# PROBABILIDAD INICIAL Y ÉXITO PROBABILÍSTICO[1]

VALERIANO IRANZO
Fundación General de la Universitat de València (España)
Valeriano.Iranzo@uv.es

## Resumen

Una cuestión controvertida en la teoría bayesiana de la confirmación es el estatus de las probabilidades iniciales. Aunque la tendencia dominante entre los bayesianos es considerar que la única constricción legítima sobre los valores de dichas probabilidades es la consistencia formal con los teoremas de la teoría matemática de la probabilidad, otros autores –partidarios de lo que se ha dado en llamar "bayesianismo objetivo"– defienden la conveniencia de restricciones adicionales. Mi propuesta, en el marco del bayesianismo objetivo, recoge una sugerencia de Wesley Salmon. No obstante, mientras que para Salmon las probabilidades iniciales se justifican en tanto incorporan la *frecuencia de éxito* de las hipótesis del pasado, el factor decisivo en mi opinión es el *éxito probabilístico* de las hipótesis del pasado. Así, la asignación de una probabilidad inicial a una hipótesis $h$ será correcta si refleja el éxito probabilístico de hipótesis del pasado semejantes a $h$. En esta línea, propongo una regla para fijar las probabilidades iniciales de las hipótesis, y acabo discutiendo algunas objeciones generales en contra de mi planteamiento.

PALABRAS CLAVE: Probabilidad; Bayesianismo; Confirmación; Frecuencialismo; Comparación (entre teorías).

## Abstract

One of the most controversial issues for Bayesian theory of confirmation concerns the status of prior probabilities. Orthodox Bayesianism claims that logical consistency with the theorems of mathematical theory of probability is the only constraint on prior probabilities, while some authors claim for further requirements. I agree with the latter standpoint. My proposal is inspired by Wesley Salmon's frequentist justification of prior probabilities. According to him, prior probabilities are the best assessments of past scientific hypotheses' success. I think that historical information is relevant here, but I propose to focus on a particular sort of success –i.e., "probabilistic success". I claim that a particular prior for $h$ is correct insofar as it reflects the probabilistic success of past scientific hypotheses similar to $h$. I also suggest a rule to set the priors for contemporary hypotheses. The closing paragraphs are devoted to some general objections against this proposal.

KEY WORDS: Probability; Bayesianism; Confirmation Theory; Frequentism; Theory-choice problem.

[1] La investigación conducente a este trabajo ha recibido subvención del Ministerio de Ciencia e Innovación español a través del proyecto FFI 2008-01169/FISO. El presente artículo desarrolla algunas de las ideas expuestas en Iranzo (2008). Partes de él fueron discutidas en las universidades de Bolonia y Santiago de Compostela. Mi agradecimiento a la audiencia en tales ocasiones, así como a los árbitros de *Análisis Filosófico*, por sus atinados comentarios.

Las dificultades planteadas por el llamado método hipotético-deductivo[2] han dado lugar en las últimas décadas a diversos enfoques sobre la contrastación de hipótesis. El bayesianismo, cuyo punto de arranque se encuentra en los escritos de F. P. Ramsey y B. De Finetti, es uno de los que goza de mayor aceptación en nuestros días. Esta corriente sostiene que la probabilidad no es más que el "grado de convicción" (*credence* en inglés) que un sujeto tiene respecto a un enunciado determinado. De ahí que se la considere una concepción subjetivista o "personalista" de la probabilidad, ya que, en estricto sentido bayesiano, atribuir una probabilidad –a un enunciado, una hipótesis, o un suceso–, no tiene contenido a menos que haga referencia a un sujeto.

En este artículo pretendo abordar un problema bien conocido en la tradición bayesiana como es el estatus y justificación de las probabilidades iniciales. Aunque es verdad que no en todas las situaciones el cálculo de la probabilidad inicial tiene por qué ser problemático, sí que lo es en los casos más interesantes desde un punto de vista filosófico. El bayesianismo ortodoxo –B. De Finetti, L. Savage, C. Howson y P. Urbach– confía, además, en el poder de la evidencia empírica para neutralizar las diferencias en las asignaciones de probabilidad inicial de distintos sujetos, sin que haga falta introducir otras constricciones que la consistencia con los teoremas de la teoría matemática de la probabilidad.[3] Pienso, sin embargo, que ésta no es una perspectiva adecuada en el contexto de la evaluación/justificación de hipótesis científicas.[4] Para un bayesiano las probabilidades iniciales son tan subjetivas como cualesquiera otras, pues todas son estados creenciales del sujeto; pero eso no significa que cualquier asignación de probabilidades iniciales sea admisible con tal de que no entre en contradicción lógica con el aparato matemático. Tras comentar el peso que corresponde a la probabilidad inicial en el Teorema de Bayes, señalaré diversas estrategias para limitar el relativismo en la asignación de tales probabilidades. Desarrollaré mi propuesta a partir de la sugerencia de Wesley Salmon de justificar las probabilidades iniciales de las hipótesis actuales en relación a las frecuencias de éxito obtenidas por

[2] Véase Cassini (2003) y Salmon (2005, cap. 4).

[3] *Scientific Reasoning: The Bayesian Approach*, de C. Howson y P. Urbach, constituye la defensa canónica de esta posición entre los filósofos de la ciencia. Este texto cubre una amplia problemática y se ha ido actualizando a lo largo de los años en tres ediciones (1989, 1993 y 2006). Citaré la segunda edición por ser la más completa de las tres. Una introducción concisa, pero igualmente recomendable, a la filosofía de la ciencia bayesiana es Strevens (2005).

[4] Huber (2005) vindica esta tesis en contra del bayesianismo ortodoxo, si bien en una línea distinta a la que aquí defenderé.

las hipótesis científicas del pasado. La innovación por mi parte reside, primero, en interpretar el éxito del pasado en términos probabilísticos; y en segundo lugar, en proponer un algoritmo para fijar la probabilidad inicial de las hipótesis actualmente en liza.

## 1. El Teorema de Bayes y la confirmación de las hipótesis científicas

Este teorema, que debe su nombre a su descubridor, el reverendo Thomas Bayes (1702-1761), es una consecuencia *deductiva* de los axiomas de la teoría de la probabilidad. Conviene advertir, pues, que cualquier interpretación de la probabilidad que pretenda ser consistente con su definición matemática debe aceptarlo, sea bayesiana o no bayesiana. Convenientemente interpretado el Teorema de Bayes permite calcular el apoyo confirmacional que una evidencia $e$, un resultado experimental, por ejemplo, confiere a una hipótesis $h$. En su forma más simple se plantearía así:

$$p\,(h \mid e) = \frac{p\,(h) \cdot p\,(e \mid h)}{p\,(e)} \qquad (1)$$

$p\,(h \mid e)$ probabilidad condicional de $h$, esto es, la probabilidad de $h$ dado que $e$ ocurre;
$p\,(h)$ probabilidad inicial de $h$, o sea, la probabilidad de $h$ descontando la evidencia $e$ que, supuestamente, confirma a $h$;
$p\,(e \mid h)$ verosimilitud (*likelihood*) de $h$ en relación a $e$;[5]
$p\,(e)$ probabilidad inicial de $e$, esto es, la probabilidad de $e$ sin contar con $h$.

Naturalmente, $p\,(e \mid h)$ no estará definida cuando $p\,(e) = 0$. Por lo demás, la idea que subyace al teorema es bastante elemental: $e$ confirma $h$ si y sólo si la probabilidad de $h$ dado $e$ es mayor que la probabilidad inicial de $h$, o sea, sii $p\,(h \mid e) > p\,(h)$. Confirmación equivale, en suma, a *incremento de probabilidad*. Entonces,

$e$ *confirma* a $h$ sii $p\,(h \mid e) > p\,(h)$
$e$ *disconfirma* a $h$ sii $p\,(h \mid e) < p\,(h)$
$e$ *es neutra* respecto de $h$ sii $p\,(h \mid e) = p\,(h)$

[5] Usualmente los textos de estadística traducen *likelihood* por verosimilitud. Seguiré aquí esta convención, aunque no debe confundirse la *likelihood* bayesiana con la *truthlikeness* popperiana, que también suele traducirse por verosimilitud.

A veces se incluye un término que refiere al "conocimiento de fondo". Se pretende así destacar que la confirmación no ocurre en el vacío y que las estimaciones de probabilidad son relativas a un cuerpo dado de conocimientos.[6] Nótese, sin embargo, que el bayesianismo es incompatible con el holismo. Se ha repetido hasta la saciedad que no es posible extraer consecuencias observacionales de una hipótesis teórica sin contar con otros enunciados adicionales, como hipótesis auxiliares o enunciados que describen las condiciones iniciales del experimento. Sin embargo, a pesar de admitir esto, para los bayesianos es factible, en principio, deslindar el efecto de la evidencia sobre un elemento del edificio teórico, esto es, sobre una hipótesis concreta, del impacto que aquélla pueda tener sobre el resto de enunciados adicionales que intervienen, explícita o implícitamente, en la predicción. Con otras palabras, se puede calcular el grado en que $h$ se ve confirmada o disconfirmada por una evidencia particular $e$, independientemente del efecto confirmacional que $e$ tenga sobre un bloque más amplio de enunciados en el que $h$ esté incluida.[7]

La aplicación del teorema de Bayes en ciertas situaciones no reviste gran dificultad. Veamos un ejemplo sencillo. Supóngase que el 10% de los sujetos de una población padece una enfermedad endémica. Existe una prueba para detectar la enfermedad, pero no es totalmente fiable, ya que sólo el 95% de las personas enfermas da positivo. Además, un pequeño porcentaje de los sujetos (1%) da positivo en la prueba y no tiene la enfermedad. Lo que queremos saber es cuál es la probabilidad de que una persona que haya dado positivo en la prueba esté sana. Tenemos todos los datos, excepto $p\ (e)$. Por eso aquí conviene aplicar una versión del teorema de Bayes más general que (1) –la fórmula de la "probabilidad total":

$$p\ (h_i \mid e) = \frac{p\ (h_j) \cdot p\ (e \mid h_j)}{\sum_{i=1}^{n} p\ (h_i) \cdot p\ (e \mid h)}\ j = 1, 2, \ldots, n \qquad (2)$$

Esta fórmula resulta especialmente adecuada en situaciones en las que existen hipótesis rivales, ya que nos permite comparar el apoyo confirmacional de $e$ a cada una de las hipótesis $h_i$, siempre que el conjunto de las $h_i$ constituya una *partición* del conjunto de resultados posibles, o sea, siempre que se cumplan las siguientes dos condiciones: (i) que las $h_i$

[6] Siendo $b$ el conocimiento de fondo, la fórmula es:

$$p\ (h \mid e \wedge b) = \frac{p\ (h \mid b) \cdot p\ (e \mid h \wedge b)}{p\ (\mathrm{e} \wedge b)}, \text{ siempre que } p\ (\mathrm{e} \wedge b) \neq 0.$$

[7] Una propuesta detallada en esta línea se encuentra en Strevens (2001).

sean excluyentes tomadas dos a dos; y (ii) que su conjunción agote todas las posibilidades del espacio muestral, con lo cual la suma de las probabilidades iniciales de todas las $h_i$ ha de ser igual a 1. En el ejemplo que nos ocupa sólo hay dos alternativas, tener la enfermedad ($h_1$) o estar sano ($h_2$). Obviamente, las probabilidades iniciales de $h_1$ y $h_2$ suman 1, puesto que una es la negación de la otra. Sea $e$ un resultado positivo; entonces la probabilidad de que una persona que haya dado positivo en la prueba esté sana es:

$$p\,(h_2 \mid e) = \frac{p\,(h_2) \cdot p\,(e \mid h_2)}{p\,(h_2) \cdot p\,(e \mid h_2) + p\,(h_1) \cdot p\,(e \mid h_1)} =$$

$$= \frac{0'9 \cdot 0'01}{0'9 \cdot 0'01 + 0'1 \cdot 0'95} = 0'086.$$

Es decir, entre un ocho y un nueve por ciento de los resultados positivos son personas que no están enfermas.

