# BAYESIANISMO

(Subjetivo)

El bayesianismo subjetivo es simplemente el desarrollo de la filosofía subjetiva de la probabilidad, con el agregado normativo de que los grados de creencia deben condicionalizarse según el Teorema de Bayes.

La respuesta bayesiana al desafío de Hume consiste en decir que Hume tiene razón en lo siguiente: *Dadas las premisas de un argumento inductivo, cualquier conclusión es igualmente razonable.* (Esto es así porque las probabilidades son absolutamente subjetivas).

¿No es esto una rendición frente a Hume?

Desde cierto punto de vista, lo es. Pero el bayesiano dice que ese no es el punto importante. Durante nuestra vida, tenemos diversos grados de creencia en ciertas proposiciones. La cuestión no es si esas creencias son “racionales”. Según el bayesiano la cuestión es si, a la luz de la evidencia, **modificamos racionalmente** esos grados de creencia.

De modo que para el bayesiano, lo que un agente racional necesita en un mundo cambiante, es una forma racional de modificar sus creencias, no un seguro de que éstas son racionales en cada circunstancia. Y esa forma existe: Si uno tiene un grado de creencia **P(h)** en la proposición **h**, y aprende que la proposición **e** es verdadera, su nuevo grado de creencia en **h** debe ser **P(h/e).**

Diremos que e confirma o apoya h si $P(h/e) > P(h)$

Diremos que e disconfirma h si $P(h/e) < P(h)$

Diremos que e es neutral con respecto a h si $P(h/e) = P(h)$

En la evaluación de esa última probabilidad la herramienta fundamental es el teorema de Bayes, que en su versión más simple dice:

$$P(h/e) = \frac{P(h)P(e/h)}{P(e)}$$

Se observa que aparecen probabilidades “a priori”, P(h) y P(e), que no tienen por qué ser iguales –ni siquiera parecidas- para todos los agentes racionales.

Sin embargo, como mostramos cuando hablamos del exclusivismo subjetivista de de Finetti, el bayesiano dispone de teoremas que muestran que, aparentemente, bajo condiciones plausibles, todos los agentes racionales convergerán a un mismo grado de creencia si se les presenta la misma evidencia.

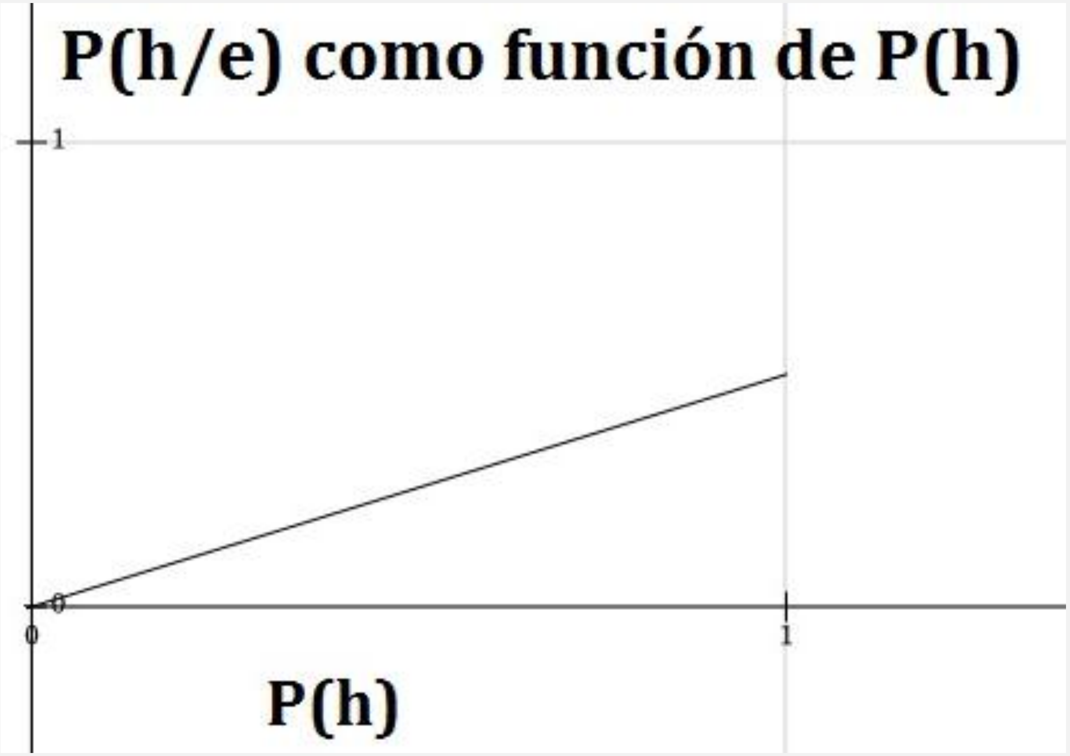
# Diversas formas del teorema de Bayes.

