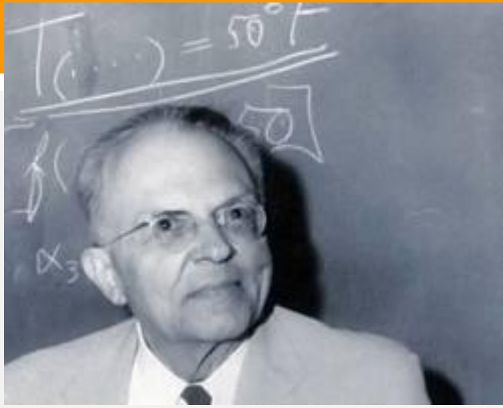
Lógica inductiva (basada en la teoría lógica de la probabilidad)



Rudolf Carnap (1891-1970)

(De pie y descubran sus cabezas, por favor)

RSE: Regla de inducción por simple enumeración para SIL (Lógica inductiva Científica)

RSE: Sean F y G propiedades cualesquiera, la *probabilidad inductiva* de que el próximo objeto observado que sea F, sea también G, aumenta con el número de instancias ya observadas de F que son G, asumiendo que no hay contrainstancias (un F que no es G) y que sentencias de la forma ‘x es F’ y ‘x es G’ son lógicamente independientes.

RSE alternativa:
RSE-C (para Lógica Contrainductiva).

RSE-C: La *probabilidad inductiva* de que el próximo F observado sea G disminuye a medida que el número de instancias observadas en las cuales los F hayan sido G aumenta (asumiendo que no hay contrainstancias (un F que no es G) y que sentencias de la forma ‘x es F’ y ‘x es G’ son lógicamente independientes).

RSE alternativa:
RSE-R (para Lógica «Inductiva» Randómica).

RSE-R: La *probabilidad inductiva* de que el próximo F observado sea G no depende del número de instancias observadas en las cuales los F hayan sido G (asumiendo que no hay contrainstancias (un F que no es G) y que sentencias de la forma ‘x es F’ y ‘x es G’ son lógicamente independientes).

(p.181)

Lenguaje

Partimos de un conjunto de objetos o individuos y un número finito de propiedades básicas que los objetos puede o no poseer.

El dominio D es el conjunto de objetos y propiedades.

El lenguaje formal para el dominio tendrá:

- nombres propios, un nombre para cada objeto y
- letras de predicado, cada una de las cuales expresa una propiedad básica.

- Las sentencias atómicas son del tipo:

Fa que significa ‘a es F’.

(una sentencia atómica siempre adscribe una propiedad a un individuo).

- El lenguaje contiene además las conectivas lógicas proposicionales: $\vee$, $\wedge$, $\neg$.

- Las propiedades son lógicamente independientes.

Definición. Dos propiedades F y G son *lógicamente independientes* si proposiciones de la forma ‘Fx’ y ‘Gx’ son lógicamente independientes.

(para cualquier x, son posibles los cuatro juegos de valores de verdad de Fx y Gx).

Q-propiedades

Definición. Una ***Q-propiedad*** de un dominio es una propiedad compleja expresada por una conjunción que contiene exactamente una ocurrencia (negada o no) de cada letra de predicado del lenguaje.

Q-propiedades de D

En lo que sigue supongamos D ={a,b, F,G}

Ejemplo: En D hay 4 Q-propiedades

$Fx \wedge Gx$ $\qquad Fx \wedge \neg Gx$ $\qquad \neg Fx \wedge Gx$ $\qquad \neg Fx \wedge \neg Gx$

Observaciones:

- Cada objeto tendrá exactamente una Q-propiedad.
- Si D tiene n propiedades básicas, entonces D tiene 2^n Q-propiedades.

Descripciones de estado

Definición. Una ***descripción de estado*** de un lenguaje es una conjunción que especifica una Q-propiedad para **cada** objeto del dominio correspondiente.

Si el dominio tiene n propiedades básicas y m objetos entonces hay 2^{mn} descripciones de estado del lenguaje.

L_D : 16 descripciones de estado.

Ejemplos:

$Fa \wedge \neg Ga \wedge Fb \wedge Gb$

$Fa \wedge Ga \wedge Fb \wedge \neg Gb$

Cada conjunción de una descripción de estado es lógicamente independiente de las otras y no es lógicamente falsa.

