

JUDEA PEARL
Y DANA MACKENZIE

EL LIBRO DEL PORQUÉ

La nueva ciencia de la causa y el efecto

Traducción de
GONZALO GARCÍA

PASADO & PRESENTE

PASADO & PRESENTE
BARCELONA

INTRODUCCIÓN

MEJOR EL CEREBRO QUE LOS DATOS

Siempre que una ciencia ha prosperado, ha prosperado gracias a sus propios símbolos.

AUGUSTUS DE MORGAN (1864)

Este libro narra la historia de una ciencia que, a pesar de haber transformado la manera en la que distinguimos los hechos de la ficción, apenas ha recibido la atención del público general. Las consecuencias de esta nueva ciencia ya están impactando en facetas cruciales de nuestra vida y tienen el potencial de afectar a un número aún mayor, desde el desarrollo de nuevos fármacos hasta el control de la política económica, desde la educación y la robótica hasta el control de armas y el calentamiento global. Pese a la diversidad de estos conjuntos de problemas y su carácter en apariencia inconmensurable, y por extraño que pueda parecer, la nueva ciencia los abarca todos bajo un marco unificado que prácticamente no existía hace dos décadas.

La nueva ciencia no tiene un nombre especialmente atractivo. Por mi parte, al igual que muchos de mis colegas, la llamo sencillamente «inferencia causal». Tampoco se caracteriza por una tecnología especialmente avanzada. De hecho, la tecnología ideal que la inferencia causal aspira a emular se halla en el interior de nuestra mente. Hace decenas de miles de años, los seres humanos empezaron a darse cuenta de que ciertas cosas causan otras cosas, y que alterar lo primero puede modificar lo segundo. Ninguna otra especie lo ha comprendido; no, desde luego, en la misma medida que nosotros. Este descubrimiento dio origen a las sociedades organizadas, a los pueblos y las ciudades, y con el tiempo a la civilización de raíz científica y tecnológica de la que hoy disfrutamos. Y todo ello por hacer una pregunta simple: ¿Por qué?

La clave de la inferencia causal es tomarse en serio esta pregunta. Partimos de suponer que el cerebro humano es la herramienta más avanzada jamás concebida para la gestión de las causas y los efectos. Nuestros cerebros almacenan una cantidad increíble de conocimiento causal que, añadiendo los datos, podríamos utilizar para dar respuesta a algunas de las cuestiones más acuciantes de nuestro presente. Con una ambición aún mayor, cuando de verdad comprendamos la lógica que subyace al pensamiento causal, podremos emularla en los ordenadores modernos para crear un «científico artificial». Este robot inteligente sería capaz de descubrir fenómenos aún desconocidos, daría explicación a dilemas científicos aún pendientes, diseñaría nuevos experimentos y seguiría incrementando el conocimiento causal a partir de la observación de nuestro entorno.

Pero antes de entrar en conjeturas sobre tal clase de evolución futurista es importante entender qué logros corresponden ya de hecho a la inferencia causal. Exploraremos de qué manera ha transformado el pensamiento científico en prácticamente todas las disciplinas que trabajan con datos y de qué modo está a punto de cambiar nuestras vidas.

La nueva ciencia aborda preguntas aparentemente tan directas como las siguientes:

- ¿Qué eficacia posee un determinado tratamiento a la hora de prevenir una enfermedad?
- ¿Nuestras ventas han subido a consecuencia de la nueva ley tributaria o quizá de nuestros anuncios?
- ¿Cuál es el coste de la atención sanitaria atribuible a la obesidad?
- ¿El historial de contrataciones puede demostrar que un empleado se rige por criterios de discriminación sexual?
- Estoy a punto de renunciar a mi trabajo. ¿Debo hacerlo?

Estas preguntas tienen en común el interés por las relaciones de causa y efecto, según se refleja en palabras como «prevenir», «a consecuencia», «atribuible», «criterios» y «debo». Son palabras corrientes en la lengua cotidiana y nuestra sociedad no cesa de formular tal clase de preguntas. Aun así, hasta hace muy poco tiempo, la ciencia no nos proporcionaba medios para expresarlas, no digamos ya para darles contestación.

La principal aportación de la inferencia causal a la humanidad ha sido, con gran diferencia, convertir esta desatención científica en una

cosa del pasado. La nueva ciencia ha engendrado un lenguaje matemático simple que permite expresar tanto las relaciones causales que ya conocemos bien como aquellas que deseamos investigar. La capacidad de expresar esta información en forma matemática ha producido una gran diversidad de métodos poderosos y fundamentados que combinan nuestro conocimiento con los datos y responden a preguntas causales como las cinco mencionadas más arriba.

Personalmente he tenido la suerte de participar en esta evolución científica durante el último cuarto de siglo. He contemplado cómo se forjaban sus avances en los pupitres de los estudiantes y los laboratorios de investigación, y he oído resonar sus hitos en conferencias científicas alejadas del foco de la atención pública. En estos momentos, cuando entramos en la era de una inteligencia artificial fuerte y muchos pregonan las posibilidades infinitas de los macrodatos (*Big Data*) y el aprendizaje profundo (*deep learning*), creo que resulta oportuno —y emocionante— desvelar a las lectoras y lectores algunos de los caminos más innovadores en los que la nueva ciencia se está adentrando, cómo afecta esto a la ciencia de los datos, y las múltiples maneras en que ello cambiará nuestras vidas en el siglo XXI.

El calificativo de «nueva ciencia» quizá lleve al lector a enarcar las cejas con escepticismo. Tal vez esté pensando que esta supuesta novedad no es tal, sino tan antigua como Virgilio, que en el 29 a. C. ya afirmaba: «Afortunado el que ha podido comprender las causas de las cosas». O quizá cifre su nacimiento en el momento en que los fundadores de la estadística moderna, Francis Galton y Karl Pearson, descubrieron por primera vez que los datos de población pueden arrojar luz sobre cuestiones científicas. Por desgracia, no aprovecharon la ocasión para dar importancia a la causalidad y, como se verá en las secciones históricas de este libro, hay mucho que contar sobre este desafortunado error. Aun así a mi modo de ver el impedimento más grave ha sido el abismo que separa el vocabulario en el que expresamos las preguntas causales y el vocabulario en el que tradicionalmente hemos comunicado las teorías científicas.