Primera forma:

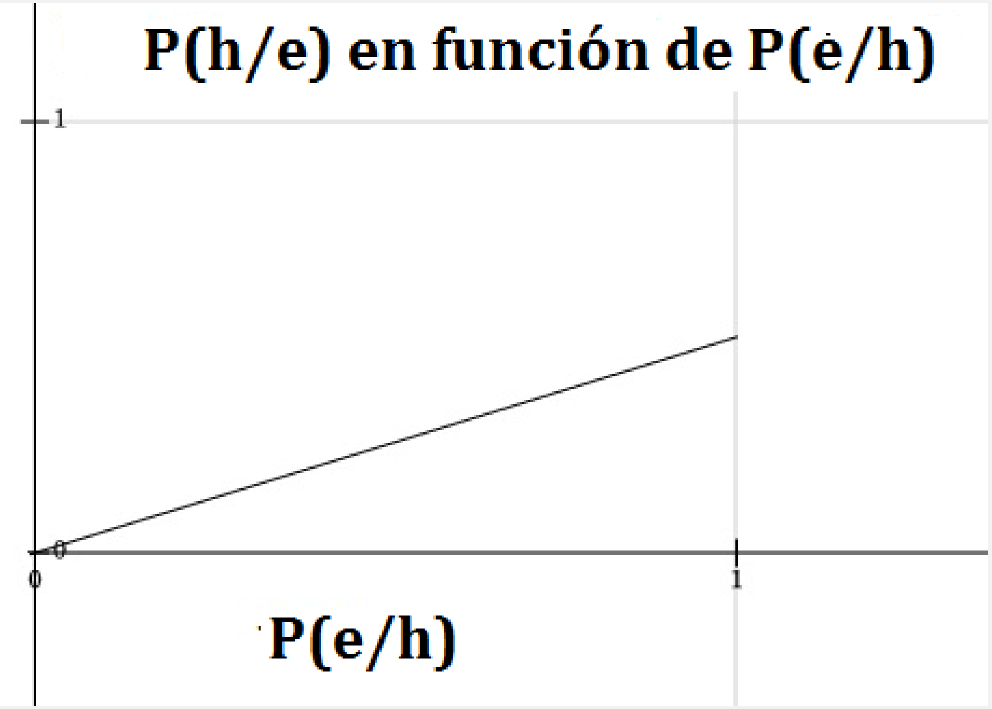
$$P(h/e) = \frac{P(h)P(e/h)}{P(e)}$$

P(h/e) es la probabilidad a posteriori de h, P(e/h) se llama «verosimilitud» de la hipótesis, P(e) es la probabilidad de los datos.

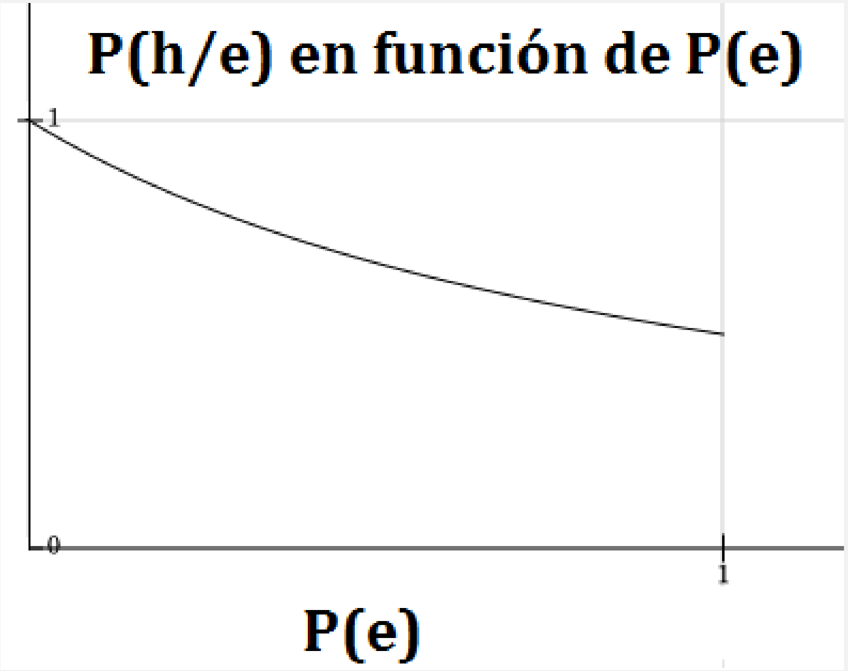
La probabilidad a posteriori de h (dejando todo lo demás igual) crece con la probabilidad a priori de h.



La probabilidad a posteriori de h (dejando todo lo demás igual) crece con P(e/h).



La probabilidad a posteriori de h (dejando todo lo demás igual) decrece con P(e).




Observación: mirando la forma 1 del teorema de Bayes parecería que P(h/e) $\rightarrow \infty$ cuando P(e) $\rightarrow$0. Esto no es así, como se hará evidente al ver la segunda formulación del teorema.

