Desinformación en la era digital: una oportunidad para la educación matemática

Trabajo Fin de Máster

Máster universitario en formación del profesorado de educación secundaria obligatoria y bachillerato, formación profesional y enseñanzas de idiomas

Especialidad de Matemáticas

Universidad Nacional de Educación a Distancia

Facultad de Educación

Emanuel G. Mompó Pavesi

Dirigido por: Antonio Pérez Hernández

Curso 2021/22

Índice

Listado de figuras

Resumen

En este Trabajo Fin de Máster se presentan diversos tipos de campañas, engaños y manipulaciones que son comunes en las redes (e incluso fuera de ellas), ante las cuales los adolescentes pueden estar expuestos por la elevada compenetración de las redes sociales en sus vidas. Frente a éstos, se presenta una serie de argumentaciones de corte matemático y científico que los profesores pueden mostrar y enseñar a sus alumnos para que se armen de herramientas contra la desinformación.

En el capítulo 1 se hace un breve repaso del contexto actual en el que se transmite y difunde información y contenido, algunas de las formas de desinformar que existen, la importancia de la cultura matemática frente a ciertas manipulaciones, y se brinda una motivación para este trabajo.

En el capítulo 2 se presentan algunos temas de actualidad donde el conocimiento sobre matemáticas es importante, pues, o bien son el fundamento de éstos, o bien permiten entenderlos o interpretarlos de forma algo menos superficial. Primero se habla de fenomenología no lineal y de ejemplos no intuitivos, luego de los esfuerzos políticos que se realizan de cara a ilegalizar u obstaculizar la encriptación de datos, así como de algunas controversias relacionadas con la obligación del uso de mascarillas.

En el capítulo 3 se plantea una serie de engaños, o «bulos», donde las matemáticas pueden ser utilizadas para desmontarlos o quitarles relevancia. Concretamente, se habla de cómo la estadística y las simplificaciones de la realidad pueden ser mal utilizadas, del terraplanismo y, en general, de la viabilidad de las conspiraciones masivas.

En el capítulo 4 se amplía el ámbito de las matemáticas al de la ciencia en general, y se hace un repaso del funcionamiento básico de la investigación científica, pero también se habla de cómo algunas de sus virtudes son utilizadas para desacreditar el conocimiento científico en un contexto posmoderno, fomentando la proliferación de teorías pseudocientíficas.

Por último, en el capítulo 5 se da conclusión al presente trabajo, comentando algunas de las puestas en marcha de varias de las ideas propuestas a lo largo de éste, y que además lo inspiraron. Adicionalmente se comentan las perspectivas de futuro de estas propuestas y posibles líneas de trabajo.

Adicionalmente se incluyen dos anexos. En el anexo A se incluyen algunos engaños que pueden encontrarse en las redes y en la vida real. En el anexo B se presenta formalmente algunos de los conceptos a los que se alude durante el cuerpo del trabajo, con objeto de ser un respaldo teórico para el profesor o lector.

Capítulo 1

Introducción

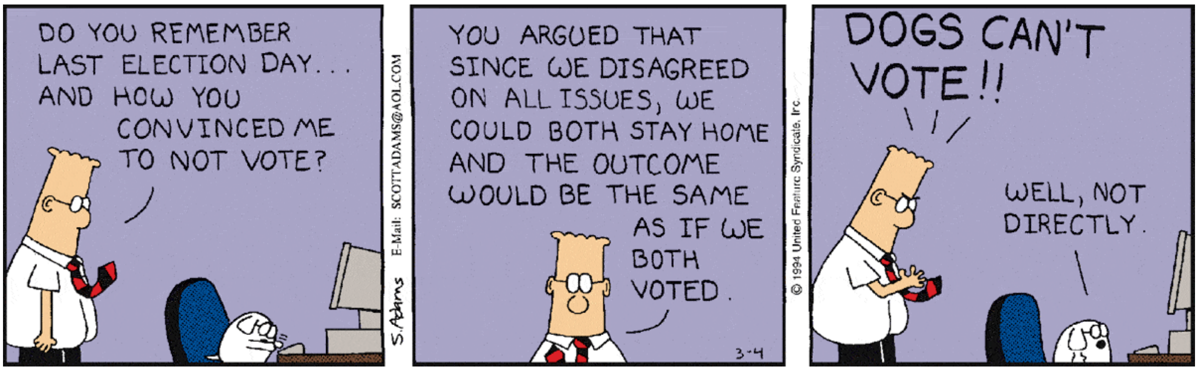

Fig. 1.1 Influencia indirecta. Fuente: https://dilbert.com/strip/1994-03-04.

Internet y las redes sociales han cambiado el paradigma en el que se difunde la información: «todos» tenemos voz y un sitio donde poder ser escuchados. En consecuencia tenemos mayor libertad a la hora de escoger el tipo de contenido que queremos consumir, así como la propia fuente de esa información.

Este mayor abanico ofrecido por internet, sin embargo, tiene sus debilidades. Las dos más básicas son, por una parte, que no toda la información es *correcta* (o fidedigna), y, por otra, que las propias fuentes podrían estar promoviendo sus propios intereses (Diethelm and McKee, 2009).

Las redes sociales añaden algunas debilidad adicionales a las anteriores: la identidad virtual está íntimamente ligada a la identidad real de una persona, y distintas herramientas permiten perfilar a los usuarios (BBC Mundo, 2018; Kleinman, 2018). Es decir, una fuente malintencionada puede escoger dirigir la información a personas concretas, atendiendo a sus características individuales. Además, la facilidad con la que se replica la información en redes sociales hace que información falsa pueda ser difundida con suficiente velocidad y alcance como para volverse perniciosa e, incluso, anteponerse a información real (Shin et al., 2018; Sommariva et al., 2018; Waszak et al., 2018).

La yuxtaposición de todas estas cuestiones es especialmente relevante respecto a los adolescentes de hoy día, oportunamente tildados como *nativos digitales*, pues emplean parte considerable de su tiempo en redes sociales y a partir de edades muy tempranas. Dicho de otro modo, en un periodo de sus vidas donde las bases de su cosmovisión aún están siendo establecidas.

A esto cabe añadir que, bajo ciertas métricas, las generaciones más recientes de estudiantes están en peores condiciones que sus equivalentes en generaciones pasadas (Bratsberg and Rogeberg, 2018; Flynn and Shayer, 2018), y estas últimas no son inmunes a las nuevas herramientas de influencia (BBC, 2018; Shuster, 2016).

Así pues, será de interés social (y democrático) formar a los estudiantes sobre estos nuevos paradigmas de manipulación, muchas veces sutil, e intentar brindarles algunas de las herramientas y conocimientos necesarios para discernir la calidad de la información que consumen.

1.1 La desinformación y algunas de sus formas

Las vías en las que se puede desinformar y manipular son muchas. Algunas son bastante clásicas (como las falacias argumentales –también llamadas sofismas, en referencia a los sofistas de la antigua Grecia–), y otras han sido habilitadas (o potenciadas) por la emergencia de nuevas tecnologías y el fácil acceso a ellas.

Entre las argucias más clásicas tenemos el uso de información falsa (v.g. estadísticas sin ninguna clase de respaldo, las alusiones a ciertas autoridades inexistentes o referencias ambiguas), la interpretación incorrecta de datos (especialmente cuando se habla de probabilidades condicionadas), las falacias argumentales (por citar algunas de las más comunes, se tiene la afirmación del consecuente, la negación del antecedente, la falsa dicotomía[1], el hombre de paja, o la apelación a la autoridad –*ad verecundiam*–), las falsas analogías o analogías sesgadas[2], o la asociación incorrecta de datos. Algunos ejemplos se pueden encontrar en el Anexo A.

El avance tecnológico ha traído consigo nuevas herramientas para desinformar, algunas de uso deliberado y otras no. Por ejemplo, se puede crear falsa evidencia[3] o difundir «noticias» de forma artificial y masiva mediante el uso de *bots* (Shao et al., 2018). Por su parte, también hay ciertos efectos que surgen de manera natural por el uso de algoritmos de recomendación, como la formación de comunidades que funcionan a modo de «cámaras de eco» (Cinelli et al., 2021), las cuales a su vez potencian el problema de desinformación (Choi et al., 2020).

Además se están desarrollando tecnologías que, en el futuro, podrán acrecentar los problemas de falsas evidencias, como es el caso de los *deepfakes*[4] (Westerlund, 2019).

[1]Ejemplo relativamente reciente: «Comunismo o Libertad».
[2]Ejemplo de comparativa deshonesta: Zulet (2021).
[3]Imágenes alteradas presentadas como originales en plataformas de rápida difusión social: Diariocrítico (2020).
[4]Ejemplo disponible en https://www.youtube.com/watch?v=AmUC4m6w1wo.

1.2 El rol de la cultura matemática

En la era de la información masiva y el mundo virtual (internet, videojuegos, redes sociales, criptomonedas, o los «metaversos»), la típica afirmación de que «las matemáticas están por todas partes» se vuelve más veraz: las matemáticas, efectivamente, lo impregnan todo. Al menos, todo lo relacionado con el mundo digital.

Cuando un alumno afirma que lo que se le enseña en clases de matemáticas no sirve para nada[5], es porque ciertamente desconoce cómo son utilizadas en su propio entretenimiento y beneficio, o, incluso más relevante, en su perjuicio, para lo cual a veces tampoco hace falta alejarse demasiado de las matemáticas elementales.

Por ejemplo, datos numéricos correctos pueden ser presentados de forma tergiversada para defender una postura pseudocientífica. O para cosas tan mundanas como conseguir que una persona contrate una tarifa de la luz aparentemente más barata que la que ya tiene pero que, a fin de cuentas, acabará siendo más cara.

Una cierta cultura matemática (y científica) es beneficiosa para evitar algunos engaños, ser más conscientes de otros, y tener una idea de cómo funcionan a nuestro favor (y en contra, alimentando los sesgos personales) algunos de los algoritmos que hacen más cómoda nuestra vida virtual (que hoy día se relaciona mucho con la vida «analógica»). Al fin y al cabo, podemos entender que el fin más grande de la educación y la enseñanza es formar ciudadanos plenos que sean capaces de defender la cultura democrática.

1.3 Objetivo y motivación de este trabajo

Este trabajo se realiza con tres objetivos en mente:

- Facilitar al profesorado diversos temas actuales donde las matemáticas son relevantes directa o indirectamente, pero que puedan llegar a ser de interés para los propios alumnos.
- Proponer al profesorado varios tipos de actividades o ejercicios a realizar con los alumnos para que estos últimos puedan interiorizar las ideas, además de intentar acercarles a la forma de pensar que tendría un investigador.
- Acercar al profesorado de matemáticas algunos aspectos de la filosofía de la ciencia que no son propios de las matemáticas.

Estos objetivos están motivados, entre otras cosas, por los puntos siguientes:

[5]Es posible que en algunos aspectos tengan razón, pero no por los motivos adecuados. Después de todo, como reza el dicho anglosajón, «a broken clock is right twice a day».

- La postura sobre cómo enseñar o trabajar en clases de matemáticas promovida por asignaturas del máster, en particular en Matemáticas en la Era de los Computadores. A su vez también por artículos como el de Lepage (2021).
- La aparente falta de interés general, observado durante el prácticum, por parte del alumnado hacia las matemáticas. Posiblemente porque no conocen su utilidad y omnipresencia en el mundo que les rodea.
- Las metodologías docentes observadas, de nuevo, durante el prácticum, no parecen fomentar la participación activa del alumnado.
- El salto generacional entre los profesores con los que trabajé durante el prácticum y sus alumnos potenciaba que no tuviesen ejemplos atractivos, al menos para los alumnos, del uso de las matemáticas.
- La falta de cultura matemática y científica hace que proliferen ideas y engaños que son perniciosos para las personas y la sociedad en general.
- Las matemáticas no son una ciencia experimental: la validez de los postulados no depende del peso de la evidencia. Luego hay problemas metodológicos a los que un matemático no está expuesto y podría no ser consciente de su existencia.

Capítulo 2

La cultura matemática y asuntos de actualidad

En este capítulo se presentarán a modo informativo varios temas actuales donde las matemáticas son importantes o donde un poco de cultura matemática puede ser beneficiosa a la hora de entenderlos.

Esto se hace con la esperanza de brindar al profesor más respuestas al usual comentario de que las matemáticas «no sirven para nada». Concretamente, respuestas que se salgan de lo clásico o el usual historicismo de los libros de texto, y que sean sobre cuestiones que a día de hoy estén en desarrollo.

2.1 Fenómenos poco intuitivos

Posiblemente el reto más grande de las matemáticas (y la física) en el siglo XXI sea el de «resolver» los problemas no lineales. Además de no disponer de una metodología general de resolución, a diferencia de los problemas lineales, ofrecen una casuística muy rica y muchas veces poco intuitiva, de nuevo, a diferencia de los problemas lineales.

El interés por presentar algunos fenómenos poco intuitivos es que, en ocasiones, se presentan argumentos reduccionistas sobre dinámicas complejas (como puede ser en economía) que parecen no dar lugar a alternativas. Pero la realidad a veces es más complicada[1].

2.1.1 Organización y orden espontáneos

Un ejemplo bastante curioso de fenómeno no lineal es el de la *resonancia coherente*. Hay ciertos sistemas oscilatorios que, una vez perturbados con ruido, empiezan a oscilar o vibrar a una cierta frecuencia privilegiada (pero modulada por la amplitud del ruido) en lugar de hacerlo simplemente al azar (Pikovsky and Kurths, 1997).

[1] «Hay mas cosas en el cielo y en la Tierra, Horacio, de las que sueña tu filosofía» –Hamlet.

Este fenómeno es relevante en el campo de la neurociencia, pues el tipo de neuronas que se consideran para crear modelos matemáticos de la corteza cerebral lo presentan (Lee et al., 1998). Concretamente, para que esas neuronas se pongan en funcionamiento (es decir, que «disparen» pulsos eléctricos) o bien necesitan ser excitadas con la aplicación de un potencial (mediante un electrodo), o bien hace falta que perciban suficiente ruido en su entorno. En cualquier caso, por sí solas no funcionan y requieren de ayuda externa.

Sin embargo, cuando hay muchas de estas neuronas en colectividad intentando imitar las conexiones de la corteza cerebral, se encuentra que el pequeño ruido generado por cada una de ellas, en conjunto, es suficiente como para que todas ellas entren en funcionamiento, sin necesidad de una ayuda externa (Mejias and Wang, 2021).

Otro ejemplo relacionado con osciladores que puede observarse es el de los estados quimera (Abrams and Strogatz, 2004). Cuando muchos osciladores se interconectan de cierta forma, aunque inicialmente todos ellos estén desincronizados, por momentos muchos de ellos (aunque no todos) se sincronizarán. Y así como surge espontáneamente el orden, también se perderá. Pero esto ocurrirá de manera intermitente.

Estos ejemplos anteriores sirven para ver que, de manera natural, el orden puede emerger espontáneamente, sin intervención externa. Es decir, no hace falta que haya un ente que diseñe el comportamiento, pues surge por sí solo gracias a la mera agregación de «individuos». O, visto de otra forma, cuando se reúnen muchas personas pueden ocurrir cosas no previstas.

Con un enfoque opuesto hay trabajos donde se plantean sistemas de individuos o «agentes» con una predisposición natural a organizarse, como es el clásico modelo de Vicsek et al. (1995), pero con algún factor externo (el cual que podemos controlar) que dificulta la organización, con el fin de estudiar cómo se producen las transiciones entre la organización y la desorganización. Estas cuestiones pueden ser de interés si, por ejemplo, deseamos controlar una flota de drones que tengan pocos sensores y poca capacidad de procesamiento (Li et al., 2017).

2.1.2 Efectos de memoria

Cuando un sistema está fuera del equilibrio, y bajo ciertas condiciones, su evolución no está completamente descrita por las variables termodinámicas. Es decir, en las mismas condiciones de temperatura, presión y volumen, dos sistemas «idénticos» pueden tener una evolución transitoria diferente hacia el equilibrio. En esos casos se habla de efectos de memoria, pues los sistemas «recuerdan» su pasado, o preparación, hasta el inicio del experimento.

Uno de estos efectos de memoria se conoce como efecto Kovacs (1964). Este ocurre cuando, en un experimento donde el sistema está evolucionando hacia su estado de equilibrio, se cambian las condiciones de manera prematura para que el sistema se encuentre ya en sus condiciones de equilibrio, pero, sin embargo, el sistema omite ese cambio de condiciones y, momentáneamente, se aleja del nuevo equilibrio para luego acercarse (Lasanta et al., 2019).

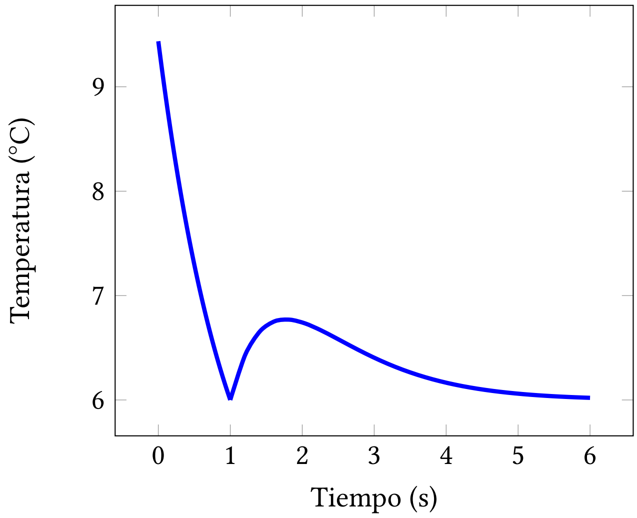

Fig. 2.1 Representación del efecto Kovacs. Evolución de la temperatura de un sistema que, en principio, debería disminuir hasta los 4 °C a los que está establecido el termostato, pero que, en el instante $t = 1$ s, se cambia el termostato a 6 °C con la esperanza de que la temperatura del sistema no evolucione. Sin embargo, en lugar de quedarse estancada en el valor instantáneo, primero se aleja hasta $t = 2$ s y luego retoma el valor que debería tener.

Algunos economistas consideran que un sistema económico también puede mostrar este efecto. Por ejemplo, cuando la economía se retrae y los gobiernos deciden estimularla para frenar esa contracción, por un tiempo, la economía continúa en contracción hasta un momento que vuelve a recuperarse. O al revés, inicialmente se fortalece pero, pasado un tiempo, se contrae hasta el punto en que se encontraba en el momento de aplicar los estímulos.

Otro efecto poco intuitivo es el efecto Mpemba (Mpemba and Osborne, 1969). *Grosso modo*, este se observa cuando un sistema alcanza el equilibrio (se enfría) más rápido si su temperatura inicial es «caliente» en lugar de «templada». En el caso particular del agua, cuando colocamos dos recipientes idénticos en un congelador, y el que estaba inicialmente más caliente se congela antes que el que estaba inicialmente a una temperatura intermedia.

Este último efecto tiene interés práctico, pues puede ser aprovechado para acelerar el calentamiento (o enfriamiento) de un sistema, como podría ser un horno industrial (Gal and Raz, 2020; González-Adalid Pemartín et al., 2021).