Para apreciar la profundidad de esta brecha imaginemos a qué dificultades se enfrentaría un científico que quisiera expresar algunas relaciones causales obvias; por ejemplo, que la lectura B de un barómetro registra la presión atmosférica P . Es fácil expresar esta relación en una ecuación como $B = kP$, donde k es alguna constante de proporcionalidad. Actualmente las reglas del álgebra nos permiten reescribir la mis-

ma ecuación de múltiples formas, por ejemplo $P = B/k$, $k = B/P$ o $B - kP = 0$. Todas significan lo mismo: que si conocemos dos de las tres cantidades, las que sean, la tercera está determinada. Ninguna de las letras k , B o P goza de una situación de privilegio matemático sobre las demás. Pero entonces, ¿cómo podemos expresar la firme convicción de que la presión causa los cambios del barómetro, y no al revés? Y si no podemos expresar siquiera algo tan sencillo, ¿cómo confiar en expresar las otras muchas convicciones causales que carecen de fórmulas matemáticas, como sería por ejemplo que el canto del gallo no provoca la salida del sol?

El profesorado que tuve en la universidad tampoco podía hacerlo, pero nunca se lamentó por ello. Apostaría a que tampoco se quejó ninguno de los maestros que usted haya tenido. Ahora entendemos por qué: nunca les mostraron un lenguaje matemático de causas, ni les revelaron tampoco sus beneficios. De hecho hay que reprochar a la ciencia que, a lo largo de tantas generaciones, haya desatendido el desarrollo de tal lenguaje. Todo el mundo sabe que apretar un interruptor enciende o apaga una luz, o que una tarde sofocante aumentará las ventas de la heladería local. ¿Por qué los científicos no han dado con fórmulas que describan hechos tan obvios, como sí han hecho con las leyes básicas de la óptica, la mecánica o la geometría? ¿Por qué han permitido que estos hechos languidezcan entre la simple y desnuda intuición, desprovistos de los útiles matemáticos que han permitido que otras ramas de la ciencia florezcan y maduren?

La respuesta es, en parte, que las herramientas científicas se crean para responder a necesidades científicas. Y precisamente porque no hallamos problema en manejar todo lo relativo a los interruptores, los helados y los barómetros, hasta ahora no se ha percibido la necesidad de una maquinaria matemática específica. Pero cuando la curiosidad científica se ha incrementado y hemos empezado a plantear preguntas causales en situaciones complejas —legales, de negocios, médicas, de acción gubernamental—, nos hemos encontrado que carecíamos de las herramientas y los principios que debía proporcionarnos una ciencia madura.

No es infrecuente que en la ciencia se produzcan tal clase de despertares tardíos. Por ejemplo, hasta hará unos cuatro siglos, la gente vivía feliz con su pericia natural para manejar las incertidumbres de la vida cotidiana, desde cruzar una calle a meterse en una trifulca, con los posibles riesgos de cada caso. Solo después de que se inventaran com-

plejos juegos de azar —en ocasiones, diseñados con mimo para invitarnos a tomar decisiones erróneas—, matemáticos como Blaise Pascal (1654), Pierre de Fermat (1654) y Christiaan Huygens (1657) consideraron necesario desarrollar lo que hoy conocemos como teoría de la probabilidad. Igualmente, matemáticos como Edmond Halley (1693) y Abraham de Moivre (1725) solo empezaron a examinar las tablas de mortalidad para calcular esperanzas de vida cuando las compañías de seguros de vida pidieron cálculos precisos sobre las rentas vitalicias. Un tercer ejemplo: cuando los astrónomos reclamaron predicciones rigurosas del movimiento celestial animaron a Jacob Bernoulli, Pierre-Simon Laplace y Carl Friedrich a crear una teoría de errores que nos ayudara a separar las señales del ruido. Todos estos métodos fueron predecesores de la estadística actual.

Irónicamente, la necesidad de contar con una teoría de la causalidad empezó a emerger al mismo tiempo que la estadística. De hecho, la estadística moderna nació a partir de las preguntas causales que Galton y Pearson formularon sobre la herencia, y de sus ingeniosos intentos de darles respuesta mediante datos transgeneracionales. Por desgracia, fallaron en la empresa; en vez de pararse a preguntar *por qué*, optaron por declarar que eran preguntas inabordables y se centraron en desarrollar —libre de toda causalidad— una próspera disciplina llamada estadística.

Fue un momento crítico en la historia de la ciencia. Se estuvo muy cerca de aprovechar la ocasión para dotar a las preguntas causales de un lenguaje propio, pero al final no se le sacó partido. A partir de entonces se consideró que se trataba de preguntas acientíficas, que se confinaron a la clandestinidad. Pese al heroico esfuerzo del genetista Sewall Wright (1889-1988), durante más de medio siglo el vocabulario causal estuvo prácticamente prohibido. Y cuando se veta la expresión, se veta el pensamiento y se anquilosan los principios, los métodos y las herramientas.

No es preciso que un lector sea científico para haber sido testigo de esta prohibición. En los primeros cursos de Estadística se enseña a todo estudiante a repetir el mantra de que la correlación no supone causalidad. ¿Acaso no es de sentido común? El canto del gallo se correlaciona muy a menudo con el amanecer, pero no lo causa.

Por desgracia, la estadística ha convertido la perogrullada en fetiche. Se insiste en que la correlación no supone causalidad, pero no se nos dice qué es la causalidad. En el índice de materias de un manual de

estadística no hay entrada para «causa». A los estudiantes nunca se les permite aseverar que X es la causa de Y; solo que X e Y están «relacionadas» o «asociadas».¹

A consecuencia de esta prohibición, se entendió asimismo que no era necesario disponer de útiles matemáticos con los que gestionar las preguntas causales, y la estadística se centró exclusivamente en el compendio de los datos, sin interpretación. Hubo una excepción brillante, el análisis de caminos (*path analysis*), inventado por el genetista Sewall Wright en la década de 1920. Aunque fue un antecesor directo de los métodos abordados por el presente libro, el análisis de caminos fue objeto de un claro menosprecio por la estadística y comunidades afines, por lo que languideció durante varias décadas en su condición embrionaria. Lo que debería haber sido el primer paso hacia la inferencia causal quedó sin continuación hasta la década de 1980. El resto de la estadística —y las numerosas disciplinas que se guiaban por ella— permaneció en la Era de la Prohibición, sumida en el error de creer que la respuesta a todas las preguntas científicas debe buscarse en los datos gracias a las ingeniosas estrategias de una «minería» especializada (*data mining*).

En buena medida, esta historia *datocéntrica* nos sigue persiguiendo hoy. En nuestra época se presupone que solventaremos todos nuestros problemas gracias a los macrodatos. En las universidades proliferan los cursos de «ciencia de datos» y las empresas de la «economía de los datos» ofrecen empleos lucrativos a los «científicos» especializados. Sin embargo, confío en que este libro convencerá al lector de que los datos carecen de toda inteligencia. Nos dirán que las personas que han tomado un medicamento se han recuperado antes que las que no lo han hecho, pero no pueden explicar por qué. Pudiera ser, por ejemplo, que todos los que tomaron el fármaco lo hubieran hecho porque podían permitírselo y de no haberlo ingerido se habrían recuperado con la misma rapidez.