Definición. Una ***lógica inductiva (probabilista)*** para un dominio y un lenguaje es un conjunto de reglas que determina un único valor de probabilidad para cada par de sentencias <p,q> del lenguaje, esto es, un número $x \in \mathbb{R}$ tal que $Pr(p/q)=x$.

Dada una proposición p del lenguaje, su probabilidad incondicional P(p) se puede calcular a partir de una asignación de probabilidades incondicionales a las descripciones de estado del lenguaje. Además, la probabilidad condicional de una par de proposiciones, P(p/q) es calculable a partir de las probabilidades incondicionales de cada una.

(Meta)Teorema 1. Toda asignación de valores a una descripción de estado de un lenguaje que sume 1, determina una lógica inductiva (en sentido muy amplio) para el dominio del lenguaje.

Probabilidad / probabilidad condicional

$P(p) =_{def.} P(p/q)$ donde □p.

Esto es requerido desde una filosofía lógica de la probabilidad.

Veamos que en este marco tenemos la posibilidad de crear lógicas que respondan a :

SIL: la *probabilidad inductiva* de que el próximo objeto observado que sea F, sea también G, aumenta con el número de instancias ya observadas de F que son G, asumiendo que no hay contrainstancias.

CIL: la *probabilidad inductiva* de que el próximo objeto observado que sea F, sea también G, disminuye con el número de instancias ya observadas de F que son G, asumiendo que no hay contrainstancias.

RIL: la *probabilidad inductiva* de que el próximo objeto observado que sea F, sea también G, es independiente del número de instancias ya observadas de F que son G.

Veamos entonces que asignando probabilidades a las 16 descripciones de estado (de forma que sumen 1) obtenemos lógicas inductivas que respetan RSE, RSE-C y RSE-R:

1. $Fa \wedge Ga \wedge Fb \wedge Gb$	9. $Fa \wedge Ga \wedge \neg Fb \wedge \neg Gb$
2. $Fa \wedge \neg Ga \wedge Fb \wedge \neg Gb$	10. $\neg Fa \wedge \neg Ga \wedge Fb \wedge Gb$
3. $\neg Fa \wedge Ga \wedge \neg Fb \wedge \neg Gb$	11. $\neg Fa \wedge \neg Ga \wedge \neg Fb \wedge Gb$
4. $\neg Fa \wedge \neg Ga \wedge \neg Fb \wedge \neg Gb$	12. $Fa \wedge \neg Ga \wedge \neg Fb \wedge Gb$
5. $Fa \wedge Ga \wedge Fb \wedge \neg Gb$	13. $Fa \wedge \neg Ga \wedge \neg Fb \wedge \neg Gb$
6. $Fa \wedge \neg Ga \wedge Fb \wedge Gb$	14. $\neg Fa \wedge \neg Ga \wedge Fb \wedge \neg Gb$
7. $Fa \wedge Ga \wedge \neg Fb \wedge Gb$	15. $\neg Fa \wedge Ga \wedge \neg Fb \wedge \neg Gb$
8. $\neg Fa \wedge Ga \wedge Fb \wedge Gb$	16. $\neg Fa \wedge \neg Ga \wedge \neg Fb \wedge Gb$

Descripciones isomorfas

Definición. Dos descripciones de estado son isomorfas si y solo si una se puede transformar en la otra por medio de una sucesión de intercambios de sus nombres propios, intercambiando todas las ocurrencias de un nombre propio por ocurrencias del otro.

Ejemplo:

Fa ∧ ¬Ga ∧ ¬Fb ∧ Gb es isomorfa a **¬Fa ∧ Ga ∧ Fb ∧ ¬ Gb**

Observe que ambas descripciones contienen la información de que un individuo tiene F y no tiene G y otro individuo no tiene F y tiene G.

No son isomorfas a: **Fa ∧ Ga∧ ¬Fb ∧ Gb** (que contiene la información de que un individuo tiene F y G y otro no tiene F y tiene G)

Medida de isomorfismo

Definición. La *medida de isomorfismo de una descripción de estado d* (notación: I(d)) es el número de descripciones isomorfas a d, incluyendo también a d.

Ejemplo:

Supongamos el dominio U={a,b,F,G,H}

- $Fa \wedge Ga \wedge Ha \wedge Fb \wedge Gb \wedge Hb$ tiene medida de isomorfismo 1.
- $\neg Fa \wedge Ga \wedge Ha \wedge Fb \wedge Gb \wedge Hb$ tiene medida de isomorfismo 2.

