hay infinitos, y la pregunta evidente es: ¿qué modelos usar en una epidemia? El modelo SIR fue el primero en levantar la mano, por razones claras, es el modelo que distribuye una población en tres categorías: Susceptibles, Infectados y Recuperados. Fácil, ¿no? Es un modelo basado en ecuaciones con derivadas, propuesto en 1926 por el escocés Anderson Gray McKendrick para estudiar la malaria. Hay muchas variantes, siendo el modelo SEIR (con la E de Expuestos) el más presente, pudiendo analizar a nivel teórico los efectos de las medidas restrictivas en el resto de variables. Este modelo tiene el problema evidente que si no tienes una buena estimación de los infectados (la población I), no obtenemos la información más relevante: cuánta gente puede estar inmunizada.

Los datos nos hablan, todos los días, a todas horas. Otra cosa es que no queramos oírlos, o seamos miopes a ellos y los veamos borrosos. Siempre me gustó ese concepto: las matemáticas son las gafas que nos faltan para ver con nitidez el mundo que nos rodea

Puesto que en tiempo real el modelo SIR parecía tener sus debilidades, salieron de golpe todos los modelos matemáticos que podía aportar cada familia. Podríamos destacar tres



grandes categorías: las cadenas de Markov, las series temporales y el *machine learning*.

Las **cadenas de Markov** tienen origen a principios del siglo XX, dentro de la teoría de la probabilidad. El nombre viene de Andrei Markov cuyo interés al crear las cadenas no tenía conexión con ningún fin práctico, salvo en otra de sus pasiones: la poesía. Es un método basado en probabilidades pasadas, para ponderar la probabilidad de un evento en el presente o futuro. Es un método muy aplicado para realizar predicciones como en meteorología, los buscadores de internet como Google o, claro está, en epidemias.

Las **series temporales** se explican por sí mismas: son series de datos en un intervalo de tiempo. Podemos marcar el comienzo del uso de estas técnicas en los años 20 del siglo XX, con George Udny Yule y con el *crack* del 29, donde saber cómo evoluciona una variable que se ve afectada por cientos de ellas tomó una gran importancia en economía (cuando empezaría la econometría). En estos modelos se pueden aplicar decenas de versiones: aditivos, multiplicativos, autorregresivos, con distintas frecuencias, etc. Este ha sido la base del modelo que utilicé para predecir la curva de fallecidos de abril a mediados de mayo, con mejor aproximación que el resto de métodos.

El **machine learning** es la herramienta más reciente y la menos matemática. Aunque Geoffrey Hinton introdujera el concepto por 1986, la aplicación y desarrollo han sido posteriores. Hay multitud de métodos, todos basados en el aprendizaje desde el entrenamiento de la máquina, es decir, de los datos pasados.

Con todos estos modelos, nos encontramos con distintas predicciones, cada una con un error distinto. Ahora, otra pregunta es ¿qué medir? ¿Qué variable? ¿Cuáles son los datos que más información nos dan? Los datos nos hablan, todos los días, a todas horas. Otra cosa es que no queramos oírlos, o seamos miopes a ellos y los veamos borrosos. Siempre me gustó ese concepto: las matemáticas son las gafas que nos faltan para ver con nitidez el mundo que nos rodea. No sé si tatuarme la frase o dejarla aquí, pero el concepto me parece acertado. Ahora, más que nunca, se habla de datos por todos lados. Tenemos los números cada día de infectados, fallecidos, recuperados, personas en la UCI, hospitalizados... y todo esto, por países, por comunidades, por provincias, incluso desagregados por edades. Tenemos muchos datos y, por tanto, mucho que estudiar. Vamos a ponernos esas gafas y entender qué

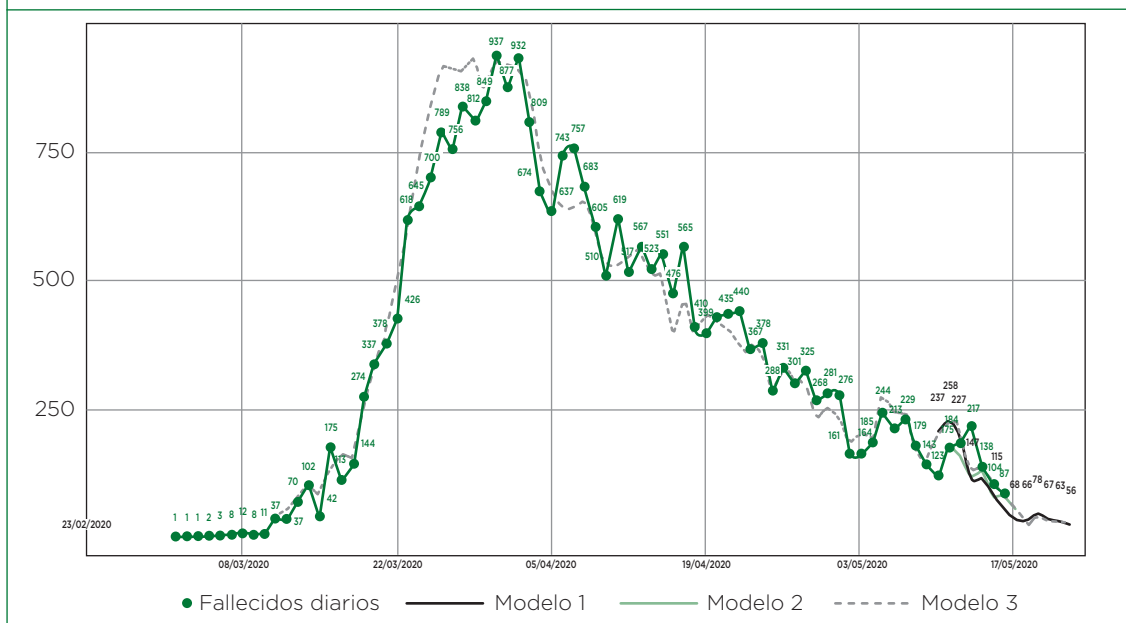
datos son claves en esta Guerra Mundial contra el SARS-CoV-2.