El asunto se complica cuando tratamos de abordar ejemplos más interesantes para la filosofía de la ciencia. Es un lugar común afirmar que las observaciones efectuadas durante el eclipse de sol de 1919 dieron un respaldo notable a la Teoría General de la Relatividad (TGR). La mecánica newtoniana también predecía una deflexión de los rayos de luz, pero los valores obtenidos por A. Eddington estaban mucho más cerca de la predicción derivada de TGR. Por eso se tomaron como evidencia decisiva a favor de la tesis einsteiniana de que el espacio-tiempo posee una curvatura variable y una estructura métrica semi-rienmanniana, precipitando el abandono de la concepción euclídea del espacio tiempo. Traduciendo el episodio al vocabulario bayesiano diríamos que la probabilidad de TGR después de las observaciones de Eddington (OE) es mayor que la de TGR sin contar con ellas, o sea, $p$ (TGR | OE) > $p$ (TGR). Eso significa que OE confirma TGR, por el criterio bayesiano de confirmación expuesto anteriormente.[8]

Ahora bien, del mismo modo que antes nos interesaba conocer la probabilidad de estar sano dado que la prueba ha resultado positiva, supongamos que no queremos saber solamente si OE confirma o no a TGR, sino también cuál es la probabilidad de TGR dado que se ha cumplido la predicción y OE es verdadera. Para calcular $p\,(h \mid e)$, o sea $p$ (TGR | OE), tal como hicimos en el ejemplo de la enfermedad endémica, estamos obligados a dar un valor a la probabilidad inicial de TGR. Sin embargo, hay

[8] Véase Mayo (1996) para una discusión detallada del episodio. Incidentalmente, la interpretación de esta autora es opuesta al bayesianismo.

una diferencia entre ambos ejemplos que no puede pasarse por alto. En el primero conocemos la prevalencia de la enfermedad en la población y sabemos también en qué medida falla la prueba. Esto nos da los valores de la probabilidad inicial y la verosimilitud, respectivamente. Además, podemos aplicar el teorema de la probabilidad total directamente, puesto que no hay más que dos hipótesis: tener la enfermedad o no tenerla. En una situación de rivalidad entre hipótesis teóricas, en cambio, no es tan fácil precisar las alternativas, teniendo en cuenta que éstas deben ser "exhaustivas". Sabemos que TGR y la mecánica newtoniana (MN) son rivales, pero, ¿por qué habríamos de pensar que estas dos alternativas agotan todas las posibilidades? Un modo de generar una partición del espacio muestral es plantear la rivalidad a tres bandas: TGR, MN, y la negación de ambas, o sea, ¬TGR ∧ ¬MN. Pasar por alto esta última posibilidad, a la que los bayesianos se refieren usualmente como la "*catch-all hypothesis*", y que aquí llamaré la "hipótesis residual", equivaldría a suponer que no hay más alternativas teóricas posibles que TGR y MN, lo cual resulta, cuanto menos, presuntuoso. Por otra parte, si TGR y MN son incompatibles, $p$ (TGR) + $p$ (MN) + $p$ (¬TGR ∧ ¬MN) = 1, con lo cual ya tenemos la partición que necesitábamos. No obstante, es difícil hacerse una idea de qué significa realmente ¬TGR ∧ ¬MN. No parece una hipótesis científica en un sentido genuino. Apenas afirma nada sobre el mundo, salvo que éste no es ni como TGR ni como MN dicen que es. ¿A qué apelar para averiguar $p$ (OE | ¬TGR ∧ ¬MN)? En cuanto a la tentación de retornar al teorema de Bayes en su versión (1), no está claro que sirva de mucho: evitaríamos el problema de calcular el valor de la hipótesis residual introduciendo otro, como es el de fijar la probabilidad inicial de $e$, en este caso, la probabilidad de que la desviación de la luz sea la que dice TGR. Tendríamos entonces que vérnoslas no con una, sino con dos probabilidades iniciales, la de $h$ y la de $e$.

Así pues, aparte de las dificultades específicas que surgen en caso de optar por (1) o por (2), determinar la probabilidad inicial de la evidencia o la de la verosimilitud de la hipótesis residual, respectivamente, el algoritmo bayesiano no podrá funcionar bajo ninguna de ambas formulaciones, si no se aclara previamente cómo se fija la probabilidad inicial de la hipótesis en cuestión.

## 2. El estatus de las probabilidades iniciales

Se ha dicho ya que para los bayesianos ortodoxos los axiomas/teoremas de la teoría de la probabilidad son la única constricción que el sujeto debe respetar en su asignación de probabilidades iniciales. Se trata de una condición puramente coherentista, en tanto es la consistencia lógica con aquéllos el único criterio que determina la corrección o incorrección de una determinada distribución de probabilidades iniciales. Así, puede ocurrir que diferentes sujetos atribuyan probabilidades iniciales muy distintas a la misma hipótesis, sin que ninguno de ellos viole los axiomas. Por tanto, ninguno se equivoca y todos están justificados en sus "*credences*". Esto no es problemático porque, según el bayesianismo ortodoxo, a largo plazo la evidencia neutraliza las diferencias entre las probabilidades iniciales haciendo que los valores de la probabilidad final converjan, de modo que la acumulación de evidencia ha de conducir finalmente al consenso, a pesar de las discrepancias en el punto de partida.[9]

Aquí hay que introducir matizaciones importantes. El bayesianismo intenta ser una aproximación formal, algorítmica, a esa vaga idea que sería "aprender de la experiencia". El recurso bayesiano más conocido para calcular el efecto acumulativo de la evidencia es el llamado "Principio de Condicionalización".[10] Siendo $e_1, e_2, e_3,\ldots, e_n$, sucesivos enunciados observacionales, este principio nos dice que al calcular $p\ (h \mid e_2)$ hemos de tomar como "probabilidad inicial" $p\ (h \mid e_1)$. La probabilidad condicionada a la nueva evidencia, $e_2$, tiene en cuenta el efecto acumulado sobre $p\ (h)$ por toda la evidencia anterior ($e_1$, en este caso). Al reevaluar la probabilidad de $h$ conforme la evidencia aumenta se introducen como probabilidades iniciales los sucesivos valores obtenidos para la probabilidad condicionada de $h$, $p\ (h \mid e_1)$, $p\ (h \mid e_1 \wedge e_2)$, $p\ (h \mid e_1 \wedge e_2 \wedge e_3)$,… Es verdad que el peso relativo de la probabilidad inicial de $h$ en relación al valor de la probabilidad condicionada va disminuyendo a medida que aumenta el número de términos $e$. Pero que el peso de $p\ (h)$ disminuya no implica que la acumulación de evidencia vaya a resolver por sí sola las discrepancias sobre las probabilidades iniciales. En repetidas ocasiones se ha señalado que las demostraciones matemáticas de la convergencia a un límite asumen condiciones ideales bastante alejadas de las que rigen en la práctica científica.[11]

[9] En Howson (2003, pp. 208 y ss.) se argumenta este punto de vista. Para una formulación matemática precisa, véase Gaifman y Snir (1982, espec. § 2). Cabe señalar que algunos teoremas de convergencia sólo pueden probarse asumiendo el axioma de la "aditividad enumerable", ya que no basta con la aditividad finita. Sobre los recelos de Bruno De Finetti respecto a la aditividad enumerable, y sobre la inocuidad de ésta para la interpretación bayesiana de la probabilidad, véase Gillies (2000, pp. 65-69).

[10] Para una discusión más detallada del Principio de Condicionalización, véase Howson y Urbach (1993, cap. 6).

[11] Véase, por ejemplo, Earman (1992, cap. 6) y Jaynes (2004, pp. 126 y ss.).

Con la idea de constreñir en mayor grado los valores admisibles para las probabilidades iniciales se han ensayado diversas alternativas. Las posiciones radicalmente objetivistas no dejan ningún margen de libertad al sujeto para elegir las probabilidades iniciales. En esta línea, el Principio de Razón Insuficiente, defendido en su día por J. Bernoulli y por el marqués de Laplace, y renombrado como Principio de Indiferencia por J.M. Keynes, ha sido refinado dando lugar a reglas más sofisticadas como el principio de Máxima Entropía de E.T. Jaynes.[12] Otros autores han optado por introducir diversas restricciones para, si no eliminar, sí al menos suavizar el margen de variabilidad aceptable en la probabilidad inicial. Aquí cabe destacar el "personalismo atemperado" de A. Shimony (1970) y la particular interpretación de la teoría de la decisión de I. Levi (1980). Ambos insisten en que cualquier hipótesis seriamente propuesta ha de tener la oportunidad de imponerse sobre sus rivales, contando con un curso de experiencia favorable, en un lapso de tiempo razonable. Por eso piensan que se deben excluir las estimaciones de probabilidad inicial

[12] Al lanzar una moneda contamos con dos resultados posibles: cara y cruz. Dependiendo de si pensamos que la moneda está cargada o no, preferiremos una distribución de probabilidad u otra. Pero si no tenemos ninguna razón para creer que la moneda está cargada, consideraremos que los resultados posibles son *equiprobables*, o sea, $p$ (cara) = $p$ (cruz) = $1/2$. Optamos entonces por una distribución de probabilidad *uniforme*, y al actuar así no hacemos más que aplicar el Principio de Indiferencia, según el cual $p_k = 1/n$, siendo $n$ el número de resultados posibles y $p_k$ la probabilidad que debe asignarse a cada uno de ellos. El Principio de Máxima Entropía (PME) es una generalización del Principio de Indiferencia, ya que este último sólo es aplicable a variables aleatorias discretas. La entropía de una distribución de probabilidad es inversamente proporcional a su contenido informativo. Para una variable discreta $X$, la entropía $H$ de cada distribución es:

$$H(X) = -\sum_{i=1}^{n} p_i \log p_i$$

Para una variable discreta, la distribución de probabilidad de mayor entropía, la menos informativa, es la distribución uniforme, con lo cual el Principio de Indiferencia se convierte en un caso particular de PME. En el caso de variables aleatorias contínuas, donde el número de resultados posibles puede ser infinito y no enumerable, surgen complicaciones a la hora de aplicar PME sin violar el axioma de la aditividad enumerable. Sobre el aparato matemático involucrado, véase Jaynes (2004, cap. 12). En cuanto a la justificación intuitiva de PME, todo depende de lo defendible que resulte traducir la ignorancia o la carencia de información por parte del sujeto a una noción puramente formal como es la equiprobabilidad. Resulta instructivo comparar los favorables comentarios al respecto en Rosenkrantz (1977, pp. 77-81) con las críticas de los bayesianos Howson y Urbach en su (1993, pp. 413-417).

excesivamente bajas, justamente para que la evidencia empírica pueda hacer su trabajo.[13]

Nuestra intención es terciar en este debate. Para ello recuperaremos una propuesta, en mi opinión insuficientemente desarrollada, de W. Salmon (1967, 1970 y 1990), según la cual los juicios del científico sobre lo *plausible* que a él le resulta una hipótesis teórica inicialmente, esto es, sobre su probabilidad inicial, son estimaciones implícitas de ciertas *frecuencias*. Salmon pensaba que, aunque *de facto* encontremos diferencias entre los científicos en las probabilidades iniciales que atribuyen a la misma hipótesis, hay razones para distinguir entre atribuciones correctas e incorrectas en un sentido que va más allá de la mera consistencia con el aparato matemático de la teoría de la probabilidad. No se trata, pues, de proponer un camino alternativo para mostrar que el desacuerdo en las probabilidades iniciales será indefectiblemente neutralizado por la experiencia. La idea de Salmon es que el valor de la probabilidad inicial que un sujeto asigna a una hipótesis será correcto o incorrecto en tanto refleje la frecuencia de éxito que han tenido las hipótesis científicas del pasado. Veamos esto con más detalle.