Descripción de estructura

Definición. Una *descripción de estructura* es una disyunción de todas las descripciones isomorfas a una dada. (Si una descripción de estado es isomorfa solamente a sí misma, es también una descripción de estructura).

Ejemplo en U:

$(\neg Fa \wedge Ga \wedge Ha \wedge Fb \wedge Gb \wedge Hb) \vee (Fa \wedge Ga \wedge Ha \wedge \neg Fb \wedge Gb \wedge Hb)$

Lógica inductiva científica (SIL).

Axiomas de probabilidad + PSL (Principio de lógica científica):

Para cada descripción de estado d,

$P^{S}(d) = 1 / (l(d) \times g)$

donde g es el número de descripciones de estructura.

Lógica contrainductiva (CIL).

Axiomas del cálculo + PCL (Principio de lógica contrainductiva):

Para cada descripción de estado d,

$P^{C}(d) = I(d) / g_c$

donde g_c es la suma de las medidas de isomorfismo de todas las descripciones de estado.

Lógica inductiva randómica (RIL).

Axiomas del cálculo + PRL (Principio de lógica randómica):

Para cada descripción de estado d,

$P^R (d) = 1 / g_R$

donde g_R es el número de descripciones de estado.

SIL (científica)	$P^S (d) = 1 / (I(d) \times g)$ g es el núm. de descripciones de estructura.
CIL (contrainductiva)	$P^C (d) = I(d) / g_c$ g_c es la suma de las medidas de isomorfismo de todas las descripciones de estado.
RIL (Randómica)	$P^R (d) = 1 / g_R$ g_R es el número de descripciones de estado.

Descripción de Estado d.	Medida de Isomorf. I(d)	P(d) en SIL	P(d) en CIL	P(d) en RIL
1. $Fa \wedge Ga \wedge Fb \wedge Gb$	1	2/20	1/28	1/16
2. $Fa \wedge \neg Ga \wedge Fb \wedge \neg Gb$	1	2/20	1/28	1/16
3. $\neg Fa \wedge Ga \wedge \neg Fb \wedge \neg Gb$	1	2/20	1/28	1/16
4. $\neg Fa \wedge \neg Ga \wedge \neg Fb \wedge \neg Gb$	1	2/20	1/28	1/16
5. $Fa \wedge Ga \wedge Fb \wedge \neg Gb$	2	1/20	2/28	1/16
6. $Fa \wedge \neg Ga \wedge Fb \wedge Gb$	2	1/20	2/28	1/16
7. $Fa \wedge Ga \wedge \neg Fb \wedge Gb$	2	1/20	2/28	1/16
8. $\neg Fa \wedge Ga \wedge Fb \wedge Gb$	2	1/20	2/28	1/16

Descripción de Estado d.	Medida de Isomorf. I(d)	P(d) en SIL	P(d) en CIL	P(d) en RIL
9. $Fa \wedge Ga \wedge \neg Fb \wedge \neg Gb$	2	1/20	2/28	1/16
10. $\neg Fa \wedge \neg Ga \wedge Fb \wedge Gb$	2	1/20	2/28	1/16
11. $\neg Fa \wedge \neg Ga \wedge \neg Fb \wedge Gb$	2	1/20	2/28	1/16
12. $Fa \wedge \neg Ga \wedge \neg Fb \wedge Gb$	2	1/20	2/28	1/16
13. $Fa \wedge \neg Ga \wedge \neg Fb \wedge \neg Gb$	2	1/20	2/28	1/16
14. $\neg Fa \wedge \neg Ga \wedge Fb \wedge \neg Gb$	2	1/20	2/28	1/16
15. $\neg Fa \wedge Ga \wedge \neg Fb \wedge \neg Gb$	2	1/20	2/28	1/16
16. $\neg Fa \wedge \neg Ga \wedge \neg Fb \wedge Gb$	2	1/20	2/28	1/16

$$P^S(Gb\,/Fb) = \frac{P^S\,(Fb \wedge Gb)}{Pr^S\,(Fb)} = \frac{1/10 + 3(1/20)}{2(1/10) + 6(1/20)} = 1/2$$

