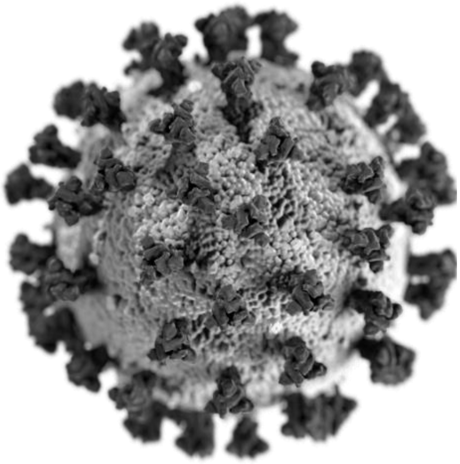
La curva que teníamos que aplanar



o de cómo intentamos domesticar la naturaleza.

Felipe E. Ramírez
Profesor de matemáticas

En estos largos días de confinamiento junto al **coronavirus** se ha infiltrado en nuestros hogares una compañera: la curva. Un objeto matemático que hemos convertido en el enemigo a batir. La batalla era (y es) de todos contra la curva a la que había que doblegar.

Nunca antes el "*aplanamiento*" de una curva, un término que no usamos los matemáticos, ha pasado a convertirse en el objetivo de una particular y falaz batalla del Hombre contra un enemigo. En este caso el enemigo era el virus responsable de la covid-19 y el objetivo aplanar la curva. ¿curva de qué?. En esta ocasión, nos centraremos en la curva y dejamos su doblegamiento para otro momento.

En las ficciones cinematográficas sobre contiendas bélicas solemos oír a los generales lanzar un mensaje a sus tropas: "¡Tomen la colina!" que como paradigma de la irracionalidad bélica insufla en los infantes una orden no siempre fácil de cumplir.

En las siguientes páginas trataremos de explicar el significado de *la colina* que los generales de esta batalla –los responsables sanitarios- nos han pedido a los infantes –nosotros- que tomemos. Nuestra intención es explicar por qué surge una curva de un fenómeno epidemiológico y por qué aplanarla es una victoria en la batalla.

Datos: la materia prima

Todo comienza con los datos. Éstos como venimos diciendo son fundamentalmente de dos tipos: el número de fallecidos diarios y el número de nuevos infectados (y detectados) diarios. Hay más datos, ya lo hemos descrito, como son el número de pacientes que cada día ingresan en la UCI o el número de pacientes que son dados de alta.

Si nuestro interés es conocer cómo evoluciona una epidemia debemos incidir en cualquiera de los dos primeros. El número de nuevos casos señala la incidencia del contagio del virus; su registro nos indicará día a día como progresa su expansión en una determinada población. El número de fallecidos diarios nos describe el daño real de la enfermedad; no es que el caer enfermo de la enfermedad no sea relevante, es que los fallecidos determinan la maldad del virus ¿Siempre?.

Si utilizamos un símil bélico la maliciosa **efectividad** de un arma puede medirse con el número de víctimas no mortales que su aplicación inflige en el enemigo o con el número de muertos que ocasiona en las líneas enemigas. No siempre es mejor disponer de un arma que “mate” muchos enemigos que de otra que infrinja muchas bajas. La WWI fue paradigmática en este sentido: la máxima “*mejor que matar, es herir*” se convirtió en el *modus operandi* de los bandos en litigio. Cada soldado herido en el frente generaba enormes costes sanitarios, problemas de intendencia, traslados, curaciones, alimentación, medicamentos, etc. En cambio, un muerto en batalla solo genera el coste de su entierro.

Algo muy parecido sucede con las epidemias. El **Ébola** famoso virus dañino donde los haya, mortal y sanguinario es un virus “bueno” en el sentido anterior: la persona infectada, el huésped, desarrolla enseguida los síntomas, quedan postrados en cama y no puede ir aireando al virus; en cuestión de días causan el fallecimiento del huésped y con él, el del virus. Sería muy difícil que una epidemia de Ébola pudiera contagiar a tantas personas como está haciendo el SARS-Cov-2.

