

LA PROBABILIDAD Y LA CAUSALIDAD

Sergio F. Martínez Muñoz

I. INTRODUCCION

Uno de los problemas centrales en la teoría del conocimiento y en la filosofía de la ciencia es la relación entre la probabilidad y la causalidad. La elucidación de esta relación está íntimamente ligada al desarrollo y evaluación de teorías o modelos de la explicación científica, a discusiones acerca del método y la racionalidad en la ciencia, y al planteamiento de una serie de cuestiones metafísicas en la ciencia moderna. La elaboración de un nexo profundo entre los conceptos de causa y probabilidad principia con la revolución científica del siglo XVII y ha continuado como parte de la consolidación de los presupuestos filosóficos que han guiado el desarrollo de la ciencia y la filosofía desde entonces. En esta introducción daremos un breve resumen de la historia de los conceptos de causa y probabilidad. Ambos conceptos tienen una venerable historia y hacer ver aspectos importantes de esta historia nos parece necesario para dar una presentación que no prejuzgue de entrada la relación que se da entre la causalidad y la probabilidad.

Ya en Aristóteles se encuentra la noción de *endoxos* (que coincide en gran medida con la noción medieval de *probabilis*) utilizada para referirse a un tipo de creencias que dependía de la información del sujeto. Se encuentra además en Aristóteles el concepto de una «regla general», tomada de la medicina, con la que se hace referencia a aquello que ocurre en la mayoría de los casos. La elaboración de estas ideas por los escépticos resulta en una serie de distinciones de niveles cualitativos de creencias que son consideradas como el punto de partida para cualquier decisión en la vida práctica (cf. Byrne, 1968) y que constituyen el origen del concepto de probabilidad tal y como lo conocemos hoy. Por otra parte, debates filosóficos acerca del concepto de causalidad se encuentran ya entre los filósofos presocráticos. Aristóteles sintetiza una larga tradi-

ción cuando distingue cuatro diferentes tipos de causa (razones o principios explicativos). El rechazo de la identificación de principios explicativos y causas es una de las características más distintivas de la «revolución científica» que tiene lugar alrededor del siglo XVII. Este rechazo se encuentra estrechamente vinculado a una de las constantes distintivas de la ciencia moderna, el énfasis en la búsqueda de mecanismos como principios explicativos.

Una de las intuiciones básicas de la noción moderna de causa es que de alguna manera la causa necesita sus efectos. Supongamos que yo golpeo la puerta con mis nudillos, es de esperar (en condiciones normales) que un sonido típico va a ocurrir como consecuencia. Similarmente, si suelto un objeto pesado en el aire esperamos que se irá rápidamente al suelo. Hay una gran cantidad de teorías filosóficas que tratan de elaborar esta intuición central del concepto de causa, la intuición de que una relación causal (una relación entre una causa y su efecto) involucra una conexión necesaria. A partir de Hume, sin embargo, estas teorías se consideran muy sospechosas y por lo menos no dignas de formar parte de un programa empirista. Hume arguye que no es posible inferir racionalmente la existencia de una conexión necesaria entre causa y efecto y considera que todo lo que estamos justificados a decir es que una relación causal es una sucesión que ejemplifica una regularidad (caracterizada por las propiedades de contigüidad, prioridad temporal y conjunción constante). Hay muchos problemas con esta propuesta. La luz se prende generalmente si presiono el interruptor, pero esto no sucede cuando hay una mala conexión o no hay electricidad. Este problema del enfoque humeano se denomina con frecuencia el *problema de la dependencia del contexto*. Otro problema con el enfoque de Hume es el *problema de la causa común*. Los efectos de una causa común van a estar regularmente unidos sin que uno cause el otro. La explosión de una bomba en una ciudad causa que se rompan muchos vidrios y que tiemble el suelo, pero que se rompan los vidrios no causa que tiemble el suelo. El problema consiste en distinguir el caso de una relación de conjunción constante que es el resultado de una causa común y el caso en que se trata de una relación genuina de causa y efecto. Finalmente, otro problema central para la teoría de Hume es el problema de la «similaridad», examinado en detalle por Russell, y que en la actualidad se conoce más por el nombre de *problema de la ambigüedad*. Muchas veces la noción de causa se aplica en el lenguaje científico a sucesos singulares. Un tipo de problemas muy importante para el desarrollo de nuestra concepción de causa proviene de preguntas en la medicina como: ¿qué queremos decir con «causa de muerte» y cómo clasificamos las causas de muerte para escribir un certificado de defunción? Independientemente de la solución que le demos a este tipo de problemas (ampliamente debatidos en el siglo XIX) es claro que queremos aplicar la noción de «causa de muerte» a individuos concretos. Queremos responder preguntas como ¿de qué murió Juan? y sólo posteriormente hacer estadísticas. De otra manera las estadísticas correrían el