$$P^S(Gb\,/\,Fa \wedge Ga \wedge Fb) = 2/3$$

$$\mathbf{P^S(Gb\,/\,Fa \wedge Ga \wedge Fb) > P^S(Gb\,/Fb)}$$

$$P^C(Gb\,/Fb) = \frac{P^C\,(Fb \wedge Gb)}{P^C\,(Fb)} = \frac{1/28 + 3(2/20)}{2(1/28) + 6(2/28)} = 1/2$$

$$P^C(Gb/Fa \wedge Ga \wedge Fb) = \frac{P^C(Fa \wedge Ga \wedge Fb \wedge Gb)}{P^C\,(Fa \wedge Ga \wedge Fb)} = \frac{1/28}{1/28 + 2/28} = 1/3$$

$$\mathbf{P^C(Gb\,/\,Fa \wedge Ga \wedge Fb) < P^C(Gb\,/Fb)}$$

$$P^R(Gb\,/\,Fb) = \frac{4(1/16)}{8(1/16)} = 1/2$$

$$P^R(Gb\,/\,Fa \wedge Ga \wedge Fb) = \frac{1/16}{2/\,16} = 1/2$$

$$\mathbf{P^R(Gb\,/\,Fa \wedge Ga \wedge Fb) = P^R(Gb\,/Fb)}$$

Carnap trabajó durante mucho tiempo en lógicas inductivas. Lo que acabamos de presentar es una adaptación de uno de sus sistemas más simples.

Veremos ahora las características principales de su sistema más elaborado que fue publicado póstumamente.

A Carnap no le interesaban todas las lógicas «inductivas» sino solo las que se comportan como SIL, dado que solo estas pueden soportar el proyecto empirista y justificar la inducción científica.

Además, al llevar adelante su proyecto Carnap debe tener en cuenta el devastador argumento GRUE de Goodman, por lo que necesita hacer algunas consideraciones semánticas.

Construcción del lenguaje L

Se considera un conjunto de objetos de los que se puede hablar en L, a los que se llama individuos. Se denotan con constantes individuales a_1,a_2, …

Constantes individuales diferentes denotan individuos diferentes.

$a_i = a_j$ es una contradicción si i es distinto de j.

Modalidades

Una modalidad es una clase de propiedades. Por ejemplo:

- Color (rojo, negro, amarillo, etc.)
- Forma (esférico, cúbico, etc.)
- Sustancia (de hierro, de piedra, etc.)
- Edad en años (0 , 1, etc…)

Familia de propiedades –predicados primitivos

Una familia de propiedades es un conjunto de propiedades que pertenecen a **una** modalidad, son excluyentes y exhaustivas.

Ejemplo:

{rojo, violeta, azul, verde, *otro color*}

{amarillo, no amarillo}

{de menos de un año, de un año, de dos años,… }

Los predicados primitivos de L denotan los elementos **de una** familia de propiedades. Son denotados como F_1, F_2, …

Bloqueo del argumento GRUE

La forma en que Carnap ha definido los predicados primitivos impide la creación de un predicado que corresponda a GRUE porque ese concepto es una combinación de dos modalidades.

Definición. Una *sentencia atómica* es un predicado primitivo seguido de una constante de individuo.

Ejemplo: ‘F_4a_7’ significa ‘a_7 tiene la propiedad F_4’.

Definición. Una *muestra* es un conjunto finito de individuos.

Definición. Una descripción de una muestra es una conjunción de sentencias atómicas una para cada individuo de la muestra.

Ejemplo: $F_4a_7 \wedge F_1a_2$ es una descripción de la muestra $\{a_7, a_2\}$

El conjunto vacío se considera una muestra, y una descripción del conjunto vacío es cualquier sentencia analítica.

Axiomas

Carnap va a intentar formalizar una lógica verdaderamente inductiva. Para eso agregará a los axiomas de la probabilidad los siguientes. En lo que sigue, y dado que Carnap trabaja desde la teoría lógica de la probabilidad, P(A) significará P(A/T) donde T es analítica.

Axiomas

Sea S una descripción de muestra.

- **Regularidad**: P(S) >0.
- **Simetría:** P(S) no cambia si se permutan constantes individuales. Ejemplo: $P(F_2a_1 \wedge F_5a_2) = P(F_2a_2 \wedge F_5a_1)$
- **Relevancia instancial**: $P(F_ia_m/F_ia_n) > P(F_ia_m)$ para todo i, m, n.
- **Condición λ**: Si a no aparece en S, entonces $P(F_ia/S)$ depende solo del número de individuos que aparecen en S y del número de individuos que según S tienen F_i. Ejemplo: $P(F_1a_3 /F_1a_1 \wedge F_2a_2) = P(F_1a_3 /F_1a_1 \wedge F_3a_2)$.