2.2 Algoritmos de clasificación y predicción

En las plataformas de contenido digital hay un interés comercial detrás de hacer recomendaciones a sus usuarios, pues haciendo sugerencias adecuadas satisfacen y retienen al cliente.

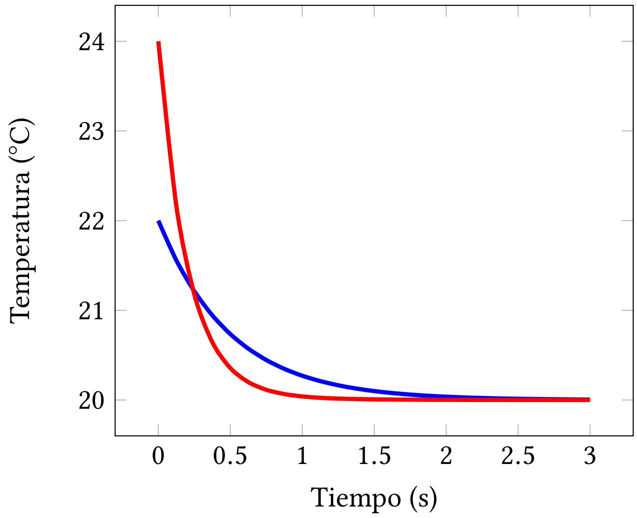

Fig. 2.2 Representación del efecto Mpemba. Evolución de la temperatura de dos sistemas idénticos, pero inicialmente a temperaturas diferentes, donde el sistema más caliente (curva roja) alcanza antes la temperatura de equilibrio (20 °C, establecidos por un termostato) que el sistema inicialmente más frío (curva azul).

Una forma muy sencilla de hacerlas es mostrando el contenido más consumido. Esto no requiere más que ordenar el material por número de visualizaciones o popularidad. Pero esto es una recomendación impersonal, guiada por el consumo de masas. Es decir, el cliente no es el protagonista de la sugerencia. Es muy diferente a hablar con otra persona conocedora tanto del contenido como del cliente, que pueda hacer recomendaciones personalizadas.

Una forma de personalizar las sugerencias podría ser tener un equipo de expertos que atiendan a los clientes. Sin embargo hoy en día esto sería impracticable por el elevado número de usuarios que tienen las distintas plataformas de contenidos. Pero en ese inconveniente hay una virtud: la disponibilidad masiva de información. Esto hace que sea posible utilizar algoritmos matemáticos que perfilen a los usuarios, obtengan patrones de consumo, hagan asociaciones entre perfiles similares, y puedan predecir qué contenido puede ser de mayor interés para un usuario en particular (Davidson et al., 2010; Gomez-Uribe and Hunt, 2015). Y todo esto, hacerlo con *todos* los usuarios.

Aunque estos algoritmos son muy útiles, no hay que perder de vista que son una herramienta que, y omitiendo el interés particular de las empresas que los utilizan, nos satisfacen aprendiendo de nosotros. Es decir, somos nosotros quienes moldeamos sus elecciones (y ellos nos ahorran tiempo de búsqueda). Esto puede crear un sesgo muy elevado y reducir mucho el espectro de contenidos que nos ofrece.

En otro contexto, que no es el de contenido audiovisual, sino en el de las noticias[2], ese sesgo puede acotar nuestra percepción de los acontecimientos, además de amplificar ciertos contenidos en

[2] Notemos que servicios como Facebook funcionan como agregadores de contenido y noticias.

particular. Adicionalmente, si una plataforma ofrece también espacio de discusión, se acrecenta la formación de «cámaras de eco», como se menciona en el capítulo 1.

Para evitar que surjan problemas como los tratados por O'neil (2016), hay que ser conscientes de que estos algoritmos no son magia[3], y que nosotros mismos podemos influenciarlos para romper, aunque sea ligeramente, con el sesgo que de forma natural les introducimos. Pero para eso es necesario tener unas nociones básicas sobre las ideas en las que se fundamentan las técnicas que utilizan.

En el anexo B.1 se describen tres de los algoritmos elementales en el tratamiento de datos[4]: descomposición en valores singulares (para reducir la dimensionalidad de la variedad que contiene los datos), análisis de componentes principales (para detectar la relación que mejor describe la variabilidad de los datos), y k-medias (para catalogar los datos).

Por último, en el ámbito de las predicciones, se puede señalar la existencia de algoritmos algo más propensos a generar problemas éticos, como es el caso de la predicción de criminalidad o, mejor dicho, de la vigilancia predictiva (*predictive policing*) (Mohler et al., 2015). En el caso que se cita, se trata de hacer pronósticos de los lugares en los que podría ocurrir un crimen con el fin de poder enviar un coche patrulla, teniendo la esperanza de que la mera presencia de la policía sea suficiente como para que no ocurra ese crimen.

En particular, con este algoritmo se han hecho ensayos prácticos y, aunque tiene una gran probabilidad de error, ha mejorado bastante los resultados de las herramientas usuales de la policía. Por supuesto, como se dice previamente, este algoritmo puede incurrir en problemas de sesgo. Concretamente en tendencias a la discriminación por cultura o color de piel.

2.3 La obligación de uso de mascarilla como método de control social

Con el advenimiento de la pandemia COVID-19 diversos gobiernos optaron, entre otras medidas, por hacer obligatorio el uso de mascarilla. Decisión que, en gran parte, se ha mantenido (o vuelto a establecer) debido a que se ha demostrado que la vía principal de trasmisión del virus es a través aerosoles. Es decir, respirando.

A raíz de esto han surgido diversos grupos de corte ideológico generalmente común que claman que tal medida atenta contra las libertades individuales y que es una forma de control social y opresión por parte de los gobiernos. Por si no fuera poco, también hay que pagar por esas mascarillas con el dinero propio[5].

¿Qué puede tener esto que ver con matemáticas? Otra vez más, por el uso de algoritmos. Concretamente los algoritmos relacionados con el reconocimiento facial. En la vida de los adolescentes (y no

[3]En referencia a la frase de Arthur C. Clarke: «Any sufficiently advanced technology is indistinguishable from magic».

[4]Aunque se describen en el anexo, cabe la posibilidad de explicar las ideas elementales a alumnos de 2ºBachiller, pues ya están familiarizados con el concepto de matriz.

[5]Igual que la ropa de vestir, sobre la cual, si bien no hay una ley a nivel estatal o comunitario obligando su uso, el no usarla en una ciudad suele acarrear una multa. Además de que ya existen ordenanzas municipales restringiendo el nudismo.

adolescentes) éstos son casi omnipresentes: plataformas como TikTok deben detectar la presencia de un rostro en la pantalla para poder aplicar filtros visuales (como agrandar los ojos o añadir maquillaje virtual); el propio bloqueo del teléfono móvil debe poder reconocer a su dueño; Facebook, en nuestras fotos con otras personas, hace sugerencias de a quién etiquetar o, lo que es lo mismo, identifica a nuestros contactos.

Es decir, ya existe la tecnología (o al menos la base tecnológica) que permite identificar y monitorizar a las personas a través de cámaras (o a través de las redes). Esto, por ejemplo, y según la intención, puede llegar a ser un problema si se utiliza para identificar manifestantes y monitorizarlos posteriormente. Parafraseando a Edward Snowden, la democracia actual la hemos logrado a través de *ideas* revolucionarias, las cuales, sin un mínimo de privacidad, no podrían haber prosperado.

Así pues, paradójicamente y con un carácter contrario al de los grupos previamente mencionados, las mascarillas pueden ser un elemento subversivo, como ha ocurrido en las protestas de Hong Kong de 2019 (Cheung and Hughes, 2020; Hernandez and Scarr, 2019). Además de proteger contra distintos tipos de gas que la policía puede lanzar para disuadir o dividir a los manifestantes, se ha comprobado que, en general, las mascarillas reducen la efectividad de los algoritmos de reconocimiento facial (Ngan et al., 2020)[6].

2.4 La encriptación y su «prohibición»

En la línea de la sección anterior, otra herramienta que se utiliza para proteger la información o garantizar la privacidad es la encriptación de datos. Con los avances del álgebra, la aritmética modular y la informática (concretamente la ciberseguridad), disponemos de algoritmos y prácticas de cifrado que hacen que sólo los que conozcan las claves de encriptación puedan tener acceso a la información real[7].

Sin ir más lejos, las compras por internet dependen por completo de ello. Como poco, debemos facilitarle nuestros datos bancarios a los vendedores (en los que confiamos) para que puedan cobrar, esperando que, en el camino, nadie robe esos datos.

Cualquier tipo de identificación en una página web también depende de ello, pues al fin y al cabo la información que brindamos, antes de ser contrastada con la base de datos de usuarios de la misma, tiene que pasar por distintos nodos (como *routers* o repetidores) y, en principio, terceros podrían verla.

Algunos programas de mensajería móvil disponen de esquemas de encriptación para las conversaciones, como es el caso de Signal, que trae la encriptación por defecto. Otro programa que dispone de ella es Telegram, pero en éste debe activarse la seguridad manualmente. Más recientemente, WhatsApp ha contratado la licencia de uso del sistema de Signal.

[6] Cabe señalar que estas pruebas han sido realizadas con algoritmos de desbloqueo facial. Es decir, sólo tienen que decidir si la cara que observan se corresponde a la del dueño del dispositivo o no. En un caso de vigilancia masiva, los algoritmos deben comparar con una base de datos, por tanto cabe esperar que las mascarillas sean aún más efectivas contra éstos.

[7] No es imposible que un tercero pueda romper la seguridad, pero no es viable en términos del tiempo que requeriría.

Las conversaciones cifradas pueden ser muy útiles para organizar protestas en un estado policial (Aiken, 2020). Pero en los últimos tiempos se han ido haciendo esfuerzos por limitar o ilegalizar el uso de la encriptación. Esta, aunque tiene más recorrido histórico en EE.UU., es una postura que se ha ido importando en la Unión Europea (Huertas Cerdeira, 2020), encarnada recientemente en el lema «seguridad a través de la encriptación y seguridad *a pesar* de la encriptación», promovida con la intención de protegernos del terrorismo y proteger a los niños de las redes de pedofilia.

Cabe señalar que prohibir o limitar (en el sentido de dejar «puertas traseras») la encriptación punto a punto no tiene mucho sentido, pues sería equivalente a prohibir o limitar parte de las propias matemáticas. Por tanto es una estrategia que sólo afectaría a la población general, reduciendo su seguridad y privacidad en las redes, mientras que un agente malintencionado podría utilizar igualmente los algoritmos de encriptación actuales, dado que no son más que matemáticas[8] (y algunos matices técnicos adicionales).

[8]Un ejemplo de ello es el algoritmo RSA de encriptación asimétrica o de cable pública, explicado en al Anexo B.2.

Capítulo 3

Las matemáticas contra algunos engaños

En este capítulo se presenta una serie de engaños que son más comunes de lo que cabría esperar, y una forma de abordarlos desde el punto de vista de las matemáticas. Se procurará brindar ideas con un cierto enfoque pedagógico accesible a estudiantes de último curso de ESO o de Bachiller, que se materializará en forma de actividades a realizar o explicaciones simplificadas que mantengan la esencia de la argumentación.

3.1 Simplificando la realidad: algunos usos erróneos de la estadística

El mundo y la realidad ciertamente son bastante complejos y, para poder estudiarlos y entenderlos un poco mejor, los catalogamos y simplificamos. En cierta manera, «quitamos la paja» filtrando aquello que es irrelevante, al menos aparentemente.

Las simplificaciones son muy útiles (de lo contrario nadie las utilizaría[1]), pues nos brindan nociones generales y, además, es posible argumentar en base a ellas. Pero también puede ocurrir que un detalle sobresimplificado de la realidad no sea suficiente fundamento para un argumento. Por ejemplo, esto puede ocurrir cuando hay varios postulados compatibles con esos detalles pero no entre ellos.

La estadística (en sus distintas ramas) es una de las herramientas que se puede utilizar para percibir la realidad de una forma más sencilla y, además, ofrece medios para cuantificar cuánto podemos confiar en las simplificaciones. Pero, al igual que cualquier herramienta, mal utilizada puede llevar a conclusiones erróneas.

A continuación tratamos algunas situaciones comunes donde la estadística está incorrectamente utilizada para sacar conclusiones.

[1] Esto podría interpretarse como un ejemplo de falacia *ad populum*.

3.1.1 Correlaciones espurias

Uno de los usos incorrectos de la estadística es el establecimiento de una relación causal cuando lo único de lo que se dispone es de una correlación entre datos. Es decir, cuando hay dos (o más) fenómenos que, en conjunto, parecen obedecer un cierto patrón, y concluimos que una cosa causa la otra.

Un ejemplo típico es afirmar que el calentamiento global está causado por la disminución del número de piratas en los mares (ver Fig. 3.1). Pero esto no tiene por qué ser así, pues puede ser una mera coincidencia, como es el caso del ejemplo en cuestión.

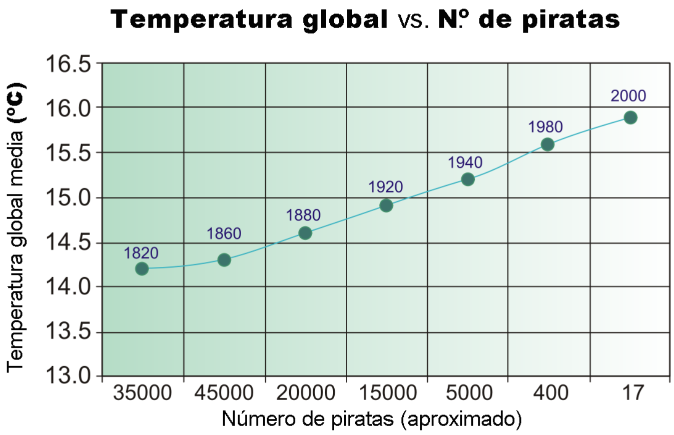

Fig. 3.1 Evolución en los dos últimos siglos de la temperatura media frente al número de piratas en los mares.

Otro ejemplo de correlación espuria es aquel en el que los fenómenos observados, si bien presentan un patrón parecido, no son causa el uno del otro. Sino, mas bien, que hay otros fenómenos que los causan y, de alguna manera, no los estamos teniendo en consideración. Es el caso del incremento en ventas de comida orgánica y del número de casos diagnosticados de autismo en EE.UU. (ver fig. 3.2). En este caso hay varios cofactores, como puede ser el aumento mismo de población (unos 25 millones en esa década), unido a cambios en la dieta general de la población, así como un mayor número de pruebas o diagnósticos de autismo (y no necesariamente una mayor prevalencia).

Aún con todo, en ciertas circunstancias, cabe la posibilidad de que haya una relación causal (la cual habría que comprobar o justificar, y una correlación sólo es una pista de que podría haber algo interesante). Por ello este tipo de relaciones se prestan a ser utilizadas para hacer pseudociencia y justificar algunas líneas de pensamiento.

3.1.1.1 Ejercicios

Dado que calcular correlaciones y su significancia estadística no es parte del programa, se plantean una serie de ejercicios más bien intuitivos para los propios alumnos.

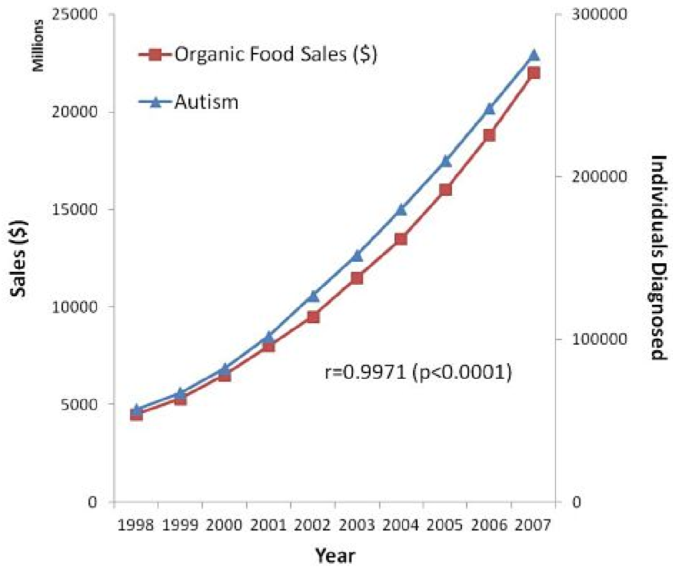

Fig. 3.2 Ventas de comida orgánica y número de casos diagnosticados de autismo, durante la década de 1998 a 2007. Fuente: https://www.datasciencecentral.com/profiles/blogs/hilarious-graphs-and-pirates-prove-that-correlation-is-not.

Tarea I: Pensar o hallar dos fenómenos que no tengan relación causal aparente pero que tengan una tendencia en común. Por ejemplo, tasa de divorcios y consumo medio de leche.

Tarea II: Pensar o hallar dos fenómenos que tengan una tendencia en común y que sea uno consecuencia del otro, e intentar encontrar un motivo.

Tarea III: Intentar hallar varios fenómenos que tengan una tendencia en común y que, además, puedan tener alguna relación causal. Debatir en clase la posibilidad de que alguno de ellos sea la causa del resto. Por ejemplo, el aumento de la riqueza *per cápita* en un país puede explicar el aumento de consumo eléctrico en los hogares, gasto en ocio, gasto en ropa, o inversión sanitaria, entre otros.

Una fuente de este tipo de relaciones es la página web https://tylervigen.com/spurious-correlations, así como un apartado más concreto: https://tylervigen.com/discover.

3.1.1.2 Estableciendo una relación causal

Después de insistir en que una correlación no implica causalidad, quizá sea pertinente señalar cómo se puede establecer un vínculo causal.

Las herramientas para hacerlo dependen del área de conocimiento en la que se trabaje, pero no suele bastar con una correlación de datos. En esencia se trata de realizar una deducción argumental a partir del conocimiento ya establecido, pero eso no siempre es algo alcanzable porque puede no ser posible controlar todos los factores involucrados en la fenomenología.

En las ciencias experimentales se hacen contrastes de hipótesis, con diseños de experimento que incluyen tanto grupos de control como de prueba, así como enmascaramiento de los grupos (como el doble ciego, para evitar cualquier origen de sesgos de confirmación). Aunque un contraste de hipótesis sea una herramienta puramente estadística, es este control de las condiciones del experimento lo que hace que, en caso de haber un efecto, pueda considerarse que no sea una relación espuria.

En las ciencias físicas se suelen establecer unos primeros principios con mucho valor explicativo, afianzados a base de experimentación (ver capítulo 4), y a partir de ellos, o mediante ellos, se deducen relaciones causales entre fenómenos. O, a falta de una deducción completa, se establece una tentativa de relación con cierto grado de plausibilidad.

También se pueden aplicar heurísticas de corte filosófico, como es el caso de la navaja de Ockham. La cual, si bien no es un recurso científico, representa esa mentalidad: a igualdad de poder explicativo, una teoría simple es preferible a una compleja. En el caso particular de las matemáticas, viene a significar que es preferible una teoría con pocos axiomas, o teoremas que requieran la menor cantidad de hipótesis posible.