Exponencial y el país de las maravillas

Hay curvas por todos lados. A mí mismo me han llamado el “chico de la curva”, no sé si será por los kilos que estoy cogiendo (ya los soltaré por algún sitio). Es un tema complejo y, sobre todo, dramático, pero si algo nos han enseñado nuestros mayores es a aprender de la experiencia y a recuperarnos de los golpes de la vida. Entonces, vamos a aprender de los errores cometidos, de lo que se ha comunicado mal o no se ha entendido bien. Nos fijamos en las curvas de datos sanitarios. Empecemos primero por el inicio de la pandemia, el momento **exponencial**.

Al principio de un contagio, todas las cifras suben rápido, se duplican, se triplican e incluso se cuatuplican en un solo día, y eso no puede ser para siempre. A eso le llamamos el factor de contagio, en epidemiología se conoce como R_0 (el número de personas que se estiman que contagia cada persona infectada). Si ese factor de contagio fuese constante durante toda la pandemia, estaríamos perdidos, y estaríamos

FIGURA 1. SERIE TEMPORAL SOBRE LOS FALLECIDOS DE LA COVID-19 EN ESPAÑA



Fuente: Santi García Cremades.

hablando de una función exponencial. Se dice mucho la palabra “virus” y la palabra “viral”. Pero, ¿cómo es algo viral? ¿Crece tanto el virus como pensamos? Hablemos de exponencial. Sí, de la función exponencial, que tanta lata nos da en el Instituto. Si hay algo que es útil en Matemáticas es dar certeza de lo que uno puede intuir. Si hablamos de la magnitud de la función exponencial, no hay nada como poner un ejemplo visual para hacernos una idea. Es un ejemplo clásico, pero nos vamos a sorprender. Si tienes hipo, tranquilo que se te va a quitar.

Coge un folio. Coge el primer folio en blanco del libro, adelante, no estoy mirando. El grosor del folio es aproximadamente de 0,1 milímetros. Vamos a intentar doblarlo ahora. Una vez. Dos veces. Tres veces... Ya cuesta más. Se creía que ocho veces es lo máximo que podrías doblar el papel. Pues es mentira. La americana Britney Gallivan, cuando era tan solo una estudiante de secundaria, en 2002, demostró una fórmula que determinaba el tamaño del papel para poder doblarlo n veces, y hasta realizó la prueba doblando un papel doce veces. El récord actual es de trece dobleces (gente aburrida, qué sé yo). Vamos a ver qué anchura va cogiendo el folio según lo van doblando. Es sencillo, a cada doblado se dobla la anchura anterior, así que es una función exponencial en base 2. ¿Cuánto crees que llegaría el folio que tienes en la mano si pudieras doblarlo 22 veces (que por poder, no se puede, la física y sus límites)? Pues sería más alto que la Torre Eiffel, que mide 300 m de altura.

Aún tienes hipo, parece... A ver, ¿y si doblamos 42 veces el folio? Ahora la anchura es igual a 2 potencia 42 por la anchura del folio. Coge calculadora, que de cabeza no vas a poder. Te espero. Tendríamos ahora un total de 4.398.046.511.104 capas, más de 4 billones de capas (con “b”), equivalente a 439.805 kilómetros. Son muchísimos kilómetros, si dejásemos ese folio en el suelo, llegaríamos hasta a la Luna, que está a 384.400 km aproximadamente. Si quieres exponencialmente, podrás conseguirle a la persona amada la Luna, con tan solo 42 pasos. Y ya para rematar, si doblamos el folio 103 veces, el grosor sería mayor que el diámetro del Universo observable (93.000 millones de años luz) ...