riesgo de ser tan arbitrarias como las zoologías fantásticas de la antigüedad. Supongamos que Juan murió de un disparo al corazón. ¿Cómo podríamos formular este simple enunciado en la teoría de Hume? Hume y muchos filósofos después de él han tratado de responder a este problema diciendo que la relación causal se establece entre cosas similares. El disparo en el corazón fue la causa de la muerte si la muerte fue inmediatamente posterior al disparo y disparos similares (disparos al corazón) tienen siempre como consecuencia la muerte del que los recibe. Para explicar la muerte de Juan, Hume reformularía el contenido epistémico de enunciados singulares causales (enunciados que describen una relación de causa y efecto entre sucesos singulares) en términos de una sucesión constante de tipos de sucesos. Russell fue el primero en recalcar la importancia de la dificultad de definir la noción de similaridad requerida por la teoría de Hume. ¿Cuándo decimos que dos disparos al corazón son similares? ¿Cuándo entran en el mismo ángulo y penetran en la misma región del corazón y son disparados con una pistola del mismo tipo? Esto no sería suficiente para establecer que *siempre* un disparo al corazón sea seguido de la muerte de individuos similares a Juan. Podría ser que Juan (o una persona similar a Juan) se hubiera muerto un segundo antes del disparo de un ataque al corazón. O podría ser que el hermano gemelo de Juan no se muriera de recibir el tipo de disparo que mató a Juan porque este hermano gemelo podría estar conectado a un corazón artificial.

La teoría de Hume requiere que reinterpretemos un enunciado causal singular como refiriéndose a una clase de relaciones entre sucesos similares. El problema es que es ambigua con respecto a qué clase se tendría que evaluar para el enunciado causal singular. Si en el caso de Juan incluimos en la clase pertinente a su hermano conectado a un corazón artificial entonces deberíamos concluir que el disparo no es la causa de su muerte. La intuición de Hume en este punto es que deberíamos referir un enunciado causal a una clase *pertinente*. El gran problema es cómo elucidar esta noción de pertinencia. Si decimos que la clase de referencia (que incluye todos los sucesos similares que queremos tomar en cuenta) debe incluir todos los sucesos causalmente pertinentes (y no otros) introducimos una definición circular.

Los problemas de la teoría de Hume son muchos y están en gran medida todavía con nosotros como problemas de la filosofía contemporánea. Es importante recalcar, sin embargo, que una de las contribuciones más relevantes de la filosofía de Hume para la filosofía posterior fue el planteamiento del papel central de la relación entre causalidad y probabilidad para la epistemología. Antes de Hume había por una parte «probabilidad» (*probabilis*) en el sentido de opinión, y «ciencia» o *episteme*, esto es, conocimiento cierto. Estos dos tipos de creencia eran considerados cualitativamente diferentes. Esta idea persiste en Hume, pero él ya toma en serio la idea de que las probabilidades a través de la acumulación de evidencia pueden llevarnos gradualmente al conocimiento, y

ésta es una tesis central para su proyecto epistemológico. Así, según Hume, todos los razonamientos acerca de cuestiones de hecho se basan en la relación de causa y efecto, y esto requiere que elucidemos la noción de causa. Sin embargo, no hay ningún argumento racional que nos lleve de la causa al efecto, y por tanto se plantea el problema de cuál es la base empírica de la creencia en relaciones de causa y efecto. La probabilidad para Hume consistiría precisamente en la teoría que serviría como el nexo entre la definición de causa como conjunción constante a la que ha sido llevado por su análisis naturalista y la evidencia empírica de conjunciones casi constantes.

Hume no desarrolla esta idea, y las interpretaciones tradicionales de Hume ignoran esta sugerencia en la filosofía de Hume que sólo va a ser retomada por la filosofía en el siglo XX. El desarrollo de la idea de Hume requiere el desarrollo de un cálculo cuantitativo de las probabilidades que si bien parece no haber sido conocido directamente por Hume, empezó a desarrollarse a finales del siglo XVI.

Concluimos esta sección mencionando brevemente otros intentos por definir el concepto de causa. El mismo Hume, inmediatamente después de formular su famosa definición de causa en términos de sucesión constante (sección VII de la *Investigación acerca del entendimiento humano*) sugiere otra definición de causa muy diferente pero que él parece considerar como una manera de decir lo mismo: «Si el primer objeto no hubiera sido, el segundo nunca hubiera existido». Esta definición contrafáctica ha sido desarrollada recientemente por una serie de autores. El *locus* clásico de este tipo de propuestas es Lewis (1973, reimpreso en Lewis, 1986). Hay también una serie de teorías que parten de la intuición de que el concepto de causa debe analizarse en términos de condiciones necesarias y/o suficientes para la ocurrencia de un suceso. Una de las propuestas más elaborada de este tipo es la de J. L. Mackie (1974). Este libro ofrece también un estudio detallado de otras propuestas similares. Finalmente están todas aquellas teorías de la causalidad en las que no se acepta que una relación causal sea reducible a regularidades. Ducasse (1924), Cartwright (1983 y 1989), Woodward (1986), son una muestra de propuestas de este tipo. Finalmente, hay también intentos como el de Russell de mostrar que los filósofos han sido vagos e inconsistentes en su uso del concepto de causa y que lo mejor es evitar su uso por completo. Me parece que Ducasse tiene razón cuando en 1924 opina que la oposición de Russell a las discusiones filosóficas del concepto de causa se basan en una confusión entre el concepto de causa que está en discusión y la noción de ley empírica como conjunción constante.