Una y otra vez, en la ciencia como en los negocios, vemos situaciones en las que los datos no son suficientes. Los entusiastas de los macrodatos, en su mayoría, aun siendo relativamente conscientes de sus limitaciones, siguen aspirando a una inteligencia datocéntrica como si no hubiera acabado ya la Era de la Prohibición.

Ahora bien, como ya he dicho antes, en las últimas tres décadas las cosas han sufrido una transformación radical. En la actualidad, gracias a modelos causales muy elaborados, los científicos pueden abordar

problemas que antaño se consideraban insolubles, cuando no ajenos a toda investigación científica. Hace tan solo un siglo, por ejemplo, no se habría considerado científico plantear si fumar supone un peligro para la salud. La simple mención de las palabras «causa» o «efecto» habría levantado una tormenta de objeciones en cualquier revista estadística digna de tal nombre.

Incluso dos décadas atrás, formularle a un estadístico una pregunta como: «¿La aspirina me ha cortado el dolor de cabeza?» habría despertado casi el mismo asombro que interesarse por si creía en el vudú. Por citar a un querido compañero de profesión, habría podido ser «el tema de una conversación informal, pero no de un estudio científico». En nuestros días, por el contrario, los epidemiólogos, los expertos en ciencias sociales e informática, y al menos algunos de los economistas y estadísticos más adelantados plantean tal clase de preguntas de forma rutinaria y les dan respuesta con precisión matemática. A mi modo de ver, es un cambio plenamente revolucionario. Por eso me atrevo a hablar de la Revolución Causal, una conmoción científica que incluye, en vez de descartar, nuestra innata capacidad cognitiva de comprender causas y efectos.

La Revolución Causal no ha sucedido en el vacío. Se apoya en un secreto matemático cuya descripción más precisa es «cálculo de la causalidad» y que responde a algunos de los problemas más difíciles jamás planteados sobre las relaciones de causa y efecto. Estoy impaciente por desvelar este cálculo, no solo porque la historia de su evolución es tan turbulenta como intrigante, sino sobre todo porque espero que algún día se desarrollará su potencial en plenitud, más allá de lo que alcanzo siquiera a imaginar... ¿Quizá incluso por obra de alguna lectora o lector de este libro?

El cálculo de la causalidad precisa de dos lenguajes: diagramas causales para expresar lo que sabemos y un lenguaje simbólico, similar al álgebra, para expresar lo que deseamos saber. Los diagramas causales son simples dibujos de puntos y flechas que compendian el conocimiento científico disponible. Los puntos representan «variables», que se corresponden con intereses cuantificables. Por su parte las flechas representan relaciones causales, constatadas o hipotéticas, entre esas variables; nos dicen a qué variables «escucha» otra determinada variable. El trazo, la comprensión y el uso de estos diagramas resultan extremadamente fáciles, y el lector los encontrará por docenas en las páginas de este libro. Si sabe orientarse con un mapa de calles unidirec-

cionales, comprende los diagramas causales y puede contestar a la clase de preguntas que se han planteado al principio de esta introducción.

Aunque en este libro (como de hecho en mis últimos treinta y cinco años de investigación) he optado por los diagramas causales, esta herramienta no es el único modelo causal posible. A algunos científicos (por ejemplo los econométricos) les gusta trabajar con ecuaciones matemáticas; otros (como los estadísticos radicales) prefieren enumerar puntos de partida que en apariencia resumen la estructura del diagrama. Independientemente del lenguaje, el modelo debería describir (con unas u otras cualidades) el proceso que genera los datos; en otras palabras, las fuerzas de causa y efecto que actúan en nuestro entorno y dan forma a los datos.

En paralelo a este «lenguaje del conocimiento», de carácter diagramático, tenemos asimismo un «lenguaje de las preguntas», simbólico, que expresa aquellas cuestiones a las que les buscamos respuesta. Por ejemplo, si nos interesa el efecto de un fármaco (F) a lo largo de una vida (V), la duda podría formularse simbólicamente como: $P(V | do(F))$. En otras palabras, ¿cuál es la probabilidad (P) de que un paciente típico sobreviva V años si se le da ese fármaco? La pregunta describe lo que los epidemiólogos calificarían de intervención o tratamiento y se corresponde con lo que medimos en un ensayo clínico. En muchos casos también desearíamos comparar $P(V | do(F))$ con $P(V | do(sin F))$, que se corresponde con los pacientes «de control», que no han recibido el tratamiento. El operador do^* significa que nos ocupamos de una intervención activa, no de una observación pasiva. En la estadística clásica no existe nada remotamente parecido a este operador.

Es necesario invocar un operador de intervención $do(D)$ para asegurarnos de que los cambios observados en los años de vida (V) se deben al fármaco en sí y no se confunden con otros factores que tienden a acortar o alargar la vida. Si en lugar de intervenir nosotros dejamos que el propio paciente decida si toma el medicamento o no, esos otros factores podrían influir en su decisión y las diferencias percibidas en los años de vida ya no se deberían tan solo al hecho de tomar o no el fármaco. Imaginemos, por ejemplo, que solo los enfermos terminales tomaron el fármaco. Sin duda estas personas serían distintas a las que no tomarían el medicamento y al comparar los dos grupos se refleja-

* Literalmente, «hacer». Para no alterar las fórmulas, dejaremos este operador en su forma original en inglés. (*N. del t.*)

rían diferencias en la gravedad de la dolencia, antes que en el efecto del fármaco. Por el contrario, si sean cuales sean las condiciones previas se obliga a los pacientes a utilizar el fármaco, o se les impide hacerlo, se borrarían todas las diferencias preexistentes y dispondríamos de una comparación válida.

En lenguaje matemático, la frecuencia observada en los años de vida V entre los pacientes que han tomado voluntariamente el fármaco se escribe $P(V | F)$, lo que es la probabilidad condicional estándar que aparece en los manuales clásicos de la estadística. Aquí se expresa que la probabilidad (P) de una determinada esperanza de vida V está condicionada a ver que el paciente toma el fármaco F . Nótese que $P(V | F)$ puede resultar del todo distinto a $P(V | do(F))$. La diferencia entre *ver* que alguien toma un fármaco y *hacer* que lo tome es fundamental y explica por qué no llegamos a la conclusión de que el descenso de un barómetro está causando la tormenta que se avecina. Ver que el barómetro desciende incrementa la probabilidad de la tormenta, pero en cambio hacer que baje no afecta a esa probabilidad.