¡Vaya poder, la exponencial! El coronavirus, sin embargo, parece crecer de manera mucho

más lenta que la exponencial, gracias a los sistemas estatales de Salud Pública y a la OMS, gracias a que tenemos defensas inmunitarias y a que tomamos prevenciones (que no es lo mismo que despertar el caos). Dejemos lo exponencial únicamente en el terreno matemático, por favor. Por poner un ejemplo, si cada contagiado hubiera contagiado a 2 personas cada día, en 33 días estaríamos más de 8 mil millones de personas contagiadas, más gente que hay en la Tierra. Entonces, vale, NO es exponencial, el factor de contagio es dinámico. Primer aprendizaje: crecíamos rápido, pero no exponencialmente.

Confinamiento “rules”

Hablemos ahora del confinamiento, aunque prometo no hablar del “Resistiré”. Según estadísticas de movilidad que sacó Google en marzo y abril, estuvimos entre los 4 países del mundo que más han respetado el confinamiento. Y eso tiene un efecto claro: el famoso **aplanamiento de la curva**. Está claro que necesitábamos aplanar la curva para no saturar el sistema sanitario y poder atender a los casos más graves. Hemos conseguido, entre todos, que unas cifras que crecían muy rápido (no exponencialmente, pero rápido) se ralentizaran. Esto se puede medir a través de una medida clásica en estadística, la curtosis. La curtosis mide cuánto de afilada o de aplanada es la curva que tenemos. Pues si nos confinamos el día 14 de marzo (el día Mundial de las Matemáticas, te recuerdo), el día 24 de marzo vimos un efecto directo en esta medida. La curtosis pasó de decirnos que estábamos en una curva afilada a un momento aplanado de la curva de infectados o fallecidos diarios. Por tanto, *QED*, queda demostrado, conseguimos aplanar la curva. Primero crecer rápido, y luego lento.

Ahora toca el momento clave, la luz al final del túnel. La curva no dejó de crecer hasta el 31 de marzo de 2020, lo que conocemos como el **punto máximo de la curva**. En modelos epidemiológicos, se entiende que sería el punto de inflexión de la pandemia, “de ahí p’abajo”. Bueno, eso es una medida a ojímetro, pero ¿podemos saber cuánto queda de estas cifras tan duras? Difícil, pues la curva no es simétrica, no

es una campana de Gauss. Además, el momento alto de las cifras de infectados y fallecidos se ha mantenido más tiempo del que nos gustaría, es una consecuencia de aplanar la curva, que la meseta dura más en el tiempo. Pero la bajada está ahí, la tendencia decreciente también está y las gafas de las matemáticas hacen ver una lucecita al final. Y ahora nos queda un momento clave: la victoria final.

Al final de esa primera ola, estamos ya desesperados, y hay algo más que el coronavirus en nuestras mentes: ¿cómo está afectando esto a nivel económico y psicológico? Aquí entra en juego un concepto que no tiene más de 100 años de formalismo matemático: ecuaciones simultáneas. Como decía Forrest Gump, “la vida es como una red: todo está conectado”. Vale, no es de Forrest, pero es una forma de entender el todo. En una red, cualquier movimiento afecta al total de los nodos de la red. Sin embargo, hay cosas que no se ven afectadas por cambios en la red, y otras que no. Por ejemplo, las horas de Sol afectan la actividad social de un lugar, la actividad social afecta a la economía, a la salud mental, y viceversa: la salud mental afecta a la economía, a la actividad social. Sin embargo, que tú trabajes más o que estés más feliz, poco lo importa al Sol. En este caso, la variable “horas de Sol” sería una variable exógena, una variable pasota, a su bola, y el resto son endógenas, como la red que decía Forrest Gump (en mi fantasía).

En epidemiología, casi todas las variables son endógenas, por tanto, cualquier decisión, tanto restrictiva como permisiva, tiene un efecto en TODAS las variables sanitarias, económicas y sociales. De ahí la complejidad del tema. Estamos en una red, donde si salvas vidas por COVID, causas pérdidas en otros factores. El confinamiento ha provocado una gran depresión económica y social, y para afrontar esos problemas, provocamos otras olas pandémicas.

Olas van, olas vienen

Es natural bajar la guardia, en situación estable, cuando ya no se habla de datos dramáticos, buscamos también la vida pre-pandémica, la llamada “nueva normalidad”. Y este es



un problema para los matemáticos, pues el movimiento de poblaciones es un movimiento casi impredecible, lo que conocemos como movimiento caótico. En julio y agosto tuvimos el principio de una segunda ola. Bueno, antes habría que definir ola: cambio de tendencia positiva (de tendencia decreciente a creciente, o de tendencia constante a creciente) de la curva de contagios. Y esto ocurrió, y no se volvió a tomar ninguna restricción útil hasta octubre. Ahí hay un problema de asesoramiento matemático claro. No se supo remarcar un umbral de decisión en 3 meses, parecía como si no hubiéramos aprendido de la primera ola. Y, si bien es verdad que las olas pandémicas no tienen nada que ver, tampoco son tan diferentes estadísticamente hablando.

La clave está siempre en saber qué medir, cómo medir y, posteriormente, saber interpretar esas medidas. Algo estamos aprendiendo, seguro, sobre todo el gran valor que tiene la ciencia y, en particular, la ciencia más básica de todas: las **Matemáticas**. ●