El señalamiento de esta confusión está en el fondo de una serie de planteamientos modernos acerca de la relación entre causalidad y explicación. Torretti formula claramente el origen de la confusión (1991, 270):

Hay una tendencia en filosofía a considerar los diversos estados de un proceso [gobernado por una ecuación diferencial] como constituyendo una ca-

dena causal en la que los estados sucesivos están unidos uno con otro como causa y efecto. Esta tendencia no generaría problemas si aquellos que la siguen se abstuvieran de usar el término «causa» en su sentido pre-científico ordinario. Tal abstinencia, sin embargo, va contra su fin, que consiste precisamente en presentar las ecuaciones diferenciales de la física matemática como el medio apropiado para comprender las conexiones causales en la naturaleza. (Traducción mía.)

II. LA TEORIA MATEMATICA DE LA PROBABILIDAD

La revolución científica del siglo XVII es sobre todo el principio de una manera de hacer ciencia basada en un nuevo ideal de objetividad que asume que la tarea de la ciencia es el descubrimiento de «mecanismos», y no de los principios causales o «principios activos» que son en última instancia responsables de los procesos naturales pero que están fuera del alcance de nuestro intelecto. En la tradición aristotélica, los límites de nuestra percepción son los límites de nuestras capacidades cognitivas. Lo que no podemos percibir con nuestros sentidos está fuera del alcance del conocimiento científico. Para la filosofía natural del siglo XVII, más allá de nuestra percepción sensorial hay mecanismos inteligibles que constituyen el punto de partida objetivo de las teorías científicas. Pero así como un reloj de cuerda puede tener la misma carátula que un reloj de péndulo, los intentos por descubrir los mecanismos subyacentes en los fenómenos naturales sólo pueden ser confiables hasta cierto punto. Los filósofos naturales comparten con los jueces, los mercaderes y los aficionados a juegos de azar, el problema de decidir en circunstancias de incertidumbre. El cálculo de probabilidades surge como un intento por cuantificar esta incertidumbre y por formular claramente el sentido de la antigua creencia escéptica de que «la probabilidad es la guía en la vida».

Jakob Bernoulli es el primero de una serie de matemáticos que tratan de reducir «el arte de la conjetura» a un cálculo matemático. El teorema de Bernoulli (publicado en 1713) es el primer intento por relacionar las probabilidades con frecuencias observadas de sucesos en casos en donde no es plausible asumir que todos los casos posibles son igualmente probables. Esto libera al cálculo de las probabilidades de su asociación con juegos de azar y permite la aplicación del cálculo a situaciones de interés científico y filosófico. En notación moderna, Bernoulli demostró el siguiente teorema:

$$\lim_{n \rightarrow \infty} P(|p - m/n| < \epsilon) = 1, \text{ para } \epsilon \text{ tan pequeño como se desee,}$$

donde p = proporción «verdadera»
 n = número de pruebas
 m/n = la proporción observada.

El teorema de Bernoulli asume que la probabilidad *a priori* p se conoce. Nos permite computar el número de observaciones requeridas

para hacer que la probabilidad de m/n esté en el intervalo $[p - \epsilon, p + \epsilon]$ con una cierta probabilidad. Bernoulli parece haber creído que su teorema podía constituirse en un método para resolver el problema de la inferencia, esto es, el problema de encontrar la probabilidad de la conjetura que m/n es igual a una probabilidad desconocida p , «la probabilidad inversa», dada una razón observada de m/n . Este problema, sin embargo sólo fue resuelto por Bayes y posteriormente por Laplace con mayor generalidad, si bien el significado de esta solución es todavía hoy el problema central en la teoría de la inferencia estadística.

Laplace, en su trabajo de 1774 *De la probabilité des causes* y en una serie de escritos posteriores, y en particular en su famoso tratado de 1814 sobre las probabilidades, propone por primera vez un método general para estimar *las probabilidades de las causas* de un suceso observado. Éste es el inicio de toda una tradición en los fundamentos de la ciencia de concebir el concepto de causa como íntimamente ligado al concepto de probabilidad, y que está muy alejada de la noción de causa presupuesta por muchos filósofos aún en la actualidad. La idea básica detrás del método de Laplace es simple: el uso del teorema de Bayes con una distribución inicial uniforme sobre las causas. Laplace concibió el problema de las probabilidades inversas como un tipo especial de promedio: si un suceso puede ser el resultado de n causas diferentes, la probabilidad de que en un caso específico se deba a una causa C_m es igual a la probabilidad de E dada C_m dividida por la suma de las n probabilidades condicionales de E dado C_i :