Esta confusión entre ver y hacer ha resultado en una multitud de paradojas, y en este libro nos ocuparemos de algunas. Un mundo carente de $P(V | do(F))$ y regido tan solo por $P(V | F)$ sería muy extraño, sin lugar a dudas. Por ejemplo, los pacientes evitarían ir al médico, para reducir la probabilidad de caer enfermos de gravedad; las ciudades despedirían a los bomberos para limitar la incidencia de los fuegos; los médicos recetarían un fármaco a hombres y mujeres, pero no a un paciente cuyo sexo no constara; etc. Resulta difícil creer que hace menos de tres décadas la ciencia actuaba de hecho en un mundo así: el operador *do* no existía.

Uno de los logros culminantes de la Revolución Causal ha sido explicar cómo se pueden predecir los efectos de una intervención sin haberla llevado a cabo de hecho. Pero esto habría sido imposible si, para empezar, no hubiéramos definido un operador *do* que nos permita formular la pregunta adecuada y, en segundo lugar, no hubiéramos concebido una forma de simularlo por medios no invasivos.

Cuando la cuestión científica que nos interesa implica pensar en retrospectiva, recurrimos a otro tipo de expresión exclusiva del razonamiento causal, que se denomina «contrafactualidad». Supongamos por ejemplo que Juan se ha tomado el fármaco F y ha fallecido un mes más tarde; queremos averiguar si el fármaco le ha provocado la muerte. Para responder a esta duda debemos imaginar un escenario en el

que Juan estaba a punto de tomar el medicamento pero ha cambiado de opinión. En tal caso, ¿habría sobrevivido?

Si nos fijamos en la estadística clásica, una vez más, se limita a acumular datos, por lo que ni siquiera nos proporciona un lenguaje con el que formular la pregunta. La inferencia causal nos permite la notación y, más importante aún, nos ofrece una solución. Al igual que cuando se predice el efecto de intervenciones (según el ejemplo que planteábamos antes), en muchos casos podemos emular un pensamiento humano retrospectivo gracias a un algoritmo que recoge cuanto sabemos sobre el mundo que observamos y genera una respuesta sobre el mundo contrafactual. Esta «*algoritmación* de los contrafactuales» es otra de las perlas que ha descubierto la Revolución Causal.

El razonamiento contrafactual, que se ocupa de los *¿y si...?*, quizá enoje a algunos lectores o lectoras, que lo tendrán por acientífico. Sin duda, la observación empírica no puede ni confirmar ni rebatir las respuestas a tal clase de preguntas. Ahora bien, nuestra mente no cesa de formular valoraciones muy fiables y reproducibles sobre lo que podría ser o haber sido. A nadie se le escapa, por ejemplo, que si esta mañana el gallo no hubiera cantado, el sol habría salido igualmente. Este consenso surge del hecho de que los contrafactuales no son el fruto de un capricho, sino que reflejan la estructura misma de nuestro modelo del mundo. Dos personas que comparten el mismo modelo causal también compartirán todas las valoraciones contrafactuales.

Los contrafactuales son las piezas básicas de construcción tanto de la conducta moral como del pensamiento científico. La capacidad de reflexionar sobre las propias acciones pasadas e imaginar escenarios alternativos está en la base del libre albedrío y la responsabilidad social. *Algoritmar* los contrafactuales invita a las máquinas pensantes a beneficiarse también de esta capacidad, a participar de esta forma de reflexionar sobre el mundo, exclusiva (hasta ahora) de los seres humanos.

La referencia a las *máquinas pensantes* del último párrafo es deliberada. Por mi parte llegué a este tema como científico informático, cuando trabajaba en el campo de la inteligencia artificial (IA). Esto implica dos diferencias con respecto a la mayoría de los colegas con los que comparto la arena de la inferencia causal. En primer lugar, en el mundo de la IA, no se puede decir que un tema se ha comprendido a fondo hasta que se le puede enseñar a un robot mecánico. Por eso se verá que hago hincapié, una y otra vez, en la notación, el lenguaje, el vocabulario y la gramática. Soy obsesivo, por ejemplo, con el tema de

si podemos expresar una determinada afirmación en un lenguaje dado y si una afirmación se sigue de otras. Es sorprendente cuánto se puede aprender por el mero hecho de atender a la gramática de las formulaciones científicas. Mi hincapié en el lenguaje procede asimismo de la profunda convicción de que el lenguaje da forma a nuestros pensamientos. No hay forma de responder a una pregunta que no se puede plantear, y no se puede plantear una pregunta si se carece de las palabras necesarias. Como estudioso de la filosofía y la ciencia informática, mi atracción por la inferencia causal deriva en buena medida de la emoción de ver cómo un lenguaje científico desatendido pasa de la cuna a la madurez.

Mi formación previa en el aprendizaje de las máquinas me ha proporcionado un incentivo más para estudiar la causalidad. A finales de la década de 1980 llegué a la conclusión de que el hecho de que las máquinas no pudieran entender las relaciones causales era, quizá, el mayor de los obstáculos en su adquisición de una inteligencia equiparable a la humana. En el último capítulo de este libro volveré a mis raíces para que exploremos juntos las consecuencias de la Revolución Causal en el campo de la inteligencia artificial. Creo que una IA fuerte es un objetivo alcanzable, y que no le debemos tener miedo precisamente porque la causalidad forma parte de la solución. Un módulo de razonamiento causal dotará a las máquinas de la capacidad de reflexionar sobre sus errores, señalar las deficiencias de programación, actuar como entidades morales y conversar con los seres humanos, con naturalidad, al respecto de nuestras decisiones e intenciones.

UN ESQUEMA DE LA REALIDAD

En nuestros días, todo el mundo ha oído sin duda términos como «conocimiento», «información», «inteligencia» y «datos», y alguien quizá tenga dudas al respecto de sus diferencias o cómo interactúan entre sí. Aquí propongo añadir otro concepto en el mismo puchero, el de «modelo causal», y me parece razonable que el lector se pregunte si el resultado final no será aún más confuso.

¡No lo será! De hecho, anclará en un contexto concreto y significativo los conceptos escurridizos de ciencia, conocimiento y datos y nos permitirá ver cómo los tres colaboran para generar respuestas a cues-

tiones científicas difíciles. La Figura I.1 esquematiza un «motor de inferencia causal» que quizá serviría para que una futura inteligencia artificial manejara el razonamiento causal. Es importante advertir que no se trata tan solo de un esquema de cara al futuro, sino también de una guía sobre cómo funcionan los modelos causales en las aplicaciones científicas de nuestros días y cómo interactúan con los datos.

El motor de inferencia acepta tres clases de entradas —Premisas, Interrogantes y Datos— y genera tres tipos de elementos de salida. El primero es una decisión de Sí/No al respecto de si, en teoría, el interrogante formulado se puede responder con el modelo causal existente (presuponiendo que los datos son perfectos e ilimitados). Si se responde que Sí, el motor de inferencia produce entonces un Estimando. Se trata de una fórmula matemática que podemos imaginar como una receta para generar la respuesta a partir de cualesquiera datos hipotéticos, siempre que se pueda disponer de ellos. Por último, cuando el motor de inferencia haya recibido la entrada de Datos, usará la receta para generar una Estimación de respuesta real, junto con los cálculos estadísticos de cuánta incertidumbre incluye esa estimación. Esta incertidumbre da cuenta del volumen del conjunto de datos y del hecho que puede ser incompleto y contener errores.