Dejando a un lado el apoyo evidencial directo que pueda tener una hipótesis, los científicos manifiestan marcadas preferencias. Tienden a rechazar las hipótesis *ad-hoc*, las que resultan difíciles de encajar con los resultados obtenidos en campos anejos y favorecen las que son compatibles con el cuerpo de conocimientos aceptados, las más simples, las que poseen mayor poder unificador, en suma las mejores desde un punto de vista explicativo. Esto no significa que la evidencia experimental

[13] Por mor de la argumentación encuadramos a ambos autores bajo idéntico epígrafe ("objetivistas moderados"), aunque las diferencias son apreciables. En la propuesta de Shimony solamente intervienen funciones de probabilidad, mientras que Levi introduce la verdad y la falsedad de la hipótesis como *utilidades* epistémicas, de modo que la utilidad de considerar una hipótesis (respecto al valor de un parámetro) como verdadera o como falsa es la base para eliminar las distribuciones de probabilidad inicial que se alejan excesivamente de la distribución uniforme. Aparte de esto ambos autores coinciden en que una hipótesis no merece nuestra atención sólo porque sea una posibilidad lógica. Así, Levi insiste en que sólo cabe tomar en consideración las hipótesis que son "serias posibilidades", mientras que Shimony apela a las hipótesis que han sido efectivamente planteadas en la práctica científica. Por otro lado, en el marco de la teoría de la decisión se han propuesto modelos que intentan disolver el problema planteado aquí. Desde esta perspectiva ya no se trata de discutir la justificación objetiva de las probabilidades iniciales, sino de adoptar una posición "eliminacionista", definiéndolas operacionalmente en términos de estrategias y utilidades. Véase por ejemplo, Nau (2005), donde la probabilidad inicial como grado subjetivo de convicción es reemplazada por la noción de "probabilidad neutral respecto al riesgo".

desempeñe un papel puramente accesorio, claro está. Pero en las primeras fases de la investigación, cuando la hipótesis aún es una apuesta falta de respaldo evidencial determinante, los científicos están más dispuestos a invertir esfuerzos en investigar, desarrollar y contrastar aquellas hipótesis que en principio les resultan más plausibles. Si diferenciamos varias dimensiones en la evaluación de las hipótesis teóricas, como pueden ser la búsqueda, la confirmación y la aceptación, en correspondencia con las distintas fases de la investigación científica,[14] los juicios sobre la plausibilidad de una hipótesis pertenecerían a la primera dimensión. Con otras palabras, tales juicios son una evaluación informal del grado en que la hipótesis satisface las virtudes teóricas mencionadas, y funcionan como un filtro que selecciona, de entre las posibles alternativas compatibles con la evidencia y con los conocimientos aceptados, las que resultan más prometedoras. Se trata de una primera criba que persigue seleccionar las hipótesis que se vayan a comportar mejor en la posterior confrontación con la evidencia experimental.

A menudo las estimaciones de plausibilidad son juicios intuitivos, ya que el propio científico difícilmente puede explicar con detalle en qué basa su apreciación, y mucho menos por qué determinados valores como la simplicidad, en sus diversas dimensiones (matemática, ontológica, teórica), el alcance explicativo, la fertilidad, la compatibilidad o afinidad con hipótesis o explicaciones ya aceptadas en otros ámbitos, la belleza, incluso, deben ser favorecidos frente a otros.[15] Que tales juicios sean intuitivos, no significa necesariamente que sean irremediablemente subjetivos como si de juicios de gusto se tratase. La imagen romántica de la ciencia, en la que el protagonista es el genio científico que trabaja en aislamiento, sirve de poco para entender la ciencia contemporánea. La educación de los científicos es un proceso reglado que pretende familiarizar al individuo no sólo con un conocimiento consolidado, sino también con estrategias específicas para mejorar e incrementar dicho conocimiento. El proceso educativo del investigador consigue su meta justamente

[14] Véase Niiniluoto (2007).

[15] Las hipótesis o teorías científicas no tienen por qué satisfacer los diversos valores epistémicos implicados en las estimaciones de plausibilidad en la misma medida. En un contexto con diversas alternativas, la más simple, en cuanto al aparato matemático que introduce, puede ser la menos fértil, o tal vez, la más difícil de encajar en el cuerpo de conocimientos ya aceptados. En este sentido cabe señalar los intentos de abordar la cuestión con el aparato formal de la teoría de la revisión de creencias. Un ejemplo es Arló-Costa (2006). Inspirándose en los planteamientos de I. Levi, este autor propone una ampliación de la dimensión axiológica y trata de dar cuenta de la racionalidad epistémica en términos de maximización y no de optimización.

cuando éste aplica pertinentemente aquellas estrategias. Las pautas y criterios que se transmiten de una generación a otra han de ser bastante homogéneos para permitir una actividad cooperativa en alto grado como la ciencia,[16] y recogen el saber acumulado durante siglos respecto a las estrategias investigadoras que han resultado *más exitosas*. Visto así, los juicios de plausibilidad no responden a la pura idiosincrasia personal, o al feliz azar de estar vinculado a una u otra escuela, sino que, en términos generales, recogen el éxito, o fracaso, que han tenido las hipótesis anteriores de la historia de la ciencia. Salmon sostiene (1970 y 1990) que las probabilidades iniciales, por su parte, traducen a un valor numérico el contenido de dichos juicios. En palabras de este autor, las probabilidades iniciales son "las mejores estimaciones de las *frecuencias con que ciertas clases de hipótesis tuvieron éxito*" (1990, p. 187; la cursiva es mía). Esto no es sino una justificación *frecuencialista* de las probabilidades iniciales, ya que éstas se entienden como juicios implícitos acerca de la tasa relativa de éxitos obtenida mediante diversos tipos de hipótesis.

Ciertamente, el problema que nos ocupa aquí no son las hipótesis del pasado. La cuestión es más bien cómo justificar, y cómo calcular si es posible, la probabilidad inicial de una hipótesis que es objeto de discusión en la actualidad. Pues bien, aquí el frecuencialismo invita a una inducción simple del éxito pasado al futuro: en la historia de la ciencia se han sucedido diversos tipos de hipótesis; las hipótesis que han poseído ciertas características (simplicidad, poder explicativo, etc.) han llevado en mayor proporción al éxito que las que no las poseen; entonces, si $h_1$ tiene esas características y $h_2$ no, $h_1$ merece mayor probabilidad inicial que $h_2$. La historia de la ciencia puede verse, según Salmon, como "una crónica de nuestra experiencia previa en la formulación de hipótesis y en la teorización, de nuestro aprendizaje sobre qué tipos de hipótesis funcionan y qué otros no." (1970, p. 86). Así, si el curso de la historia muestra que las hipótesis más atractivas desde un punto de vista explicativo, las que resultan más plausibles en principio, han ido acompañadas de éxito con mayor frecuencia que las que no lo son, queda justificada tanto la asignación de una mayor probabilidad inicial a las primeras, como los propios criterios de plausibilidad invocados por los científicos para preferirlas, siempre que no seamos totalmente escépticos respecto a la inducción. El elemento decisivo es, por tanto, la *frecuencia histórica de éxitos*.

[16] Alguien podría invocar aquí al modesto empleado de la oficina de patentes de Berna como ilustre contraejemplo. Dejando de lado que sólo una pequeña parte de la investigación se produce en el terreno de la física teórica, es claro que el contexto en el que se lleva a cabo la investigación hoy en día es muy diferente al de la época del joven Einstein.

En cuanto a su radicalidad, la propuesta de Salmon queda, en principio, más cerca de un objetivismo fuerte, que de un objetivismo moderado al estilo de Shimony o Levi. Si las frecuencias en cuestión son hechos del mundo, como es un hecho, por ejemplo, que la proporción de niños y niñas entre los recién nacidos es, aproximadamente, un 48% y un 52%, respectivamente, sólo serán correctas las distribuciones de probabilidad inicial que reflejen tales frecuencias dentro de un margen de error aceptable. De la propuesta objetivista de Salmon, sin embargo, se deriva una justificación *histórica*, o si se quiere, puramente inductiva para la probabilidad inicial, pues lo que decide son los resultados –el éxito– acumulados hasta ahora. Es por esto que el enfoque de Salmon, aunque va más allá de lo que aquí hemos llamado "objetivismo moderado", plantea una perspectiva distinta a la de un objetivismo basado en reglas como el Principio de Máxima Entropía antes mencionado. De todos modos, Salmon siempre planteó su posición sobre las probabilidades iniciales como una mera conjetura razonable sin darle mayor concreción. Si queremos que resulte realmente iluminadora, hemos de afinar más, comenzando por aclarar qué ha de entenderse por 'éxito'. Nuestra tarea a partir de aquí será, pues, precisar la sugerencia de Salmon, aunque ello nos lleve a alejarnos de una justificación estrictamente frecuencialista.

Antes de continuar, cabe hacer constar que lo que sea plausible según el juicio científico no tiene por qué ser equivalente a lo que resulte probable.[17] Y aunque lo fuera, tampoco está claro que la probabilidad inicial pudiera recoger todo el contenido de aquél. La plausibilidad de una hipótesis o de una explicación teórica también parece depender de lo razonable o previsible que la hipótesis haga a un *explanandum* en particular. Si quisiéramos traducir este componente a un valor probabilístico, seguramente deberíamos reflejarlo en $p(e \mid h)$ y no en la probabilidad inicial. Por otro lado, quizá una noción matemática de probabilidad no agote los usos del predicado "probable" que sirven de base, y legitiman, nuestras expectativas inductivas.[18] En este trabajo, empero, asumiremos que el grado de plausibilidad que se conceda a una hipótesis científica, cuando aún no se dispone de evidencia confirmatoria o disconfirmatoria directamente relacionada con ella, se puede traducir a un valor probabilístico, y que el factor determinante que encapsula esto es la probabilidad inicial, la probabilidad de $h$ sin contar con $e$. Si en verdad no hubiera ninguna conexión relevante entre lo plausible que resulta $h$ al científico,

[17] *Cf.* Spohn (1988).