3.1.2 Simplificaciones excesivas

Usualmente los datos «brutos» están influenciados por muchos factores y eso hace que sea dificil entender cómo esos factores afectan a las observaciones. Por ejemplo, en la figura 3.2, se tratan valores absolutos de ventas de comida orgánica, y este número bien podría estar directamente asociado al crecimiento poblacional. *Grosso modo*, si se duplica la población, cabe esperar que se dupliquen las ventas. Pero, como se clama previamente, podría haber un cambio de dieta generalizado (sea por circunstancias socioeconómicas o por avances científicos y tecnológicos).

Una forma de eliminar el efecto del tamaño de la población es buscar una observación relativa, como puede ser el gasto medio (o *per cápita*). Si este se mantiene constante a lo largo de los años, quiere decir que sólo es relevante el tamaño de la población; en caso contrario hay que buscar otros motivos.

Una práctica muy habitual es comparar esas observaciones relativas. Este tipo de comparativas, realizadas por alguien inexperto o malintencionado, pueden incurrir en errores metodológicos. Por ejemplo, la evolución anual del gasto medio en ventas de comida orgánica debería tener en cuenta la inflación monetaria. O, en otro ejemplo, una comparación entre países debería tener en cuenta los diferentes valores de las monedas de cada uno, entre otros aspectos.

En las discusiones económicas, particularmente en las redes sociales, un concepto muy recurrido es el salario bruto anual medio a la hora de comparar la riqueza entre países y, en consecuencia, apoyar distintos modelos económicos, especialmente los de corte neoliberal. Es un recurso sencillo y directo, a diferencia de otros como el «esfuerzo fiscal», erróneamente representado por el engañoso[2] índice de Frank[3].

Sin embargo, comparar únicamente medias puede ser una sobresimplificación de la realidad, pues nos puede interesar no sólo la riqueza general de un país, sino también qué tan equilibrado sea el reparto de la riqueza. De hecho, si los salarios se distribuyesen como una normal, conocer la media no sería suficiente. Dicho de otro modo, incluso la distribución más sencilla y común requiere conocer más de un parámetro.

Aunque encontrar información sobre la varianza o la desviación típica respecto al salario medio es algo más complicado que encontrar éste último, hay otro tipo de estadísticos en relación a esta cuestión que son más accesibles, como bien puede ser la mediana de los salarios. Es decir, el valor que divide a la población en dos: la mitad que cobra como mucho esa cantidad y la mitad que cobra por encima.

La mediana tiene interés porque es un estadístico más estable que la media muestral frente a los valores atípicos. Lo cual viene a significar que si hay un individuo con unas ganancias inusualmente elevadas, la media aumentará, pero la mediana no.

Para que los alumnos (de 2ºESO en adelante) puedan visualizar estas diferencias, se propone el ejercicio que sigue. Cabe recalcar que la ambigüedad de significado de algunas preguntas es deliberada, con el propósito de ilustrar que algunas valoraciones tienen un importante sesgo personal.

> **Tarea I:** Tenemos información de los sueldos en varias empresas (la cual viene representada en la figura 3.3). A simple vista, ¿en qué empresa se gana «más»?
>
> **Tarea II:** Calcular el salario total (el dinero ganado por el total de empleados) y comprobar si este dato se corresponde con la respuesta anterior.
>
> **Tarea III:** Calcular el sueldo medio. De nuevo, ¿en qué empresa se gana «más»?
>
> **Tarea IV:** ¿Y calculando la mediana de los salarios?
>
> **Tarea V (Debate):** ¿Qué empresa parece más rica? ¿Cuál más equitativa? ¿De qué manera es posible usar la información de la media, la mediana, y el total, para apoyar estas elecciones? ¿Representan a los valores reales de manera fiel?

Por último, para dar un sentido de utilidad y realidad al ejercicio anterior, se propone otro ejercicio:

> **Tarea VI:** ¿Es posible encontrar países que, en media, sean más ricos que otros pero con mayor desigualdad? Si es que sí, ¿En te apoyas para afirmarlo?

Por ejemplo, en https://worldpopulationreview.com/country-rankings/median-income-by-country podemos encontrar que en Reino Unido, de media, los salarios son ligeramente mayores que en Bélgica.

[2]Algunos fallos conceptuales vienen descritos por Rodríguez (2020), y uno de sus usos falaces analizado por Cabrales (2021).

[3]Una de las personas conocidas que utilizan este concepto es el economista Juan Ramón Rallo, además de ser uno de los motivos argüidos por diversos «youtubers» para justificar su traslado de España a Andorra.

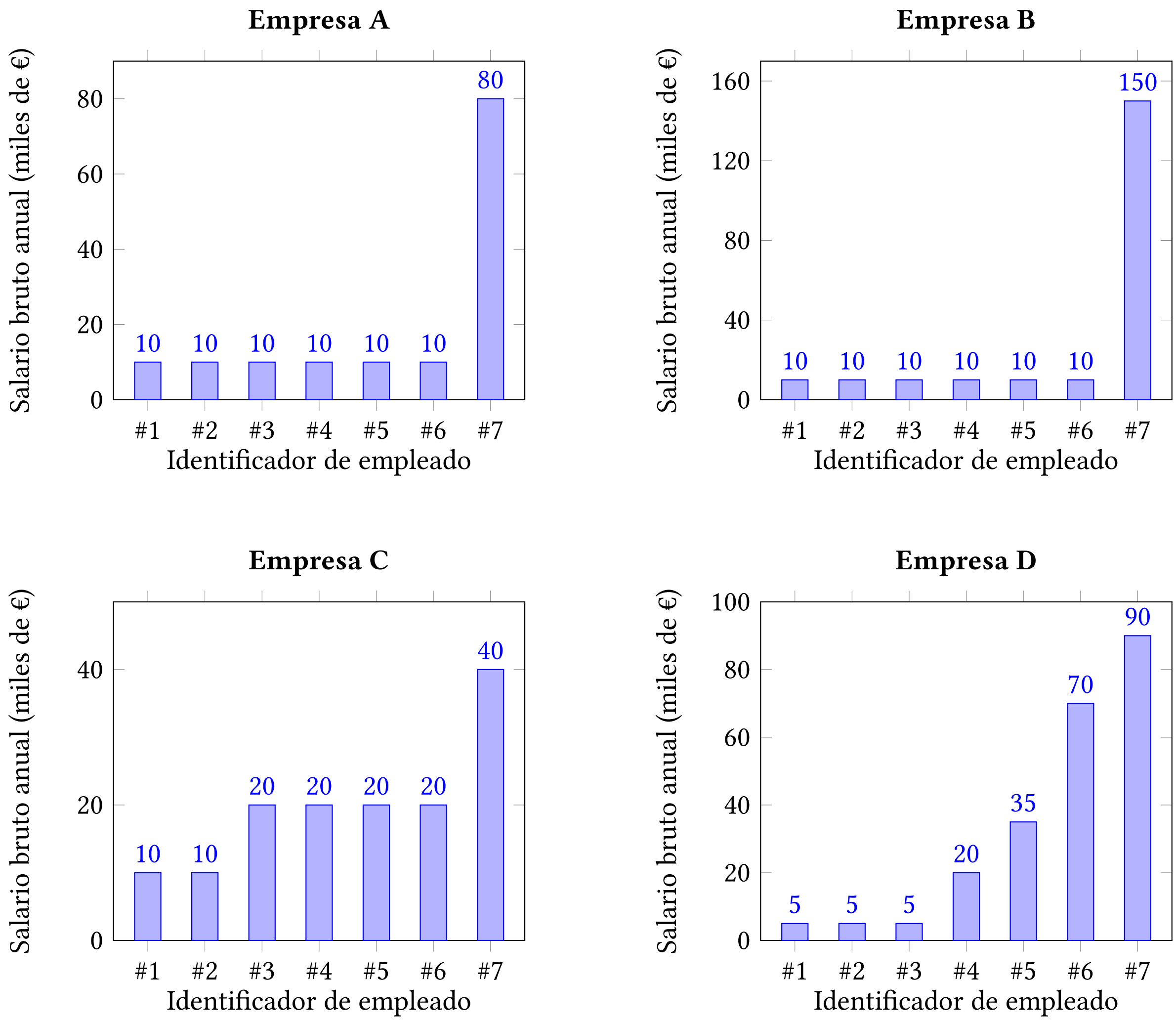

Fig. 3.3 Distribución de salarios en cuatro empresas ficticias.

Sin embargo, la mediana en Bélgica es bastante superior a la de Reino Unido. Dicho de otro modo, hay más personas con un salario mayor en Bélgica.

3.1.3 Interpretación errónea de la información

El contexto actual nos brinda ejemplos de uso erróneo (y posiblemente malintencionado) de datos por parte de algunos colectivos en su afán por desacreditar el uso de vacunas. Nos vamos a centrar en dos, que son a su vez ejemplos de la *paradoja de Simpson*, y como ejercicio en el aula se puede plantear una discusión sobre por qué ocurre y cómo desmontarla. De hecho, puede ser un buen ejercicio para explicar cómo se relaciona el teorema de la probabilidad total con las probabilidades condicionadas (en el anexo B.3 se incluye una explicación formal).

3.1.3.1 Primer ejemplo

Es posible encontrarse con la siguiente afirmación

> «Si en las UCIs de los hospitales hay ingresadas por motivos relacionados a COVID-19 tantas personas vacunadas como no vacunadas, entonces las vacunas no sirven»

Primero habría que señalar que el que hayan personas vacunadas en las UCIs es natural, pues las vacunas tienen un cierto grado de efectividad. Además, argumentando de manera reduccionista, si toda la población estuviera vacunada, sólo habrían personas vacunadas en las UCIs.

Pero esto ciertamente no responde a por qué esa situación no pone en entredicho la efectividad de las vacunas. Para ello podemos ilustrar la situación o podemos realizar una reducción al absurdo.

Una ilustración de la circunstancia sería hablar de una población de 100 personas, donde 90 están vacunadas y 10 no lo están. Si hay 5 vacunados en la UCI y 5 no vacunados, hay igual cantidad de ingresados de cada tipo, pero sólo el 5.6% de los vacunados han necesitado ingresar, a diferencia del 50% de los no vacunados.

En cambio, una reducción al absurdo podría ir como sigue: Sean x e y las probabilidades de que alguien vacunado y no vacunado, respectivamente, ingresen en una UCI. Que las vacunas no sirvan, como sostienen esos colectivos, significa que $x \geq y$. Por otra parte, sea p el porcentaje de población vacunada, y N el número de personas. El número de personas vacunadas ingresadas, por tanto, será $N \cdot p \cdot x$, mientras que el de no vacunadas será $N \cdot (1-p) \cdot y$. Puesto que se observa la misma cantidad de cada colectivo en las UCIs, podemos afirmar que $N \cdot p \cdot x = N \cdot (1-p) \cdot y$. Simplificando, que $p \cdot x = (1-p) \cdot y$, y, como se sostiene que $x \geq y$, llegamos a $p \cdot y \leq p \cdot x = (1-p) \cdot y$. Es decir, que $p \leq 1-p$, y, en consecuencia, que $p \leq 0.5$.

Dicho de otro modo, si los antivacunas tienen razón en sus afirmaciones, se deduce que menos del 50% de la población está vacunada. Pero esto no es consistente con *la realidad*, donde, en las fechas en las que esto ocurre, más del 80% de la población está vacunada con la pauta completa (Ministerio de Sanidad, 2022).

3.1.3.2 Segundo ejemplo

Con los datos brindados por el instituto nacional de estadística de Inglaterra (Office for National Statistics, 2021), se ha empezado a afirmar lo siguiente:

> «Hay una mayor mortalidad por COVID-19 entre gente vacunada que no vacunada»

Cabe destacar que esta afirmación se realiza al agregar los datos del rango de edad entre los 10 y los 59 años, pues esa tasa de mortalidad no aparece de manera explícita en los datos mencionados.

La explicación en este caso es un poco más elaborada, pues es una mezcla entre el ejemplo anterior y las desigualdades que pueden haber dentro de un grupo (como vienen ilustradas en la sección 3.1.2).

Al unificar distintos grupos de edad en uno único, las circunstancias de unos se entremezclan con las de los otros. La población mayor está vacunada casi en su totalidad, y a la vez tiene un mayor riesgo de muerte (en general). Por contra, el porcentaje de población joven vacunada es inferior, y además los jóvenes tienen un menor riesgo de muerte. Sin embargo, comparando entre rangos de edad no tan amplios, la mortalidad es inferior entre vacunados respecto a los no vacunados.

Así, al unificar la información (lo que equivale a asumir que la edad no es relevante), puede parecer que la mortalidad entre los vacunados es mayor que entre los no vacunados. Pero, de nuevo, es un argumento sesgado y no representativo de la efectividad de las vacunas.

3.2 Terraplanismo

Si bien es sabido *hace un par de milenios* (Pirazzini, 2020) que la tierra no es plana (a diferencia de lo que ocurre en algunas novelas de Terry Pratchett –ver Fig. 3.4–), en la última década se ha visto el nacimiento paradójico de la creencia de que, en realidad, sí lo es (Hancock, 2018). También nació la creencia en una conspiración para ocultar «la verdad» sobre este asunto, pero ese aspecto se tratará de manera general en el apartado 3.3.

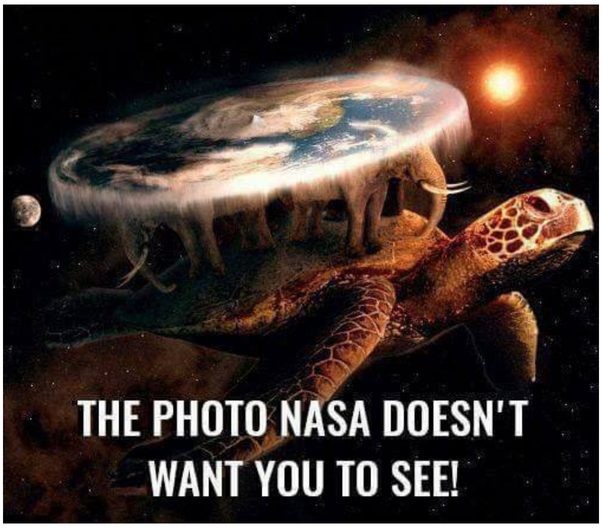

Fig. 3.4 Una tierra plana. Fuente: https://www.reddit.com/r/discworld/.

Una de las justificaciones más antiguas de las que hay registro de que el planeta Tierra no es plano es un compendio de observaciones realizadas, entre otros, por Aristarco de Samos y Eratóstenes de Cirene. Por su parte, Aristarco estima la distancia a la que se encuentra el Sol de la Tierra, mientras que Eratóstenes usa esa evidencia y, junto a una serie de ideas inteligentes (Brown and Kumar, 2011), estima las dimensiones de nuestro planeta. Pero, como parte del proceso de estimación, se encuentra con evidencia a favor de la no planaridad de la Tierra.

Los pasos clave de la justificación son los siguientes:

1. Los rayos del Sol inciden en paralelo sobre la Tierra.
2. Dos objetos en la misma longitud pero diferente latitud y de la misma altura proyectan «al mismo tiempo» sombras de diferente longitud.
3. En consecuencia, la Tierra no puede ser plana.

Con la intención de acercar las justificaciones a los estudiantes, vamos a presentarlas en forma de actividades que los propios estudiantes pueden realizar. Parte de la actividad se tendrá que realizar con el apoyo de las redes sociales (o contactando con otro centro educativo ubicado en otra latitud geográfica que quiera hacer la misma actividad).

3.2.1 Actividad I – El Sol está *muy* lejos

Aristarco fue capaz de estimar (con mayor o menor acierto) tanto la distancia al Sol como a la Luna en relación al diámetro de la Tierra (Berggren and Sidoli, 2007). Aunque saber que el Sol está muy lejos tanto de la Tierra como de la Luna, y que la distancia de la Tierrra a la Luna es mayor que el diámetro de la Tierra son datos relevantes de cara a la validez del argumento de Eratóstenes, llegar a esas conclusiones requiere de observaciones de eclipses lunares y solares, lo cual se aleja de las posibilidades observacionales de un aula o un alumno.

En cambio, vamos a darnos la licencia de asumir que la Luna está más lejos que cualquier sitio de la Tierra al que podamos llegar por vía terrestre, pero vamos a intentar justificar que el Sol está suficientemente lejos. Con estos dos hechos se puede concluir que los rayos del Sol inciden en paralelo sobre la Tierra.

Medir la distancia al Sol de manera directa es imposible con las herramientas que pueda tener un alumno en casa, pero sí que puede aproximarse el ángulo formado por dos objetos que observemos en la esfera celeste.

Tarea: Pensar en dos objetos relevantes que puedan verse en el cielo al mismo tiempo.

Cabe esperar que sean el Sol y la Luna. Éstos, junto con la Tierra, forman un triángulo (ver Fig. 3.5).

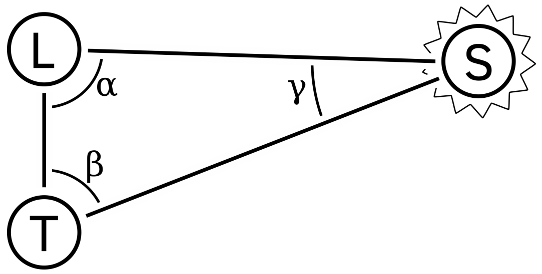

Fig. 3.5 Triángulo cuyos vértices recaen en la Tierra, la Luna y el Sol. *No está a escala.*

De este triángulo, en principio, sólo podemos conocer el ángulo asociado al vértice de la Tierra (denotado por β en la Fig. 3.5).

Pregunta: ¿Podría medirse el ángulo β formado entre esos dos objetos usando un transportador de ángulos?[4]

Pero conocer ese ángulo, por sí solo, no es suficiente. Hay que conocer el ángulo asociado al vértice de la Luna (α en la Fig. 3.5).

Tarea: Hallar algún aspecto conocido sobre la Luna que pueda indicarnos la posición del Sol respecto a ella y la Tierra. Usar la Fig. 3.6 como inspiración.

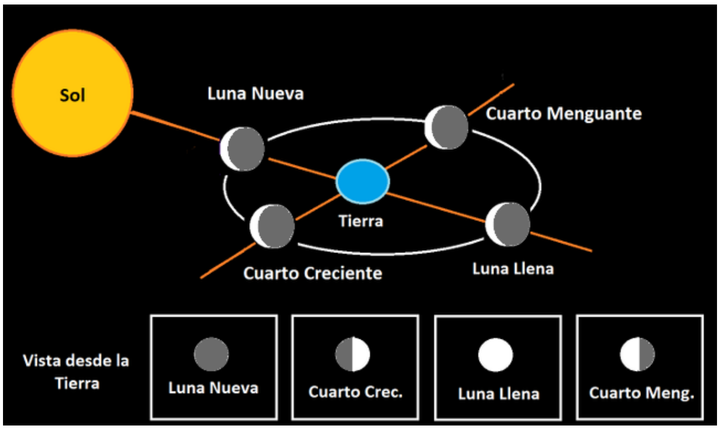

Fig. 3.6 Posiciones de la Luna respecto a la Tierra y el Sol, y sus respectivas fases vistas en el cielo nocturno. *No está a escala.*

Como podemos ser capaces de presenciar la Luna en cuarto menguante o en cuarto creciente y, vista su iluminación, podemos concluir que el ángulo asociado a la Luna, α, es aproximadamente 90º.