Sobre los axiomas

P cumple **regularidad** si ningún F_i es infinitamente preciso.

La **simetría** implica que las constantes individuales no portan información acerca de qué propiedades tiene el individuo que denotan.

La **relevancia instancial** es absolutamente central: representa la posibilidad de aprender de la experiencia.

La **condición** λ es una suposición común: normalmente tomamos la frecuencia relativa de una propiedad de una muestra como lo que es relevante para predecir si otros individuos tendrán esa propiedad.

Teorema λγ (Carnap, publicado en 1980)

Si P satisface los axiomas de la probabilidad y regularidad, simetría, relevancia instancial y la condición λ, y k>2, entonces existe λ>0 y

$\gamma_1,\ldots,\gamma_k \in (0,1)$ tales que:

$$P(F_i a/S) = \frac{n_i + \lambda\gamma_i}{n+\lambda} \quad \text{para todo } i = 1,\ldots,k$$

- donde a es cualquier constante individual que no aparece en la descripción de muestra S.
- n es el número de constantes individuales en S.
- n_i es el número de individuos que según S tienen F_i.

Ejemplo: Supongamos $\gamma_1 = 1/4$ y $\lambda=2$ entonces

$$P(F_1a_4 / F_1a_1 \wedge F_2a_2 \wedge F_3a_3) = \frac{1+{}^{2}/_{4}}{3+2} = 3/10$$

Cuando k=2 el teorema requiere asumir más que los axiomas vistos.

¿Qué es γ_i?

$$P(F_i a/S) = \frac{n_i + \lambda\gamma_i}{n+\lambda}$$

Supongamos que no tenemos ninguna muestra. En ese caso, nuestra muestra es vacía, $n=n_i=0$, y como la descripción de $\emptyset$ es una sentencia analítica T, obtenemos

$$P(F_i a/T) = \frac{0 + \lambda\gamma_i}{0+\lambda} = \gamma_i = P(F_i a)$$

O sea, γ_i es la probabilidad a priori de que algo tenga F_i.

Laplace supuso que es ½. Keynes mostró que esto no siempre es así.

¿Qué propone Carnap sobre γ_i?

Su propuesta no es diferente de la de Keynes.

Sugiere:

- El espacio de atributos para Fi es el espacio lógico cuyos puntos son las propiedades más específicas de la modalidad relevante.
- γ_i es la proporción del espacio de atributos que ocupa F_i.

Ejemplo: Si F_1 significa «rojo» y el rojo ocupa 1/30 de la medida de un espacio cromático, podemos asignar γ_1=1/30 (asumiendo que el objeto es monocromático)

El significado de λ.

$$P(F_i a/S) = \frac{n_i + \lambda\gamma_i}{n+\lambda} = \left(\frac{n}{n+\lambda}\right)\frac{n_i}{n} + \left(\frac{\lambda}{n+\lambda}\right)\gamma_i$$

Entonces $P(F_i a/S)$ es una mezcla de un factor empírico, n_i/n y un factor lógico, γ_i .

Cuanto mayor es λ , más peso tiene el factor lógico y más lentamente se aprende de la experiencia.

Críticas (expresadas por Hájek)

La teoría de la probabilidad lógica intenta explicar la inferencia ampliativa, pero dado que hay arbitrariedad en la elección del lenguaje y en la determinación de λ -esto es, en la elección de la función de confirmación- uno se puede preguntar qué tanto se logra el objetivo.

Respuesta: La arbitrariedad no hace que un explicatum sea inadecuado. No hay criterio de no arbitrariedad para buenos explicata.

Críticas (expresadas por Hájek)

Fine dice que los axiomas agregados por Carnap no son verdades lógicas, y lo que es más grave no podemos imponer otras condiciones que parecen tan plausibles como las de Carnap sin caer en inconsistencia.

Respuesta: Los axiomas son parte de la definición de P y por tanto, analíticos en ese sentido. Por otro lado, se discute lo que Fine (1973) dice haber hallado como inconsistencia.