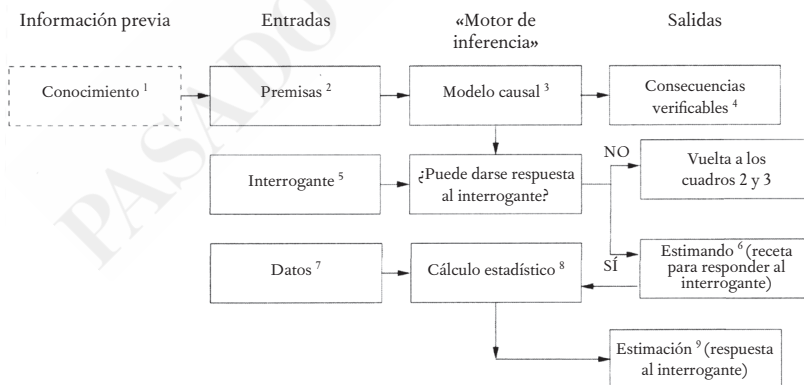


FIGURA I.1. Manera en que un «motor de inferencia» combina los datos con el conocimiento causal para generar respuestas a los interrogantes que nos interesan. El recuadro de líneas intermitentes no forma parte del motor en sí, pero es necesario para construirla. También podríamos trazar flechas entre los cuadros 4 y 9 y el cuadro 1, pero aquí he preferido simplificar al máximo el diagrama.

Para poder explorar el esquema con mayor profundidad he numerado los recuadros del 1 al 9, y procederé a examinarlos en el contexto de la pregunta: «¿Cuál es el efecto del fármaco F en una esperanza de vida V ?».

1. Con «Conocimiento» aludo a las huellas de la experiencia que el agente razonador ha tenido en el pasado, incluidas las observaciones pasadas, acciones pasadas, formación educativa y rumores que se hayan considerado relevantes para el interrogante planteado. El recuadro se marca con una línea de puntos para indicar que el «Conocimiento» está implícito en la mente del agente pero no se explica formalmente en el modelo.
2. La investigación científica requiere siempre de premisas simplificadoras, es decir, supuestos y postulados que el investigador cree conveniente explicitar a partir del Conocimiento disponible. Aunque la mayor parte de lo que el investigador sabe permanece implícito en su cerebro, solo las premisas ven la luz del día y se encapsulan en el modelo. En realidad se pueden leer a partir del modelo, por lo que algunos lógicos han llegado a la conclusión de que un modelo no es otra cosa que una lista de premisas. Sin embargo, los científicos informáticos no están nada de acuerdo con tal afirmación, y defienden que la forma en que las premisas se representan pueden suponer una enorme diferencia en la capacidad personal de especificarlas correctamente, extraer conclusiones e incluso ampliarlas o modificarlas a la luz de pruebas relevantes.
3. Existen varias alternativas para los modelos causales: diagramas causales, ecuaciones estructurales, proposiciones lógicas, etc. Por mi parte soy un firme partidario de los diagramas causales para casi todas las aplicaciones, esencialmente por su carácter transparente, pero también porque ofrecen respuestas explícitas a muchas de las cuestiones que deseamos preguntar. En lo que atañe a la construcción del diagrama, la definición de la «causalidad» es simple, aunque un tanto metafórica: una variable X es una causa de Y si Y «escucha» a X («atiende» a X) y determina su valor en respuesta a lo que oye. Por ejemplo, si sospechamos que los años de vida V de un paciente «escuchan» si este ha tomado el fármaco F , denominamos a F causa de V y en un diagrama causal trazamos una flecha de F a V .

Como es natural, es probable que la respuesta al interrogante sobre F y V dependa igualmente de otras variables que también se deben representar en el diagrama, junto con sus causas y efectos. (Aquí las designaremos colectivamente como Z .)

4. El patrón de escucha prescrito por los caminos del modelo causal suele tener como resultado dependencias o patrones observables en los datos. Se les da el nombre de «Consecuencias verificables» porque se pueden utilizar para verificar el modelo. Se trata de afirmaciones como «Ningún camino conecta F y V », que se traduce en la aseveración estadística « F y V son independientes»; es decir, hallar F no modifica la probabilidad de V . Si los datos contradicen esta consecuencia es una señal de que necesitamos revisar nuestro modelo. Tales revisiones requieren de otro motor que recibe las entradas desde los recuadros 4 y 7 y computa el «grado de idoneidad», es decir, hasta qué punto los datos son compatibles con las premisas del modelo. Por mor de la simplicidad, en la Figura 1.1 no he mostrado este segundo motor.
5. Los interrogantes que se plantean al motor de inferencia son las preguntas científicas que deseamos responder. Deben formularse con un vocabulario causal. Por ejemplo, ¿qué es $P(V | do(F))$? Entre los principales logros de la Revolución Causal figura haber dotado a este lenguaje de transparencia científica y rigor matemático.
6. «Estimando» es una palabra que procede del latín, con el sentido de «lo que se debe estimar». Se trata de una cantidad estadística que se calculará a partir de los datos y que, una vez calculada, es legítimo que represente la respuesta a nuestro interrogante. Aunque se escribe como una fórmula de probabilidad —por ejemplo, $P(V | F, Z) \times P(Z)$ — de hecho es una receta para responder al interrogante causal mediante el tipo de datos de los que disponemos, una vez certificada por el motor.

Es de gran importancia comprender que, en contra de las estimaciones tradicionales de la estadística, algunos interrogantes quizá no puedan recibir respuesta con el modelo causal vigente, ni siquiera después de haber reunido una cierta cantidad de datos. Por ejemplo, si nuestro modelo muestra que tanto F como V dependen de una tercera variable Z (como pudie-

ra ser aquí el estadio de una enfermedad) y no tenemos ninguna manera de medir Z , entonces la pregunta $P(V \mid do(F))$ no se puede responder. En este caso, reunir datos supone solo una pérdida de tiempo. Lo que debemos hacer es volver atrás y refinar el modelo, ya sea añadiendo otros conocimientos científicos que pudieran permitirnos calcular Z o incorporando supuestos de simplificación (a riesgo de equivocarnos), como por ejemplo que el efecto de Z en F carece de relevancia.

7. Los datos son los ingredientes que se usarán para la receta del estimando. Resulta crucial comprender que los datos no poseen ninguna comprensión de las relaciones causales. Nos hablan de cantidades como $P(V \mid F)$ o $P(V \mid F, Z)$. Es tarea del estimando decirnos cómo cocinar estas cantidades estadísticas para obtener una expresión que se basa en los supuestos del modelo y resulta equivalente, desde el punto de vista de la lógica, al interrogante causal; pongamos, $P(V \mid do(F))$.