Al tratarse de un triángulo, sabemos que los ángulos deben sumar 180º.

Tarea (opcional): Medir el ángulo β entre el Sol y la Luna cuando esta última esté en cuarto creciente.

Si se hace, se verá que este ángulo también está muy cerca de los 90º. En consecuencia el ángulo asociado al vértice del Sol es cercano a 0º.

Tarea: Intentar representar de forma más o menos fidedigna el triángulo formado por la Tierra, el Sol y la Luna. En comparación con la Luna, ¿Qué tan lejos parece estar el Sol?

Con esto ya podemos confiar en que el Sol está tan lejos que sus rayos caen sobre la Tierra de forma prácticamente paralela.

[4] Lo planteamos como pregunta pues se presentará una resolución alternativa.

3.2.1.1 Forma alternativa de estimar el ángulo entre la Luna y el Sol

Podemos concluir lo mismo si hacemos esa medición al atardecer, mientras el Sol se pone en el horizonte (ver Fig. 3.7). El ángulo asociado a la Luna seguiría siendo casi de 90° puesto que está en cuarto creciente. Pero, dado que veríamos la Luna delante de nosotros, y el Sol en el horizonte, el ángulo será, también, cercano a 90°.

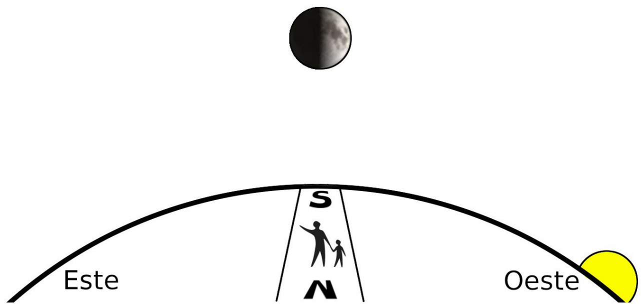

Fig. 3.7 Diagrama que representa a una persona en el hemisferio norte, orientada al sur, y lo que vería al atardecer cuando la Luna se halla en cuarto creciente.

3.2.2 Actividad II – Midiendo sombras en latitudes diferentes

Esta tarea es bastante directa y busca replicar la labor de Eratóstenes, que se encargó de conocer la longitud de las sombras que proyectaba un mismo objeto en Siena y en Alejandría en un mismo instante.

Para esta tarea los alumnos o el propio profesor deberá ponerse en contacto con otros, ubicados en una latitud diferente, para coordinar una medición simultanea de la sombra de un objeto idéntico.

> **Tarea:** Hallar objetos que puedan tener la misma altura en distintos sitios del país. En su defecto, fabricarlos.

Estos podrían ser aquellos que sigan alguna normativa oficial (como carteles de tráfico), o bien un palo de madera de dos metros.

Una vez coordinados, deberá hacerse de manera simultánea la medición de las sombras con esos objetos puestos de manera perpendicular al suelo.

> **Tarea:** Medir las sombras al mismo tiempo.

Si hay varios grados de diferencia en la latitud de los lugares donde se realizan ambas mediciones, saldrán medidas claramente diferentes.

3.2.2.1 Sobre la longitud geográfica

A ser posible, los lugares donde se midan las sombras deben estar en la misma longitud geográfica. Dada la existencia de usos horarios, la hora oficial no coincide con la hora solar.

3.2.3 Actividad III – Concluyendo que la Tierra no es plana

En esta actividad los alumnos deberán razonar por qué una Tierra plana no es consistente con la evidencia de las dos actividades anteriores. En esencia se trata de realizar una reducción al absurdo.

Para ello se pueden apoyar en un modelo a escala.

> **Tarea:** Sobre una cartulina, colocar dos bolígrafos o palillos de forma perpendicular.

Como queremos emular lo que sucede con la Tierra, sacamos ese modelo a la luz natural del exterior.

> **Tarea:** Medir las sombras cuando la cartulina está aplanada y cuando la cartulina está algo doblada.

> **Tarea:** Justificar qué circunstancia es consistente con las observaciones de la actividad II.

Con esto deberían tener ya una prueba al alcance de sus manos (al menos gran parte de ella) de que la Tierra no es plana.

3.2.3.1 Experimento con luz artificial

Es posible que alguien pruebe el modelo con luz artificial y se percate de que, sobre la cartulina plana, las sombras tienen diferente longitud.

> **Tarea:** Justificar que este experimento no es consistente con las observaciones de la actividad I.

3.2.4 ¿Cómo se hizo sin comunicación instantánea?

Los alumnos podrían desconfiar de la forma de realizar la actividad II. Por un lado, no había forma instantánea de comunicarse para medir *al mismo tiempo* las sombras en Siena y Alejandría. Por otro, estos lugares están en longitudes geográficas diferentes.

Sin embargo, la metodología de Eratóstenes fue lo suficientemente brillante como para no sufrir de esos dos problemas (Brown and Kumar, 2011). Las medidas no fueron simultáneas, pero sí que se realizaron el mismo día y a la misma *hora solar* de cada lugar. Escogiendo previamente el día y el momento se logra sortear la falta de comunicaciones en tiempo real y, por usar la misma hora solar en cada ciudad, corregimos el problema de la diferencia de longitud geográfica entre ambos sitios.

La forma de establecer la fecha fue escogiendo un día particular que no depende del calendario. Éste fue el solsticio de verano, que se puede identificar como el día que más alto llega a estar el Sol en el año. Para fijar el momento de la medición se procedió con una idea similar y se escogió el mediodía de la hora solar. Es decir, el momento del día en que el Sol está lo más alto posible en el cielo. O, visto de otro modo, cuando la sombra es lo más corta posible en cada sitio.

3.3 Las grandes conspiraciones y su plausibilidad

Otro tipo de engaño habitual en las redes es el de las teorías sobre conspiraciones. Por ejemplo, que las grandes farmacéuticas tienen vacunas para ciertas enfermedades pero que, por intereses económicos, no las sacan a la luz. O que los gobiernos obligan a vacunarnos y que, por medio de las vacunas, nos inyectan algún tipo de dispositivo de control. Otra conspiración típica es que el alunizaje fue un montaje de la NASA y los EE.UU. y que la humanidad nunca pisó realmente la Luna.

Aunque este tipo de ideas no son algo novedoso[5], ahora se ven reforzadas por un cierto descontento y desconfianza respecto a las instituciones, además de nuevas concepciones sobre las particularidades que debe tener algo para ser considerado como verdadero.

Si bien las teorías antes comentadas pueden parecer inverosímiles[6], puede ocurrir que se oculte información de manera deliberada al público general. Sin ir más lejos, en la última década se ha filtrado información sobre un programa secreto de vigilancia, Prism, organizado por varios servicios de inteligencia de EE.UU. (BBC, 2014). La existencia de este programa fue desvelada por las filtraciones de un analista, Edward Snowden, de la propia agencia de inteligencia (C.I.A.) y, posteriormente, de la agencia de seguridad nacional (N.S.A.) (BBC, 2013).

Inspirado en estos hechos, Grimes (2016) propone un modelo relativamente sencillo para estimar la probabilidad de que una conspiración falle con el paso del tiempo. En este caso, que una conspiración «falle» significa que algún participante filtra información sobre la actividad, sea de manera deliberada o accidental.

Con este modelo podemos estimar, entre otras cosas, el tiempo que haría falta para que una conspiración falle con probabilidad superior al 95%. Por ejemplo, la estimación para el montaje del alunizaje es algo inferior a los cuatro años, al igual que ocurre con la ocultación de vacunas (si se tiene en cuenta a todos los posibles participantes, y no nos restringimos únicamente a la Organización

[5]Sin necesidad de mirar demasiado al pasado, en la década de los noventa ya habían algunas teorías conspiratorias formando parte de la cultura popular, como ocurre con la idea de que un OVNI se estrelló en Roswell, EE.UU., en el año 1947, y que el ejército del país está ocultando la evidencia de tal hecho.

[6]Por ejemplo, estos últimos dos años apoyan la idea de que, en todo caso, es más rentable ofrecer las vacunas al público, y no lo contrario.

Mundial de la Salud). Dicho de otro modo, *lo más probable es que una conspiración se caiga por su propio peso.*

En las secciones que siguen se detallarán las ideas principales del modelo propuesto (con la intención de convencer al propio profesorado), así como una simplificación para hacerlo accesible a estudiantes de 2ºBachiller.

3.3.1 Ingredientes principales del modelo de Grimes (2016)

Para este modelo se asume que no hay agentes externos a la conspiración que intenten, de forma activa, descubrirla. Es decir, si se desvela, es por culpa de alguno de los participantes. Esto puede ser bien porque alguno decide filtrar datos voluntariamente o porque comete un error y se filtran igualmente. Por ello, a cada participante se le asigna una probabilidad «intrínseca», denotada por p, de que, por el motivo que sea, acabe filtrando información en un instante dado. A modo de simplificación, se asume que todos los participantes tienen la misma probabilidad de cometer fallos o filtrar información.

También, por su naturaleza dentro de una conspiración, tiene sentido asumir que las filtraciones, de ocurrir, son independientes y aisladas. Es decir, se pueden considerar *eventos raros*. Con estas premisas se puede considerar que el número de filtraciones a lo largo del tiempo es un proceso de Poisson (ver el Anexo B.4 para una descripción más detallada).

Esto hace que el tiempo necesario para que haya *al menos* una filtración siga una distribución exponencial, cuya función de densidad acumulada tiene la forma $1 - e^{-t\lambda}$, donde $\lambda > 0$ es la tasa de filtraciones que se espera en un intervalo de tiempo dado.

El valor asignado a λ se relacionará con la probabilidad intrínseca p. Pero, además, tendrá sentido considerar que, a mayor número de participantes, mayor es la posibilidad de que ocurran fallos o filtraciones. Así pues, si $N(t)$ denota al número de participantes en un instante t, se define la tasa de filtraciones como sigue: $\lambda := 1 - (1 - p)^{N(t)}$. Esto lo podemos interpretar del siguiente modo: si $N(t) = 1$ (i.e. hay un único participante), $\lambda = p$, con lo cual la tasa de filtraciones representa la probabilidad del propio participante de filtrar información; pero si $N(t)$ es muy grande, entonces λ se acerca a 1 y la filtración está asegurada, salvo que la conspiración sea perfecta y se tenga que $p = 0$.

Como cabe esperar, la probabilidad intrínseca p no es algo que se pueda medir directamente. Sin embargo se puede calibrar utilizando información de varias conspiraciones que han quedado al descubierto[7].

Por otra parte, para el número de participantes en un instante dado, $N(t)$, es posible estimar su valor inicial según los datos y, dependiendo de la naturaleza de la conspiración, asumir un cierto tipo de evolución. En el trabajo citado se contemplan tres tipos: no hay cambios en el número (es una conspiración que requiere un mantenimiento constante), los participantes van muriendo por causas naturales, o los participantes van siendo activamente eliminados.

[7] En caso de que la información fuese parcial, se hacen sobrestimaciones de ciertos datos. Esto es una decisión conservadora dado que, en última instancia, aumenta las probabilidades de que los conspiradores tengan éxito según el modelo.

3.3.2 Simplificación para 2ºBachiller

Los procesos estocásticos en tiempo continuo, aparte de estar bastante lejos del currículum, pueden ser poco intuitivos[8]. Así que añadiremos tres simplificaciones para poder abordar el modelo con herramientas de 2ºBachiller, incluso aunque ello haga que se pierdan algunas características del modelo original.

La primera simplificación es considerar que el tiempo es discreto. Por ejemplo, contemplando el tiempo en términos de días en lugar de instantes infinitesimales. Es decir, que cada día «se decide» si un participante filtra información o no, con una cierta probabilidad $p > 0$. La segunda simplificación es establecer un límite de días a contemplar, d.

Bajo estos supuestos, el número de filtraciones que hace *un participante*, X, durante un periodo de d días se distribuye como una binomial de d pruebas con probabilidad de «éxito» p. Es decir, consideramos como «éxito» el que haya una filtración. En estos términos, la probabilidad de que una conspiración se desvele por culpa de ese participante en particular es $\mathbb{P}(X > 0)$, o, lo que es lo mismo, $1 - \mathbb{P}(X = 0) = 1 - \binom{d}{0}p^0(1-p)^{d-0} = 1 - (1-p)^d$.

El tercer supuesto es que hay N participantes durante toda la conspiración. Para cada uno de ellos aplica el párrafo anterior. Así pues, el número de filtraciones que pueda tener la conspiración será la suma de todas las filtraciones asociadas a cada participante. Esta nueva variable aleatoria también se distribuye como una binomial con la misma probabilidad de éxito p, pero con $d + d + \overset{N)}{\ldots} + d = Nd$ pruebas. Es consecuencia, la probabilidad de que la conspiración se descubra es $1 - (1-p)^{Nd}$.

Esta simplificación, aparte de apoyarse en un tipo de distribución que es parte del currículum, se presta a realizar algunos experimentos sencillos con los alumnos en el propio aula de clase, así como una serie de ejercicios.

3.3.2.1 Ejercicios

En la línea del trabajo de Grimes (2016), podemos plantear a los alumnos que resuelvan las siguientes cuestiones.

> **Problema I:** Supongamos que hay $N = 5$ personas guardando un secreto y que, cada uno de ellas, tiene una probabilidad hacerlo público de $p = 0.05$.
>
> 1. En un plazo de cuatro días, ¿Es totalmente seguro que alguien comente el secreto con quien no debe?
> 2. Calcular la probabilidad de que el secreto salga a la luz en, como mucho, uno, tres, diez, y veinte días.

[8] Si en cada «instante de tiempo» hay una probabilidad no nula de que ocurra un fallo, es natural pensar que deberían ocurrir fallos constantemente si consideramos el tiempo como un continuo.

3. ¿Cuántos días hacen falta para que, con probabilidad del 99%, la conspiración se descubra?

Problema II: Estamos interesados en poner en marcha un proyecto y ocultarlo durante unos $d = 100$ días, para entonces publicarlo y que tenga mucho impacto mediático (por ejemplo, el desarrollo de un videojuego). Asumiendo que cada miembro del proyecto tiene una probabilidad de filtrar información de $p = 0.01$, nos preguntamos qué tan viable es esta estrategia de marketing.

1. Calcular la probabilidad de que hayan filtraciones a prensa según la cantidad de miembros del proyecto. Probar con $N = 1, 2, 3, 4, 5$. ¿Es una estrategia viable?
2. ¿Y en 50 días?

3.3.2.2 Experimento

En la línea del primer problema, podemos hacer grupos de cinco personas en clase y darle un dado de veinte caras[9] a cada uno de los alumnos, o a cada grupo.

Cada grupo representará una conspiración, y la decisión de filtrar información de cada uno de los miembros se decidirá con una tirada de dado: si sale la cara 20, filtra información y la conspiración falla; en caso contrario la conspiración continúa. Aunque en última instancia no sea relevante, por respetar la analogía de los días, todos los miembros deben haber lanzado un dado antes de repetir tirada.

En estas condiciones, tenemos $N = 5$ y $p = \frac{1}{20} = 0.05$. Con lo cual podemos poner a prueba los resultados obtenidos previamente. Es importante que cada grupo anote en qué momento ha ocurrido su filtración.

Tarea I: Simular como mucho tres días de conspiración (tres rondas de lanzamiento de dados). En aproximadamente la mitad de los grupos debería haber una filtración.

Tarea II: Las conspiraciones que siguen vigentes, continuar simulando hasta los seis días (otras tres rondas de lanzamiento de dados). Al final, del total de grupos, unos cuatro de cada cinco habrán sido descubiertos.

Tarea III: Repetir las dos tareas anteriores desde cero.

Tarea IV: Debatir si tiene sentido simular un total de veinte días.

Podemos aprovechar la información recogida y proponer lo siguiente:

[9] Como los que se utilizan en juegos de rol y se venden en tiendas de juegos de mesa.

Tarea V: Entre todos, recopilar los días que tardó cada grupo en que su conspiración fallara. Hacer un histograma y compararlo con la Fig. 3.8[10].

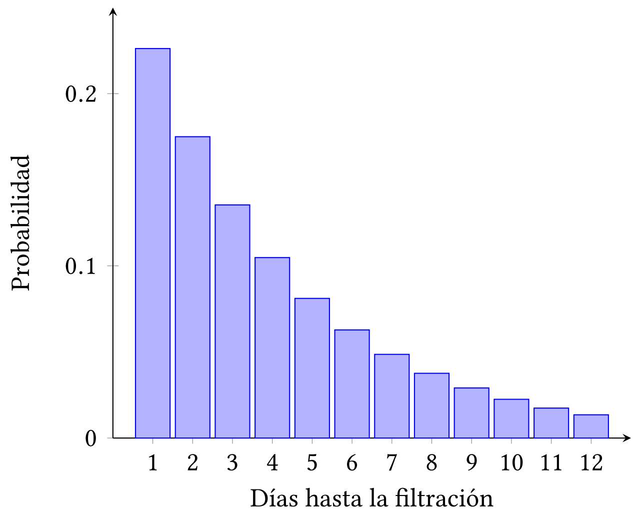

Fig. 3.8 Función de probabilidad del número de días que tomará en filtrarse una conspiración de $N = 5$ personas y con probabilidad de filtración individual de $p = 0.05$.

Por último, y para emular los pasos del artículo de Grimes (2016), podemos experimentar cambiando el número de participantes en una conspiración con el pasar de los días.

Tarea VI: El número de participantes irá reduciéndose: El primer día serán cinco, el segundo día serán tres, el tercero serán dos, y a partir de entonces quedará un único miembro. ¿Cuántos días pasan hasta que la conspiración se destapa?

Tarea VII: Repetir la tarea VI varias veces más y, entre todos y con la información recopilada, realizar un histograma sobre la cantidad de días que hicieron falta para que falle la conspiración. ¿Ha cambiado respecto al anterior?

[10]Dado el bajo número de grupos y experimentos, es posible que la figura y el histograma difieran ligeramente, especialmente en la escala vertical. Sin embargo, de manera aproximada, el número de conspiraciones fallidas al primer día deberían ser alrededor del doble del número al cuarto día.

Capítulo 4

La ciencia y el método científico contra otros engaños

Algunos engaños, para volverse más atractivos, se arman de un discurso pseudocientífico, usualmente cargado de nula experimentación (o con fallos metodológicos y carentes de criterio), afirmaciones descontextualizadas de expertos, o directamente falsas atribuciones a expertos, afirmaciones de falsos expertos, estudios aislados, evidencia anecdótica, y sesgos de confirmación, entre otros recursos.

Usualmente, una vez desarmados los portavoces de estos engaños, también se recurre a una postura posmoderna aparentemente inocente, que es que «todas las opiniones son igual de válidas», lo que, en estas circunstancias, se puede interpretar como una forma elegante de aplicar alguna de las ideas de Schopenhauer (2006).