Segunda forma:

Si $P(h_1 \vee \ldots \vee h_n) = 1$ y $(h_i \rightarrow \neg h_j)$ para todo $i \neq j$ y $P(h_i)$, $P(e)>0$ entonces:

$$\boldsymbol{P(h_k/e) = \frac{P(e/h_k)P(h_k)}{P(e/h_1)P(h_1) + \cdots + P(e/h_n)P(h_n)} = \frac{P(e/h_k)P(h_k)}{\sum P(e/h_i)P(h_i)}}$$

Esta forma se demuestra como está en las notas usando el teorema de probabilidad total.

Utilizando la segunda forma sobre la partición {h,¬h} y dividiendo numerador y denominador del miembro derecho entre P(e/h) se obtiene la

Tercera forma:

$$P(h/e) = \frac{P(h)}{P(h) + \frac{P(e/\neg h)}{P(e/h)} P(\neg h)}$$

Esta forma es tal vez la más importante desde el punto de vista de la inferencia inductiva.

Si la expresamos así:

$$P(h/e) = \frac{P(h)}{P(h) + \frac{P(e/\neg h)}{P(e/h)}(1 - P(h))}$$

observamos que todo el impacto evidencial de e está en el factor

$$\frac{P(e/\neg h)}{P(e/h)}$$

que se llama *tasa de verosimilitud* o *factor de Bayes*.

Tenemos que $P(h/e) = f\left(P(h), \frac{P(e/\neg h)}{P(e/h)}\right)$

La probabilidad a posteriori de h es una función creciente de la probabilidad a priori de h y decreciente de la tasa de verosimilitud. O sea, para un valor dado de la tasa de verosimilitud, la probabilidad a posteriori de h se incrementa con su probabilidad a priori, y para un valor dado de la probabilidad a priori de h la probabilidad a posteriori de h es mayor cuanto menos probable es e relativamente a ¬h que a h.

## Bayesianismo sobre teorías deterministas

Si h implica e, o sea, si $\vDash(h \rightarrow e)$ entonces

a) $P(h \wedge e) = P(e/h)P(h) = P(h)$

b) $P(h \wedge \neg e) = P(\neg e/h)P(h) = 0$

c) Si además ni h ni e tienen probabilidad nula, se deduce que $P(e/h) = 1$ y por lo tanto $P(h/\neg e) = 0$.

La interpretación bayesiana de esto es que h recibe la máxima desconfirmación cuando es refutada. Además una vez que una teoría ha sido refutada, ninguna evidencia posterior podrá confirmarla, a menos que se revoque la evidencia que refuta. Porque si e' es otra observación consistente con e y si $P(h/\neg e) = 0$, entonces

$P(h/\neg e \wedge e') = 0$

Cuando se muestra que una consecuencia lógica de una teoría es verdadera, esto se considera una confirmación de la teoría. Esto tiene una explicación simple en el bayesianismo: Supongamos que ni h ni e tienen probabilidades extremas y que $\vDash(h \rightarrow e)$. Entonces P(e/h) =1 por lo que

$$P(h/e) = \frac{P(h)}{P(e)}$$

Y dado que $0<P(e)<1$ y $P(h)>0$, se tiene que $P(h/e) > P(h)$.

Otra característica que el bayesianismo puede explicar es que sucesivas confirmaciones por consecuencias lógicas eventualmente disminuyen su fuerza confirmatoria.

Antes de ver el desarrollo demostremos que si $\models(h \to e)$, $P(h)=P(h \wedge e)$:

$$P(h \wedge e) = P(h)P(e/h)=P(h).1 =P(h)$$

Sea $e_1, e_2, \ldots, e_n$ una sucesión de consecuencias lógicas de h.

$P(h/e_1\wedge \ldots\wedge e_{n-1}) = P(h\wedge e_n/e_1\wedge\ldots\wedge e_{n-1}) =$

$$= \frac{P(h\wedge e_1\wedge\cdots\wedge e_{n-1}\wedge e_n)}{P(e_1\wedge\cdots\wedge e_{n-1})} = \frac{P(h\wedge e_1\wedge\cdots\wedge e_{n-1}\wedge e_n)}{P(e_1\wedge\cdots\wedge e_{n-1})} \cdot \frac{P(e_1\wedge\cdots\wedge e_{n-1}\wedge e_n)}{P(e_1\wedge\cdots\wedge e_{n-1}\wedge e_n)} =$$

$= P(h/e_1\wedge\ldots\wedge e_n).\ P(e_n/e_1\wedge\ldots\wedge e_{n-1})$

Por lo que mostramos antes, como h implica todos los e, $P(h/e_1\wedge\ldots\wedge e_n) \geq P(h/e_1\wedge\ldots\wedge e_{n-1})$. Entonces $P(h/e_1\wedge\ldots\wedge e_n)$ es una sucesión monótona creciente acotada y por lo tanto tiene límite L. Entonces

$L= \lim P(h/e_1\wedge\ldots\wedge e_n) = \lim P(h/e_1\wedge\ldots\wedge e_{n-1})$.

Se deduce que $\lim P(h/e_1\wedge\ldots\wedge e_n) - P(h/e_1\wedge\ldots\wedge e_{n-1}) = 0$ o sea, que el incremento en el apoyo evidencial tiende a 0.