$$P(C_m/E) = P(E/C_m) / \sum_i P(E/C_i)$$

Ésta es la formulación del teorema de Bayes por Laplace. Laplace utiliza esta fórmula como punto de partida para un análisis detallado de las consideraciones probabilistas de una situación dada que nos permite inferir si lo dado es el resultado de la aplicación de una causa (o sistema de causas) *constante*, o si por el contrario lo dado se explica simplemente como el resultado de una serie de fluctuaciones azarosas que típicamente provendrían de errores en la medición. El trabajo de Laplace en astronomía es una aplicación de este método. Laplace utilizó este método para inferir que las irregularidades de los movimientos de Júpiter y de Saturno eran producto de su interacción mutua, así como para formular su teoría de las mareas y para resolver otra serie de problemas que constituyen el núcleo de la corroboración empírica del programa newtoniano de la mecánica celeste en el siglo XIX. Una vez que Laplace hizo ver la utilidad de su método en la mecánica celeste pasó a hacer ver la importancia de este método en otra serie de disciplinas científicas. La diversidad de las aplicaciones consideradas por Laplace puede apreciarse en su famoso prefacio a su tratado de 1814. Laplace aquí recalca la importancia de la posibilidad de aplicar su método a las «ciencias morales» y da una serie de ejemplos que están íntimamente ligados al desarrollo de la interpre-

tación clásica de las probabilidades (la posibilidad de analizar matemáticamente los testimonios en un juicio, por ejemplo).

El trabajo de Laplace se ha desarrollado a lo largo de líneas muy diferentes de investigación en el siglo XX. Uno de estos desarrollos es el tema de las teorías probabilistas de la causalidad que presentaremos en el apartado IV. Otras líneas de desarrollo que se basan en el trabajo de Laplace son el programa bayesiano en inferencia estadística y una serie de modelos en la teoría de las decisiones. Un libro reciente de inspiración laplaciana es el de Suppes (1984).

III. EXPLICACION, CAUSALIDAD Y PROBABILIDAD

El problema de encontrar un modelo filosófico satisfactorio de la explicación científica es uno de los grandes temas de la filosofía contemporánea de la ciencia. Cuando se formuló por primera vez explícitamente un modelo de la explicación científica, en un famoso trabajo de Hempel y Oppenheim en 1948, se asumía implícitamente que el mundo era determinista y se asumía que las explicaciones eran argumentos. Estas dos suposiciones han sido expuestas y severamente criticadas en una serie de trabajos posteriores. Por un lado la filosofía fue tomando conciencia de la importancia de las implicaciones filosóficas de la mecánica cuántica y de otras teorías científicas que parecen describir un mundo no determinista que incluye descripciones probabilistas irreducibles. De alguna manera las probabilidades parecen ser parte esencial de una descripción correcta del mundo en que vivimos. Por otra parte, la suposición de que las explicaciones son (o pueden reconstruirse como) argumentos confronta una serie de problemas. Es de destacarse el problema que las explicaciones tienen un orden temporal mientras que los argumentos no. Un eclipse puede deducirse de ciertas leyes astronómicas y de ciertas condiciones iniciales. Pero sólo en el caso de que las condiciones iniciales antecedan al eclipse hablamos de una explicación. En 1965, Hempel presenta un modelo de explicación estadística. Hempel requiere que una explicación estadística-inductiva, lo que él llama una explicación IS, sea un argumento inductivo que haría el *explanandum* (el enunciado a explicar) predecible con alta probabilidad inductiva, dado el *explanans* (las premisas del argumento-explicación). Es fácil encontrar argumentos inductivos con premisas compatibles cuyas conclusiones se contradicen mutuamente. A esta propiedad de los argumentos inductivos, Hempel la llamó «la ambigüedad de la explicación estadística». Hempel resolvió el problema de la ambigüedad requiriendo que las explicaciones inductivas sean relativas a clases de referencia. Este modelo es insatisfactorio por una serie de razones. Una de las críticas más contundentes es la de Alberto Coffa (1974). Coffa muestra que la relativización epistémica de las explicaciones inductivas a clases de referencia no permite entender el sentido en que un argumento inductivo es una explicación. Esto hizo ver

muy claramente que argumentos inductivos no podían resolver un problema básico para cualquier candidato a modelo de la explicación.

Críticas al modelo estadístico-inductivo de Hempel y a otros similares llevaron a muchos filósofos a la conclusión de que era necesario introducir algún tipo de postulado causal para formular un modelo de la explicación estadística. La íntima relación entre causalidad y explicación ha sido recalcada en una serie de trabajos de Wesley Salmon.

Salmon (elaborando una idea de Reichenbach) parte del análisis de la relación causal implícito en la física contemporánea (en la teoría de la relatividad especial en particular). El ejemplo paradigmático de Salmon es el proceso causal típico de la teoría de la relatividad especial: un rayo de luz. Según esta teoría, ninguna cosa puede viajar a mayor velocidad que la luz. Pero, ¿a qué se refiere este enunciado? Indudablemente que se refiere a objetos materiales como bolas de billar, planetas, fotones etc. Pero no se refiere a las sombras por ejemplo. Las sombras pueden viajar a cualquier velocidad. La luz de un reflector giratorio que cae sobre una pared puede viajar a cualquier velocidad. Todos estos ejemplos, como Salmon hace notar, comparten la propiedad de que son incapaces de transmitir mensajes. Esta observación es el punto de partida de Salmon para formular una distinción entre *procesos causales* y *pseudo-procesos*. Los procesos causales, pero no los pseudo-procesos pueden ser marcados y transmitir la marca. Otro concepto central de la teoría de Salmon es el de *interacción causal*. Las interacciones causales ocurren cuando un proceso causal intersecta a otro produciendo una modificación en su estructura.