Este es un *bichito* silencioso, no se advierte y daña al enemigo (nosotros) desgastando sus sistemas sanitarios. Como algunas armas de la WWI este virus desgasta al enemigo causándole penosos costes en *bajas*, pero no “*muchas* muertes”. El número de **nuevos contagios** es el auténtico índice de la maldad del SARS-Cov-2: no mata tanto como el Ébola, una especie de bomba atómica, pero llega al punto de confinar a su enemigo (nosotros) en las trincheras -nuestros hogares- inutilizando todos sus mecanismos de defensa: la economía, la sociedad o la libertad.

Por tanto, tenemos que fijarnos en el número de nuevos casos detectados en un día. ¿Por qué en un día? Por cuestiones exclusivamente de efectividad. No es muy interesante hacer el **conteo** de nuevos casos por hora. ¿Qué se ganaría utilizando esta unidad de medida salvo aumentar el flujo de datos ya complejos por sí mismos? Así las cosas, debemos asumir que los datos que recopilemos serán siempre una sucesión de puntos discretos diarios que se recogen en la Tabla 1

Día	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Infectados	12	15	19	23	24	27	67	79	92	120	170	189	210	250	370
Día	16	17	18	19	20	21	22	23	24	25	26	27	28		
Infectados	450	560	800	920	1200	1600	1870	2300	2500	3400	4200	5100	6090		

Tabla 1 (datos ficticios)

Allá por el siglo XVII, en una época convulsa de la historia europea (¿cuál no lo ha sido?) en la que trascurrió la violenta Guerra de los Treinta Años que esquilmó a la población europea, dos matemáticos franceses Pierre Fermat y René Descartes y de forma independiente observaron que los datos se podían representar utilizando un eje horizontal y otro perpendicular. Con este artificio -hoy estudiado por todos nuestros alumnos- dieron comienzo a una nueva geometría: la analítica, glorioso hito en la historia de la Humanidad, aunque no es momento este de abundar en estos hechos.

Baste decir que desde entonces todos nos hemos acostumbrado a referirnos a esos ejes como **cartesianos** (de Des-**cartes**). Las matemáticas también son humanas y clasistas: Descartes fue mucho más preminente filosófica y mediáticamente que Fermat. Los ejes **cartesianos** podrían haber trascendido como ejes ***fermatianos,*** pero eso sencillamente no sucedió. Es otra larga historia.

René Descartes
(1596-1650)

Pierre Fermat
(1601-1665)

El caso es que desde entonces nos dedicamos a representar los datos en **un eje horizontal** donde por ejemplo llevamos los días de la epidemia y en un eje vertical que utilizamos como escala del número de nuevos casos detectados cada día. Esta simple idea es crucial en la formación matemática de todos los ciudadanos de nuestro mundo. Posiblemente nuestro sistema educativo no le dedique el suficiente tiempo curricular para que sea aprehendido por nuestros alumnos: como es un instrumento matemático muy sencillo y ampliamente utilizado nos olvidamos de la trascendencia que tiene.

Bien, toda vez que disponemos de un cúmulo de datos y un sistema de representación, hagámoslo.

Los poli(demonios)

Su aspecto se parecerá al siguiente gráfico que hoy en día obtenemos con hojas de cálculo o programas matemáticos. Pero no lo olvides descansa sobre una idea que tiene cuatrocientos años.

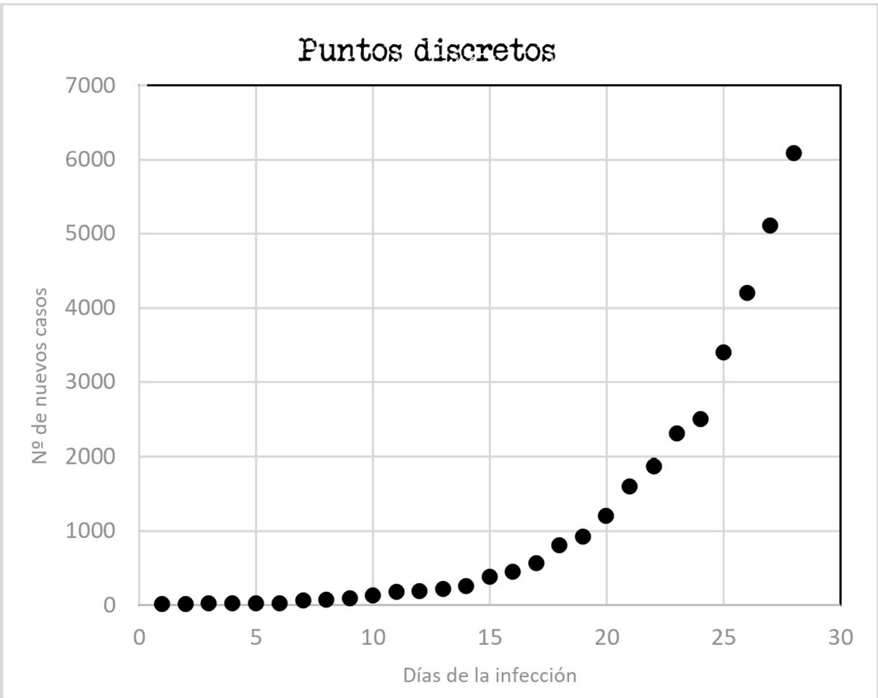
Supongamos que los datos que recogemos durante los primeros veintiocho días de la epidemia del covi-19 en un lugar cualquiera son los recogidos en la Tabla 2:

Día	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Infectados	12	15	19	23	24	27	67	79	92	120	170	189	210	250	370
Día	16	17	18	19	20	21	22	23	24	25	26	27	28		
Infectados	450	560	800	920	1200	1600	1870	2300	2500	3400	4200	5100	6090		

Tabla 2

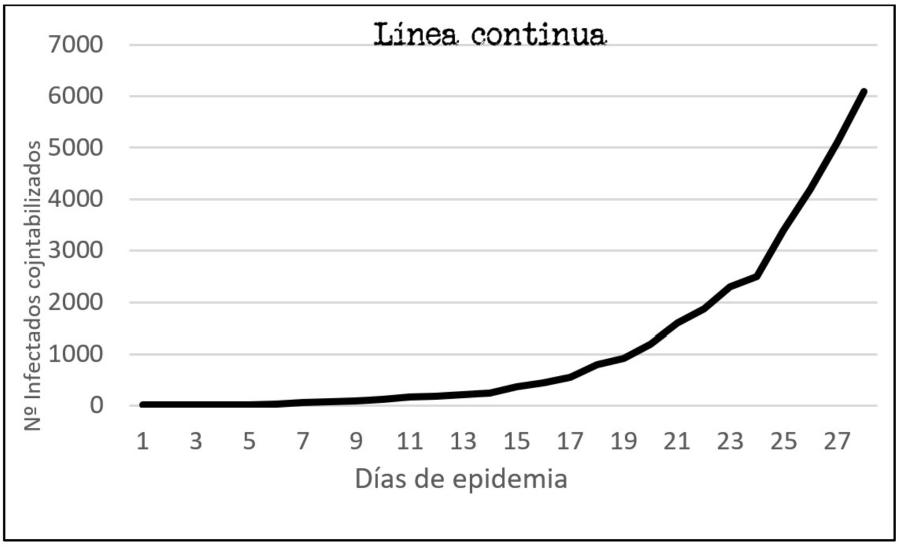
Esta es la información en bruto de la que partimos. Aquí se encuentra el **fundamento** que diría un famoso y mediático cocinero.

Para comenzar los datos los representamos en los ejes cartesianos, obteniendo un conjunto de puntos **discretos**, aislados unos de otros. Insistimos por importante: en el eje horizontal se representan los días y en el vertical el número de nuevos casos registrados en ese día. ¡ya disponemos de una primera *radiografía* de la curva de la infección! Pero sólo son puntos y las curvas son **continuas**. Mira si no la Gráfica 1.



Gráfica 1

El problema surge en cómo conseguir "**unir**" esos puntos de un modo realista. Y no solo eso. Unirlos es un juego de niños. Basta con tomar un **lápiz y *juntar*** los distintos puntos, como en la Gráfica 2.