(Se deduce además que $\lim P(e_n/e_1\wedge\ldots\wedge e_{n-1}) =1$ o sea, que las ulteriores consecuencias lógicas están fuertemente confirmadas por las anteriores, después de un cierto punto). Este resultado no nos dice a partir de qué punto deja de ser valioso acumular evidencia a través de consecuencias lógicas de h.

Otra característica de la confirmación por consecuencias lógicas que el bayesianismo puede explicar es que en muchas circunstancias hay categorías específicas de consecuencias que tienen una capacidad de confirmación propia y limitada. Reconocemos que, sin importar qué tan frecuentemente se repita un experimento, sus resultados solo pueden confirmar una teoría general en una medida limitada y que cuando se ha agotado la capacidad del experimento para generar evidencia confirmatoria a través de la repetición, se busca apoyo de otros experimentos cuyos resultados son predichos por otras partes de la teoría.

Sea h una hipótesis general y sea $h_{rest}$ una restricción de h. Por ejemplo, una restricción de la teoría gravitatoria es la hipótesis de que los cuerpos caen con aceleración constante en las cercanías de la tierra. **Siempre h implica $h_{rest}$** y por esto $P(h) \leq P(h_{rest})$. (Si $h_{rest}$ es mucho menos especulativa que h su probabilidad frecuentemente será mucho mayor).

Consideremos una sucesión de consecuencias $e_1, e_2, \ldots, e_n$ de h que son también consecuencias de $h_{rest}$. Si esas consecuencias se verifican, tanto h como $h_{rest}$ resultan confirmadas, y sus probabilidades posteriores son dadas por el teorema de Bayes.

$P(h/ e_1 \wedge e_2 \wedge \ldots \wedge e_n) = \frac{P(h)}{P(e_1 \wedge e_2 \wedge \cdots \wedge e_n)}$

$P(h_{rest}/ e_1 \wedge e_2 \wedge \ldots \wedge e_n) = \frac{P(h_{rest})}{P(e_1 \wedge e_2 \wedge \cdots \wedge e_n)}$

Igualando el factor que aparece en ambas se obtiene

$P(h/ e_1 \wedge e_2 \wedge \ldots \wedge e_n) = \frac{P(h)}{P(h_{rest})} P(h_{rest}/e_1 \wedge e_2 \wedge \ldots \wedge e_n)$

Como el valor máximo de $P(h_{rest}/e_1 \wedge e_2 \wedge \ldots \wedge e_n)$ es 1, se tiene que $P(h/e_1 \wedge e_2 \wedge \ldots \wedge e_n)$ nunca puede superar al valor de $\frac{P(h)}{P(h_{rest})}$. Por lo tanto la probabilidad a priori de $h_{rest}$ determina una cota a qué tanto la evidencia implicada por ella puede confirmar h.

# La probabilidad de la evidencia: Macbeth

# La probabilidad de la evidencia: Macbeth

MACBETH

Hablad si sabéis. ¿Quiénes sois?

BRUJA Iª

¡Salud a ti, Macbeth, Barón de Glamis!

BRUJA 2ª

¡Salud a ti, Macbeth, Barón de Cawdor!

BRUJA 3ª

¡Salud a ti, Macbeth, que serás rey!

¡Esperad, imperfectas hablantes, decid más!
Por la muerte de Sinel soy Barón de Glamis,
mas, ¿cómo de Cawdor? El Barón de Cawdor vive
y continúa vigoroso; y ser rey
traspasa el umbral de lo creíble,
tanto como ser Cawdor. Decid de dónde
os ha llegado tan extraña novedad o por qué
cortáis nuestro paso en este yermo
con proféticos saludos.

ANGUS

Venimos a darte las gracias en nombre del rey
y a conducirte a su presencia,
no a recompensarte.