[18] Cabe reseñar aquí como alternativa la teoría "baconiana" de la probabilidad defendida en Cohen (1989).

por un lado, y la probabilidad inicial que éste le atribuye, por otro, la argumentación que viene a continuación perdería todo su fuste. Buscamos una justificación objetiva de las probabilidades iniciales en un contexto particular como es el de la comparación y contrastación de hipótesis científicas. Pero esa justificación depende de que se cumplan dos condiciones básicas: que haya un consenso mínimo respecto de lo que hace más o menos plausible la hipótesis con independencia del respaldo evidencial directo que tenga, y en segundo lugar, que la estimación de plausibilidad sea traducible, aun con cierta pérdida de contenido, a una asignación de probabilidad inicial.

## 3. Una propuesta para fijar las probabilidades iniciales

Dado que el éxito tiene que ver con la consecución de un objetivo, un modo de precisar la sugerencia de Salmon es repasando los diferentes objetivos en relación a los cuales podría decirse que las hipótesis han resultado exitosas.

La identificación del éxito con la consecución de la verdad, un objetivo clásico atribuido a la teorización científica, daría lugar, en mi opinión, a una versión inaceptable del frecuencialismo. Pensar que las hipótesis del pasado explicativamente más atractivas han resultado verdaderas en mayor porcentaje que las demás, supone un fuerte compromiso de partida con el realismo científico (frecuencia de éxitos = frecuencia de teorías *verdaderas*). A esto se podría objetar que al identificar éxito con verdad estamos fijando una meta quimérica, bien porque es inalcanzable, bien porque no podríamos saber siquiera cuándo la hemos alcanzado si es que eso llega a ocurrir. La objeción no deja de tener su razón de ser, ya que la verdad de las hipótesis *teóricas*, en las que el problema de las probabilidades iniciales es más agudo (v. *supra* sección 1), es precisamente el caballo de batalla entre realistas y antirrealistas. Por eso no es conveniente que la justificación basada en el éxito pasado haya de tomar partido en esta polémica. El argumento para justificar las probabilidades iniciales debiera ser compatible con el hecho de que ninguna de las hipótesis teóricas postuladas hasta el momento sea verdadera en sentido estricto. Desde esta premisa, la verdad no es, pues, el tipo de éxito que buscamos.

Otra alternativa es apelar al *éxito observacional*. Las hipótesis más atractivas desde un punto de vista explicativo han contado con cierto respaldo evidencial, y a largo plazo han soportado mejor la confrontación con la experiencia que las hipótesis que no lo fueron. Podría replicarse que de poco sirve tomar como referencia el éxito de hipótesis que resistieron los embates de la contrastación *temporalmente*. Qué más da si contaron con el beneplácito de la comunidad científica y fueron aceptadas durante mucho tiempo; si al final fueron falsadas concluyentemente, y por tanto, rechazadas, no parece que guiarse por su mérito explicativo sea fiable a fin de cuentas. Al favorecerlas estaremos priorizando hipótesis que, a lo sumo, tienen perspectivas de resistir por más tiempo que el resto, pero que pueden acabar siendo falsadas como las demás.

Esta línea argumentativa tal vez resulte tentadora para los popperianos. Según ellos la falsación es el sino de toda hipótesis científica. Los bayesianos, sin embargo, enfocan el asunto de otra forma: las hipótesis científicas son más o menos probables y los valores 0 y 1 son casos extremos. Esto no impide a un bayesiano dar cuenta de la falsación concluyente. Cuando hay una conexión deductiva entre la hipótesis y la evidencia, $h \models e$, entonces $p\,(e \mid h) = 1$. Consiguientemente, $p\,(\neg e \mid h) = 0$. Y bien, si ocurre $\neg e$, $p\,(h \mid \neg e) = 0$.[19] Así que $h$ es refutada por $\neg e$ terminantemente. Se trata, además, de una falsación definitiva porque la acumulación de evidencia posterior ya no puede tener efecto alguno sobre $h$. Al aplicar el Principio de Condicionalización y tomar $p\,(h \mid \neg e)$ como la probabilidad inicial, volverá a aparecer un factor igual a cero en el numerador y la probabilidad condicionada seguirá siendo igual a cero.

Aunque nada hay en el bayesianismo ortodoxo que excluya esta posibilidad, debe hacerse notar que la refutación concluyente sólo es posible cuando la probabilidad inicial de la hipótesis es cero o cuando hay una conexión *deductiva* entre hipótesis y evidencia –lo que se traduce en verosimilitudes con valores extremos: $p\,(e \mid h) = 1$ y $p\,(\neg e \mid h) = 0$. Sin embargo, la potencia de la teoría de la confirmación bayesiana estriba precisamente en que subsume la confirmación hipotético-deductiva como un caso particular. Y de hecho, cuando $p\,(h) \neq 0$, y $0 < p\,(e \mid h) < 1$, no puede ocurrir que $p\,(h \mid \neg e) = 0$. Si además de esto se cumple que $p\,(h) \neq 1$, tampoco ocurrirá que $p\,(h \mid e) = 1$. O sea que si $p\,(h)$ y $p\,(e \mid h)$ están entre cero y uno, entonces $p\,(h \mid e)$ y $p\,(\neg h \mid e)$ nunca podrán valer cero, con lo cual $e$ no puede refutar concluyentemente ni $h$ ni $\neg h$. Es de suponer que las hipótesis ya descartadas poseerán una probabilidad baja, igual o cercana a cero, a la luz de la evidencia actualmente disponible. En todo caso, lo que no parece acertado desde el bayesianismo es entender la contrastación de hipótesis como un proceso que establece un salto cualitativo entre lo falsado (probabilidad igual a cero absoluto) y lo que no lo está.

[19] $p\,(h \mid \neg e) = \dfrac{p\,(h) \cdot p\,(\neg e \mid h)}{p\,(h) \cdot p\,(\neg e \mid h) + p\,(\neg h) \cdot p\,(\neg e \mid \neg h)}$. Como $p\,(\neg e \mid h) = 0$, $p\,(h \mid \neg e)$ también valdrá cero.

Más bien estaríamos ante un rango de valores probabilísticos que fluctúan en función de la evidencia que vamos obteniendo, salvo en aquellos casos particulares en que ha habido una refutación concluyente en el sentido antes expuesto.[20]

Hechas estas aclaraciones, queda una alternativa más consecuente con el bayesianismo que consistiría en entender el éxito en términos de probabilidad: las hipótesis más plausibles, más atractivas desde un punto de vista explicativo, son también más probables, que las demás. La plausibilidad/probabilidad inicial de las hipótesis recién propuestas guarda entonces íntima relación con la probabilidad atribuida a las hipótesis del pasado. La lectura frecuencialista de las probabilidades iniciales dice que cuando un científico emite un juicio sobre la plausibilidad inicial de una hipótesis $h$ en un momento $t$, lo que hace es extrapolar el éxito obtenido en el pasado por hipótesis semejantes a $h$. De acuerdo con nuestra interpretación tener éxito es sinónimo de elegir las hipótesis condicionalmente más probables. La justificación del juicio científico dependerá del grado en que la estimación del valor de la probabilidad inicial de $h$ se aproxime a la probabilidad, *condicionada a la evidencia*, de las hipótesis del pasado semejantes a $h$ en cuanto a su mérito explicativo.

Pero, ¿a qué evidencia habremos de referirnos? ¿a la que había en la época en que las hipótesis fueron planteadas y desarrolladas, o a la disponible actualmente? Gracias a un instrumental cada vez más sofisticado, capaz de obtener información de ámbitos de la realidad antes desconocidos, el registro observacional actual corrige y acrecienta el de cualquier época pasada, tanto en amplitud como en profundidad. Aunque se puede discutir si ha habido progreso teórico en la ciencia, cuestionar que ha habido un progreso en el ámbito observacional está fuera de lugar. Así pues, referir la probabilidad condicionada de las hipótesis del pasado a la evidencia ahora disponible equivale, en la práctica, a adoptar la mejor perspectiva que podemos tener, aun a sabiendas de que el cuerpo evidencial constantemente se está modificando.

Recuérdese que la posición de Salmon se propuso como alternativa al subjetivismo radical acerca de las probabilidades iniciales. Estamos de acuerdo con él en que no basta con la consistencia con los teoremas de la teoría de la probabilidad. También compartimos la idea, en buena lógica bayesiana, de que la asignación de probabilidades iniciales debe tener en cuenta toda información que pueda ser relevante. Aunque aquí los avatares de las hipótesis del pasado resultan decisivos, nosotros preferimos

[20] No faltan quienes piensan que esto es un defecto del bayesianismo a la hora de dar cuenta de la "lógica" de la justificación en la ciencia. Véase Kelly y Glymour (2004).

hacer hincapié en el "éxito probabilístico", en vez de en la "frecuencia de éxito". El matiz es importante ya que, como veremos después, apelar al éxito probabilístico de las hipótesis del pasado ni implica apelar a una frecuencia, ni es incompatible con ello. Pero nuestra propuesta requiere mayor concreción. ¿Qué significa exactamente que una hipótesis del pasado ha tenido "éxito probabilístico"? ¿cómo justificar las asignaciones de probabilidad inicial, las preferencias manifiestas de los científicos en sus apreciaciones de plausibilidad, apelando a otras probabilidades? ¿cómo determinar qué valores para las probabilidades iniciales son admisibles a partir del éxito probabilístico? Intentaré responder a estas preguntas.

Sea $H$ el conjunto de todas las hipótesis $h_i$ formuladas por los científicos a lo largo de la historia y $E$ la evidencia disponible *en la actualidad*. Sea $L$ un subconjunto de $H$ que contiene todas las hipótesis explicativamente valiosas y $L'$ su complementario (el subconjunto de hipótesis que no lo son). Lo que sigue son tres modos de entender el éxito probabilístico de las hipótesis del pasado:

(a) Sea $\kappa$ un número real positivo tal que $0 < \kappa < 1$ que designa la probabilidad de la hipótesis más probable de $L'$. Entonces, para cualquier $h_i \in L$, $p\,(h_i \mid E) > \kappa$.

(b) Sea $\lambda$ un número racional positivo tal que $0 < \lambda < 1$. Sean $f_L$ y $f_{L'}$ las frecuencias relativas de hipótesis de $L$ y $L'$ respectivamente cuya probabilidad es mayor que $\lambda$, esto es:

$$f_L = \frac{|\{h_i \in L: p\,(h_i \mid E) > \lambda\}|}{|L|} \quad \text{y} \quad f_{L'} = \frac{|\{h_i \in L': p\,(h_i \mid E) > \lambda\}|}{|L'|}$$

Entonces, $f_L > f_{L'}$.

(c) Sean $m_L$ y $m_{L'}$ la probabilidad media de las hipótesis de $L$ y $L'$, respectivamente, condicionada a la evidencia actualmente disponible, esto es:

$$m_L = \frac{\sum_{hi \in L} p\,(h_i \mid E)}{|L|} \qquad m_{L'} = \frac{\sum_{hi \in L'} p\,(h_i \mid E)}{|L'|}$$

Entonces, $m_L > m_{L'}$.

Tenemos, pues, tres formulaciones distintas de la idea general de que las $L$-hipótesis son más probables a la luz de la evidencia actual que las $L'$-hipótesis. Nótese que las desigualdades son, en todos los casos, estrictas. Por eso cada una de estas conjeturas respecto a los valores probabilísticos de las hipótesis del pasado establece una condición *suficiente* al menos para justificar una asignación de probabilidad inicial más alta a una $L$-hip que a una $L$'-hip. Discutiré a continuación cuál de las formulaciones es preferible.