En este capítulo se repasan algunos aspectos de la práctica científica que distinguen a la ciencia de otras fuentes de saber, y que, en los contextos pertinentes, la ponen irremediablemente por encima. Lo cual hace que no todas las opiniones sean igual de válidas.

4.1 La ciencia no es una religión

Otra afirmación que suele acompañar a la anterior es que uno escoge *creer* en la ciencia y el trabajo científico, al igual que alguien cree en su religión y lo que puedan decir sus mensajeros.

Realizar esta apreciación es natural. Al fin y al cabo, establecemos modelos y teorías que parecen describir el mundo y la realidad en ciertos contextos y circunstancias. Además, puesto que han sido contrastados con experimentos y evidencia, podría decirse que incluso los «aceptamos». Es decir, aceptamos esos modelos, así como aceptamos la evidencia reportada por los científicos.

Una teoría científica tiene un contexto determinado, y busca describir y explicar una serie de fenómenos concretos. Su validez no viene determinada por el número de personas que crean en ella, sino por la experimentación. La cual, por un lado, inspira a las teorías y, por otro, pone a prueba sus predicciones.

Además, las teorías científicas no son absolutas: pueden haber distintas conviviendo al mismo tiempo (como la teoría de cuerdas o la teoría de la gravedad cuántica de bucles, o, en otro ámbito, las distintas interpretaciones de la mecánica cuántica), pueden unas perecer frente a otras (como la teoría de vórtices del átomo, cuando se descartó la existencia del éter luminífero), y pueden extenderse (como la teoría de la relatividad general recoge y extiende la teoría de la gravitación de Newton).

La evidencia, por su parte, gana valor al reproducirse. Un estudio aislado no significa mucho, pero múltiples estudios, siguiendo una metodología correcta y obteniendo las mismas observaciones, ganan valor por el propio peso que acumulan. Darle tanto valor a la reproducibilidad de un experimento, aunque pueda parecerlo, no se trata de una falacia *ad populum*. Esto es porque se sigue una postura escéptica: una hipótesis no se acepta, sólo se descarta. Dicho de otro modo, no se crea verdad, se filtra aquello que resulta inverosímil.

Estos son algunos de los motivos por los que el conocimiento científico no puede tratarse como una creencia. Las teorías no se aceptan o rechazan por lo que diga una autoridad[1], no son perfectas (en el sentido de que sean verdad por imposibilidad de ser falsas), y la evidencia anecdótica o metodológicamente torpe no pueden ser consideradas un sustento para una teoría.

En las teorías pseudocientíficas, en cambio, se aceptan las afirmaciones de algún experto en particular, especialmente si apoyan un cierto sesgo, y se da valor a la anécdota, omitiendo la evidencia general. Esto último se ve alimentado tanto por teorías conspiratorias sobre la comunidad científica a la vez que por la creencia posmoderna de que el conocimiento científico es una construcción social[2].

4.2 El revisionismo científico

La investigación científica y el mundo de las publicaciones no es perfecto (Sokal and Bricmont, 1999), pero esa imperfección se puede aliviar a base de reproducción y comprobación de experimentos, incluyendo la posibilidad de revisión del conocimiento previo.

El revisionismo es especialmente relevante en las ciencias de la vida, como lo son la medicina, la farmacia o la biología. Controlar las condiciones de un experimento suele ser más complicado que en las ciencias físicas, al igual que es difícil conocer la casi totalidad de variables que intervienen en un fenómeno.

La consecuencia de esto último es que puede ocurrir que en ocasiones surjan resultados erróneos y que, cuando son correctos, tengan un alcance bastante limitado[3].

[1]Por ejemplo, A. Einstein rechazaba los postulados de la mecánica cuántica –«Dios no juega a los dados»–, pero posteriormente se convirtió en una de las teorías físicas más verificadas por los experimentos.

[2]Alan Sokal suele invitar a dejar de creer en cualquier teoría sobre la gravedad y saltar por una ventana.

[3]En el caso de la medicina o la farmacia, la evidencia a favor de una hipótesis no hace que la contraria sea automáticamente inverosímil. Por ejemplo, al contemplar distintos tratamientos.

4.2.1 Las vacunas y el autismo

Uno de los casos famosos de revisionismo científico está relacionado con el artículo de Wakefield et al. (1998), en el cual se sostenía que la vacuna del sarampión provocaba autismo en los niños vacunados con ésta.

Es este artículo el primer sustento científico a los colectivos antivacunas. Es decir, hay un artículo de investigación científica que da validez a la postura o decisión de rechazar las vacunas. Además publicado en The Lancet, nada menos que una de las revistas más importantes de investigación en medicina.

Pero es a partir de ahí que la postura antivacunas pasa a ser pseudocientífica. Tal artículo tiene errores metodológicos (por ejemplo, no hay grupos de control), además de no seguir una serie de buenas prácticas, motivo este último por el cual el artículo fue retractado por la revista The Lancet (Plotkin et al., 2009). Además, el artículo no pudo ser reproducido posteriormente. En todo caso la evidencia apunta a que no hay relación entre la vacuna del sarampión y el autismo (Hornig et al., 2008).

Pero el peso de la evidencia general se omite, al igual que las revisiones, y así se dispone de información desactualizada pero que coincide con una postura deseada. Lo que no deja de ser un sesgo de confirmación.

4.2.2 La homeopatía y la medicina alternativa

La homeopatía y distintas medicinas alternativas han ido ganando popularidad en los últimos tiempos. Posiblemente por una conjunción entre la apelación a que todo aquello que sea natural es necesariamente «bueno» (y su *recíproco*: todo lo que no sea natural es «malo») y que pueden surgir ciertos recelos hacia las grandes farmacéuticas.

Esta popularidad es tal que distintas farmacias le dan una visibilidad bastante importante a los productos homeopáticos (ver fig. 4.1). Sin embargo es una popularidad infundada, al menos científicamente, pues hasta el momento la evidencia sólo apunta a que los tratamiento homeopáticos tienen nula efectividad (Ernst, 2002). Más concretamente, que no ofrecen ventajas respecto a un placebo. Dicho de otra manera, *creer* que se está haciendo algo relevante es igual de efectivo (y más barato). Pero a veces no es suficiente con creer (Salas, 2016).

Como decimos, parte del atractivo de las medicinas alternativas es que no son parte de la medicina convencional o tradicional. Lo cual también es una postura paradójica: si las medicinas alternativas pasasen la prueba del rigor científico, pasarían a formar parte de la medicina convencional.

Esto es porque la medicina convencional se cataloga como tal porque se apoya en la investigación científica y no en el sesgo de confirmación. Por tanto se limita mayoritariamente a aplicar aquellos tratamientos que estén avalados por estudios. Por contra, si un tratamiento no se recomienda es, o bien porque se ha comprobado que es peor en algún sentido, o bien porque aún no se ha estudiado lo suficiente.

Fig. 4.1 Farmacias que ofrecen productos homeopáticos a pesar de ser un despropósito.

Una consecuencia de esta forma de proceder es que, aunque podrían haber otros tratamientos incluso más efectivos que los conocidos, no se aplican si no han sido comprobados primero. Un ejemplo llamativo de esto es que, antiguamente, los esguinces de tobillo se trataban con reposo, pero actualmente se ha visto que es más efectivo realizar movimientos controlados. Esto es un caso más donde el revisionismo científico es una fortaleza.

4.3 El contexto y las condiciones son importantes

Las ideas «sencillas» se pueden descontextualizar y aplicar incorrectamente, pero el propio peso de su fama puede ser suficiente para que sean aceptadas como correctas. Un ejemplo de esto lo veremos en forma de réplica simplista a una de las consecuencias del cambio climático.

Por otra parte, ideas válidas pueden ser tergiversadas lo suficiente como para apoyar mensajes contrarios. Veremos un ejemplo de esto a la hora de reportar resultados de investigación al gran público.

4.3.1 Cambio climático: aumento del nivel del mar y el principio de Arquímedes

Un asunto muy actual y que enfrenta a muchos colectivos es el relacionado con el cambio climático. Concretamente, la postura respecto a la influencia de la actividad humana en la aceleración de esos cambios.

Entre las predicciones de lo que puede ocurrir con el cambio climático hay una que es bastante impactante: el aumento del nivel de los mares, que ha incrementado su ritmo de crecimiento durante las últimas décadas (Cazenave and Llovel, 2010). Una de las fuentes de agua para ese incremento es el propio hielo continental.

Son variados los esfuerzos por desacreditar la postura general de la comunidad científica, que es la de que la actividad humana tiene una gran influencia en el cambio climático (Houghton et al., 1990) y que, entre otras consecuencias, aumentaría el nivel de los mares, retrasando bruscamente la línea costera tierra adentro.

Uno en particular es llamativo por su sencillez, pues sólo requiere conocimientos básicos de física. Mejor aún, si uno no tiene esos conocimientos, se puede realizar un experimento en casa o en el aula de clase.

El argumento clama que no puede haber aumento del nivel de los mares, y eso se debe al principio de Arquímedes:

> «Un cuerpo total o parcialmente sumergido en un fluido en reposo experimenta un empuje vertical hacia arriba igual al peso del fluido desalojado»

Más concretamente, la observación que mereció su afamado «¡Eureka!»:

> «Un objeto flotante desplaza un volumen de fluido de peso equivalente»

La consecuencia es que los casquetes polares desplazan un volumen de agua con un peso equivalente y, como, al descongelarse, el agua proveniente del casquete debería tener el mismo volumen que el volumen desplazado, no habría elevación del nivel del agua.

La bondad de este argumento es que tiene una comprobación visual: echar un hielo en un vaso con agua e ir comprobando que el nivel del agua no cambia al derretirse (representado en la Fig. 4.2).

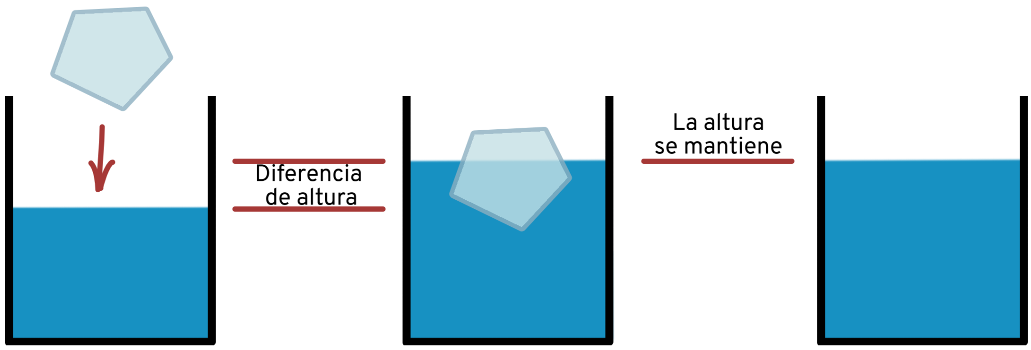

Fig. 4.2 Representación del experimento. Al echar un hielo que flote libremente, el nivel del agua en el recipiente aumenta. Al derretirse, en cambio, se mantendrá el nivel.

Entonces, ¿Cómo es posible que gran parte de la comunidad científica defienda algo que claramente es erróneo? Esto puede ser un buen debate de clase, y hasta puede dar lugar a un poco de experimentación.

Como se dice previamente, una teoría científica tiene un contexto. En este caso es que el objeto tiene que estar flotando, lo que ocurriría si el objeto se trata de un iceberg. Pero no todo el hielo que se derretiría proviene de icebergs: existe el *permafrost* de Siberia y el hielo continental de la Antártida.

Otro error más sutil, que se aplica más bien al desarrollar el argumento, es el de suponer que el hielo y el agua desplazada tienen la misma composición, con lo cual en el cambio de fase no tienen la misma densidad y, en consecuencia, el volumen del hielo una vez líquido no coincide exactamente con el volumen de agua salada inicialmente desplazada.

4.3.2 La efectividad de las mascarillas en interiores

Recientemente, el diario digital elEconomista.es (2022) anunciaba los resultados de un artículo de Oswin et al. (2022) subido al repositorio *medRxiv*.

La noticia del diario se titula «Un estudio sugiere que el coronavirus **apenas** puede contagiar suspendido en el aire en entornos cerrados», y la primera línea es «Las mascarillas por la calle o el miedo a bares y restaurantes podrían tener una buena razón para terminarse». Otra de las frases que se destacan es «**Pasados unos segundos** en el aire en un entorno como un bar, **el virus pierde gran parte de su eficacia**».

Aunque dispone de algunas afirmaciones a modo de descarga de responsabilidad (v.g. «Los resultados aún deben ser probados fuera de laboratorio y se hicieron antes de la aparición de ómicron»), el mensaje general que se intenta transmitir es que apenas hay riesgo de contagio en entornos cerrados. De hecho, eso mismo se afirma: «la investigación señala que tanto entornos abiertos como cerrados no son tan peligrosos de contagio».

De ser realmente así, es un giro bastante grande respecto a la postura de la comunidad científica en general. Podría ser así y ser correcto, pero también es posible que haya algún problema. Para evitar caer en la *falacia del ataque al mensajero*, si queremos juzgar el trabajo de Oswin et al. (2022) hay que acudir a el, en lugar de basarse únicamente en la noticia.

Cabe destacar que *medRxiv* no es una revista, sino un repositorio de *pre-prints*, y el material allí disponible no tiene por qué haber superado una revisión por pares.

El estudio como tal se centra en estudiar la evolución de la efectividad para el contagio de una única unidad del virus cuando es transportada por una gota de un diámetro cien veces mayor, y cómo se ve afectada esa evolución según la humedad relativa del ambiente. En principio el artículo no parece tan ambicioso como hace parecer la noticia, pues sólo estudian aerosoles formados por unas 10 gotas.

En el apartado de discusión del artículo se afirman varias cosas interesantes: «Under all conditions measured, the majority of SARS-CoV-2 is inactivated within 10 minutes of aerosolization», es decir, aunque un único virus pierde gran parte de su efectividad en pocos segundos, no es hasta pasados diez minutos que la colectividad (de 10 gotas o virus) pierde casi por completo su efectividad; «Further research is required to determine for how long the remaining fraction persists and how this may depend on the viral load in the aerosol», por tanto no se sabe si los resultados pueden extenderse a cargas víricas mayores; y «The importance of elucidating of the role of pH in the survival of SARS-CoV-2 in the aerosol phase cannot be understated», con lo cual también reconocen algunos de los límites del estudio.

Sin embargo, quizá las afirmaciones más contrarias a lo que se transmite en la noticia sean las siguientes:

> «Elevation of CO_2 levels within a room is taken as a clear sign of occupancy and poor ventilation. There has been increasing discussion surrounding the use of CO_2 monitors as a means of determining the relative risk of COVID-19 transmission in various settings. **The data from this study give further credence to this approach**. Not only is elevated CO_2 an indication of a densely occupied, poorly ventilated space, but it could also be indicative of an environment in which SARS-CoV-2 is more stable in the air.»

4.4 Los límites de la ciencia

Es necesario concluir este capítulo insistiendo en que el conocimiento científico tiene un ámbito de acción. En ese ámbito de acción es naturalmente superior a otras fuentes de saber. Pero fuera de esos contextos es posible que carezca de sentido.

Por ejemplo, si alguien se pregunta por el significado de la vida, la ciencia no es capaz de brindar una respuesta. Puede dar nociones sobre cómo surge la vida, pero qué sentido o significado tiene es un problema de otras áreas del saber, como la filosofía o la religión.

Otra circunstancia en la que pierde valor es cuando se descontextualizan las distintas áreas de la ciencia. Por ejemplo, no tiene mucho sentido hablar de la conciencia humana como una topología, como hacen ciertos autores, pues una topología es una estructura matemática bastante concreta y se debe cumplir una serie de axiomas poder ser calificada como tal. Más grave es sacar conclusiones, válidas en el contexto pertinente, sin haber verificado que se cumplan las condiciones.

También es posible que la ciencia no sea suficiente para resolver ciertas decisiones. Es decir, cabe la posibilidad de que decisiones incompatibles entre sí sean igual de científicamente válidas. Pero si hay que escoger una vía en particular, hay que ampliar el registro de valoraciones (v.g. económicas, morales o éticas).

Un caso así son las estrategias con las que se podía afrontar la actual pandemia. Ciertamente, por lo que conocemos de la evolución de los virus, la humanidad no iba a desaparecer. Si sólo nos preocupa que la humanidad siga existiendo, la estrategia de la inacción está validada por la evidencia histórica de epidemias anteriores. Pero, si nos preocupamos de cuánta gente podría perecer en el camino, ya hay que contemplar otras estrategias.

Por último, hay que mencionar que existen cuestiones que exceden las capacidades actuales de experimentación (o quizá siempre las excedan), incluso cuando se trata del propio ámbito de la ciencia. Sin ir más lejos, en las ciencias físicas, hasta hace relativamente poco tiempo no podíamos detectar exoplanetas u ondas gravitacionales. Es decir, incluso aunque tuviese sentido que existiesen, no podíamos corroborarlo con seguridad.

Capítulo 5

Conclusiones y perspectivas futuras

La relevancia de las redes sociales en la vida de los adolescentes no se puede poner en duda. Ocupan su tiempo, forman e informan su visión de la realidad, influyen en cómo se relacionan entre ellos y con el mundo que les rodea. Esto podría potenciarse en el futuro, o quizá deflagrarse hasta cierto punto, pero las relaciones humanas y la transmisión de información a través del mundo virtual han surgido y no desaparecerán.

Como sociedad recién estamos pudiendo observar y juzgar algunas de las consecuencias de este primer contacto masivo y sin filtrar –es casi salvaje– con las redes sociales. Son plataformas con una potencialidad muy grande, pero también brindan algunas posibilidades nefastas si se utilizan con intenciones predatorias.

Los alumnos adolescentes se encuentran en un estadio de desarrollo cognitivo, formativo y personal que los vuelve candidatos ideales para ser influenciados mediante exposición masiva de información tergiversada. Aunque en las redes hay divulgadores de ciencia que pueden influenciar positivamente en ellos, también hay comunicadores con un discurso elaborado, aparentemente legítimo pero en última instancia viciado y perjudicial.

La educación obligatoria debe trabajar para formar un pensamiento crítico en los alumnos y fomentar la cultura científica. Como parte de esa labor debe conectar los conocimientos que se transmiten con el mundo que rodea al alumnado, con el doble fin de adquirir cultura y de poder «palpar» su utilidad.

De este Trabajo Fin de Máster se pueden extraer las siguientes conclusiones:

- La dinámica de las redes sociales brinda oportunidades didácticas en el ámbito de las matemáticas: no se trata de resolver problemas de hace milenios[1], sino problemas que alumnos pueden encontrarse en su día a día.

- Además son una oportunidad donde se puede «experimentar» con matemáticas.

[1] Salvo cuando se trata de terraplanismo.