Una explicación científica para Salmon consiste en citar por lo menos una parte de los procesos e interacciones causales que resultan en el suceso que queremos explicar. Nótese que para Salmon la explicación de un tipo (clase) de relación causal es derivativa. La teoría de Salmon es en primer lugar un modelo que explica procesos causales singulares.

No es difícil aceptar que Salmon ha logrado con su teoría capturar una serie de aspectos importantes del concepto de causalidad tal y como este concepto surge en la mecánica clásica y el electromagnetismo. Pero parece realmente dudoso que Salmon pueda capturar con su modelo una serie de explicaciones causales que se alejan de su paradigma. Sin ir muy lejos, el tipo de explicación que se da en la teoría general de la relatividad de por qué una partícula se mueve de la manera como lo hace no hace referencia a un mecanismo de transferencia de marcas sino a la *estructura* (afín y métrica) del espacio-tiempo. Y un poco más lejos, un tipo de explicaciones que Salmon reconoce que escapan a su modelo son las explicaciones de correlaciones distantes que la mecánica cuántica predice y que han sido corroboradas empíricamente. Estas correlaciones no son meramente un tipo muy especial de interacción que no parece adecuarse al modelo de Salmon (como Salmon ha sugerido en varias oportunidades). Este tipo de correlaciones cuestiona el punto de partida mismo del modelo de Salmon, la distinción entre procesos causales y

pseudo-procesos. El modelo de Salmon confronta una serie de problemas en otras áreas de la física. Una serie de artículos críticos del modelo de Salmon y de propuestas alternativas pueden encontrarse en Kitcher y Salmon (1989). Una vez que nos alejamos de la física las dificultades con la teoría de Salmon son todavía más aparentes. Uno de los temas centrales en la biología moderna es el tema de la causalidad implícita en procesos evolutivos. Si bien hay una gran controversia al respecto, es claro que sea cual sea el resultado de la controversia, la noción de causalidad en juego tiene poco que ver con los procesos causales a lo Salmon (cf. por ejemplo, Sober, 1984).

IV. TEORIAS PROBABILISTAS DE LA CAUSALIDAD

La condición de conjunción constante en la teoría humeana de la causalidad implica que la probabilidad del efecto dada la causa es máxima: $P(E/C)=1$. En donde E representa el efecto, C la causa y $P(E/C)$ la probabilidad condicional de E dado C definida por la fórmula $P(E/C)=P(E \wedge C)/P(C)$. $P(X)$ representa la probabilidad del suceso X. Como el mismo Hume lo reconoce, la mayoría de veces sólo tenemos una conjunción constante aproximada. El problema que se plantea entonces a todo filósofo que siga las prescripciones epistemológicas de Hume es explicar el concepto de causa a partir de conjunciones constantes imperfectas. En esta sección nos interesa presentar una serie de propuestas que tratan de incorporar, en la discusión del problema filosófico acerca de la naturaleza de la causalidad, los desarrollos de la teoría matemática de la probabilidad y su aplicación como método para el establecimiento de causas (un tema que hemos bosquejado en el apartado II). En este contexto de ideas, Suppes (1970) sugiere una estrategia atractiva. Suppes parte de la intuición central que guía el método probabilista para la detección de causas, que la causa debe de aumentar la probabilidad del efecto, esto es $P(E/C)>P(E)$. Una manera simple de ver que esta condición por sí sola no puede caracterizar una relación causal es notar que esta relación es simétrica. Si $P(E/C)>P(E)$ entonces se sigue como una mera consecuencia de la definición de probabilidad condicional que $P(C/E)>P(C)$. Si queremos caracterizar una relación causal en términos de probabilidades es necesario entonces agregar condiciones adicionales. El núcleo de la teoría de Suppes puede resumirse en las siguientes tres condiciones.

C causa E si y sólo si:

- i) $P(E/C)>P(E)$
- ii) No hay un suceso F tal que $P(E/CF)=P(E/F)$
- iii) C precede temporalmente a E.

La condición (ii) es la condición de filtrado de una causa común. La condición (iii) postula la asimetría requerida para distinguir la causa del efecto. Estas condiciones no son del todo inofensivas. Supongamos

por ejemplo que un suceso B causa C y además que $P(A/BC)=P(A/C)$. Por ejemplo, la estricnina causa que el corazón se detenga, por lo que la probabilidad de muerte *dado* que el corazón se detiene y suficiente estricnina es la misma que la probabilidad de muerte *dada* suficiente estricnina. Esta conclusión es problemática porque si un suceso anterior determina completamente un suceso posterior entonces todos los sucesos intermedios tendrían que ser considerados como causalmente ineficaces. Pero no parece correcto considerar al hecho que el corazón se detiene como una mera causa aparente de muerte.