Tengamos en consideración que el mismo concepto de los estimandos (y, de hecho, toda la parte superior de la Figura 1.1) no existe en los métodos tradicionales del análisis estadístico. Ahí, estimando e interrogante coinciden. Por ejemplo, si nos interesamos por el porcentaje de personas que han vivido V y han tomado el fármaco F , nos limitamos a escribir la cuestión como $P(F \mid V)$. La misma cantidad sería nuestro estimando. Esto ya especifica qué porcentaje de los datos conviene estimar y no necesita de ningún conocimiento causal. Por esta razón, a algunos estadísticos, todavía hoy, les resulta sumamente difícil comprender que ciertos conocimientos residan fuera de las competencias de la estadística o entender por qué los datos por sí solos no pueden compensar la ausencia de conocimiento científico.

8. La estimación es lo que sale del horno. Sin embargo, su carácter es meramente aproximado debido a otra verdad constatada sobre los datos: la población puede ser infinita, en teoría, pero la muestra siempre será finita. En nuestro ejemplo de trabajo la muestra consta de los pacientes que deseamos estudiar. Incluso si los elegimos al azar siempre hay cierta posibilidad de que los porcentajes medidos en la selección no sean representativos de los que se dan en la población en su conjunto. Por suer-

te, la disciplina de la estadística se beneficia de las técnicas avanzadas del aprendizaje de las máquinas y nos ofrece muchas, pero muchas formas de manejar esta incertidumbre: para suavizar los datos escasos se recurre a menudo a modelos paramétricos y semiparamétricos, métodos de probabilidad máxima y cálculos de la propensión.

9. A la postre, si nuestro modelo es correcto y los datos, suficientes, tendremos una respuesta al interrogante causal, del tipo: «El fármaco F incrementa la vida V de los pacientes diabéticos un 30 % (con un margen de error de ± 20 %)». ¡Hurra! Esta respuesta se incorporará a nuestro conocimiento científico (recuadro 1) y, si las cosas no han salido como esperábamos, quizá nos sugiera mejoras para nuestro modelo causal (recuadro 3).

Este diagrama de flujo quizá parezca complicado, a primera vista; quizá el lector se esté preguntando si es verdaderamente necesario. A fin de cuentas, en nuestra vida corriente somos más o menos capaces de hacer valoraciones causales sin tener que acomodarnos a un proceso tan complicado y, desde luego, sin necesidad de aplicar la matemática de las probabilidades y los porcentajes. Nuestra intuición causal, por lo general, se basta para manejar la clase de incertidumbres que podemos hallar en la gestión de la casa y la familia o incluso en la vida profesional. Pero si queremos que un robot necio aprenda a pensar causalmente, o hacer avanzar las fronteras del conocimiento científico donde no podemos contar con la guía de la intuición, entonces resulta imprescindible disponer de un procedimiento como este, cuidadosamente estructurado.

En todo este proceso quiero destacar en especial el papel de los datos. En primer lugar téngase en cuenta que solo reunimos datos después de haber planteado el modelo causal, haber afirmado qué interrogante científico deseamos responder y haber derivado el estimando. Es una diferencia clara con respecto al modelo estadístico tradicional que, como hemos recordado más arriba, ni siquiera cuenta con un modelo causal.

Pero el mundo científico actual presenta un nuevo desafío cuando se trata de razonar con sensatez sobre las causas y efectos. Aunque en el ámbito científico la conciencia de que se requiere un modelo causal haya dado un salto prodigioso, sin embargo muchos investigadores en

materia de inteligencia artificial preferirían ahorrarse la fase penosa de construir o adquirir un modelo causal y basar todas las tareas cognitivas exclusivamente en los datos. Se tiene la esperanza —en nuestros días, en general, tácita— de que siempre que surjan cuestiones causales los propios datos nos guiarán hasta las respuestas correctas.

Soy declaradamente escéptico con respecto a esta tendencia porque sé que los datos por sí mismos son del todo incapaces de comprender las causas y efectos. Por ejemplo, la información sobre los efectos de una acción o intervención simplemente no figura en los datos básicos, salvo que se reúnan mediante una manipulación experimental controlada. Por el contrario, cuando contamos con un modelo causal a menudo podemos predecir el resultado de una intervención y ello a partir de datos no manipulados ni intervenidos.

La defensa de los modelos causales resulta todavía más necesaria cuando aspiramos a responder a interrogantes contrafactuales del tipo: «¿Qué habría pasado si hubiéramos actuado de otro modo?». Analizaremos los contrafactuales con especial detalle porque son los interrogantes más exigentes para cualquier inteligencia artificial. Además se encuentran en el centro mismo de los avances cognitivos que nos hicieron humanos y de las capacidades imaginativas que han posibilitado la ciencia. También aclararemos por qué todo interrogante sobre el mecanismo por el que las causas transmiten sus efectos —la pregunta de «¿Por qué?» por antonomasia— es de hecho, aunque quizá no lo parezca, una cuestión contrafactual. Así pues, si queremos que en algún momento los robots sepan responder a un «¿Por qué?» o por lo menos comprender qué significa la pregunta, deberemos pertrecharlos con un modelo causal y enseñarles a resolver interrogantes contrafactuales, como en la Figura 1.1.

Otra ventaja de los modelos causales en comparación con la minería de datos y el aprendizaje profundo es la adaptabilidad. Obsérvese que en la Figura 1.1 el estimando se computa exclusivamente a partir del modelo causal, antes de examinar las especificidades de los datos. Esto dota al motor de inferencia causal de una adaptabilidad suprema porque el estimando computado resulta válido para cualquier dato compatible con el modelo cualitativo, independientemente de las relaciones numéricas existentes entre las variables.

Para ver qué importancia reviste esta adaptabilidad compárese este motor con un agente de aprendizaje —en este ejemplo, un ser humano, pero en otros podría ser quizá un algoritmo de aprendizaje pro-

fundo o quizá una persona que utiliza un algoritmo de aprendizaje profundo— que intente aprender solamente a partir de los datos. Cuando observe el resultado V de muchos pacientes a los que se ha dado el fármaco F , nuestra persona puede predecir la probabilidad de que un paciente con las características Z pueda sobrevivir V años. Ocurre que la trasladan a otro hospital, en otro barrio de la ciudad, donde las características de la población (dieta, higiene, hábitos de trabajo) son distintas. Incluso si las nuevas características se limitan a modificar las relaciones numéricas existentes entre las variables apuntadas, no tendrá más remedio que reciclarse, empezar de cero y aprender una nueva función de predicción. Un programa de aprendizaje profundo no da para más: solo sabe unir datos y funciones. En cambio, si nuestra persona contara con un modelo de cómo funciona el fármaco y la estructura causal se mantiene intacta en la nueva ubicación, entonces el estimando obtenido durante la fase de instrucción no perdería su validez. Se podría aplicar a los nuevos datos para generar una nueva función de predicción específica para la nueva población.