- Los engaños pueden ser elaborados y complejos, o muy sencillos pero convincentes al fin y al cabo. Para más inri, y lo que es aún más grave, es que por desconocimiento podríamos estar alimentándolos sin siquiera ser participantes activos.
- Hay que fomentar un cierto escepticismo, pero con sentido crítico. Es decir, partiendo desde la cultura matemática y científica.
- También hay que transmitir en qué contextos y circunstancias tiene valor el conocimiento científico, pues no se trata de una fe o creencia religiosa.

Algunas de las propuestas o temáticas tratadas en el presente trabajo fueron puestas en marcha, aunque de forma más primitiva, durante el prácticum. Véase:

- Se comentaron las ideas detrás de cómo funcionan los filtros de programas como TikTok (detección y monitorización de rostros).
- Se comentó también que, conociendo esos detalles, se pueden buscar formas de hacerlos fallar. A veces puede ser algo aparentemente obvio (como una mascarilla) y otras veces algo irrisorio como añadir una línea de luz.
- Se hizo que los alumnos manipulasen datos para que comprueben cómo los cambios afectan a la media muestral, que se inventasen datos que cumplieran algunas condiciones, y que comprobasen por sí mismos que, a veces, las simplificaciones pueden ser un poco engañosas.
- Para ejemplificar que, en ocasiones, hay que añadir consideraciones adicionales para resolver un problema «real», se plantearon modificaciones de ejercicios de sistemas de ecuaciones lineales, como sistemas de inecuaciones lineales (concretamente problemas de programación lineal entera), donde no había unicidad de solución, y para conseguirla había que considerar algo como el coste asociado a cada solución (para maximizarlo o minimizarlo).

Las posibles líneas en las que se puede desarrollar este trabajo, por su parte, son:

- Ampliar el repertorio de engaños actuales que pueden encontrarse en las redes.
- Asimismo, teniendo en cuenta el campo de las matemáticas en la que puedan ser pertinentes, diseñar actividades de experimentación que pongan en relieve la utilidad de las matemáticas.
- Compatibilizar las actividades con el currículum de matemáticas.
- En particular, compatibilizarlas con el currículum en la enseñanza obligatoria, diseñando a su vez actividades que puedan ser aprovechadas por alumnos de 4ºESO.
- Aprovechar la transversalidad de la ciencia para diseñar actividades y experimentos realizables con la colaboración de profesores del resto de áreas.

No obstante, esta línea de trabajo quizá no deba ser atendida como una actividad eventual en una asignatura concreta y, más bien, se deba tener un espacio dedicado a ello (Horn and Veermans, 2019).

Referencias

Abrams, D. M. and Strogatz, S. H. (2004). Chimera states for coupled oscillators. *Physical Review Letters*, 93(17):174102.

Aiken, S. (2020). Digital resistance: security & privacy tips from Hong Kong protesters. Disponible en https://medium.com/crypto-punks/digital-resistance-security-privacy-tips-from-hong-kong-protesters-37ff9ef73129.

BBC (2013). Edward Snowden was NSA Prism leak source. Disponible en https://www.bbc.com/news/world-us-canada-22836378.

BBC (2014). Edward Snowden: Leaks that exposed US spy programme. Disponible en https://www.bbc.com/news/world-us-canada-23123964.

BBC (2018). Vote Leave's targeted Brexit ads released by Facebook. Disponible en https://www.bbc.com/news/uk-politics-44966969.

BBC Mundo (2018). 5 claves para entender el escándalo de Cambridge Analytica que hizo que Facebook perdiera US$37.000 millones en un día. Disponible en https://www.bbc.com/mundo/noticias-43472797.

Beckman, D., Chari, A. N., Devabhaktuni, S., and Preskill, J. (1996). Efficient networks for quantum factoring. *Physical Review A*, 54(2):1034.

Berggren, J. L. and Sidoli, N. (2007). Aristarchus's on the sizes and distances of the sun and the moon: Greek and arabic texts. *Archive for History of Exact Sciences*, 61(3):213–254.

Bratsberg, B. and Rogeberg, O. (2018). Flynn effect and its reversal are both environmentally caused. *Proceedings of the National Academy of Sciences*, 115(26):6674–6678.

Brown, R. A. and Kumar, A. (2011). A new perspective on eratosthenes' measurement of the earth. *The Physics Teacher*, 49(7):445–447.

Buhler, J. P., Lenstra, H. W., and Pomerance, C. (1993). Factoring integers with the number field sieve. In *The development of the number field sieve*, pages 50–94. Springer.

Cabrales, A. (2021). ¿Por qué no se debe invocar el esfuerzo fiscal como argumento si uno quiere que los impuestos sean bajos? Disponible en https://nadaesgratis.es/cabrales/por-que-no-se-debe-invocar-el-esfuerzo-fiscal-como-argumento-si-uno-quiere-que-los-impuestos-sean-bajos.

Cao, W., He, L., Cao, W., Huang, X., Jia, K., and Dai, J. (2020). Recent progress of graphene oxide as a potential vaccine carrier and adjuvant. *Acta Biomaterialia*, 112:14–28.

Cazenave, A. and Llovel, W. (2010). Contemporary sea level rise. *Annual Review of Marine Science*, 2:145–173.

Cheung, H. and Hughes, R. (2020). Why are there protests in Hong Kong? All the context you need. Disponible en https://www.bbc.com/news/world-asia-china-48607723.

Choi, D., Chun, S., Oh, H., Han, J., Kwon, T., et al. (2020). Rumor propagation is amplified by echo chambers in social media. *Scientific reports*, 10(1):1–10.

Cinelli, M., Morales, G. D. F., Galeazzi, A., Quattrociocchi, W., and Starnini, M. (2021). The echo chamber effect on social media. *Proceedings of the National Academy of Sciences*, 118(9).

Davidson, J., Liebald, B., Liu, J., Nandy, P., Van Vleet, T., Gargi, U., Gupta, S., He, Y., Lambert, M., Livingston, B., et al. (2010). The youtube video recommendation system. In *Proceedings of the fourth ACM conference on Recommender systems*, pages 293–296.

Diariocrítico (2020). Polémico fotomontaje usado por Vox con la Gran Vía de Madrid llena de ataúdes. Disponible en https://www.diariocritico.com/im-presentable/vox-manipulacion-fotografia-gran-via-madrid-ataudes.

Diethelm, P. and McKee, M. (2009). Denialism: what is it and how should scientists respond? *The European Journal of Public Health*, 19(1):2–4.

elEconomista.es (2022). Un estudio sugiere que el coronavirus apenas puede contagiar suspendido en el aire en entornos cerrados. Disponible en https://www.eleconomista.es/actualidad/noticias/11564099/01/22/Un-estudio-sugiere-que-el-coronavirus-apenas-puede-contagiar-suspendido-en-el-aire-en-entornos-cerrados.html.

Ernst, E. (2002). A systematic review of systematic reviews of homeopathy. *British Journal of Clinical Pharmacology*, 54(6):577–582.

Flynn, J. R. and Shayer, M. (2018). IQ decline and Piaget: Does the rot start at the top? *Intelligence*, 66:112–121.

Gal, A. and Raz, O. (2020). Precooling strategy allows exponentially faster heating. *Physical Review Letters*, 124(6):060602.

Géron, A. (2019). *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems*. O'Reilly Media.

Gomez-Uribe, C. A. and Hunt, N. (2015). The netflix recommender system: Algorithms, business value, and innovation. *ACM Transactions on Management Information Systems (TMIS)*, 6(4):1–19.

González-Adalid Pemartín, I., Mompó, E., Lasanta, A., Martín-Mayor, V., and Salas, J. (2021). Slow growth of magnetic domains helps fast evolution routes for out-of-equilibrium dynamics. *Physical Review E*, 104:044114.

Grimes, D. R. (2016). On the viability of conspiratorial beliefs. *PloS one*, 11(1):e0147905.

Hancock, J. (2018). La Tierra plana, una teoría de la conspiración construida con memes y grupos de Facebook. Disponible en https://historia.nationalgeographic.com.es/a/terraplanismo-edad-media_14991.

Hernandez, M. and Scarr, S. (2019). Coordinating chaos: The tactics protesters use to fortify the frontlines. Disponible en https://graphics.reuters.com/HONGKONG-EXTRADITIONS-TACTICS/0100B0790FL/index.html.

Horn, S. and Veermans, K. (2019). Critical thinking efficacy and transfer skills defend against 'fake news' at an international school in finland. *Journal of Research in International Education*, 18(1):23–41.

Hornig, M., Briese, T., Buie, T., Bauman, M. L., Lauwers, G., Siemetzki, U., Hummel, K., Rota, P. A., Bellini, W. J., O'Leary, J. J., et al. (2008). Lack of association between measles virus vaccine and autism with enteropathy: a case-control study. *PloS one*, 3(9):e3140.

Houghton, J. T., Jenkins, G. J., and Ephraums, J. J. (1990). Climate change.

Huertas Cerdeira, V. (2020). Encryption: Council adopts resolution on security through encryption and security despite encryption. Disponible en https://www.consilium.europa.eu/en/press/press-releases/2020/12/14/encryption-council-adopts-resolution-on-security-through-encryption-and-security-despite-encryption/.

Kleinman, Z. (2018). Cambridge Analytica: The story so far. Disponible en https://www.bbc.co.uk/news/technology-43465968.

Kovacs, A. J. (1964). Transition vitreuse dans les polymères amorphes. etude phénoménologique. In *Fortschritte der hochpolymeren-forschung*, pages 394–507. Springer.

Lasanta, A., Reyes, F. V., Prados, A., and Santos, A. (2019). On the emergence of large and complex memory effects in nonequilibrium fluids. *New Journal of Physics*, 21(3):033042.

Lee, S.-G., Neiman, A., and Kim, S. (1998). Coherence resonance in a hodgkin-huxley neuron. *Physical Review E*, 57(3):3292.

Lepage, G. P. (2021). Active learning in a graduate quantum field theory course. *American Journal of Physics*, 89(3):317–323.

Li, H., Feng, C., Ehrhard, H., Shen, Y., Cobos, B., Zhang, F., Elamvazhuthi, K., Berman, S., Haberland, M., and Bertozzi, A. L. (2017). Decentralized stochastic control of robotic swarm density: Theory, simulation, and experiment. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4341–4347. IEEE.

Licastro, D., Rajasekharan, S., Dal Monego, S., Segat, L., D'Agaro, P., and Marcello, A. (2020). Isolation and full-length genome characterization of sars-cov-2 from covid-19 cases in northern italy. *Journal of virology*, 94(11):e00543–20.

López Frías, D. (2021). Guerra civil entre los antivacunas: el bulo del grafeno enfrenta a los negacionistas. Disponible en https://www.elperiodico.com/es/sociedad/20211013/guerra-civil-antivacunas-bulo-grafeno-12234908.

Marsland, S. (2011). *Machine learning: an algorithmic perspective*. Chapman and Hall/CRC.

Mejias, J. F. and Wang, X.-J. (2021). Mechanisms of distributed working memory in a large-scale network of macaque neocortex. *BioRxiv*, page 760231.

Miguel Trula, E. (2020). El moderno "príncipe nigeriano", la estrella de Instagram arrestada por un fraude de 390 millones. Disponible en https://magnet.xataka.com/un-mundo-fascinante/moderno-principe-nigeriano-estrella-instagram-arrestada-fraude-390-millones.

Ministerio de Sanidad (2022). Estrategia de vacunación COVID-19 en España. Disponible en https://www.mscbs.gob.es/profesionales/saludPublica/ccayes/alertasActual/nCov/vacunaCovid19.htm.

Mohler, G. O., Short, M. B., Malinowski, S., Johnson, M., Tita, G. E., Bertozzi, A. L., and Brantingham, P. J. (2015). Randomized controlled field trials of predictive policing. *Journal of the American statistical association*, 110(512):1399–1411.

Mpemba, E. B. and Osborne, D. G. (1969). Cool? *Physics Education*, 4(3):172.

Ngan, M. L., Grother, P. J., Hanaoka, K. K., et al. (2020). Ongoing face recognition vendor test (frvt) part 6a: Face recognition accuracy with masks using pre-covid-19 algorithms.

Office for National Statistics (2021). Deaths involving COVID-19 by vaccination status, England: deaths occurring between 1 January and 31 October 2021. Disponible en https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/bulletins/deathsinvolvingcovid19byvaccinationstatusengland/deathsoccurringbetween1januaryand31october2021.

O'neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown.

Organization, W. H. et al. (2021). Genomic sequencing of sars-cov-2: a guide to implementation for maximum impact on public health, 8 january 2021.

Ortega, G. N. (2011). *Un curso de números*. Universitat de València.

Oswin, H. P., Haddrell, A. E., Otero-Fernandez, M., Mann, J. F., Cogan, T. A., Hilditch, T., Tian, J., Hardy, D., Hill, D. J., Finn, A., et al. (2022). The dynamics of sars-cov-2 infectivity with changes in aerosol microenvironment. *medRxiv*.

Pikovsky, A. S. and Kurths, J. (1997). Coherence resonance in a noise-driven excitable system. *Physical Review Letters*, 78(5):775.

Pirazzini, G. (2020). Terraplanismo en la Edad Media. Disponible en https://historia.nationalgeographic.com.es/a/terraplanismo-edad-media_14991.

Plotkin, S., Gerber, J. S., and Offit, P. A. (2009). Vaccines and autism: a tale of shifting hypotheses. *Clinical Infectious Diseases*, 48(4):456–461.

Rodríguez, R. (2020). El uso falaz del esfuerzo fiscal. Disponible en https://www.eldiario.es/opinion/zona-critica/presion-fiscal-esfuerzo_129_6035085.html.

Salas, J. (2016). "a mi hijo lo ha matado la incultura científica". Disponible en https://elpais.com/elpais/2016/02/24/ciencia/1456341289_969832.html.

Schopenhauer, A. (2006). *El arte de insultar*. Editorial Edaf.

Selim, S. Z. and Ismail, M. A. (1984). K-means-type algorithms: A generalized convergence theorem and characterization of local optimality. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (1):81–87.

Shao, C., Ciampaglia, G. L., Varol, O., Yang, K.-C., Flammini, A., and Menczer, F. (2018). The spread of low-credibility content by social bots. *Nature Communications*, 9(1):1–9.

Shin, J., Jian, L., Driscoll, K., and Bar, F. (2018). The diffusion of misinformation on social media: Temporal pattern, message, and source. *Computers in Human Behavior*, 83:278–287.

Shuster, S. (2016). The U.K.'s Old decided for the Young in the Brexit vote. Disponible en https://time.com/4381878/brexit-generation-gap-older-younger-voters/.

Sokal, A. and Bricmont, J. (1999). *Imposturas intelectuales*. Paidós Barcelona.

Sommariva, S., Vamos, C., Mantzarlis, A., Đào, L. U.-L., and Martinez Tyson, D. (2018). Spreading the (fake) news: exploring health messages on social media and the implications for health professionals using a case study. *American Journal of Health Education*, 49(4):246–255.

Strogatz, S. (2017). Einstein's first proof. In *The Best Writing on Mathematics 2016*, pages 122–131. Princeton University Press.

Thomas, M. (2000). Albert einstein and ld: An evaluation of the evidence. *Journal of Learning Disabilities*, 33(2):149–157.

Van Kampen, N. G. (1992). *Stochastic processes in physics and chemistry*, volume 1. Elsevier.

Vicsek, T., Czirók, A., Ben-Jacob, E., Cohen, I., and Shochet, O. (1995). Novel type of phase transition in a system of self-driven particles. *Physical Review Letters*, 75(6):1226.

Wakefield, A. J., Murch, S. H., Anthony, A., Linnell, J., Casson, D. M., Malik, M., Berelowitz, M., Dhillon, A. P., Thomson, M. A., Harvey, P., et al. (1998). RETRACTED: Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children.

Wang, Z. and Rousseau, R. (2021). Covid-19, the yule-simpson paradox and research evaluation. *Scientometrics*, 126(4):3501–3511.

Waszak, P. M., Kasprzycka-Waszak, W., and Kubanek, A. (2018). The spread of medical fake news in social media–the pilot quantitative study. *Health Policy and Technology*, 7(2):115–118.

Westerlund, M. (2019). The emergence of deepfake technology: A review. *Technology Innovation Management Review*, 9(11).

Xu, M., Zhu, J., Wang, F., Xiong, Y., Wu, Y., Wang, Q., Weng, J., Zhang, Z., Chen, W., and Liu, S. (2016). Improved in vitro and in vivo biocompatibility of graphene oxide through surface modification: poly (acrylic acid)-functionalization is superior to pegylation. *Acs Nano*, 10(3):3267–3281.

Zulet, Í. (2021). Los menas y la abuela: la campaña de derecha dura de Vox para Madrid que ya investiga la Fiscalía. Disponible en https://www.elespanol.com/espana/madrid/20210420/abuela-campana-derecha-vox-madrid-denunciada-fiscal/575193442_0.html.

Anexo A

Ejemplos de engaños

En este anexo se recopilan varios engaños o «bulos» que podemos hallar en la calle o por internet.

A.1 Pseudociencia

A.1.1 Concentración antivacunas

Esta ocurrió en Callao, Madrid, el 18 de diciembre de 2021. En las figuras A.1 y A.2 se pueden apreciar algunos de los mensajes que presentaban.

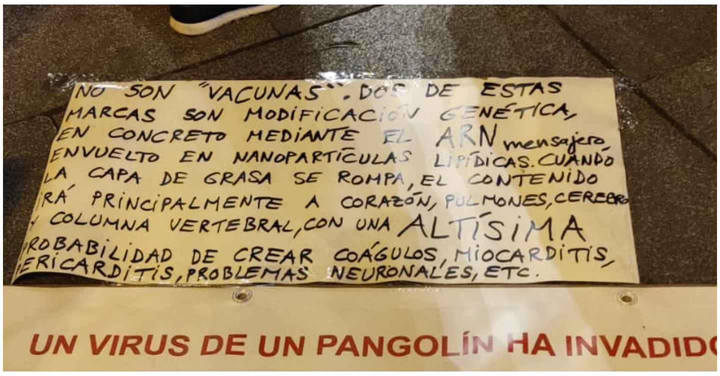

Fig. A.1 Transcripción: *No son "vacunas", dos de estas marcas son modificación genética, en concreto mediante el ARN mensajero, envuelto en nanopartículas lipídicas. Cuando la capa de grasa se rompa, el contenido irá principalmente a corazón, pulmones, cerebro, y columna vertebral, con una altísima probabilidad de crear coágulos, miocarditis, pericarditis, problemas neuronales, etc.* Imagen cedida por M. Gutiérrez Vela.

Concretamente, en la figura A.1 se aprecia una gran mezcla de conceptos, además de una falacia *non sequitur*: lo que claman no contradice que se trate de vacunas. También hay una interpretación errónea de información (en relación a la «altísima» probabilidad), así como una asociación incorrecta de datos (los distintos problemas que mencionan están asociados a otro tipo de vacuna).