Críticas (expresadas por Hájek)

Una proposición universal en un universo infinito siempre recibe confirmación 0 sin importar la evidencia (finita). Muchos encuentran que esto es contraintuitivo porque las leyes de la naturaleza aparentemente son universales y se pueden confirmar con evidencia finita.

Respuesta: Hay formas conocidas de modificar los sistemas de Carnap para evitar este resultado. De modo que es una objeción al explicata particular, no al proyecto general de la lógica inductiva. Además, el explicata de Carnap es útil en conjuntos finitos.

Críticas (expresadas por Hájek)

Desde Goodman sabemos que la lógica inductiva debe ser sensible al significado de los predicados, lo que sugiere fuertemente que una aproximación puramente sintáctica como al de Carnap está condenada al fracaso.

Respuesta: En el sistema básico los predicados se organizan en familia que pertenecen a una misma modalidad, y el γi depende de la extensión de los predicados. Estos son requisitos semánticos, no sintácticos, de modo que Carnap no nos dio «una aproximación puramente sintáctica» y su lógica inductiva es sensible a los significados de los predicados.

Críticas (expresadas por Hájek)

Hallar un lenguaje canónico parece un sueño absurdo, al menos si uno quiere analizar la «probabilidad lógica» de cualquier argumento sea en ciencia sea en la vida cotidiana.

Respuesta: Carnap tenía como una de sus convicciones filosóficas centrales que no existe nada como un «lenguaje canónico». Tal vez lo que Hájek quiera decir es que ningún lenguaje formalizado puede expresar la evidencia total disponible en una situación real. Eso lleva a la próxima objeción.

Críticas (expresadas por Hájek)

El criterio de evidencia total va de la mano del positivismo y de una epistemología fundacionista según la cual existen «porciones» últimas y determinadas de la experiencia. Pero tal vez el aprendizaje no pase por proposiciones que expresan esas «porciones» de experiencia sino que involucre un cambio en las probabilidades subjetivas sobre una partición, sin que ninguna parte de la partición se vuelva cierta.

Críticas (expresadas por Hájek)

Respuesta: Para que E sea un buen explicatum de la evidencia de una persona lo único que se necesita es que E sea lo suficientemente probable como para ser tratada como cierta para los propósitos de la persona y que E no se infiera de otras proposiciones bajo consideración.

Esto no supone ninguna «porción última y determinada de la experiencia».

Pearl (1990):

La evidencia no proposicional se puede representar en la lógica inductiva carnapiana con una variable que se trata como una proposición.

Críticas (expresadas por Hájek)

Si las creencias deben estar basadas en las probabilidades lógicas, deben ser relativizadas a una proposición de evidencia E. Pero ¿qué será E? La recomendación de Carnap es que E debe ser la evidencia total de uno. Pero cuando se va más allá de ejemplos artificiales, no está claro cómo se define. Supongamos que he mirado el lanzamiento de una moneda que ha caído con la cara hacia arriba. ¿Es «La moneda cayó con la cara hacia arriba» mi evidencia total? En realidad he aprendido muchas otras cosas. Entonces mi evidencia total es la conjunción infinita de todas esas proposiciones, y eso no puede representarse en el sistema de Carnap. Que solo trabaja con proposiciones finitas.

Críticas (expresadas por Hájek)

Respuesta: Efectivamente, el concepto en lenguaje ordinario de «evidencia total» es vago. Se *explica* para una situación dada usando el lenguaje formal. La explicación es buena si captura los aspectos relevantes para el problema que nos interesa. Lo que citamos en argumentos reales siempre se puede expresar con una sentencia.

Críticas (expresadas por Hájek)

En el sistema de Carnap el grado de confirmación de una hipótesis depende del lenguaje en el que se expresa la hipótesis. Pero el progreso científico suele traer cambios en el lenguaje de la ciencia (por ejemplo, agregado de nuevos predicados). De manera que el progreso de la ciencia puede derrumbar cualquier teoría de la confiramción particular. Aquí la serpiente se muerde la cola, porque se supone que la probabilidad lógica podía explicar la confirmación de las teorías científicas.

Críticas (expresadas por Hájek)

Respuesta: Como los teoremas de la lógica inductiva son analíticos, esta no supone nada al respecto de la ciencia.

Y es falso que «el progreso de la ciencia» pueda cambiar P(H/E), como debería ser obvio. A la luz de nueva evidencia E’, se evaluará P(H/E’) y eso es todo.