Muchos interrogantes científicos tienen un aspecto distinto cuando se los examina «con la lente de la causalidad», y para mí ha supuesto un placer jugar con esta lente, cuyo potencial se ha ido enriqueciendo con nuevas perspectivas y nuevas herramientas a lo largo de los últimos veinticinco años. Espero —y de hecho creo— que los lectores y las lectoras de este libro compartirán mi placer. Así pues pasaré a cerrar esta introducción con un adelanto de los próximos alcances que aparecerán en este libro.

El capítulo 1 ensambla los tres peldaños de observación, intervención y contrafactuales para armar la Escalera de la Causalidad, que es asimismo la metáfora central del presente libro. También expondrá los principios básicos del razonamiento con diagramas causales, nuestra herramienta principal de modelado, de modo que el lector dispondrá de lo necesario para dominar el razonamiento causal; de hecho habrá llegado mucho más lejos que varias generaciones de científicos de datos que los habían intentado interpretar con una lente carente de modelo, pasando por alto las distinciones sobre las que arroja luz la Escalera de la Causalidad. El capítulo 2 narra la extravagante historia de cómo la disciplina de la estadística se cegó a sí misma a la causalidad, lo que ha implicado consecuencias de gran alcance en todas las ciencias que dependían de datos. También narra la historia de uno de los grandes héroes de este libro, el genetista Sewall Wright, que en la

década de 1920 trazó los primeros diagramas causales y durante muchos años fue uno de los pocos científicos que osaron tomarse en serio la causalidad.

El capítulo 3 refiere una historia igualmente curiosa: cómo me convirtió a la causalidad trabajar en la IA y, en particular, en las redes bayesianas. Este fue el primer instrumento que permitió a los ordenadores pensar en «tonos de gris»; durante un tiempo creí que tenían la llave que desencadenaría la IA. Hacia finales de la década de 1980, sin embargo, concluí que me había equivocado, y este capítulo refiere el viaje de la profecía a la apostasía. Aun así, las redes bayesianas siguen siendo una herramienta muy importante para la IA y siguen encapsulando buena parte de la fundamentación matemática de los diagramas causales. Además de presentar una introducción asequible, regida por la causalidad, al teorema de Bayes y el método de razonamiento bayesiano, el capítulo 3 entretendrá al lector con ejemplos de aplicación de las redes de Bayes en la vida real.

El capítulo 4 habla de la principal contribución de la estadística a la inferencia causal: el ensayo controlado aleatorio (RCT, por sus siglas inglesas). Desde una perspectiva causal, el RCT es un instrumento de creación humana para desvelar el interrogante $P(V \mid do(F))$, que es una propiedad de la naturaleza. Su objetivo esencial es desasociar las variables de interés (digamos, F y V) frente a otras variables (Z) que en caso contrario tendrían influencia sobre ambas. Eliminar las distorsiones o la «confusión» que estas variables agazapadas generan ha supuesto un problema durante todo un siglo. Este capítulo conduce al lector hasta una solución sorprendentemente sencilla al problema general de los factores de confusión, que el lector podrá comprender con diez minutos lúdicos de trazar caminos en un diagrama.

El capítulo 5 versa sobre un momento especialmente fecundo de la historia de la causalidad —más aún, de la historia de la ciencia—, cuando los estadísticos lidiaban con la cuestión de si fumar causa cáncer de pulmón. Como no podían usar su instrumento favorito, el ensayo controlado aleatorio, les resultaba difícil ponerse de acuerdo en la respuesta, e incluso en cómo dar sentido a la pregunta. El debate del tabaco pone de relieve la estricta importancia de la causalidad. Millones de vidas se perdieron o acortaron porque los científicos carecían de una metodología o un lenguaje adecuados para dar respuesta a los interrogantes causales.

Si el capítulo 5 abordaba cuestiones de especial gravedad, el 6 con-

fío en que nos devolverá una sonrisa. Es un capítulo de paradojas: la paradoja de Monty Hall, la paradoja de Simpson, la paradoja de Berkson y alguna más. Esta clase de paradojas clásicas tienen su gracia como rompecabezas, pero también incluyen una faceta más seria, sobre todo cuando se las contempla desde la perspectiva causal. De hecho, casi todas representan un choque con la intuición causal que, por lo tanto, sirve para revelar la anatomía de esta intuición. Constituían una señal de alarma que debería haber llevado a los científicos a comprender que la intuición humana se basa en una lógica causal, no estadística. Creo que el lector disfrutará de esta presentación novedosa de sus viejas paradojas favoritas.

Los capítulos 7 a 9, por último, llevan al lector a emprender un ascenso vertiginoso por la Escalera de la Causalidad. Empezaremos en el capítulo 7, con preguntas sobre la intervención; contaré cómo mis estudiantes y yo pasamos veinte años batallando para automatizar las respuestas a los interrogantes de tipo *do*. Al final tuvimos éxito, y este capítulo explica el meollo del «motor de inferencia causal» que genera la respuesta Sí/No y el estimando de la Figura 1.1. Analizar este motor dará al lector el poder de detectar ciertos patrones en el diagrama causal que ofrecen respuestas inmediatas al interrogante causal. Estos patrones reciben los nombres de ajuste de puerta trasera, ajuste de puerta delantera y variables instrumentales, y, en la práctica, son los verdaderos caballos de tiro de la inferencia causal.

El capítulo 8 le conducirá hasta lo alto de la escalera con el estudio de los contrafactuales. Se considera que estos suponen una parte fundamental de la causalidad desde por lo menos 1748, cuando el filósofo escocés David Hume propuso la siguiente definición (algo tortuosa) de causalidad: «Podemos definir una causa como un objeto seguido por otro, donde a todos los objetos similares al primero siguen objetos similares al segundo. O, en otras palabras, donde de no haber existido el primer objeto, el segundo no habría llegado a existir». David Lewis, filósofo de la Universidad de Princeton, que falleció en 2001, apuntó que en realidad Hume estaba proporcionando dos definiciones, y no una; la primera, de la regularidad (es decir, a la causa sigue el efecto, de forma regular) y la segunda del contrafactual («de no haber existido el primer objeto...»). Mientras que los filósofos y científicos han prestado su atención sobre todo a la definición de la regularidad, Lewis planteaba que la definición del contrafactual es más similar y próxima a la intuición humana: «Entendemos que una causa es algo que provoca una

diferencia, y la diferencia que provoca debe ser distinta a lo que habría sucedido sin ella».