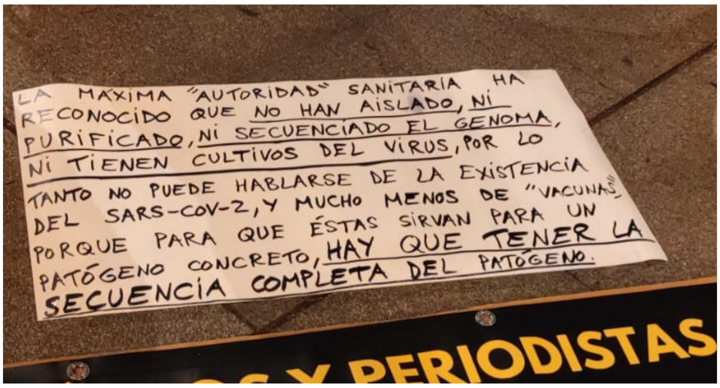

Fig. A.2 Transcripción: *La máxima "autoridad" sanitaria ha reconocido que no han aislado, ni purificado, ni secuenciado el genoma, ni tienen cultivos del virus, por lo tanto no puede hablarse de la existencia del SARS-CoV-2, y mucho menos de "vacunas", porque para que éstas sirvan para un patógeno concreto, hay que tener la secuencia completa del patógeno.* Imagen cedida por M. Gutiérrez Vela.

Mientras tanto, en la figura A.1, aluden a (pero no especifican) la máxima autoridad sanitaria y unas supuestas declaraciones. Si la autoridad es la Organización Mundial de la Salud, no es posible encontrar esas declaraciones. La organización, en cambio, ofrace pautas para secuenciar el genoma (Organization et al., 2021), cosa que ya se ha realizado (Licastro et al., 2020).

A.1.2 El magnetismo inducido por las vacunas

A las afueras de Torrelodones puede observarse un cartel publicitario afirmando tener una explicación para el supuesto magnetismo provocado por las vacunas. Es pertinente señalar que el supuesto

Fig. A.3 Valla publicitaria antivacunas, en la autopista A-6 de Madrid, cerca de Torrelodones. Fuente: López Frías (2021)

magnetismo que se popularizó en las redes era una cuestión de higiene.

Si se indaga en el historial de la teoría, se puede hallar un informe con diversos compromisos metodológicos (dependen de una única muestra, cuyo origen, además, es desconocido). La relación con las vacunas es circunstancial, pues en realidad uno de los portavoces de la teoría tiene una fijación contra un nanomaterial, que es el óxido de grafeno.

En principio, es cierto que el óxido de grafeno puede inducir algún problema de salud, pero hay mejoras de diseño respecto a su biocompatibilidad (Xu et al., 2016), convirtiéndolo en un potencial encapsulamiento para las vacunas (Cao et al., 2020).

A.2 Cadenas de mensajes

A.2.1 Un virus creado en China

Un mensaje que circula por las redes, particularmente a través de WhatsApp, y que es una falsa atribución a un experto, es la siguiente:

> **Impactante Declaración.**
>
> Profesor japonés de fisiología y medicina, Tasuku Honjo, causó revuelo y asombro en todas las redes sociales y en todos los medios de comunicación hoy cuando dijo que este coronavirus no era natural; Si fuera natural, no afectaría así al mundo entero. Porque la temperatura varía de un país a otro dependiendo de la especie; Si fuera natural, solo afectaría a países con la misma temperatura que China. En cambio, se propaga en un país como Suiza al igual que en las zonas desérticas; Si fuera natural, se propagaría en lugares fríos pero moriría en lugares cálidos. Estudié animales y virus durante más de cuarenta **(40)** años. No es natural. Ese virus está hecho y es completamente artificial. Trabajé en un laboratorio de Wuhan en China durante cuatro **(04)** años y conozco muy bien a todo el personal de ese laboratorio en China. Yo los llamé a todos después del accidente de la corona. Pero todos sus teléfonos han estado fuera de servicio durante meses, y ahora se sabe que todos estos técnicos de laboratorio están muertos.
>
> Basado en todo mi conocimiento y en todas las investigación previas que se hicieron, puedo decir con cien por ciento **(100%)** de certeza que el virus **Corona 19**, no es un virus natural. No es un virus de murciélagos. China lo creó de manera intencional y premeditada. Si lo que digo hoy resulta ser incorrecto ahora, o incluso después de mi muerte, el gobierno de China, puede retirar mi **Premio Nobel**. Pero China miente y todo su terrible gobierno comunista debe ser castigado severamente y con todo el peso de la ley penal internacional, porque esta terrible, imperdonable y nefasta gran verdad, algún día será revelada a todas las personas en todo el mundo entero.
>
> https://en.m.wikipedia.org/wiki/Tasuko_Honjo
>
> **Por favor proporcione este mensaje a toda la gran mayoría de las personas porque es necesario que todo el mundo sepa lo que nos está pasando a todos.**

Curiosamente, el propio enlace de Wikipedia menciona que Tasuku Honjo nunca ha dicho eso, y que se lamenta que le hayan atribuido esas declaraciones.

A.2.2 Filomena contra las democracias

El siguiente extracto también fue reenviado a través de WhatsApp. Hace referencia a expertos inexistentes en pos de establecer un argumento de autoridad y menciona un compuesto químico que no es más que agua, por tanto una ofuscación deliberada. Aunque utiliza varios recursos típicos de este tipo de cadena, como una sensación de alarmismo, parece intentar promover el uso de mascarillas.

> **¡IMPORTANTE!** Se ruega **máxima** difusión.
>
> La reciente nevada ha sido un evento único en muchos puntos de la España peninsular. Pero no sólo por lo inusual que es tal fenómeno meteorológico. Diversos estudios independientes han detectado **altas concentraciones de óxido de dihidrógeno** en la nieve. Esta sustancia ha sido utilizada en el pasado por diversos **gobiernos autoritarios**, pues el contacto con dicha sustancia **vuelve dóciles y manipulables a las personas**.
>
> El Dr. Javier Ortega de la Universidad de Wildenshire, especialista en la materia, comenta que la vía histórica de suministro de este compuesto ha sido la red de abastecimiento de agua potable. No obstante, señala que existe evidencia de que el óxido de dihidrógeno tiene un **efecto más potente si el contacto es por vía respiratoria**.
>
> Por otra parte, Santiago Conde, analista químico en Henron, empresa dedicada al suministro y distribución de distintos químicos, sugiere que una nevada puntual es la vía perfecta para suministrar de manera masiva tal sustancia por vía respiratoria. "Al ser algo tan inesperado, es hasta natural que la población salga y se exponga sin saberlo". Por suerte, también señala que el uso de **mascarillas** certificadas, que todos disponemos en casa a raíz de la pandemia COVID-19, **son efectivas para bloquear el paso de este agente a nuestras vías respiratorias**. Una vez preguntado por cómo llegó la sustancia a la nieve, responde "es un método conocido, basta con fumigar a la población desde las nubes". Aunque suene extraño, esta práctica la podemos ver en un día cualquiera en forma de líneas rectas de una especie de humo que surcan los cielos. Esas líneas son conocidas como **chemtrails**.
>
> La periodista Natividad Díaz comenta que "allí donde ha sido utilizado ha tenido un **impacto social devastador**". Alarmada, dice que este hecho puede ser uno de tantos otros **atentados contra la democracia occidental**. Insinúa una relación con eventos recientes: "en el reciente asalto al Capitolio en Estados Unidos se detectaron trazas de esa sustancia en el colectivo de manifestantes".
>
> **Difunde** esta información. **Nuestra salud**, democracia y libertad, están **en riesgo**. Con vuestra ayuda y la de Dios, saldremos adelante.
>
> Atentamente,

Dra. Cayetana Choclán, del hospital Nuestra Señora.

A.2.3 La promesa nigeriana

Una estafa que antes llegaba por correo y ahora, en una versión más actualizada, puede llegar a través de los servicios de mensajería instantánea es la del príncipe nigeriano que, destripado de su dinero, pide ayuda (una pequeña inversión) para poder recuperarlo, prometiendo una recompensa por la asistencia.

El siguiente texto es una de las tantas versiones en español que se pueden encontrar.

> Estimado Señor,
>
> PROPUESTA COMERCIAL: TRANSFERENCIA DE US$41.5 M (CUARENTA Y UN MILLONES, QUINIENTOS MIL DOLARES).
>
> Primero, yo debo solicitar su confianza más estricta en esta transacción, esto en virtud de su naturaleza confidencial y secreta. Usted ha sido presentado a mi por un familiar de confianza con un mutuo contacto en la Cámara Nigeriana de Comercio. Sin embargo, él ignora la naturaleza y magnitud de lo que estoy por proponerle ahora...
>
> Yo soy uno de los hijos del antiguo jefe de estado nigeriano, el general Sani Abacha. Deseo hacerlo partícipe de un negocio de transferencia de dinero que involucra una cifra de 41.5 millones de dólares.
>
> Después de la muerte de mi padre, mi madre, Mrs Marian Abacha, con la ayuda de oficiales de confianza del Banco Central de Nigeria, pasó de contrabando esta suma de dinero, a través de un diplomático de un país europeo. Este fondo se depositó en secreto en la bóveda de una compañía financiera que espera la presentación de los detalles de propiedad que incluyen el código de acceso y la factura de transporte vía aérea del diplomático. Los fondos pueden moverse igualmente de Europa a cualquier otro país que mantenga correspondencia con el Banco Central de Nigeria.
>
> Debo decirle, antes de continuar, que desde la muerte de mi padre, el gobierno de mi país ha sometido a nuestra familia a un solitario encierro que es el más traumático castigo que pueda infligirse a cualquiera. Nuestros movimientos se han restringido y todos nuestros documentos de viaje han sido secuestrados. Las cuentas de mi padre y el dinero de nuestra familia en cuentas de bancos locales y extranjeros, han sido confiscados por el actual gobierno, con la ayuda de algunos gobiernos extranjeros donde están depositados estos fondos.
>
> Mi hermano mayor, Mohammed, está enfrentando un juicio, acusado de la muerte de prominentes nigerianos. Esta imputación del gobierno es basado en acusaciones por actos hechos durante el gobierno de mi padre.

> Ante el miedo del descubrimiento de este fondo de parte de nuestro gobierno, nosotros estamos solicitando un socio que proporcione una cuenta fiable para el dinero, para cuando podamos viajar fuera de nuestro país.
>
> Estamos de acuerdo en darle a cambio el 30% del total de los fondos, por su ayuda. Eso incluye cualquier gasto en el que usted incurra durante esa transacción, después de su culminación exitosa.
>
> Por favor, note de nuevo que esta transacción es estrictamente confidencial y como tal debe guardarse.
>
> Nos gustaría participar también con nuestro fondo, en algún negocio en conjunto, una vez que hayamos ganado la libertad para viajar fuera del país.
>
> Yo espero sus sugerencias y positiva contestación a través de mi email confidencial (abacha_s_a@hushmail.com).
>
> Suyo fielmente,
>
> Mallam Sadiq Abacha

También hay encarnaciones reales de la propia estafa, como la reportada por Miguel Trula (2020).

A.3 Falsas historias

A.3.1 Albert Einstein y las matemáticas

Una de las afirmaciones que de vez en cuando pueden llegar por correo o mensaje, o simplemente encontrarla en alguna red social, es que Einstein era mal alumno o que suspendía matemáticas (ver fig. A.4), y que aún así fue capaz de revolucionar la física del momento.

Aunque puede ser un relato esperanzador, está lejos de la realidad. Lo que sí es posible es que, en la Universidad, haya descuidado los estudios de matemáticas en favor de los de física. Aún así, hay indicios de que, en la adolescencia, pudo desarrollar de manera independiente una posible demostración del teorema de Pitágoras, como relata el matemático Stephen Strogatz (2017).

A.3.2 Albert Einstein y el desarrollo tardío

Otra de las creencias alrededor de Einstein es que tuvo un desarrollo lento en su infancia (ver fig A.5).

Esta creencia, aunque tiene bastante más difusión que la anterior sobre las matemáticas (especialmente en la comunidad del desarrollo cognitivo), no parece estar sustentada en evidencia real (Thomas, 2000).

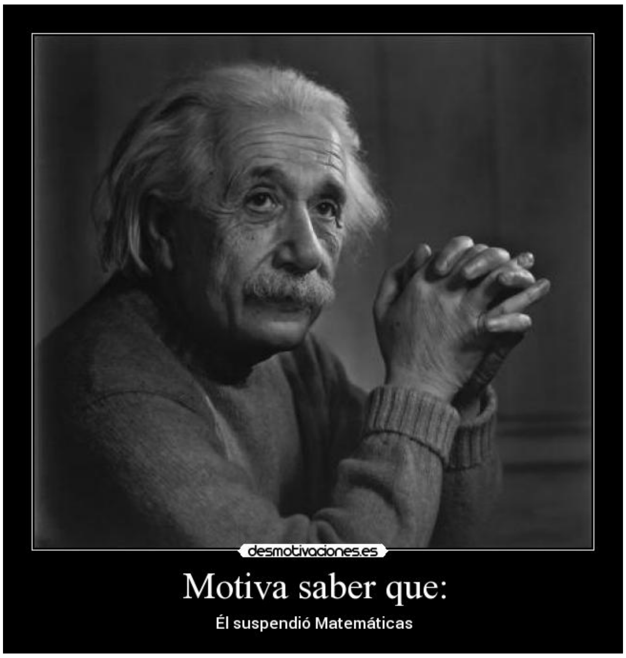

Fig. A.4 Cartel señalando que Albert Einstein suspendía matemáticas.

Fig. A.5 Cartel señalando que Albert Einstein tuvo un desarrollo tardío.

Anexo B

Conceptos formales

En este anexo se describen de manera detallada algunos de los conceptos matemáticos a los que se alude dentro del cuerpo del trabajo.

B.1 Algunos algoritmos elementales en clasificación de datos

Presentaremos tres herramientas clásicas pero importantes en el ámbito del *big data* y el aprendizaje automático (*machine learning*) (Géron, 2019; Marsland, 2011). La primera es la descomposición en valores singulares de una matriz, la cual es la pieza clave del análisis de componentes principales, que presentamos en segundo lugar pues es utilizada para hallar un subespacio vectorial de menor dimensión que aquel donde «viven» los datos multidimensionales. En tercer lugar se presenta el algoritmo de k-medias, que es utilizado en aprendizaje no supervisado siempre que tenga sentido hablar de distancia entre datos. Es decir, se utiliza para agrupar (y catalogar) datos por sus características intrínsecas.

B.1.1 Descomposición en valores singulares (factorización SVD)

La diagonalización de matrices reales cuadradas $n \times n$ es una herramienta muy potente para el análisis de datos cuando existe una jerarquía de valores propios. Dentro de las matrices diagonalizables, existe una familia que tiene propiedades especialmente fortuitas: las matrices reales simétricas, que son ortogonalmente diagonalizables con valores propios reales. Si, además, son semidefinidas positivas, los valores propios serán no negativos. Con lo que tenemos los ingredientes para una jerarquía de valores propios $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \geq 0$ (si la matriz es no nula, alguna de estas desigualdades será estricta).

Sin embargo, los datos suelen estar recogidos en matrices $n \times m$. Con lo cual nada de lo anterior es aplicable, al menos, de manera directa. Aquí es donde entra en juego la descomposición en valores singulares, que existe para toda matriz real.

Consideremos una matriz $A \in \mathbb{R}^{n \times m}$. Entonces la matriz $M = AA^t \in \mathbb{R}^{n \times n}$ será simétrica y semidefinida positiva, con $\operatorname{rank} M = r \leq \text{mín}\{n, m\} =: \hat{k}$. Luego es ortogonalmente diagonalizable con valores propios $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_r > \lambda_{r+1} = \ldots = \lambda_n = 0$. A partir de estos valores, se definen los *valores singulares* de A como $\sigma_i = \sqrt{\lambda_i}$, con $1 \leq i \leq r$.

Sea $\Sigma \in \mathbb{R}^{n\times m}$ la matriz diagonal que tiene a la $\hat{k}$-tupla $(\sigma_1, \sigma_2, \ldots, \sigma_r, 0, \ldots, 0)$ por diagonal principal. Tendremos que existirán dos matrices ortogonales $U \in \mathbb{R}^{n\times n}$ y $V \in \mathbb{R}^{m\times m}$ tales que $AV = U\Sigma$, ó, equivalentemente, $A = U\Sigma V^t$.

La descomposición en valores singulares de una matriz A con $\operatorname{rank} A = r$ permite aproximar fácilmente esta matriz por otra A_k con un rango $k < r$. Esto consiste en aprovechar que $A = U\Sigma V^t$ para construir una nueva matriz $A_k = U\Sigma_k V^t$, donde Σ_k se construye a partir de Σ sustituyendo σ_i por 0 cuando $k < i \leq r$. Es decir, filtramos los valores singulares menos relevantes. Esta aproximación cumple que, para cualquier matriz B del tamaño de A y con rango k, $\|A - A_k\| \leq \|A - B\|$, independientemente de la norma matricial. Es decir, A_k es la mejor aproximación a A con rango k.

B.1.2 Análisis de componentes principales

Sea $X = (x_{ij})$ una matriz real de tamaño $n \times k$, donde cada fila representa una repetición independiente de un experimento (de las cuales hay n) y cada columna está asociada a alguna característica u observable de interés (de los cuales hay k).

El análisis de componentes principales busca establecer una base ortonormal $\mathbb{B} = (b_1, \ldots, b_k)$ de $\mathbb{R}^k$ (i.e. el espacio de los observables) donde b_i representen las direcciones de mayor varianza en los datos. Es decir, que b_1 recoja la mayor información posible, seguido de b_2, etc., en orden decreciente.

El primer paso es estandarizar la información de cada una de las distintas características. Para ello definimos $u_j := \frac{1}{n}\sum_{i=1}^{n} x_{ij}$. Es decir, u_j es la media muestral de la j-ésima característica. Definimos $\widetilde{X} := (\widetilde{x}_{ij})$, con $\widetilde{x}_{ij} = x_{ij} - u_j$. O sea, a las columnas de X le restamos el valor medio del observable correspondiente.

El segundo paso es construir la matriz de covarianza de $\widetilde{X}$:

$$C := \frac{1}{n-1}\widetilde{X}^t\widetilde{X}, \tag{B.1}$$

que es una matriz $k \times k$ que, si denotamos $C = (c_{ij})$, tenemos que $c_{jj} = \frac{1}{n-1}\sum_{s=1}^{n}(x_{sj} - u_j)^2$. Es decir, los elementos de la diagonal son la varianza muestral de cada una de las columnas de la matriz original. Por otra parte, los elementos de fuera de la diagonal son las covarianzas entre distintas columnas.

Ahora, consideramos la descomposición en valores singulares de $\widetilde{X} = U\Sigma V^t$, donde Σ recoge en su diagonal los valores singulares de mayor a menor, y tanto $U \in \mathbb{R}^{n\times n}$ como $V \in \mathbb{R}^{k\times k}$ son ortogonales. Entonces tendremos que $C = \frac{1}{n-1}(U\Sigma V^t)^t U\Sigma V^t = \frac{1}{n-1}V\Sigma U^t U\Sigma V^t = \frac{1}{n-1}V\Sigma^2 V^t$, pues $\Sigma^t = \Sigma$ por ser diagonal. Es decir, tenemos la diagonalización ortogonal de C.