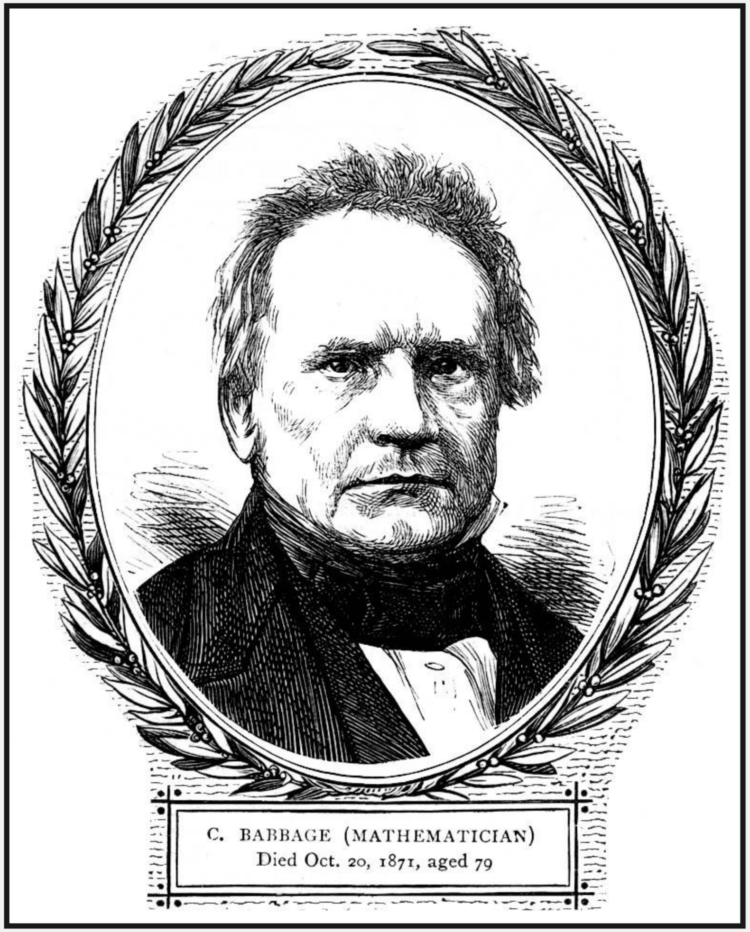
ROSS

Y, a cuenta de un honor aún más grande,
me ha mandado que te llame Barón de Cawdor.
¡Salud, nobilísimo barón, con este título,
pues tuyo es!

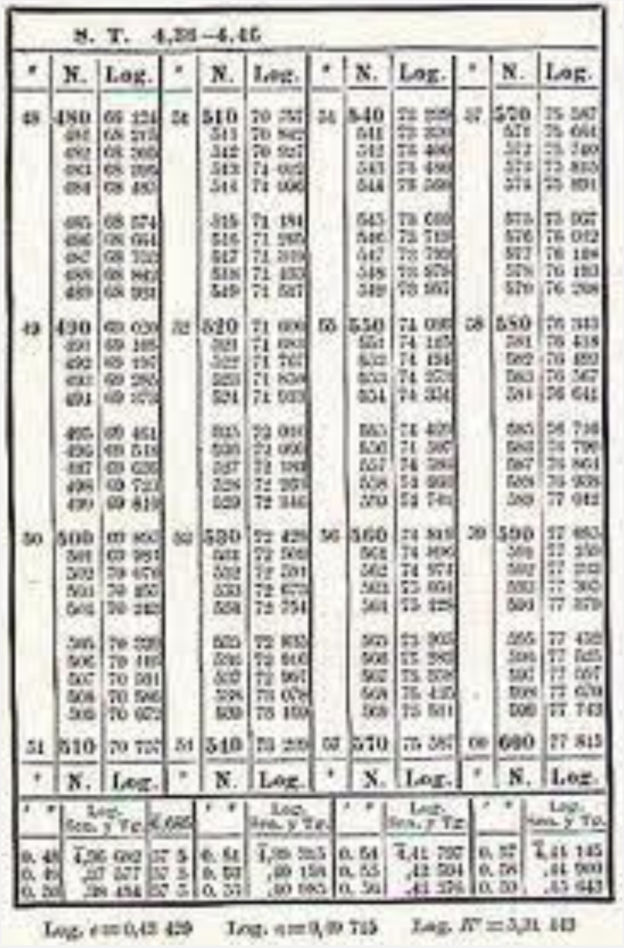
MACBETH [aparte]

Ya se han dicho dos verdades,
felices preludios a la escena gloriosa
del fin soberano. - Gracias, señores.
-[Aparte] Esta incitación sobrenatural
no puede ser mala, no puede ser buena.
Si es mala, ¿por qué me ha dado promesa de éxito
empezando con una verdad? Soy Barón de Cawdor.
Si es buena, ¿por qué cedo a esa tentación
cuya hórrida imagen me eriza el cabello
y me bate el firme corazón contra los huesos
violando las leyes naturales? Es menor
un peligro real que un horror imaginario.
La idea del crimen, que no es sino quimera,
a tal punto sacude mi entera humanidad
que la acción se ahoga en conjeturas
y sólo es lo que no es.

# Babbage (1791-1871) y las tablas de logaritmos: la probabilidad de la evidencia.



C. BABBAGE (MATHEMATICIAN)
Died Oct. 20, 1871, aged 79



S. T. 4,38—4,45

Log. e = 0,43 429   Log. π = 0,49 715   Log. R'' = 5,31 443

Babbage examinó varias tablas de logaritmos publicadas en los dos siglos anteriores a él. (Esas tablas habían sido confeccionadas calculando a mano). Encontró seis errores que se repetían en todas menos en dos.

¿Qué habrá concluido?