Un problema más de fondo en la propuesta de Suppes tiene que ver con el problema de la ambigüedad. El problema surge cuando se intenta explicar el contenido epistémico de una relación causal singular en términos de regularidades casi-constantes. Esto requiere que la relación causal singular se refiera a una clase de [relaciones entre] sucesos similares. El problema es que no parece posible encontrar un criterio no arbitrario para decidir con respecto a qué clase vamos a evaluar el contenido epistémico de la relación causal singular. Con respecto a una clase de referencia podemos llegar a concluir que hay una relación causal, pero con respecto a otra podemos llegar a concluir que no la hay. La teoría de Suppes afronta este problema de manera particularmente aguda. En primer lugar está el problema de decidir si una teoría probabilista de la causalidad se refiere a sucesos singulares o si se refiere a sucesos generales (tipos de sucesos). La muerte de Juan de un balazo al corazón es un suceso singular (localizable en un punto del espacio-tiempo). La muerte de seres humanos por balazos al corazón es un suceso general (un tipo de suceso). La relación entre enunciados causales singulares y enunciados causales generales no es simple. Esto es el núcleo del problema de la ambigüedad. Del hecho de que en general un balazo en el corazón cause la muerte no se infiere que Juan murió de un balazo en el corazón, pudo haber muerto un segundo antes de un susto. Y del hecho de que María se murió de un ataque al corazón mientras daba un paseo por el bosque no se sigue que dar paseos por el bosque sea dañino para la salud. Suppes afirma que el formalismo de su teoría se refiere tanto a sucesos generales como a sucesos singulares. Según él, el formalismo puede interpretarse como aplicándose a ambos tipos de sucesos. De ser esto cierto el problema de la ambigüedad se disolvería. Podríamos pensar que las distinciones arriba mencionadas son importantes para entender psicológicamente la situación, pero que el análisis causal es independiente de si estamos refiriéndonos a sucesos singulares o a sucesos generales. Suppes posteriormente ha argüido que la ontología de sucesos es inadecuada para la formulación de una teoría de la causalidad y que el análisis debe formularse en términos de «variables causales». Esta postura simplemente evita el problema filosófico. El problema no es simplemente si los sucesos son o no una ontología adecuada (estoy de acuerdo con Suppes en que muchas veces no lo son). El problema es en el fondo el lugar que ocupa la información estadística como evidencia empírica para una teoría de la causalidad.

La ambigüedad en la interpretación del formalismo (aplicada a sucesos o procesos o cualquier otro tipo de ontología que sea sensible a la distinción que hay que hacer entre dos tipos de entes) no puede sostenerse. La causalación entre sucesos singulares es obviamente transitiva, asimétrica y no reflexiva. Pero la relación causal entre eventos generales no es asimétrica ni transitiva. La tensión tiende a causar problemas psicológicos y los problemas psicológicos tienden a causar tensión. Un disparo en el corazón causa que el corazón se pare y que el corazón se pare causa la muerte, pero un disparo en el corazón no necesariamente causa la muerte. Contrario al deseo de Suppes parece ser que su teoría puede interpretarse solamente como una teoría de la causalidad de sucesos generales. Pero entonces, ¿cuál es la relación entre la causalidad singular y la causalidad general? De alguna manera una teoría de la causalidad general tiene que conectarse con enunciados singulares causales. El problema sería cómo. No sólo nos interesa decir que el sida causa la muerte, nos interesa *entender* este enunciado en términos de las circunstancias biológicas de individuos concretos que pueden contraer la enfermedad.

Esta restricción de una teoría probabilista de la causalidad a sucesos generales genera una serie de problemas relacionados con la interpretación de la condición (iii) de prioridad temporal. La relación de prioridad temporal es clara si nos referimos a sucesos singulares que tienen un índice temporal, pero no es clara cuando nos referimos a sucesos generales. Es posible entender esta idea diciendo que E precede a C si cuando C ocurre entonces E ocurre normalmente. ¿Pero qué queremos decir con «normalmente»? Recordemos que la idea detrás de todas estas teorías de la causalidad es que la causalidad es una relación objetiva, pero la «normalidad» no puede sino referirse a nuestra limitada experiencia y no necesariamente tiene que ver con lo que es objetivamente el caso. Es posible responder a estos problemas afinando la teoría de Suppes, pero esto nos lleva a otra serie de complicaciones (cf. Salmon, 1984; Eells, 1991). Suppes posteriormente ha argüido que la ontología de sucesos es inadecuada para la formulación de una teoría de la causalidad y que el análisis debe formularse en términos de «variables causales». Desde el punto de vista del problema filosófico relativo a la pregunta del contenido empírico de la relación causal esto es insuficiente. Por otro lado, las teorías de Cartwright y otras teorías probabilistas de la causalidad más recientes toman en cuenta esta idea.

