Paradoja de Hempel

Algunas ideas básicas sobre la confirmación

(a) **Principio de relevancia empírica:** Un dato empírico es relevante para una hipótesis, si y solo si constituye una evidencia favorable o desfavorable de ella, o sea, si la confirma o la refuta.

(b) **Criterio de Nicod:** Todo individuo que satisfaga el antecedente y el consecuente de un condicional universal, lo confirma.

(c) **Condición de equivalencia:** Lo que confirma (o refuta) a un enunciado, confirma (o refuta) también a todo enunciado lógicamente equivalente el primero.

Hipótesis H: $\forall x(Cx \rightarrow Nx)$

Considerando (a) y (b):

1. $Ca \wedge Na$ confirma H
2. $Ca \wedge \neg Na$ refuta H
3. $\neg Ca \wedge Na$ es irrelevante para H
4. $\neg Ca \wedge \neg Na$ es irrelevante H

Hipótesis H': $\forall x(\neg Nx \rightarrow \neg Cx)$

Considerando (a) y (b):

1. $Ca \wedge Na$ irrelevante
2. $Ca \wedge \neg Na$ refuta H'
3. $\neg Ca \wedge Na$ es irrelevante para H'
4. $\neg Ca \wedge \neg Na$ confirma H'

	$\forall x(Cx \rightarrow Nx)$	$\forall x(\neg Nx \rightarrow \neg Cx)$
1. $Ca \wedge Na$	confirma	irrelevante
2. $Ca \wedge \neg Na$	refuta	refuta
3. $\neg Ca \wedge Na$	irrelevante	irrelevante
4. $\neg Ca \wedge \neg Na$	irrelevante	confirma

Esa tabla es insostenible si aceptamos el **principio de equivalencia**, y debe ser modificada de la siguiente manera.

	$\forall x(Cx \rightarrow Nx)$	$\forall x(\neg Nx \rightarrow \neg Cx)$
1. $Ca \wedge Na$	confirma	~~Irrelevante~~ confirma
2. $Ca \wedge \neg Na$	refuta	refuta
3. $\neg Ca \wedge Na$	irrelevante	irrelevante
4. $\neg Ca \wedge \neg Na$	~~Irrelevante~~ confirma	confirma

En esta situación deberíamos aceptar que la observación de un calzoncillo rosado confirma la hipótesis de que todos los cuervos son negros (así como la hipótesis de que todos los tréboles son de cuatro hojas).

Entonces, según lo anterior, las únicas observaciones irrelevantes para la hipótesis de que todos los cuervos son negros son las de objetos negros que no son cuervos.

Esto es ya bastante poco intuitivo, pero…

Hempel muestra la siguiente equivalencia:

$$\forall x(Cx \rightarrow Nx) \equiv \forall x((Cx \vee \neg Cx) \rightarrow (\neg Cx \vee Nx))$$

De aquí surge un resultado más extraño aun:

Cualquier individuo, excepto los cuervos no negros, confirma H

De aquí que no hay observaciones irrelevantes para la hipótesis H. Esto es lo que se llama *paradoja de Hempel* o *paradoja de los cuervos*.

Principio inverso de la confirmación

Se basa en las siguientes consideraciones matemáticas: Desde el punto de vista probabilista, decimos que un dato observacional «e» confirma una hipótesis «H» en presencia de un conocimiento de fondo «k» si

$$P(H/e \wedge k) > P(H/k)$$

Principio inverso de la confirmación

Ahora bien, tenemos:

$P(H \wedge e/k) = P(H/e \wedge k).P(e/k)$ (demostración en el pizarrón)

Análogamente, **$P(H \wedge e/k) = P(e/H \wedge k).P(H/k)$**

Por lo tanto,

$P(H/e \wedge k).P(e/k) = P(e/H \wedge k).P(H/k)$

Entonces tenemos que

$$\boxed{P(H/e \wedge k) > P(H/k) \text{ si y solo si } P(e/H \wedge k) > P(e/k)}$$

Datos observacionales:

- B1: cuervo negro ($Ca \wedge Na$)
- B2: cuervo no negro ($Ca \wedge \neg Na$)
- B3: no cuervo negro ($\neg Ca \wedge Na$)
- B4: no cuervo no negro ($\neg Ca \wedge \neg Na$)

P(ser cuervo) = x   P(ser negro)=y

Probabi-lidad de	B1 (Ca $\wedge$Na)	B2 (Ca$\wedge\neg$Na)	B3 ($\neg$Ca $\wedge$ Na)	B4 ($\neg$Ca$\wedge\neg$Na)
Dado k	**xy**	**x (1 - y)**	**(1 - x) y**	**(1-x)(1-y)**
Dado H$\wedge$k	**x**	**0**	**y – x**	**1-y**

El estudio de la tabla anterior muestra que:

- B1 confirma H, ya que **$xy < x$**
- B4 confirma H, ya que **$(1-x)(1-y) < 1-y$**
- B1 confirma H en mayor medida que B4 (explicación en el pizarrón).
- B2 refuta H (obvio)
- B3 refuta H, ya que **$(1-x)y > y-x$.**

El hecho de que B3 refute H representa un gravísimo problema: B3 es un no cuervo negro, podemos expresarlo como $(\neg Ca \wedge Na)$

Tenemos la equivalencia lógica:

$$\neg(\neg Ca \wedge Na) \equiv (Ca \vee \neg Na)$$

Entonces, si Mackie está en lo cierto y B3 es un refutador, **$(Ca \vee \neg Na)$** es equivalente a la negación de un refutador de H, y dado que no hay observaciones irrelevantes, la negación de un refutador debería ser un confirmador. Pero obviamente, un individuo a que sea cuervo y no negro verifica la fórmula **$(Ca \vee \neg Na)$,** pero ese individuo obviamente refuta que todo los cuervos sean negros.

Solución al problema de Mackie según Hooker y Stove

La aceptación de que la probabilidad de B3 dado H es y-x se basa en las premisas ocultas siguientes:

1) La probabilidad de ser negro dado que todos los cuervos son negros es igual a la probabilidad de ser negro.

2) La probabilidad de ser cuervo dado que todos los cuervos son negros es igual a la probabilidad de ser cuervo.

Estas premisas se pueden expresar así:

1) $P(N/H \wedge k) = P(N/k) = y$

2) $P(C/H \wedge k) = P(C/k) = x$

La solución al problema pasa por rechazar 1) y 2), o sea, aceptar que C y/o N son relevantes para H.

En particular, si queremos que

$$P(\neg\ Ca \wedge Na\ /\ H \wedge k) > P(\neg\ Ca \wedge Na/k)$$

Deberíamos tener que

$$\mathbf{P(Ca/H)<P(Ca)}$$

$$\mathbf{P(Na/H)>P(Na)}$$

Sin embargo, aceptar

$P(Ca/H)<P(Ca)$ y $P(Na/H)>P(Na)$

expone a una fuerte crítica: La observación de un cuervo refuta H.

Stove y Hooker contestan que la aceptación de ambas desigualdades lo que implica es, simplemente, que la hipótesis H, equivalente a $\forall x\ (\neg Cx \vee Nx)$, es confirmada por un objeto que sea no cuervo o negro, y en particular, por un no cuervo negro, lo que Mackie ha aceptado.