La tercera formulación del teorema de Bayes explica perfectamente la conclusión de Macbeth y la de Babbage:

$$\frac{P(h/e)}{P(h)} = \frac{P(e/h)}{P(e)} = \frac{1}{P(h) + P(\neg h)\frac{P(e/\neg h)}{P(e/h)}}$$

$\frac{P(e/\neg h)}{P(e/h)}$, el factor de Bayes es lo que carga con la fuerza confirmatoria de la evidencia. Cuanto menor es este factor, o sea, cuanto más probable es la evidencia si la hipótesis es verdadera que si es falsa, mayor la confirmación. En el caso determinista, donde h implica e, P(e/h)=1 y la confirmación depende inversamente de P(e) o de P(e/¬h). Por esto, la información especialmente inesperada o sorprendente apoya la hipótesis con fuerza particular.

# Paradoja de Hempel

$h \equiv \forall x\ (Cx \rightarrow Nx)$

El análisis bayesiano (según Howson y Urbach) muestra dos cosas:

(1) Evidencia de la forma $\neg Ca \wedge \neg Na$ confirma h en un grado despreciable.

(2) El criterio de Nicod no es válido como principio universal de confirmación.

$$\frac{P(h/Ca\wedge Na)}{P(h)} = \frac{P(Ca\wedge Na/h)}{P(Ca\wedge Na)} \quad \text{(I)}$$

y

$$\frac{P(h/\neg Ca\wedge\neg Na)}{P(h)} = \frac{P(\neg Ca\wedge\neg Na/h)}{P(\neg Ca\wedge\neg Na)} \quad \text{(II)}$$

Trabajemos con (I):

tenemos $P(Ca\wedge Na/h) = \frac{P(Ca\wedge Na\wedge h)}{P(h)}$ (i)

$P(Na/h\wedge Ca) = \frac{P(Ca\wedge Na\wedge h)}{P(h\wedge Ca)}$ (ii)

$P(Ca/h) = \frac{P(Ca\wedge h)}{P(h)}$ (iii)

Se ve que (i) es el producto de (ii) y (iii). Entonces tenemos

$P(Ca \wedge Na/h) = P(Na/h \wedge Ca).P(Ca/h)$

Ahora bien es obvio que $P(Na/h \wedge Ca)=1$. Entonces $P(Ca \wedge Na/h) = P(Ca/h)$.

Y si suponemos que el hecho de que un objeto arbitrario sea un cuervo es independiente de la verdad de h, obtenemos

$P(Ca \wedge Na/h) = P(Ca)$.

Razonando análogamente se obtiene

$P(\neg Ca \wedge \neg Na/h) = P(\neg Na)$

Además, $P(Ca \wedge Na) = P(Na/Ca)P(Ca)$ y $P(Na/Ca) = \Sigma P(Na/Ca \wedge \Theta).P(\Theta/Ca)$

Donde $\Theta$ es la proposición: «La proporción de cuervos negros entre los cuervos es $\Theta$» (h dice que $\Theta$=1; se consideran un rango posible de valores de $\Theta$)

Tenemos que si $\Theta$ es independiente de Ca,

$P(Na/Ca) = \Sigma P(Na/Ca \wedge \Theta).P(\Theta/Ca) = \Sigma P(Na/Ca \wedge \Theta)P(\Theta)$

y al ser $P(Na/Ca \wedge \Theta)=\Theta$, combinando todo lo obtenido con el teorema de Bayes tenemos:

$$\frac{P(h/Ca \wedge Na)}{P(h)} = \frac{1}{\sum \Theta P(\Theta)}$$

$$\frac{P(h/Ca \wedge Na)}{P(h)} = \frac{1}{\sum \Theta P(\Theta)}$$

Esto dice que la razón entre las probabilidades a posteriori y a priori de h es inversamente proporcional a ΣΦP(Φ). Si inicialmente fuese muy probable que todos o casi todos los cuervos fuesen negros, ΣΦP(Φ) sería grande y Ca∧Na confirmaría muy poco h. Si por el contrario, fuese muy probable que pocos cuervos fueran negros, el factor del que venimos hablando sería pequeño y la confirmación sustancial.

También se obtiene:

$$\frac{P(h/\neg Ca \wedge \neg Na)}{P(h)} = \frac{1}{P(\neg Ca/\neg Na)}$$

Como presumiblemente hay muchísimas cosas más a las que se le aplica el predicado «no negro» que el predicado «cuervo», la probabilidad de que un objeto del que solo sabemos que no es negro, que sea un no cuervo debe ser altísima. Y entonces el último cociente es muy próximo a 1, de manera que el poder confirmatorio de un no cuervo no negro es mínimo.

# Críticas al criterio de Nicod.

Nicod: «Las hipótesis de la forma condicional universal son confirmadas por individuos que satisfacen el antecedente y el consecuente.»

Hipótesis h: «Todos las serpientes de cascabel habitan fuera del departamento de Lavalleja»

Observación: Un auto pisa una serpiente de cascabel que cruza la ruta 8, en el departamento de Canelones, a 10 metros del cartel que indica el límite entre Canelones y Lavalleja.

Juan, Pedro y Diego van a un bar y dejan sus camperas en las sillas. Salen a fumar varias veces, y al volver intercambian los lugares. Cuando se van, muy borrachos, cada uno se lleva la campera que está en la silla que ocupa, sin verificar de cuál se trata ni importarle lo que hacen los demás.