Hay otro problema serio con las propuestas de causalidad probabilista y con la teoría de Suppes en particular. Este problema, conocido como la paradoja de Simpson (después de su «popularización» en Cartwright, 1979) es una versión particularmente ilustrativa del problema del contexto. Esta paradoja fue ejemplificada por Cartwright citando un estudio estadístico acerca de los patrones de admisión en la Universidad de California en Berkeley. El estudio encontró una correlación entre ser admitido y ser hombre. La frecuencia de admisión (durante varios años) entre los hombres que solicitaban entrar a la universidad era mayor que

la frecuencia de admisión de las mujeres que solicitaban entrada. Esto por supuesto sugiere discriminación contra las mujeres. Sin embargo, un estudio más a fondo pudo explicar la correlación de manera tal que no implicaba que hubiera discriminación (esto es, que hubiera una relación causal entre ser mujer y tener menos oportunidades de ser admitida). La razón es que las decisiones de admisión son hechas en cada facultad por separado y cuando se investigaron los historiales de admisión por facultad se descubrió que no había en ningún departamento una correlación entre el sexo del solicitante y su admisión. Las mujeres tendían a solicitar entrada a las facultades en las que era más difícil entrar. En promedio, la frecuencia de admisión era menor para mujeres que para hombres, pero eso en todo caso sólo implicaría que las mujeres tenían preferencias distintas que los hombres en promedio. Este análisis parece exonerar a Berkeley de discriminación debido a la partición escogida de la clase de referencia. Si hubiéramos tratado de argumentar que no había discriminación haciendo una partición no por facultades sino sobre la base de la habilidad para patinar, entonces no exoneraríamos a Berkeley del cargo de discriminación. La diferencia reside en nuestro conocimiento previo del contexto causal. Sabemos que solicitar admisión en departamentos de moda es motivo para ser rechazado con facilidad. No estamos dispuestos a aceptar que saber patinar bien sea causa de ser rechazado. El ejemplo ilustra de una manera muy penetrante el problema del contexto. No es suficiente tratar que una partición incluya todos los factores causales, es igualmente importante que la partición no sea tampoco tan fina que genere correlaciones espurias (como en el ejemplo anterior). Pero entonces, sólo las particiones que toman en cuenta las variables causalmente significativas deben de contar en la especificación de una relación causal. Esto nos orilla a la conclusión que sólo una definición circular del concepto de causa es plausible.

Cartwright llega a la misma conclusión anterior después de analizar una serie de contraejemplos en contra de la idea intuitiva de que la causa incrementa la probabilidad de sus efectos. Un famoso ejemplo de este tipo es el siguiente. Generalmente se supone que fumar causa problemas al corazón. Esperamos entonces que la probabilidad de un ataque al corazón sea mayor si se es fumador que si no se es. Esta idea está equivocada. Incluso si es cierto que fumar causa problemas al corazón, la probabilidad de un ataque al corazón no aumentará si fumar va acompañado de una actividad que tiende a prevenir problemas del corazón, como hacer ejercicio regularmente. Su conclusión es que en este ejemplo y en muchos otros que examina la causa falla en incrementar la probabilidad de sus efectos porque en la situación descrita la causa esta correlacionada con otro factor causal que domina en sus efectos. Esto sugiere que la condición (i) en la teoría de Suppes tiene que modificarse de tal manera que tome en cuenta sólo las situaciones (contextos causales) en los que tales correlaciones con factores causales ocultos a la situación no estén presentes. Esto lleva a Cartwright (1979) a formular una teoría en

la que «C causa E» si y sólo si C aumenta la probabilidad de E en cada situación que es causalmente homogénea con respecto a E. El problema central de Cartwright consiste en explicar qué entiende por «homogeneidad causal». Esto es más fácil explicarlo si hacemos un poco más explícita la condición que constituye el núcleo de su teoría:

(CC) C causa E si y sólo si $P(E/CF) > P(E/F)$ donde F es cualquier factor causal alternativo.

Por «factor causal alternativo» se entiende cualquier suceso que causa E o no-E excluyendo a C o a los efectos de cualquier combinación de estos sucesos. La teoría de Cartwright se restringe explícitamente a formular una teoría de sucesos generales.

Cartwright abandona el objetivo de Hume y de Suppes, la *definición* del concepto de causa (de la clase de enunciados causales). La teoría de Cartwright es circular y no puede dar una respuesta al problema de Hume, como ella lo reconoce. Su teoría debe verse más bien como una teoría empírica acerca de la «relación causal». De esta manera no es sorprendente que la definición sea circular. Hay muchas definiciones en la ciencia que son circulares, por ejemplo la definición de masa en la física clásica. La teoría de Cartwright no está libre de problemas, sin embargo, y (como una serie de críticas han hecho ver) es necesario agregar una serie de condiciones cada vez más complicadas para caracterizar la noción de «homogeneidad causal». La teoría de Cartwright tiene también una serie de consecuencias inaceptables o por lo menos contra-intuitivas. Por ejemplo, la probabilidad de muerte de un ser humano es una con o sin un balazo en el corazón. Por tanto la teoría de Cartwright implicaría que la muerte no tiene causas (ya que en este caso la probabilidad de muerte E dado un balazo en el corazón F es $P(E/CF) \nless P(E/F)$, para cualquier C).

Laplace no tendría problema con la conclusión anterior. Para él (y sobre todo para Cournot) la vida de un individuo está fijada por las leyes eternas e inmutables que rigen el universo y lo único que se afecta a lo largo de la vida es la esperanza de vida. Las verdaderas causas son para Laplace las leyes universales de las que los sucesos de la experiencia se siguen necesariamente (en el sentido que se derivan matemáticamente de ellas). Pero si no aceptamos la hipótesis de un determinismo tan férreo tal conclusión es inaceptable.