La primera interpretación plantea un requisito demasiado fuerte. Pensemos en hipótesis descartadas hace mucho tiempo que fueron –y siguen siéndolo, excepto en lo tocante a su rendimiento observacional– brillantes desde un punto de vista explicativo. Su probabilidad condicionada a la evidencia hoy disponible sería baja, e incluso podría ser igual a cero. Por otro lado, en la actualidad podemos encontrar ejemplos de hipótesis que no se incluirían en $L$ y que poseen un éxito observacional notable (en mecánica cuántica, por ejemplo), lo cual incrementaría su probabilidad, a pesar de todo. Sin embargo, (a) nos obliga a sostener que cualquier $L$-hip *siempre* es más probable a la luz de la evidencia actual que cualquier $L$'-hip.

Las alternativas (b) y (c), por su parte, permiten que algunas $L$'-hip sean más probables que algunas $L$-hip, porque la comparación se establece entre las probabilidades de los elementos de $L$ y de $L$' tomados en conjunto, en vez de hacer la comparación entre las probabilidades de pares de elementos tomados aisladamente. En el caso de (b) hemos dejado el parámetro $\lambda$ sin definir. A primera vista parece razonable fijar un umbral de 0'5, porque eso significaría que la proporción de $L$-hip que son más probables que su negación es mayor que la proporción correspondiente de $L$'-hip, pues si $p\ (h_1 \mid E) > 0'5$, entonces $p\ (\neg h_1 \mid E) < 0'5$.[21] Pero aunque $f_L > f_{L'}$, podría ocurrir que $m_{L'} > m_L$. Piénsese en una situación en que la distribución de probabilidades en $L$ muestra una dispersión mayor que en $L$'. Podríamos tener mayor proporción de $L$-hip con probabilidad alta –superior a 0'5– que $L$'-hip, pero si las $L$-hip que no están por encima de 0'5 quedaran muy por debajo de este valor y todas las $L$'-hip están muy próximas a 0'5, podría ocurrir que las hipótesis de $L$', peores *qua* explicaciones, tuvieran por término medio una probabilidad más alta que las de $L$. En tal caso, a la hora de decidir qué probabilidad inicial atribuir a dos hipótesis que aún no han sido contrastadas experimentalmente, una perteneciente a $L$ y otra a $L$', no parece poder justificarse ni el juicio científico ni la asignación de una probabilidad inicial mayor a la primera que a la segunda, que es el resultado que buscamos.

[21] En Achinstein (2001, caps. 6 y 7) se defiende que $p\ (h \mid e) > 0.5$ es condición necesaria para que $e$ constituya evidencia a favor de $h$, o para que $e$ sea una buena razón para creer $h$. Aunque en el contexto de las hipótesis teóricas de la ciencia las cosas pueden ser diferentes, véase Iranzo (2007) sobre el particular.

La alternativa (c) pretende corregir esto justamente, ya que no se ve afectada por la dispersión de los valores de la probabilidad, y sí que da una base para conferir mayor probabilidad inicial a las $L$-hip que a las $L'$-hip. Para justificar tal cosa no se requiere que la probabilidad de las primeras sea alta; basta con que sea *más alta, por término medio*, que la de $L'$, incluso aunque $m_L$ sea bajo, o aunque toda $L$-hip sea mucho más improbable que su negación. La formulación (c) es bastante modesta, como puede verse. No nos obliga a comprometernos con un umbral mínimo de probabilidad; sólo nos dice que, en general, las $L$-hip han funcionado mejor en su confrontación con la evidencia que las $L'$-hip. Pero con esto basta para nuestro propósito de limitar la variabilidad de las probabilidades iniciales. Por eso pienso que (c) resulta la formulación más apropiada.

Nótese que de las tres formulaciones discutidas solamente (b) remite a una *frecuencia* de éxito (probabilístico, eso sí), pero tanto (a) como (c) se alejan de la propuesta de Salmon, ya que, en sentido estricto, no proporcionan una justificación frecuencialista de la probabilidad inicial. Por otro lado, (a), (b) y (c) son conjeturas acerca de la probabilidad de las hipótesis del pasado. En realidad son hipótesis de segundo orden, y como tales deben ser contrastables. En lo tocante a (c), desde luego, no conocemos con exactitud el valor de $m_L$ ni el de $m_{L'}$, y es dudoso que algún día podamos contar con un listado exhaustivo de todas las hipótesis científicas planteadas a lo largo de la historia y, a continuación, su probabilidad a la luz de la evidencia disponible. Lo que tal vez resulte más factible es analizar algunos casos particulares de hipótesis científicas postuladas en el pasado y tratar de calcular su probabilidad de acuerdo con el algoritmo bayesiano; al menos esto no tiene por qué ser más problemático que el cálculo de la probabilidad de una hipótesis contemporánea. El cálculo de $m_L$ y de $m_{L'}$ no exigiría en tal caso determinar las probabilidades de todos los elementos de $L$ y $L'$, bastando con una muestra representativa, de la misma forma que para calcular con un alto grado de fiabilidad la altura media de los estudiantes universitarios argentinos, pongamos por caso, tampoco es necesario medirlos a todos.

Una vez explicitado qué hemos de entender por éxito probabilístico de las hipótesis del pasado, podemos pasar a la cuestión de cómo se fija el valor de las probabilidades iniciales. La alternativa (c) nos sugiere la aplicación de una regla sencilla, a saber, atribuir como probabilidad inicial a la hipótesis objeto de discusión, $h$, el valor de la probabilidad media del conjunto al que $h$ pertenezca, sea éste $L$ ó $L'$. Así,

si $h \in L$ y $m_L = x$, entonces $p(h) = x$
si $h \in L'$ y $m_{L'} = y$, entonces $p(h) = y$

Que $x$ sea mayor, menor o igual que $y$ no puede determinarse a priori, obviamente. Pero si las preferencias de los científicos van bien encaminadas, y resultan ser indicadores fiables del éxito probabilístico de las teorías, hemos de esperar que $x > y$. La investigación histórica debería respaldar este extremo.

En la sección 2 comentamos distintas estrategias para constreñir los valores de las probabilidades iniciales. En aquel momento no nos referimos a lo razonable que puede resultar invocar un control externo que moldee las probabilidades (iniciales) del sujeto. Y es que, si sé, por el informe meteorológico de hoy, que es sumamente probable que mañana llueva, parece que mi grado de creencia sobre si va a llover mañana debería acomodarse a la previsión hecha por los especialistas. H. Gaifman acuñó el término "función experta" para referirse a este tipo de constricción (Gaifman 1988):

$$p\ (A \mid q\ (A) = x) = x)$$

Esto es, dado que $x$ es la probabilidad de que ocurra el suceso $A$ para el experto –que simbolizamos aquí como la función $q$ $(A)$–, para mí la probabilidad de que llueva también debe ser $x$. Generalizando la idea, podemos considerar una función experta cualquier función que cumpla el papel de $q$. Visto así, son funciones expertas el *Principio de Reflexión* de B. van Fraassen:

$$p_t\ (A \mid p_{t+\Delta}\ (A) = x) = x)$$

donde $p_{t+\Delta}$ $(A)$ es la función de probabilidad futura, es decir, la que el sujeto atribuirá a $A$ en un tiempo $t + \Delta$, y también el *Principio Principal* de D. Lewis:

$$p\ (A \mid ch\ (A) = x) = x)$$

donde $ch$ $(A)$ es la probabilidad objetiva (*chance* en inglés) del suceso $A$.[22]

El Principio de Reflexión y el Principio Principal refieren a sucesos. Nuestra regla refiere, en cambio, a hipótesis científicas, o sea, a proposiciones. Pero nuestra regla también es una función experta particular, ya que $m_L$ y $m_{L'}$ son probabilidades, aunque su ámbito de aplicación quede restringido a un ámbito concreto, como es el de las hipótesis científicas:

[22] van Fraassen (1984 y 1995) y Lewis (1980).

$$si\ h \in L, \rightarrow p\ (h \mid m_L = x) = x$$
$$si\ h \in L', \rightarrow p\ (h \mid (m_{L'} = y) = y$$

Es obvio que el fundamento de nuestro principio es distinto al de los anteriores. Pero también hemos de distinguirlo de un principio frecuencialista como el principio de Probabilidad Directa (véase Hájek 2009):

$$p\ (A \mid frec\ (A) = x) = x)$$

donde *frec* (*A*) es la proporción o frecuencia relativa de elementos con la propiedad *A* en una clase de referencia determinada.[23] Para cualquier frecuencia relativa $f$, $0 \leq f \leq 1$. Por eso las frecuencias relativas pueden expresarse como probabilidades. Aplicado en nuestro contexto, dicho principio afirmaría que:

$$si\ h \in L, \rightarrow p\ (h \mid f_L = x)) = x$$
$$si\ h \in L', \rightarrow p\ (h \mid f_{L'} = y)) = y$$

lo cual se asemeja bastante a la opción (b), antes discutida y descartada.

Recordemos el ejemplo de la enfermedad comentado en la sección 1. La propuesta frecuencialista de Salmon trata de asimilar un caso problemático en la aplicación del algoritmo bayesiano –el cálculo de la probabilidad inicial de una hipótesis teórica– a los casos fáciles, como el de la enfermedad, donde tenemos información exacta sobre las probabilidades iniciales y la verosimilitud. La idea es tomar la probabilidad inicial como una frecuencia en ambos casos (prevalencia de la enfermedad en la población, en un caso, frecuencia de éxito, en otro). Por lo que se acaba de decir, ésta no es nuestra estrategia. No obstante, a pesar de introducir la noción de éxito probabilístico, podríamos haber continuado en la estela frecuencialista de Salmon, apelando en nuestro caso a la *frecuencia* de éxito *probabilístico*, en la dirección apuntada por la opción (b). Pero al decidirnos por (c) descartamos que el factor relevante para constreñir las probabilidades iniciales sea alguna frecuencia, ni siquiera la frecuencia de éxito probabilístico. El asunto es que $m_L$ y $m_{L'}$ *no refieren a la frecuencia de hipótesis exitosas, sino a una media de probabilidades*. Y estas últimas probabilidades tampoco son frecuencias; son las probabilidades de las hipótesis anteriores a *h* condicionadas a la evidencia actual.

[23] Algún autor ha recogido esta idea bajo el término "calibración": las *credences* de un sujeto "bien calibrado" reflejarían frecuencias empíricas. Sobre cómo incorporar la noción de calibración al aparato bayesiano, véase van Fraassen (1983) y Lange (1999).