Por último, sea v_j la columna j-ésima de la matriz V. Al ser V ortogonal, se tiene que $\mathbb{P} = (v_1, \ldots, v_k)$ es una base ortonormal de $\mathbb{R}^k$. Puesto que los vectores de esta base están ordenados según el orden decreciente de los valores singulares recogidos en Σ, tenemos que v_i es una dirección que recoge más información (explica mayor varianza) que v_j siempre que $i < j$. Es decir, v_1 es la primera dirección principal, v_2 la segunda, etcétera.

Puesto que las primeras componentes principales son las que recogen la mayor varianza, y las últimas las que menos, podemos replicar la compresión mencionada con la descomposición en valores singulares. Es decir, podemos reinterpretar la información en un subespacio de $\mathbb{R}^k$ de dimensión inferior a k (i.e., el subespacio generado por $v_1, v_2, \dots, v_d$, con $d < k$). Por ello es que se dice que el análisis de componentes principales *reduce la dimensionalidad* de la información.

B.1.3 Algoritmo de k-medias

Cuando se dispone de un conjunto de observaciones, $x_1, x_2, \dots, x_n \in \mathbb{R}^d$, y queremos clasificarlas completamente en k familias disjuntas (es decir, hallar $S_1, S_2, \dots, S_k \subset \mathbb{R}^d$ de modo que $S_i \neq \emptyset$, $S_i \cap S_j = \emptyset$ si $i \neq j$, y $\{x_1, \dots, x_n\} = \bigcup_{i=1}^k S_i$), un algoritmo muy socorrido es el de k-medias

El objetivo del algoritmo consiste en hallar unos *centroides* $\mu_i \in \mathbb{R}^d$ que definen a los conjuntos S_i por minimización de distancia. Concretamente, se decide si una observación x_j pertenece a S_i si $i = \operatorname{argmín}_{1\le r\le k} \lVert x_j - \mu_r \rVert$. Así, se trata de hallar centroides que minimicen las distancias a ellos (según pertenencia). Es decir, hay que buscar $\mu_1, \dots, \mu_k \in \mathbb{R}^d$ que minimicen la función

$$\sum_{i=1}^{k} \sum_{j=1}^{n} \Delta_{ij} \lVert x_j - \mu_i \rVert^2 , \tag{B.2}$$

donde $\Delta_{ij} = 1$ si $x_j \in S_i$, y $\Delta_{ij} = 0$ en caso contrario. Es decir, sólo se tiene en consideración la distancia al centroide que define la categoría a la que el dato ha sido asignado.

El algoritmo es el siguiente:

1. Escoger unos centroides iniciales, $\mu_1^1, \dots, \mu_k^1$.
2. Para la etapa $t \geq 1$, construir las familias $S_1^{(t)}, \dots, S_k^{(t)}$ como sigue:
$$S_i^{(t)} = \left\{ x_j : \left\lVert x_j - \mu_i^{(t)} \right\rVert \leq \left\lVert x_j - \mu_r^{(t)} \right\rVert, \ \forall r \neq i \right\} .$$
3. Definir
$$\Delta_{ij}^{(t)} = \begin{cases} 1, & \text{si } x_j \in S_i^{(t)}, \\ 0, & \text{en otro caso.} \end{cases}$$
4. Definir unos nuevos centroides, $\mu_1^{(t+1)}, \dots, \mu_k^{(t+1)}$, como
$$\mu_i^{(t+1)} = \frac{\sum_{j=1}^n \Delta_{ij}^{(t)} x_j}{\sum_{j=1}^n \Delta_{ij}^{(t)}} .$$

Es decir, el valor medio en el conjunto $S_i^{(t)}$.

5. Volver al paso 2.

Este algoritmo converge a unos centroides $\mu_i^{(t)} \xrightarrow{t\to\infty} \widehat{\mu}_i$ que minimizan la función de coste (B.2), pero es posible que no representen un mínimo global (Selim and Ismail, 1984).

B.2 Encriptación – Algoritmo RSA

El algoritmo RSA es un sistema de encriptación de clave pública (o de encriptación asimétrica). Es decir, cuando se quiere transmitir un mensaje secreto no se utiliza la misma clave para cifrarlo que el utilizado para descifrarlo.

La ventaja ofrecida por estos sistemas es que permiten iniciar una comunicación privada con alguien con quién no dispongamos de un canal seguro previamente establecido, como es el caso de las compras por internet.

Vamos a plantear la estructura general de un algoritmo de encriptación de clave pública, la estructura particular de RSA, y luego los puntos clave que hay que conocer de aritmética modular para entenderlo.

B.2.1 Estructura básica

La idea básica de un sistema de encriptación de clave pública puede ser bastante visual. Si *Bob* quiere enviarle un mensaje a *Alice*:

1. Bob avisa a Alice de ese hecho.
2. Alice prepara un cofre con candado abierto y se guarda la llave.
3. Alice le envía el cofre abierto a Bob. Todos pueden ver el cofre.
4. Bob introduce el mensaje en el cofre y lo cierra.
5. Bob envía el cofre cerrado a Alice. Ahora nadie puede ver el contenido del cofre.
6. Alice abre el cofre con su llave y puede ver el contenido.

B.2.2 Estructura particular de RSA

En el caso particular del algoritmo RSA el proceso anterior se realiza recurriendo a la aritmética modular. Concretamente, si Bob quiere enviarle un mensaje M a Alice, donde M es un número (como el número de la tarjeta), lo que tendrá que hacer es:

1. Bob avisa a Alice de este hecho.

2. Alice escoge dos números primos (diferentes), p y q; calcula $n = p \cdot q$ y $\varphi(n) = (p-1)(q-1)$; y escoge $1 < e < \varphi(n)$ de modo que e y $\varphi(n)$ sean coprimos. Alice se guarda el valor de p y q.

3. Alice le envía el par ordenado (n, e) a Bob, que es la clave pública. Todos pueden ver los valores de n y e.

4. Bob calcula $C \equiv M^e \pmod{n}$, siempre y cuando $0 \leq M < n$.

5. Bob envía el valor de C a Alice. Todos pueden ver el valor de C pero no el de M.

6. Alice, mientras tanto, obtiene un valor $1 < d < \varphi(n)$ tal que $ed \equiv 1 \pmod{\varphi(n)}$, con el que forma la clave privada (n, d), y con ello recupera el mensaje de Bob puesto que $M \equiv C^d \pmod{n}$.

B.2.3 Fundamentos matemáticos

Primero recordemos un poco de lenguaje:

- Dado un número natural n y dos enteros a, b, decimos que $a \equiv b \pmod{n}$ si $n|(b-a)$. Dicho de otro modo, que a y b tienen el mismo resto $0 \leq r < n$ bajo división por n.

- La función φ de Euler (o función totiente) cuenta la cantidad de enteros positivos que son coprimos con uno dado. Por ejemplo, $\varphi(10) = 4$ (los números $1, 3, 7, 9$).

Luego, algunas de las propiedades básicas (Ortega, 2011):

- Para calcular $\varphi(n)$ se usan dos propiedades: i) cuando a y b son coprimos, $\varphi(ab) = \varphi(a)\varphi(b)$, y, ii) si p es primo y $k \geq 1$, $\varphi(p^k) = (p-1)p^{(k-1)}$. Junto con la descomposición en potencias de factores primos de n, se obtiene $\varphi(n)$. En particular, si p y q son primos, $\varphi(pq) = (p-1)(q-1)$.

- En el anillo $(\mathbb{Z}_n, +, \cdot)$, el número de elementos con inverso multiplicativo es $\varphi(n)$. Concretamente, son las clases asociadas a los números $0 < k < n$ que son coprimos con n.

Por último, el teorema que hace que este algoritmo funcione es el Pequeño Teorema de Fermat (PTF):

Teorema 1. *Sea p un número primo, y a un entero coprimo con p. Entonces $a^{p-1} \equiv 1 \pmod{p}$.*

B.2.4 Justificación del funcionamiento

En esencia, hay que comprobar que $M \equiv M^{ed} \pmod{n}$, para cualquier $0 \leq M < n$, con $n = pq$, e y d en las condiciones del algoritmo.

Dem. Distinguimos casos:

- Supongamos que p y q no dividen a M. Puesto que p es primo, será coprimo con M. Por el PTF, $M^{p-1} \equiv 1 \pmod{p}$. En consecuencia, $M^{\varphi(n)} = M^{(p-1)(q-1)} \equiv 1 \pmod{p}$. Por otra parte, como $ed \equiv 1 \pmod{\varphi(n)}$, tenemos que $\varphi(n)|(ed-1)$. Así pues, $ed-1$ es múltiplo de $(p-1)(q-1)$ y, por tanto, $M^{ed-1} \equiv 1 \pmod{n}$. De lo que se deduce que $M^{ed} \equiv M \pmod{p}$.

  De manera análoga se justifica que $M^{ed} \equiv 1 \pmod{q}$. Es decir, en suma obtenemos que tanto p como q dividen a $M^{ed} - M$. Puesto que son primos diferentes, $n = pq$ divide a $M^{ed} - M$. Con lo que podemos concluir que $M^{ed} \equiv M \pmod{n}$.

- Supongamos que sólo q divide a M. De manera análoga al caso anterior, podemos concluir que p divide a $M^{ed} - M$. Por otra parte, como $M^{ed} - M = M(M^{ed-1} - 1)$, tenemos que q divide a $M^{ed} - M$. Así, de nuevo, se llega a que $M^{ed} \equiv M \pmod{n}$.

- Supongamos que tanto p como q dividen a M. Entonces n divide a M. Como $M < n$, sólo es posible que $M = 0$. Es trivial que $M^{ed} \equiv M \pmod{n}$.

□

B.2.5 Fortaleza

Una tercera persona que esté observando, como podría ser *Eve*, sólo conocería la información relativa a la clave pública, (n, e), así como el mensaje cifrado, C.

Para desencriptar C le hace falta conocer d. Pero d es el inverso multiplicativo de e en $\mathbb{Z}_{\varphi(n)}$, el cual se puede obtener aprovechando la identidad de Bézout y el algoritmo de Euclides. Pero para poder trabajar en este último conjunto antes hay que conocer el valor de $\varphi(n) = (p-1)(q-1)$. Es decir, hace falta averiguar tanto p como q. Pero Eve sólo conoce $n = pq$, así que debe factorizar ese número, lo cual es una tarea que, actualmente, requiere bastante tiempo, puesto que no disponemos de ningún algoritmo que realice tal tarea y cuya complejidad computacional sea polinómica (Buhler et al., 1993)[1]. Sin embargo, esta situación es diferente si se cambia al paradigma de la computación cuántica, donde se dispone del algoritmo de Shor, que tiene complejidad polinómica (Beckman et al., 1996).

B.3 Paradoja de Simpson

La paradoja de Simpson ocurre cuando los datos agregados marcan una tendencia pero, al separarlos en categorías, marcan una tendencia contraria.

En la línea del caso comentado en la sección 3.1.3.2, en Wang and Rousseau (2021) se analiza el caso de la mortalidad por coronavirus segmentada por rangos de edad y por sexo: comparando cada

[1] En este caso concreto, la complejidad computacional describe cómo crece la cantidad de operaciones necesarias para realizar la tarea –la factorización en números primos– según crece el número a factorizar. A grandes rasgos, tener una complejidad polinómica significa que la cantidad de operaciones crece de manera proporcional al crecimiento del número.

grupo de edad se ve que hay una mayor tasa de mortalidad en hombres, pero, al comparar simplemente por sexos (unificando los grupos de edad), la tasa de mortalidad es mayor en mujeres.

Este caso se puede explicar por el hecho de que hay más mujeres de mayor edad que hombres. Entonces la tasa de mortalidad que dominará entre las mujeres será la asociada a una mayor edad, mientras que en los hombres será el de un rango de edad inferior.

Debemos entender que, cuando se observa la paradoja, hay tres variables aleatorias, X, Y, Z, involucradas. Por ejemplo, X puede representar la condición binaria de fallecido por coronavirus, Y el sexo, y Z la pertenencia a algún grupo de edad. De manera general, sin embargo, vamos a considerar que X toma valores en $\{A, \neg A\}$, Y los toma en $\{B, \neg B\}$, y Z los toma en $\{C_1, C_2, \ldots, C_n, \ldots\}$.

La paradoja se dice que ocurre cuando $\mathbb{P}(A|B \wedge C_n) \geq \mathbb{P}(A|\neg B \wedge C_n)$ para $n \geq 1$, pero, sin embargo, $\mathbb{P}(A|B) < \mathbb{P}(A|\neg B)$.

Esto ocurre cuando las variables Y y Z no son independientes. Si aplicamos el teorema de la probabilidad total, tenemos,

$$\begin{aligned}\mathbb{P}(A|B) &= \sum_{n\geq 1} \mathbb{P}(A \wedge C_n|B) = \sum_{n\geq 1} \frac{\mathbb{P}(A \wedge B \wedge C_n)}{\mathbb{P}(B)} = \sum_{n\geq 1} \frac{\mathbb{P}(A \wedge B \wedge C_n)}{\mathbb{P}(B \wedge C_n)} \frac{\mathbb{P}(C_n \wedge B)}{\mathbb{P}(B)} = \\ &= \sum_{n\geq 1} \mathbb{P}(A|B \wedge C_n)\mathbb{P}(C_n|B) \, ,\end{aligned}$$

y, de manera similar,

$$\mathbb{P}(A|\neg B) = \sum_{n\geq 1} \mathbb{P}(A|\neg B \wedge C_n)\mathbb{P}(C_n|\neg B) \, .$$

Es decir, $\mathbb{P}(A|B)$ es la media de $\mathbb{P}(A|B \wedge C_n)$ pero ponderada por $\mathbb{P}(C_n|B)$, y no por $\mathbb{P}(C_n)$ como nos podría indicar la intuición (o un uso *naïve* del teorema de la probabilidad total)[2]. Ocurriendo algo análogo para $\mathbb{P}(A|\neg B)$.

En el ejemplo, efectivamente ocurre que las variables asociadas al sexo y al grupo de edad no son independientes. Esto puede parecer erróneo (pues el tiempo y la edad avanzan independientemente del sexo), pero hay que pensar que las variables aleatorias se realizan como una elección al azar de un individuo. Es decir, al escoger a una persona al azar, la probabilidad de dar con un hombre en un determinado rango de edad es diferente a la probabilidad de dar con una mujer en el mismo rango de edad (o sea, que existen $n \geq 1$ tales que $\mathbb{P}(C_n|B) \neq \mathbb{P}(C_n|\neg B)$).

B.4 Procesos de Poisson

Un proceso de Poisson es un tipo particular de proceso estocástico. Lo cual, de manera abreviada, podemos presentar como una familia de variables aleatorias parametrizadas por el tiempo. Es decir,

[2] Si Y y Z fuesen independientes, no ocurriría la paradoja puesto que $\mathbb{P}(C_n|B) = \mathbb{P}(C_n)$.

una asignación $t \mapsto X_t$ donde, para cada tiempo $t > 0$ fijo, X_t es una variable aleatoria (Van Kampen, 1992).

Ejemplos usuales de procesos de Poisson son el número de llamadas realizadas a una centralita de teléfono o el número de decaimiento de electrones que detecta un contador Geiger a lo largo de un intervalo de tiempo.

Es decir, nos preocupa el número de veces que hemos observar algún tipo de evento en un intervalo $[0, t[$, lo cual denotaremos por $N(t)$. Asumiremos que la tasa de observaciones esperadas por unidad de tiempo es $\lambda > 0$.

Adicionalmente, estableceremos más hipótesis sobre la naturaleza de los eventos a observar. La primera es que en un intervalo de tiempo de longitud δ, la probabilidad de que ocurra un evento se comporta como $\lambda\delta$ cuando $\delta \longrightarrow 0$. La segunda es que la probabilidad de que dos o más eventos ocurran en ese mismo intervalo de longitud δ es mucho más pequeña que δ cuando $\delta \longrightarrow 0$. Por último, asumimos que los eventos que se observan en intervalos de tiempo disjuntos son independientes.

Con esto, si dividimos $[0, t[$ en muchos intervalos disjuntos de longitud $\delta = t/n$, tenemos que para $n \gg 1$ (o sea, $n \longrightarrow \infty$), las hipótesis anteriores hacen que el número de eventos que podamos observar en el intervalo $[0, t[$ se distribuya como una binomial. Es decir,

$$\mathbb{P}(N(t) = k) \approx \binom{n}{k} \left(\lambda \frac{t}{n}\right)^k \left(1 - \lambda \frac{t}{n}\right)^{n-k},$$

puesto que los k eventos se observarían en k subintervalos distintos de los n disponibles, y en cada uno con una probabilidad $\lambda\delta = \lambda t/n$.

Es decir, $N(t)$ se distribuye de manera similar a una binomial $\mathrm{B}(n, \lambda t/n)$. Además notemos que, cuando $n \longrightarrow \infty$, $n \cdot \lambda t/n = \lambda t$. Es decir, para un tiempo t preestablecido, y con $n \longrightarrow \infty$, la distribución $\mathrm{B}(n, \lambda t/n)$ se puede aproximar por una distribución de Poisson, $\mathrm{Pois}(\lambda t)$.

Así se motiva la definición de un *Proceso de Poisson de tasa* $\lambda > 0$ como todo proceso estocástico continuo en tiempo, $N(t)$, con $t \geq 0$, que tome valores enteros no negativos, y que cumpla las propiedades siguientes:

1. $N(0) = 0$,
2. La asignación $t \mapsto N(t)$ es continua por la derecha, y,
3. Para cualquier partición del tiempo, $0 \leq t_1 < t_2 < \ldots < t_n$, los incrementos $N(t_{k+1}) - N(t_k)$, con $1 \leq k \leq n - 1$, son variables aleatorias independientes de tipo $\mathrm{Pois}(\lambda(t_{k+1} - t_k))$.

Asociado a este tipo de proceso conviene definir los *tiempos de salto* como los valores $S_0 = 0$, y, para $n \geq 0$, $S_{n+1} = \inf\{t > S_n : N(t) > N(S_n)\}$. Se pueden interpretar como los tiempos en los que aumenta el número total de observaciones. Notemos que esto significa que $N(t) \geq n \iff t \geq S_n$.

Además podemos definir los *tiempos entre observaciones* como el valor $X_1 = S_1$ y las diferencias $X_n = S_n - S_{n-1}$. Estas diferencias, $X_1, X_2, \dots, X_n, \dots$, resultan ser variables aleatorias independientes e idénticamente distribuidas que siguen una exponencial $\text{Exp}(\lambda)$.

En el caso tratado en la sección 3.3, se preocupan por el comportamiento de X_1, que representa el tiempo hasta la primera filtración. Concretamente, sabiendo que $X_1 \sim \text{Exp}(\lambda)$, saben que $\mathbb{P}(X_1 < t) = 1 - e^{-\lambda t}$, y, entre otras cosas, buscan el valor $t^* > 0$ para el que se cumple $\mathbb{P}(X_1 < t^*) = 0.95$. Dicho de otro modo, el tiempo necesario para asegurar, con una confianza del 95%, que haya ocurrido una filtración de evidencia que haga peligrar la conspiración.