Gráfica 2

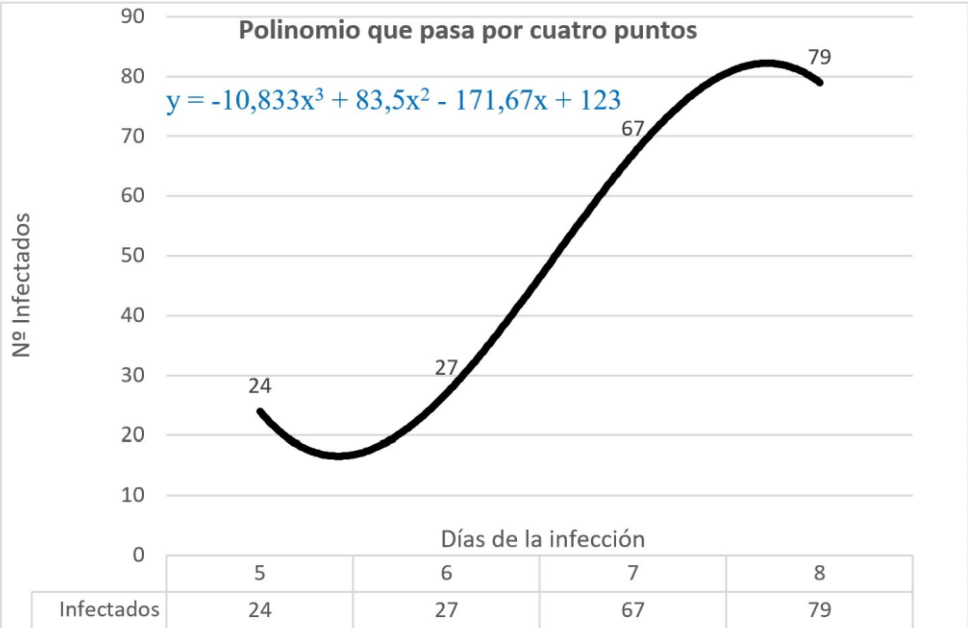
Se trata de que nuestra **línea continua** -un invento del analista, puesto que los datos son los puntos aislados- se **corresponda con la realidad**. De esta forma si somos capaces de crear esa curva podremos disponer de una herramienta de cálculo de la que podamos extraer **conclusiones veraces a futuro**. O en el lenguaje coloquial, apreciar la **tendencia** de esos datos.

Sin entrar en los detalles técnicos que son muchos y tema de los primeros cursos de grado, hay dos formas de abordar el problema que están muy entrelazados: uno es a través del **análisis matemático** y el otro del estudio de la **correlación estadística** de los puntos.

Con el análisis matemático buscaremos una curva a la que obligaremos a que **pase por todos y cada uno** de los puntos obtenidos. Si se observan los puntos del gráfico anterior se comprueba que "*encima*" de cada día sólo hay un punto, que corresponde al dato diario de infectados. En esta situación el determinante de Vandermonde entre otros mecanismos, nos garantiza que se puede construir un **polinomio** que **pasa por todos esos puntos** y sí, hablamos de uno de esos temibles objetos del álgebra elemental que padecimos en la escuela, los polidemonios como me gusta denominarlos.

Esto supone un esfuerzo titánico de cálculo si son muchos los puntos observados; por ejemplo, para analizar veinticinco días de infección tendríamos que crear un polinomio con nada menos que veinticinco términos -lo que se denomina el **grado** del polinomio-.

La Gráfica 3 recoge uno de estos polinomios creado para ajustarse a sólo cuatro datos diarios.



Gráfica 3

Dirá el lector que por tanto ya tenemos resuelto el problema. Nada más lejos de la realidad. El *quid* de la cuestión está en lo que sucede *entre esos puntos*. No siempre la forma de ellos es polinómica. Una cosa es que se pueda fabricar el polinomio y otra muy distinta es que sea esa la **forma real del fenómeno**, en este caso la evolución de los contagios. Porque se pueden utilizar otras curvas que también se adapten al fenómeno. Por tanto, se hace prioritario averiguar la forma real que manifiestan oscuramente los datos.

Y aquí es donde entra la estadística. Con la **regresión** estadística disponemos de otro instrumento matemático que nos permite "*enlazar*" los puntos aislados. En líneas generales se trata de obtener de todas las posibles curvas imaginables que se aproximen a todos los puntos, aquella que, en promedio, menos se distinga de todos los puntos a la vez.

Y esto es lo relevante: en lugar de obligar a nuestra curva a que **pase** por todos los puntos, lo cual es loable, pero puede llevar a conclusiones erróneas, buscaremos la curva que se **aproxime a todos a la vez de la mejor de las formas posibles**. ¡aunque no pase por ninguno!

Y lo mejor de todo: este método nos permite **asegurar** que el fenómeno **responde** adecuadamente **a la curva** que tenemos en mente. Pero esto lo dejamos para la siguiente entrega.

(continuará…)