Respecto al relativismo de las probabilidades iniciales, nuestra posición es afín a la de Salmon. Imaginemos dos científicos que parten de atribuciones de probabilidad inicial diferentes para $h$, y que optan por recoger evidencia experimental para confirmarla. Tras ir reajustando la probabilidad de $h$ según el Principio de Condicionalización durante un periodo de tiempo razonablemente largo, puede ocurrir que los valores converjan, es decir, que la diferencia entre los valores de $p\ (h \mid e_1 \wedge e_2 \wedge e_3 \wedge \ldots)$ obtenido por cada uno de los sujetos haya disminuido progresivamente tanto como queramos hasta quedar por debajo de una cantidad mínima. A efectos prácticos la discrepancia quedaría disuelta. Que se alcance este punto depende de varios factores, básicamente de la magnitud de la diferencia inicial y de las posibilidades de acumular más evidencia para neutralizar el peso de las probabilidades iniciales. Pero si esto no se consigue, el bayesiano ortodoxo no podrá decir sino que en condiciones ideales los valores de ambos sujetos tendrían que converger, lo que aquí resulta ser un flaco consuelo, puesto que desde su posición radicalmente subjetivista cualquier atribución inicial de probabilidades es correcta, con la única restricción de que no se viole ningún teorema de la teoría de la probabilidad. Sin embargo, tanto el frecuencialismo de Salmon como nuestra propuesta critican la concepción personalista radical de las probabilidades iniciales, al menos en el contexto científico, y aspiran a concluir algo que le está vedado al bayesiano ortodoxo, a saber, que al menos uno de los bandos ha concedido a $h$ un valor inicial muy alejado de la probabilidad media que las hipótesis semejantes a $h$ tienen a la luz de la evidencia actual (o de la frecuencia de éxitos, según Salmon). En nuestro caso sustanciar esta afirmación depende de lo preciso que sea el cálculo de $m_L$ y $m_{L'}$. Hemos aludido de pasada a la complejidad de esta tarea, pero por el momento podemos dejar esto a un lado, ya que parece prioritario discutir las objeciones que invalidarían nuestro planteamiento de raíz, con independencia de las complicaciones técnicas, por así decirlo, para llevarlo a cabo.

## 4. Discusión de algunas objeciones a nuestra propuesta

Las objeciones que discutiremos en esta sección son las siguientes:

(a) Si la probabilidad inicial atribuida por nuestra regla a cualquiera de las hipótesis que formulemos en el presente es baja, ¿no resultará demasiado sencillo encontrar evidencia para incrementar dicha probabilidad? Con tal facilidad para encontrar evidencia confirmadora, ¿no se desvirtúa por completo el proceso de contrastación empírica de las hipótesis científicas?

(b) La probabilidad inicial de una hipótesis actual requiere calcular la probabilidad, condicionada a la evidencia actual, de las hipótesis del pasado. Pero, según el teorema de Bayes, esta última probabilidad incluye a su vez la probabilidad inicial de la hipótesis antigua, que no sería sino la probabilidad de dicha hipótesis en relación a la evidencia disponible *en el momento en que fue formulada*. El cálculo de ésta última probabilidad obliga a aplicar nuestra regla de nuevo, sólo que referida a la evidencia disponible en un contexto histórico del pasado, y así sucesivamente. ¿Cómo hacer frente a esta regresión?

(c) Si la probabilidad inicial de una hipótesis actual se establece teniendo en cuenta su semejanza con las hipótesis del pasado, ¿cómo determinar la probabilidad inicial de una hipótesis novedosa que rompe con la tradición?

Veámoslas por este orden.

*(a) La confirmación de las hipótesis del presente*

Según el criterio bayesiano de confirmación (véase *supra* sección 1) la evidencia confirma la hipótesis sii $p\ (h \mid e) > p\ (h)$, la disconfirma, sii $p\ (h \mid e) < p\ (h)$, y es neutral sii $p\ (h \mid e) = p\ (h)$.

Visto así, la confirmación no depende del valor absoluto de la probabilidad de la hipótesis en relación a la evidencia, sino del incremento sobre la probabilidad inicial de la hipótesis provocado por la evidencia. De ahí que se diga a veces que el bayesianismo se compromete con una noción de *confirmación incremental*. Por otra parte, según nuestra regla, para una hipótesis virtuosa $h^*$ –una $L$-hipótesis, por tanto– que se proponga en el momento presente su probabilidad inicial es $m_L$, así que $p\ (h^*) = m_L$. Nuestra hipótesis de segundo orden es que $m_L > m_{L'}$, pero la regla para fijar la probabilidad inicial no exige un umbral mínimo para $m_L$. Ahora bien, si $m_L$ es muy baja, el efecto de prácticamente cualquier evidencia sobre $h$ sería un incremento respecto a su probabilidad inicial, con lo cual $h^*$ se vería *confirmada* de una forma cuasi-automática, fuera cual fuera $e$. En términos probabilísticos,

$$p\ (p\ (h^* \mid e) > p\ (h^*) \mid m_L \cong 0) \cong 1$$

El asunto aún puede ser peor, ya que un razonamiento parecido nos permitiría concluir no sólo que la confirmación de la hipótesis se producirá con gran probabilidad sea cual sea $e$, sino también que eso ocurrirá ¡sea cual sea $h$! Veamos por qué. Supuestamente, $m_L > m_{L'}$. Por tanto, por el mismo razonamiento debe ser más fácil que se incremente la probabilidad inicial de una mala hipótesis ya que su probabilidad inicial es aún más baja. Con lo cual, tanto si $h$ es buena como si es mala, casi cualquier $e$ será evidencia confirmadora en tanto la condicionalización sobre $e$, o sea, $p\ (h \mid e)$, tendrá como efecto sumamente probable el incremento de la probabilidad respecto de $p\ (h)$.

La objeción, de ser correcta, constituiría una reducción al absurdo de nuestra propuesta.[24] Tras replantearla en términos formales, argumentaré a continuación que sus consecuencias no son tan graves como parece a simple vista. Anticipo que la distinción entre la evidencia disponible en el presente, $E$, y cualquier otra evidencia adicional $e$ que pueda invocarse resultará crucial en mi réplica.

Nuestra regla dice que la probabilidad inicial de una $L$-hipótesis nueva, $h^*$, debe ser igual a $m_L$:

$$p\ (h^*) = m_L = \frac{\sum_{hi \in L} p\ (h_i \mid E)}{\mid L \mid} \tag{3}$$

¿Cuál será entonces la probabilidad de $h^*$ *dada la evidencia actualmente disponible* $E$? Del Teorema de Bayes (véase *supra*, sec. 1, expresión (1)), y (3) se sigue que:

$$p\ (h^* \mid E) = \frac{m_L \cdot p\ (E \mid h^*)}{p\ (E)} \tag{4}$$

En la versión del Teorema de "la probabilidad total" (véase *supra*, sec. 1, expresión (2) y también nota 19), para una partición constituida por dos hipótesis, $h^*$ y $\neg h^*$:

$$p\ (h^* \mid E) = \frac{m_L \cdot p\ (E \mid h^*)}{m_L \cdot p\ (E \mid h^*) + (1 - m_L) \cdot (p\ (E \mid \neg h^*)} \tag{5}$$

La media para fijar la probabilidad inicial de las $L$-hipótesis que se formulen con posterioridad a $h^*$ ya no sería $m_L$, puesto que la propia $h^*$ debería ser incluida en $L$. El efecto de $h^*$ sobre $m_L$ vendría dado, entonces, por la siguiente expresión:

$$m_{L+h^*} = \frac{\sum_{hi \in L} p\ (h_i \mid E) + p\ (h^* \mid E)}{\mid L \mid + 1} \tag{6}$$

[24] La objeción fue planteada por uno de los árbitros de *Análisis Filosófico*. En el texto he tratado de apurar al máximo sus consecuencias contraintuitivas.

Si las probabilidades, condicionadas a la evidencia presente, de las $L$-hipótesis del pasado fueran muy bajas, algunas incluso iguales a cero por haber sido concluyentemente refutadas, el sumatorio de las probabilidades condicionadas de todas las $L$-hipótesis del pasado podría tener un valor muy bajo, cercano a cero.[25] O sea,

$$\sum_{hi \in L} p\ (h_i \mid E) \cong 0 \qquad (7)$$

De (6) y (7) se sigue que

$$m_{L+h^*} \cdot (\mid L \mid + 1) \cong p\ (h^* \mid E)$$

Es fácil ver que $p\ (h^* \mid E)$ ha de ser bastante mayor que $m_{L+h^*}$, dado que el factor $(\mid L \mid + 1)$, que refiere al número de L-hipótesis formuladas a lo largo de la historia de la ciencia, tendrá al menos dos dígitos.[26] Pero si $p\ (h^* \mid E)$ es mayor que $m_{L+h^*}$, también será mayor que $m_L$, ya que esto ocurrirá sólo con que $p\ (h^* \mid E)$ sea ligeramente mayor que $m_L$, lo cual parece bastante lógico, pues sería absurdo pensar que las hipótesis que prefieren los científicos en el presente son más improbables en relación a la evidencia actual que las que se han descartado hace tiempo. En suma, como $p\ (h^* \mid E) > m_L$ y $p\ (h^*) = m_L$, según (3), *la hipótesis nueva $h^*$ será confirmada por E, esto es, por la evidencia disponible en el momento actual, tal como afirma el criterio básico de confirmación bayesiana* (véase *supra* p. 3). Y eso es independiente del contenido concreto de $E$.

¿Es esto grave? Pienso que no. Conviene no olvidar que el hecho de que $E$ confirme la hipótesis no implica, desde luego, que la hipótesis sea más probable que su negación, ni que debamos creer que es correcta, ni siquiera que debamos aceptarla provisionalmente. Confirmación en sentido bayesiano significa tan solo que se ha producido un *incremento*

[25] Nótese que no podemos tomar en serio la posibilidad de que dicho término valga exactamente cero, porque entonces ni siquiera podríamos plantear la objeción. El motivo de esto es que en tal caso $m_L = 0$, y entonces, dado que $p\ (h^*) = m_L$, por el teorema de Bayes tenemos que $p\ (h^*/e) = 0$. Con lo cual ya no podemos decir que cualquier evidencia será seguramente confirmadora para $h^*$; lo que debería decirse entonces es que toda evidencia ulterior será neutra ya que la probabilidad seguirá siendo 0 en cualquier caso (véase *supra*, p. 52, segundo párrafo).

[26] Aquí puede verse cómo eliminar el sumatorio no es un cambio baladí, aunque sea una cantidad infinitesimal. La razón es que si $|L|$ fuera de un orden de magnitud muy superior a $m_L$, $p\ (h^* \mid E)$ podría ser mayor que 1, lo cual viola el principio básico de que el valor de cualquier probabilidad debe pertenecer al intervalo [0, 1]. Proseguiremos suponiendo que una investigación histórica cuidadosa arrojaría valores para $m_L$ y $|L|$ compensados de tal modo que aquella posibilidad quedaría excluida.

*respecto de la probabilidad inicial*, aunque la probabilidad de $p\ (h \mid E)$ puede seguir siendo muy baja en términos absolutos, como consecuencia de que $p\ (h)$ también lo sea.[27] En términos más generales, que una evidencia confirme a una hipótesis depende exclusivamente, véase por ejemplo la expresión (4), del valor de la razón $p\ (e \mid h) / p\ (e)$. Cuando sea mayor que 1, $e$ será evidencia confirmadora para $h$. Intuitivamente, una evidencia $e$ confirma a $h$ cuando $e$ es más probable contando con la verdad de $h$ que sin contar con ella. Queda claro, pues, que el carácter confirmador o no de la evidencia es completamente independiente de la probabilidad inicial de $h$ (o sea, de $m_L$).