Hipótesis h: «Ninguno de los tres se llevó su propia campera»

Observación 1: Juan se llevó la campera de Pedro

Observación 2: Pedro se llevó la campera de Juan

(Rosenkrantz: *Inference, Method and Decision: Towards a Bayesian Philosophy of Science*, 1977)

# Problema Duhem-Quine

Se ha criticado al falsacionismo popperiano de acuerdo a las siguientes líneas: Las teorías más notables de la ciencia son «infalsables» a través de enunciados observacionales porque solo hacen predicciones en conjunto con teorías auxiliares. Si una predicción se muestra falsa, la lógica no obliga a considerar a la teoría «principal» como falsa, porque el error puede estar en las teorías auxiliares. El problema de Duhem es:

**Cuando aparecen varias teorías involucradas en la predicción de una observación que luego se muestra falsa, ¿cuál de las teorías debe considerarse falsa?**

## Respuestas al problema

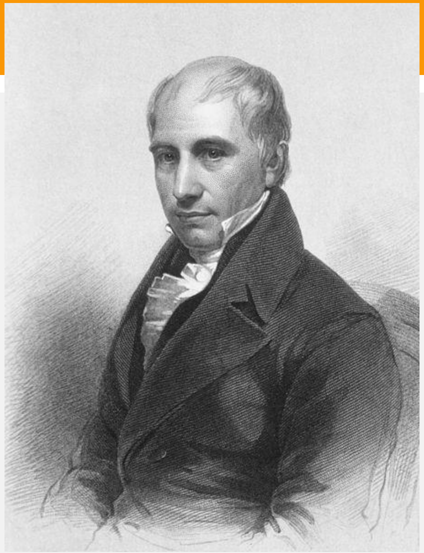
- Popper: Convencionalismo
- Kuhn: Paradigmas
- Lakatos: Programas de investigación
- La respuesta bayesiana

# William Prout (1785-1850)

Alrededor de 1815, Prout avanzó la idea de que el peso atómico de cada elemento es un múltiplo entero del peso atómico del hidrógeno. (La suposición subyacente era que toda la materia se formaba con combinaciones de algún elemento básico, y Prout creía que el hidrógeno era ese bloque elemental).

# William Prout (1785-1850)

En ese tiempo, la mayor parte de los pesos atómicos registrados no eran múltiplos del peso atómico del hidrógeno. Algunos estaban cerca de serlo, pero otros presentaban desviaciones importantes. Esto no minó la fuerte creencia de Prout en su hipótesis, porque en esos casos «le echaba la culpa» a los métodos utilizados para medir los pesos atómicos.



# Thomas Thomson (1773-1852)

Thomson, contemporáneo de Prout, -quien estaba convencido (erróneamente) que los pesos atómicos relativos del hidrógeno y del oxígeno eran 0,125-, midió 0,829 como peso atómico del boro, y lo cambió a 0,875 «porque es un múltiplo de 0,125, que es lo que parecen ser todos los átomos».

En términos bayesianos el razonamiento de Prout sobre el cloro y el de Thomson sobre el boro se puede explicar de la siguiente manera. La hipótesis T de Prout, junto con una suposición apropiada A, que establece la exactitud (dentro de ciertos límites) de las técnicas de medida, la pureza de los reactivos utilizados, etc., implica que la razón entre los pesos atómicos medidos del cloro y el hidrógeno se aproximará (en un cierto grado) a un número entero. En 1815 se reportó que esa razón era 35,83 – llamemos a esto evidencia E-, un valor que se juzgaba incompatible con la conjunción de T y A.

Las probabilidades posteriores y a priori de T y A se relacionan como sigue por el teorema de Bayes:

$$P(T/E) = \frac{P(E/T)P(T)}{P(E)} \quad \text{y} \quad P(A/E) = \frac{P(E/A)P(A)}{P(E)}$$

Para evaluar esas dos probabilidades posteriores es necesario cuantificar todos los términos de los miembros derechos de esas ecuaciones.

Consideremos P(T) y P(A). Con respecto a P(T) tenemos el testimonio de J.S.Stas, un químico belga que hizo cuidadosas medidas de pesos atómicos, que fueron muy influyentes, que nos da una razón para pensar que los químicos de la época estaban firmemente dispuestos a creer en T: «En Inglaterra la hipótesis del Dr.Prout era casi universalmente aceptada como una verdad absoluta».

(Cuando empezó a investigar el asunto, él mismo «tenía una confianza casi absoluta en la exactitud del principio de Prout».

Es menos fácil hablar sobre la confianza que Prout y sus contemporáneos tenían en los métodos para medir los pesos atómicos, pero probablemente no tenían una gran confianza en ellos dado que había muchas fuentes de error y esto se veía claramente. Había dudas sobre la pureza de los elementos utilizados, generalmente no se conocía la estructura de los compuestos de los que se extraían, y sobre todo, las medidas independientes raramente daban resultados parecidos. Por ejemplo, Thomson en 1818 reportó dos medidas independientes 2,998 y 2,66 del peso relativo al oxígeno de una molécula de ácido bórico.  A pesar de todo esto los químicos de la época tiene que haber sentido que sus medidas de pesos atómicos eran más probablemente exactas que erradas porque de otro modo difícilmente las hubieran reportado.