En 1989 (y en algunos artículos anteriores) Cartwright modifica su teoría de manera tal que las relaciones causales generales no pueden entenderse independientemente de relaciones causales singulares. Esto la lleva naturalmente a tener que introducir algún tipo de análisis de enunciados causales singulares. En estos últimos trabajos, Cartwright no sólo acepta la idea de Sober y Salmon (entre otros) de que la ciencia necesita una noción separada de ley causal sino que considera que hay que ir más allá y postular el concepto de «capacidad».

Eells en 1990 ofrece una solución alternativa a las dificultades del proyecto de una teoría probabilista de la causalidad. Eells sostiene que es necesario introducir dos teorías de la causalidad totalmente separadas, una para la causalidad singular y otra para la causalidad general. Según Eells la comparación de probabilidades condicionales (cómo en la teoría de Suppes) es apropiada para entender la causalidad general pero no lo es para entender la causalidad singular. Para desarrollar una teoría de la causalidad singular —nos dice Eells— debemos ver cómo cambia la probabilidad del suceso que se toma como el efecto en la vecindad temporal del suceso que se considera la causa (y en los tiempos intermedios entre el efecto y la causa). Intuitivamente, Eells trata de poner a vivir bajo el mismo techo a las teorías de Cartwright y de Salmon.

V. CONCLUSION

Es claro que el objetivo de Hume de dar una definición del concepto de causa sin suposiciones metafísicas ha quedado atrás. Las teorías contemporáneas de la causalidad incorporan cada vez más abiertamente suposiciones metafísicas. No veo nada de malo en esto. Hume simplemente estaba equivocado cuando pensó que se podía definir la causalidad sin recurrir a supuestos metafísicos. Sospecho sin embargo que detrás de todas estas teorías de la causalidad probabilista sigue latiendo el ideal positivista de que las probabilidades son «transparentes» epistémicamente, esto es, que sólo sirven como instrumentos intermediarios epistémicamente neutrales entre nuestra experiencia y un mundo objetivo. Lo que parece ser necesario para resolver el problema de fondo es la introducción de estructuras conceptuales que nos separen los diferentes problemas que parecen estarse confundiendo en estas teorías y que nos permitan reconstruir una noción de objetividad más acorde con las teorías científicas de finales del siglo XX. En particular, el problema de encontrar las causas a partir de las probabilidades, esto es, el problema de decidir hasta qué punto las probabilidades son un instrumento confiable para detectar las causas de sucesos, debe distinguirse del problema de definir una de las ideas regulativas cruciales de todo análisis de la objetividad científica, el concepto de causa.

BIBLIOGRAFIA

- Byrne, E. (1968), *Probability and Opinion: A Study in the Medieval Presuppositions of Post-Medieval Theories of Probability*, Hague, 1968.
- Cartwright, N. (1983), *How the Laws of Physics Lie*, Oxford University Press, Oxford-New York.
- Cartwright, N. (1989), *Nature's Capacities and their Measurement*, Clarendon Press, Oxford.
- Coffa, A. (1974), «Hempel's Ambiguity», *Synthese* 28, 141-163.
- Ducasse, C. J. (1924), *Causation and the Types of Necessity*, University of Washington Press. Reimpreso en 1969 por Dover Publications Inc.

- Eells, E. (1991), *Probabilistic Causality*, Cambridge University Press, New York.
- Fetzer, J. (ed.) (1988), *Probability and Causality*, Reidel Pub. Company, Dordrecht/Boston.
- Hempel, C. y Oppenheim, P. (1947), «Studies in the Logic of Explanation»: *Philosophy of Science* 15, 135-75 (reimpreso en Hempel, 1965).
- Hempel, C. (1965), *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*, The Free Press, New York.
- Kitcher, Ph. y Salmon, W. (1989), *Scientific Explanation*, Minnesota Studies in the Philosophy of Science vol. XII, University of Minnesota Press, Minneapolis.
- Laplace, P. S. (1774), «Mémoire sur la probabilité des causes par les événements» en *Oeuvres complètes*. Académie de Sciences, 14 vols., París, 1878-1912.
- Laplace, P. S. (1814), «Essai philosophique sur les probabilités», en *Oeuvres complètes*. V.e. A. Besio y J. Banfi, Espasa-Calpe, Buenos Aires, 1947.
- Lewis, D. (1986), *Philosophical Papers*, vol. II, OUP, Oxford.
- Mackie, J. L. (1974), *The Cement of the Universe*, OUP, Oxford.
- Salmon, W. (1984), *Scientific Explanation and the Causal Structure of the World*, Princeton University Press, Princeton.
- Sober, E. (1984), *The Nature of Selection*, The MIT Press, Cambridge, Massachusetts.
- Suppes, P. (1970), *A Probabilistic Theory of Causality*, N. Holland, Amsterdam.
- Suppes, P. (1984), *Probabilistic Metaphysics*, Basil Blackwell, New York.
- Woodward (1986), «Are Singular Explanations Implicit Covering Law Explanations?»: *Canadian Journal of Philosophy* 16, 253-280.