En el caso concreto de que la evidencia sea la evidencia disponible, $E$, el incremento cuasi-automático de $p\ (h^* \mid E)$ respecto de $p\ (h^*)$ encaja bien, por lo demás, con la práctica usual en la ciencia. Cuando formulan una hipótesis nueva, los científicos no buscan hipótesis meramente consistentes con $E$. Lo que se persigue, en principio, son hipótesis que expliquen $E$, que hagan creíble $E$, y en el caso ideal, hipótesis que mantengan una conexión deductiva con $E$. En el vocabulario probabilístico esto equivale a preferir hipótesis con una *alta verosimilitud*, o sea, hipótesis para las cuales $p\ (E \mid h_i)$ sea alta y para las que $p\ (E \mid \neg h_i)$ sea bajo, de acuerdo con (5). Por tanto, que en relación a $E$ la hipótesis novedosa sea confirmada, es algo que no debería extrañarnos. El valor de la razón $p\ (E \mid h_i) / p\ (E)$, que es lo que decide el carácter confirmador de $E$, es directamente proporcional a la verosimilitud, luego conforme aumente ésta, más fácil es que el valor sea mayor que 1, y por tanto, que $E$ confirme la hipótesis en cuestión.[28]

Así pues, bajo el razonable supuesto de que $p\ (E \mid h^*) > p\ (E)$, sí podemos concluir del razonamiento anterior que "el mero hecho de que

[27] Sea $m_L = p\ (h^*) = 0'02$, $p\ (E \mid h^*) = 1$ y $p\ (E \mid \neg h^*) = \frac{1}{2}$. Aplicando (5), $p\ (h^* \mid E) = 0'039$. Se ve cómo $E$ confirma a $h^*$, puesto que $p\ (h^* \mid E) > p\ (h^*)$, aunque $h^*$ sigue siendo altamente improbable en términos absolutos.

[28] Cuando la evidencia es conocida surge el "problema de la evidencia antigua" (*old evidence problem*). Supuestamente, si la evidencia es conocida, $p\ (E\ ) = 1$. Pero si el denominador de (4) vale uno es fácil comprobar cómo, en el mejor de los casos, o sea cuando $p\ (E \mid h) = 1$, $E$ sería evidencia neutral para $h$, ya que $p\ (h \mid E) = p\ (h)$. La evidencia conocida $E$, a pesar de ser implicada deductivamente por $h$, ni la confirma ni la disconfirma, lo cual resulta bastante contraintuitivo. Así, la precesión del perihelio de Mercurio no debería contar a favor de la Teoría General de la Relatividad por el hecho de ser conocida con anterioridad a su formulación. En Howson (2003, pp. 193-194), se argumenta, apoyándose en un ejemplo concreto que, aunque la evidencia sea conocida, $p\ (E)$ no tiene por qué ser igual a 1. Mis comentarios dan por buena una solución al problema de la evidencia antigua en esta línea.

una hipótesis $h^*$ pertenezca a $L$ determina que deba ser confirmada *por la evidencia E disponible en el momento en que se formula la hipótesis*". Cabe esperar que esto se cumpla con independencia del contenido particular que tenga $E$. La única condición es que $E$ contenga el cuerpo evidencial disponible en el momento en que se formula la hipótesis en cuestión. Y también podría ocurrir esto con una hipótesis nueva artificiosa y compleja, perteneciente a $L'$ por tanto, y construida para encajar de un modo *ad-hoc* con $E$. Su verosimilitud será elevada, y por las mismas consideraciones que hemos aducido en el caso de h*, cabe esperar que su probabilidad condicionada a $E$ sea mayor que $m_{L'}$.[29]

Ahora bien, y aquí acaba nuestra réplica a la objeción (a), lo que no podemos concluir de lo anterior, es que "el mero hecho de que una hipótesis $h^*$ pertenezca a $L$ determina que deba ser confirmada *por cualquier evidencia posible e* ". Desde luego que esto es inaceptable para cualquier teoría de la confirmación, pero es que ése no es nuestro problema. Imaginemos una consecuencia observacional $e$ de nuestra hipótesis $h^*$ que no forma parte del cuerpo evidencial aceptado $E$ y que planteamos $e$ contrafácticamente ("si realizáramos tal experimento obtendríamos $e$"). Como ya se ha dicho, que $e$ confirme o no confirme a $h^*$, dependerá exclusivamente del valor de la razón $p\ (e \mid h^*) / p\ (e)$, con independencia de cuál sea el valor de $m_L$.

Resumiré mi réplica a la objeción (a). Que la hipótesis nueva se vea confirmada por la evidencia disponible no debería preocuparnos. Además de no desentonar con la práctica científica usual, es una consecuencia natural del planteamiento que hemos defendido aquí, ya que: (i) la noción bayesiana de confirmación es puramente incremental; y (ii) la hipótesis nueva se formula para encajar con la evidencia presente. Pero de esto no se sigue, sin embargo, que cualquier evidencia ulterior vaya a ser evidencia confirmadora para la hipótesis en cuestión.

*(b) La probabilidad inicial de las hipótesis del pasado*

Si nos tomamos en serio la tarea de precisar $m_L$ y $m_{L'}$ para fijar la probabilidad inicial de una hipótesis contemporánea, hemos de calcular la probabilidad, condicionada a la evidencia actual, de otras hipótesis formuladas anteriormente. Al aplicar el teorema de Bayes nos encontramos de nuevo con una probabilidad inicial, en este caso la de las hipótesis del

[29] Un ejemplo de esto son las complejidades introducidas en el modelo ptolemaico (epiciclos, ecuantes, deferentes,…) para "salvar los fenómenos". Estos añadidos convirtieron a la hipótesis geocéntrica en una hipótesis artificiosa y poco atractiva, pero no por eso menos *verosímil*, según el significado del concepto de verosimilitud en este contexto.

pasado. Dándoles el mismo trato que a cualquier hipótesis contemporánea deberíamos decir que su probabilidad inicial remite a la plausibilidad que merecieron en el contexto histórico en el que se formularon, es decir, para los científicos de la época. Pero, siguiendo con el razonamiento, calcular y justificar la probabilidad inicial de la hipótesis antigua exige remontarnos más atrás, ya que los criterios de plausibilidad vigentes en aquella época recogerían el éxito probabilístico de hipótesis anteriores. Parece que ha de llegar un momento en que no tendremos un repertorio lo bastante amplio de casos para determinar la probabilidad inicial correspondiente. Y bien, ¿cómo fijar las probabilidades iniciales de las primeras hipótesis científicas? ¿Y si hubieran sido establecidas arbitrariamente o de un modo incorrecto?

Esta objeción apunta a una doble dificultad: la necesidad de parar la regresión al pasado y la posibilidad de que las asignaciones iniciales fueran incorrectas. Una vez hemos dejado de lado la tesis de la convergencia (véase *supra* sec. 2), tan cara a los bayesianos ortodoxos, ya no podemos responder que no hay que preocuparse por lo que sólo es un requisito formal para poner en marcha el algoritmo, porque los valores de las probabilidades atribuidas a las primeras hipótesis científicas se diluirán conforme aumente el número de hipótesis que vamos incorporando en el cálculo de $m_L$ y $m_{L'}$. Si, por el contrario, pensamos que el efecto acumulativo de la evidencia no conduce necesariamente a borrar las diferencias entre las probabilidades iniciales, ¿por qué pensar que dos sujetos que partan de valores iniciales diferentes acabarán obteniendo dos pares de valores idénticos, o muy similares, para $m_L$ y $m_{L'}$?

La pregunta por la probabilidad inicial de la primera hipótesis científica parece presuponer con una frontera precisa entre las hipótesis que merecen el calificativo de 'científicas' y las que no. Preguntar cuál fue la primera hipótesis científica de la historia seguramente carece de sentido. Aun así, hemos de admitir que remontándonos hacia atrás llegaríamos a un estadio en que las probabilidades iniciales no podían reflejar el éxito probabilístico de hipótesis científicas anteriores porque, simplemente, no existía nada anterior reconocible como ciencia. Pero esto no implica admitir que en esta situación cualquier estimación de probabilidad inicial hubiera sido correcta. En la fase que podemos llamar protocientífica, cuando todavía no hay un conjunto de procedimientos codificados y aceptados por la comunidad científica, los criterios que guían las preferencias de los científicos vienen determinados por el éxito obtenido en campos afines. Así, los conocimientos precientíficos de carácter técnico van acompañados de ciertas conjeturas respecto a los mecanismos que explican y gobiernan los procesos que ocurren a nuestro alrededor. La idea es que los criterios que se ponen en juego para comparar la probabilidad inicial de las primeras hipótesis científicas, sean cuales sean, no surgen repentinamente, sino que son criterios previamente aplicados en campos anejos. Las generalizaciones inductivas, las hipótesis analógicas y los esbozos de explicación teórica que han dado resultado en esos campos, son la base desde la que se generan unos criterios de plausibilidad.

Con esto no contradecimos lo dicho hasta aquí. El investigador, cuando hace una estimación de la probabilidad inicial de una hipótesis teórica, está recogiendo el saber metodológico acumulado por las generaciones que le han precedido. No aplica el algoritmo bayesiano para calcular la probabilidad inicial; lo que hace en realidad es juzgar la plausibilidad de la hipótesis estimando implícitamente su *semejanza* con otras hipótesis que forman parte de la tradición histórica de su disciplina. La probabilidad inicial variará en función de si la hipótesis se parece o no a las que han ido acompañadas de éxito. De esta manera la intuición del científico orienta las preferencias hacia unas hipótesis en vez de otras. Y si con el paso del tiempo no obtenemos el rendimiento esperado, son las intuiciones científicas, esto es, los criterios de lo que cuenta como una hipótesis plausible, los que habrán de ser modificados. En realidad, si describimos el proceso desde la perspectiva del científico, la plausibilidad está vinculada antes con el razonamiento analógico que con el razonamiento probabilístico en el sentido estrictamente matemático. Pues bien, en la fase protocientífica a la que nos hemos referido, la semejanza no sería respecto de hipótesis científicas anteriores, sino respecto de conjeturas aplicadas satisfactoriamente en otros campos.

*(c) La probabilidad inicial de las hipótesis novedosas*

En la objeción (a) nos ocupamos de las hipótesis *nuevas*, las que se formulan en el presente. Ahora nos referiremos a las hipótesis *novedosas*, aquellas que se formulan en el presente y que, además, destacan por su originalidad y carácter innovador.

Según nuestra regla, para fijar la probabilidad inicial de $h$ resulta imprescindible, precisar si $h$ pertenece a $L$ o a $L'$. Decíamos antes que ello depende de que $h$ satisfaga aquellas propiedades de las hipótesis y teorías que se consideran virtuosas. Nuestro modelo es simple, pero puede complicarse sin alterar las ideas básicas. Para empezar hemos apelado a una clasificación dicotómica ($L$ y $L'$) de las hipótesis científicas que seguramente es demasiado rígida. En función de la precisión que podamos dar a los elementos que constituyen la plausibilidad inicial podremos afinar más aquí. Además, no es descabellado pensar que los criterios de plausibilidad sufran variaciones de unas áreas científicas a otras. Entonces $m_L$ y $m_{L'}$ serían relativas a una disciplina o área de la ciencia. Antes dijimos que, para fijar la probabilidad inicial, la probabilidad condicionada aconsejable es la que tiene en cuenta la evidencia actual, aunque se trate de hipótesis del pasado, ya que el repertorio evidencial es ahora más amplio que nunca. Pero la ampliación del cuerpo evidencial es una búsqueda teóricamente guiada, ya que el instrumental moderno empleado a tal efecto se diseña dando por buenas ciertas teorías recientes (piénsese en la tomografía por emisión de positrones, o en los detectores de neutrinos, por ejemplo). Dichas teorías, además de contar con respaldo evidencial, fueron preferidas en las fases iniciales de la investigación según los criterios de plausibilidad vigentes en nuestra época. Por eso parece mejor recurrir, en principio, a los criterios contemporáneos, que a los que fueron operativos en otras épocas. A fin de cuentas son los que se han aplicado para preferir las teorías que subyacen a unas técnicas instrumentales muy eficaces en cuanto a la ampliación de la evidencia a nuestro alcance.