Consideremos entonces P(A)=0,6 y P(T)= 0,9.

Acá solo se intenta mostrar que hipótesis que son conjuntamente refutadas por una observación, a veces son desconfirmadas en grados muy diferentes, ilustrando la solución bayesiana al problema de Duhem. No es necesario para esto adherir a estas probabilidades asignadas aunque se verá que los resultados que se obtienen no son muy sensibles a las probabilidades a priori.

Las probabilidades posteriores de T y A también dependen de P(E), P(E/T) y P(E/A). Usando el teorema de la probabilidad total obtenemos:

$P(E) = P(E/T)P(T)+P(E/\neg T)P(\neg T)$

$P(E/T) = P(E/T\wedge A)P(A/T) + P(E/T\wedge\neg A)P(\neg A/T)$

Consideraremos que T y A son independientes o sea que, $P(A/T)=P(A)$ y $P(\neg A/T)=P(\neg A)$. Esta suposición simplifica los cálculos pero no es crucial y además parece muy correcta en principio. Hay que decir sin embargo, que incluso en este caso, esto ha sido puesto en duda.

Hecha esta suposición y teniendo en cuenta el hecho de que la conjunción $T \wedge A$ es refutada por E, y por lo tanto $P(E/T \wedge A)$ debe ser 0, obtenemos:

$$P(E/T)=P(E/T \wedge \neg A)P(\neg A).$$

Análogamente:

$$P(E/A)=P(E/\neg T \wedge A)P(\neg T).$$

$$P(E/\neg T)= P(E/\neg T \wedge A)P(A) + P(E/\neg T \wedge \neg A)P(\neg A)$$

Ahora asignaremos los siguientes números, que intentaremos justificar:

$P(E/\neg T \wedge A) = 0{,}01$

$P(E/\neg T \wedge \neg A) = 0{,}01$

$P(E/T \wedge \neg A) = 0{,}02$

La primera es la probabilidad de la evidencia si la hipótesis de Prout es falsa pero las suposiciones sobre los cálculos de los pesos atómicos son correctos. Algunos químicos decimonónicos pensaron cuidadosamente en el asunto y típicamente tomaron la teoría de una distribución aleatoria de los pesos atómicos como alternativa a la a la hipótesis de Prout. Consideremos esta teoría como alternativa.

Supongamos que haya quedado establecido como verdadero que el peso atómico del cloro está entre 35 y 36. La teoría de la distribución aleatoria asigna probabilidades iguales a que el peso atómico de un elemento caiga en cualquier intervalo de longitud 0,01. Bajo la suposición de que A es verdadera pero T falsa la probabilidad de que el peso del cloro caiga en el intervalo que va de 35,825 a 35,835 es 0,01. Se ha atribuido el mismo valor a $P(E/\neg T \wedge \neg A)$ porque si A fuese falsa debido a que, digamos, algunos reactivos fuesen impuros o hubiesen sido mal pesados, entonces suponiendo que T es falsa uno no debe esperar un sesgo de los pesos atómicos hacia ninguna parte del intervalo entre dos enteros adyacentes.

Se ha asignado a $P(E/T \wedge \neg A)$ un número bastante más alto, 0,02, porque si bien algunas impurezas en los reactivos y algunas inexactitudes en las mediciones fuesen moderadamente probables en la época, los químicos no hubieran considerado que sus técnicas eran totalmente azarosas. Entonces si la hipótesis de Prout era verdadera y las técnicas de medida imperfectas, era probable que los pesos atómicos se desviaran un poco de valores enteros pero la desviación sería menos probable cuanto más grande, de manera que la distribución de la probabilidad del peso atómico en el intervalo [35,36] no sería uniforme sino que estaría más concentrada cerca de los números enteros.

Ahora podemos hacer los cálculos aunque los valores absolutos de las probabilidades no son importantes, lo único que importa son los valores relativos. Llegaríamos a lo mismo suponiendo

$$P(E/\neg T \wedge A) = P(E/\neg T \wedge \neg A) = (1/2)P(E/T \wedge \neg A)$$

Obtenemos:

$P(E/\neg T) = (0,01)(0,6)+(0,01)(0,4)=0,01$

$P(E/T)=(0,02)(0,4)=0,008$

$P(E/A)=(0,01)(0,1)=0,001$

$P(E)=(0,008)(0,9)+(0,01)(0,1)=0,0082$

Y ahora el teorema de Bayes nos da las probabilidades a posteriori que nos interesaban:

P(T/E)=0,878 (Teníamos P(T)=0,9)

P(A/E)=0,073 (Teníamos P(A)=0,6)

Entonces la evidencia afecta la hipótesis de Prout y la suposición sobre la exactitud de las medidas en forma muy diferente: la probabilidad de la primera casi no cambia, y la de la segunda se derrumba.