Al lector le alegrará saber que ahora podemos dejar atrás las discusiones académicas y obtener un valor (o probabilidad) real para todo interrogante contrafactual, por enrevesado que sea. Resultan de especial interés las preguntas que se refieren a las causas suficientes y necesarias de los acontecimientos observados. Por ejemplo, ¿cuán probable es que la acción del demandado sea una causa necesaria del perjuicio sufrido por el demandante? ¿Cuán probable es que el cambio climático, como obra del hombre, sea una causa suficiente de una ola de calor?

El capítulo 9, por último, analiza el tema de la mediación. Tal vez se haya preguntado, cuando hablábamos sobre el trazo de flechas en un diagrama causal, si deberíamos dibujar una flecha del fármaco F a la vida V si el fármaco afecta los años de vida solo a través de su efecto sobre la presión sanguínea Z (un mediador). En otras palabras, ¿el efecto de F sobre V es directo o indirecto? Si las dos cosas son ciertas, ¿cómo evaluamos su importancia relativa? Tales cuestiones no poseen tan solo un gran interés científico, sino también ramificaciones prácticas: si entendemos el mecanismo con el que un fármaco actúa, quizá seamos capaces de desarrollar otros medicamentos con el mismo efecto que sean más baratos o tengan menos efectos secundarios.

Por fortuna veremos que esta antiquísima búsqueda de un mecanismo de mediación se ha reducido ahora a un ejercicio de álgebra, y que los científicos están usando las nuevas posibilidades de la caja de herramientas causal para resolver tales problemas.

El capítulo 10 cierra el libro volviendo al problema que en un principio me llevó a la causalidad: la automatización de la inteligencia de nivel humano (llamada en ocasiones «IA fuerte»). Creo que el razonamiento causal resulta esencial para que las máquinas se comuniquen con nosotros en nuestro propio lenguaje sobre las políticas gubernamentales, los experimentos, las explicaciones, las teorías, el recordamiento, la responsabilidad, el libre albedrío y las obligaciones; resulta esencial también para que, a la postre, tomen sus propias decisiones morales.

Si pudiera resumir el mensaje de este libro en una frase lo más concisa posible, diría: usted es más inteligente que los datos. Los datos no comprenden las causas y los efectos; las personas, sí. Espero que la nueva ciencia de la inferencia causal nos permitirá comprender mejor

cómo lo hacemos, porque no existe una forma mejor de entendernos a nosotros mismos que emulándonos. En la era de los ordenadores, esta nueva comprensión también trae consigo la expectativa de amplificar nuestras capacidades innatas de forma que podamos dotar de más sentido a los datos, sean macro o micro.

PASADO & PRESENTE

ÍNDICE

<i>Prefacio</i>	9
INTRODUCCIÓN. MEJOR EL CEREBRO QUE LOS DATOS	11
Un esquema de la realidad	21
1. LA ESCALERA DE LA CAUSALIDAD	33
Los tres niveles de la causalidad	37
El minitest de Turing	46
Sobre probabilidades y causalidad	57
2. DE LOS BUCANEROS A LOS COBAYAS:	
LA GÉNESIS DE LA INFERENCIA CAUSAL	63
Galton y la búsqueda abandonada	72
Pearson: la cólera del fanático	76
Sewall Wright, conejillos de Indias y diagramas de caminos	82
<i>E pur si muove</i> («Y sin embargo, se mueve»)	90
De la objetividad a la subjetividad: la conexión bayesiana .	99
3. DE LAS EVIDENCIAS A LAS CAUSAS: EL REVERENDO BAYES	
SE ENCUENTRA CON EL SEÑOR HOLMES	103
Bonaparte, el detective informático	104
El reverendo Bayes y el problema de la probabilidad	
inversa	105
Del teorema de Bayes a las redes bayesianas	119
Las redes bayesianas: qué nos dicen las causas de los datos .	123
¿Dónde está mi equipaje? De Aquisgrán a Zanzíbar	127
Las redes bayesianas en el mundo real	132
De las redes bayesianas a los diagramas causales	139

4.	CONFUSIÓN Y DESCONFUSIÓN, O CÓMO ELIMINAR	
	LA VARIABLE AGAZAPADA	145
	Pavor a la confusión	151
	El hábil interrogatorio a la naturaleza: por qué funcionan los RCT	153
	El nuevo paradigma de la confusión	161
	El operador <i>do</i> y el criterio de la puerta trasera	167
5.	UN DEBATE LLENO DE HUMO: DESPEJAR EL PANORAMA	177
	Tabaco: una epidemia creada	180
	La comisión del cirujano general y los criterios de Hill	189
	Tabaco para bebés	193
	Debates encendidos: la ciencia contra la cultura	196
6.	¡PARADOJAS A MOGOLLÓN!	199
	La paradoja de Monty Hall	200
	Más sesgo de colisión: la paradoja de Berkson	207
	La paradoja de Simpson	210
	La paradoja de Simpson, en imágenes	221
7.	MÁS ALLÁ DEL AJUSTE: LA CONQUISTA DEL MONTE	
	INTERVENCIÓN	229
	La vía más simple: la fórmula de ajuste de puerta trasera	230
	El criterio de puerta delantera	234
	El cálculo <i>do</i> , o la victoria de la mente sobre la materia	241
	El tapiz de la ciencia o los actores ocultos en la orquesta <i>do</i>	251
	El extraño caso (doble) del doctor Snow	254
	Colesterol bueno y malo	261
8.	CONTRAFACTUALES: MINERÍA DE LOS MUNDOS QUE PODRÍAN	
	HABER SIDO	269
	De Tucídides y Abraham a Hume y Lewis	272
	Resultados potenciales, ecuaciones estructurales y la algoritmación de contrafactuales	279
	La ventaja de ver las propias premisas	290
	Los contrafactuales y el derecho	296
	Causas necesarias, causas suficientes y cambio climático	301
	Un mundo de contrafactuales	306

9. MEDIACIÓN: LA BÚSQUEDA DE UN MECANISMO	309
Escorbuto: el mediador equivocado	312
¿Naturaleza o crianza? La tragedia de Barbara Burks	314
En busca de un lenguaje (o la paradoja de las admisiones de Berkeley)	319
<i>Daisy</i> , los gatitos y los efectos indirectos	327
La mediación en el País de las Maravillas Lineales	332
Asumir los «habría»	339
Casos prácticos de mediación	345
10. MACRODATOS, INTELIGENCIA ARTIFICIAL Y LAS GRANDES PREGUNTAS	357
Modelos causales y macrodatos	358
IA fuerte y libre albedrío	366
<i>Agradecimientos</i>	379
<i>Notas</i>	383
<i>Bibliografía</i>	387
<i>Índice alfabético</i>	415