Así pues, una hipótesis pertenece a $L$ porque satisface los criterios de plausibilidad actualmente exigidos. Éste es el respecto en el que las hipótesis, del pasado o del presente, deben juzgarse para determinar si han de ser incluidas en $L$ o en $L'$, del mismo modo que el cuerpo evidencial como referencia para las probabilidades condicionadas debe ser el actual, y no el de la época de Galileo. Tanto el cuerpo evidencial como los criterios de plausibilidad sufrirán variaciones con el tiempo. Aun así, se trata de algo tan simple como optar por lo mejor que tenemos por ahora.

Es de suponer que al incluir una hipótesis en $L$ o en $L'$ la ubicamos en el grupo al que más se asemeja. Pero entonces, ¿qué hacer ante una hipótesis científica radicalmente novedosa? ¿en qué basarnos para atribuirle una probabilidad inicial? ¿cómo valorar, en fin, la innovación científica genuina?

Cabe hacer notar primero que una hipótesis puede ser radicalmente novedosa en su contenido, pero eso no impide que haya una buena base para estimar su probabilidad inicial. Téngase en cuenta que los criterios de plausibilidad apelan a propiedades cuasi-formales o sumamente generales, de manera que hipótesis muy diferentes en cuanto a su contenido pueden ejemplificarlas. Por poner ejemplos concretos de leyes teóricas, la ley de Hooke para los resortes y la segunda ley de Mendel, también llamada ley de segregación de los caracteres independientes, pueden compartir la simplicidad matemática, y eso a pesar de que versen sobre asuntos totalmente distintos y carezcan de un contenido común. Y si esto ocurre en el caso de leyes pertenecientes a áreas de la ciencia diferentes, más razón habrá para pensar que las cosas pueden ser así cuando nos movemos en un área concreta de la ciencia. Antes de Einstein nadie había pensado en una teoría métrica de la gravitación. En cuanto a su contenido es innegable, pues, la ruptura con el modelo newtoniano. Sin embargo, no parece que ello llevara aparejado un cambio total respecto a los criterios de plausibilidad invocados (poder explicativo, capacidad unificadora, simplicidad,…). Antes bien, TGR puede verse como una superación del modelo newtoniano en este sentido (en términos de simplicidad ontológica, por ejemplo).

Pero, ¿qué ocurriría con una hipótesis revolucionaria, ya no tanto respecto a su contenido, sino en cuanto a sus propiedades cuasi-formales? La receta es atenerse a los criterios establecidos, puesto que tienen cierta justificación inductiva. De hecho, una hipótesis que entre en abierta colisión con ellos es dudoso que llegara a ser tomada en serio por la comunidad científica. En cualquier caso, su probabilidad inicial debería ser $m_{L'}$. Una situación más corriente es aquella en que la hipótesis entra en conflicto con alguno de los criterios (véase *supra* nota 15). Entonces, dependiendo del grado de desacuerdo, cabría incluirla en $L$ o en $L'$. Esto no significa que los criterios de plausibilidad sean inmutables, evidentemente. La experiencia también participa en este juego, y una baja probabilidad de partida puede aumentar como efecto de la condicionalización. Si esto ocurriera en repetidas ocasiones con distintas hipótesis, la comunidad científica modificaría sus criterios de plausibilidad, puesto que dejarían de ser indicadores fiables del éxito probabilístico. Lo que se hace aquí no es más que proyectar al futuro lo que *hasta ahora* ha dado resultado. Aun sabiendo que las cosas podrían cambiar, esto resulta bastante razonable, salvo para quien sostenga un escepticismo radical sobre la inducción, ya que no hay mejor punto de partida.

## 5. Conclusiones

Hemos retomado una propuesta de Salmon, que entiende la probabilidad inicial de una hipótesis teórica como una estimación implícita de la frecuencia de éxitos, por dos razones. Primero, porque establece una conexión entre los juicios de plausibilidad de los científicos y la probabilidad inicial, conectando así la confirmación bayesiana con la práctica científica. En segundo lugar, porque nos identificamos con la idea de que la racionalidad o justificación de la probabilidad inicial asignada a una hipótesis científica no puede depender solamente de su compatibilidad lógica con los teoremas de la teoría matemática de la probabilidad. El rendimiento obtenido por otras hipótesis anteriores también cuenta. Esto supone optar por una justificación histórica, inductiva, de la probabilidad inicial. A diferencia de Salmon, sin embargo, propongo entender el éxito en clave probabilística porque me parece que lo más consecuente para una teoría de la inferencia científica basada en la probabilidad, como el bayesianismo, es interpretar el éxito en términos a su vez probabilísticos. Hemos visto, además, que ello no tiene por qué desembocar en una justificación frecuencialista de la probabilidad inicial.

El algoritmo bayesiano se aplica sin problemas en ciertos contextos, tal como vimos en la primera sección. El planteamiento que aquí se defiende para "objetivizar" la probabilidad inicial de las hipótesis teóricas genera su propia problemática, ciertamente. Mi opinión, no obstante, es que si se dispone de un registro histórico lo bastante representativo, los obstáculos técnicos, aun no siendo despreciables, no tienen por qué abrir una brecha insalvable entre las hipótesis teóricas y las hipótesis científicas de un nivel más bajo. A pesar de todo, tal vez nuestra propuesta tropiece con dificultades de principio que impidan siquiera remontar el vuelo. En la sección precedente he sugerido cómo afrontar algunas de ellas. Aunque es posible que haya otras complicaciones de fondo aparte de éstas, quede ese asunto para ser explorado en un futuro.

## Bibliografía


Achinstein, P. (2001), *The Book of Evidence,* Oxford, Oxford University Press.

Arló-Costa, H. (2006), "Rationality and Value: The Epistemological Role of Indeterminate and Agent-Dependent Values", *Philosophical Studies,* 128, pp. 7-48.

Cassini, A. (2003), "Confirmación hipotético-deductiva y confirmación bayesiana", *Análisis Filosófico,* XXIII (1), pp. 41-84.

Cohen, L. J. (1989), *An Introduction to the Philosophy of Induction and Probability*, New York, Oxford University Press.

Earman, J. (1992), *Bayes or Bust. A Critical Examination of Bayesian Confirmation Theory*, Cambridge, The MIT Press.

Gaifman, H. (1988), "A Theory of Higher-Order Probabilities", en Skyrms, B. y Harper, W. L., (eds.), *Causation, Chance and Credence*, Dordrecht, Kluwer Academic Publishers.

Gaifman, H. y Snir, M. (1982), "Probabilities over Rich Languages, Testing and Randomness", *Journal of Symbolic Logic,* 47, pp. 495-548.

Gillies, D. (2000), *Philosophical Theories of Probability,* Londres, Routledge.

Hájek, A. (2009), "Interpretations of Probability", *The Stanford Encyclopedia of Philosophy* (Spring 2009 Edition), Zalta, E. N., (ed.), próximamente http://plato.stanford.edu/archives/spr2009/entries/probability-interpret/.

Howson, C. (2003), *Hume's Problem: Induction and the Justification of Belief,* Oxford, Oxford University Press.

Howson, C. y Urbach, P. (1993), *Scientific Reasoning. The Bayesian Approach,*, 2ª ed., La Salle, Illinois, Open Court.

Huber, F. (2005), "Subjective Probabilities as Basis for Scientific Reasoning?", *British Journal for the Philosophy of Science,* 56, pp. 101-116.

Iranzo, V. y Martínez de Lejarza, I. (2010), "Medidas de apoyo evidencial: Un análisis comparativo", *Teorema*, vol. XXIX, en prensa.

Iranzo, V. (2007), "Abduction and Inference to the Best Explanation", *Theoria,* 60, pp. 339-346.

———(2008), "Bayesianism and Inference to the Best Explanation", *Theoria,* 61, pp. 89-106.

Jaynes, E.T. (2004), *Probability Theory. The Logic of Science,* Cambridge, Cambridge University Press.

Kelly, K.T. y Glymour, C. (2004), "Why Probability Does not Capture the Logic of Scientific Justification", en *Contemporary Debates in Philosophy of Science*, Hitchcock, Ch. (ed.), Oxford, Blackwell, pp. 94-114.

Lange, M. (1999), "Calibration and the Epistemological Role of Bayesian Conditionalization", *Journal of Philosophy,* 96, pp. 294-324.

Levi, I. (1980), *The Enterprise of Knowledge*, Cambridge, Mass., The MIT Press.

Lewis, D. K. (1984), "A Subjectivist's Guide to Objective Chance", en *Philosophical Papers, vol. II*, New York, Oxford University Press, 1986, pp. 83-132.

Mayo, D. (1996), *Error and the Growth of Experimental Knowledge,* Chicago, University of Chicago Press.

Nau, R. (2005), "Bayesianism Without Priors, Acts Without Consequences", 4th International Symposium on Imprecise Probabilites and Their Applications, Pittsburgh, Pennsylvania.

Niiniluoto, I. (2007), "Evaluation of Theories", en Kuipers, Th. (ed.), *General Philosophy of Science: Focal Issues*. *Handbook of Philosophy of Science*, Amsterdam, Elsevier, pp. 175-217.

Rosenkrantz, R. D. (1977), *Inference, Method and Decision*, Dordrecht, Reidel.

Salmon, W. (1967), *The Foundations of Scientific Inference,* Pittsburgh, University of Pittsburgh Press.

———(1970), "Bayes's Theorem and the History of Science", en Stuewer, R., (ed.), *Historical and Philosophical Perspectives on Science*, Minneapolis, University of Minneapolis Press, pp. 68-86; reimpreso en Salmon (2005) con el título “Discovery and Justification”.

Salmon, W. (1990), “Rationality and Objectivity in Science or Tom Kuhn Meets Tom Bayes”, Wade Savage, C., (ed.), *Scientific Theories, Minnesota Studies in the Philosophy of Science*, vol. 14, Minneapolis, University of Minnesota Press, pp. 175-204; reimpreso en Salmon (2005) con el título “Rationality and Objectivity”.

———(2005), *Reality and Rationality*, New York, Oxford University Press.

Shimony, A. (1970), “Scientific Inference”, en Colodny, R.G. (ed.), *The Nature and Function of Scientific Theories*, Pittsburgh, University of Pittsburgh Press, pp. 79-172.

Spohn, W. (1988),“Ordinal Conditional Functions: A Dynamic Theory of Epistemic States”, en Skyrms, B. y Harper, W.L. (eds.), *Causation in Decision, Belief Change and Statistics*, vol. 2, Dordrecht, Reidel, pp. 105-134.

Strevens, M. (2001), “The Bayesian Treatment of Auxiliary Hypotheses”, *British Journal for the Philosophy of Science*, 52, pp. 515-537.

———(2005),“The Bayesian Approach to the Philosophy of Science”, en Borchert, D. M. (ed.), *Encyclopedia of Philosophy*, 2ª ed., Londres, MacMillan.

van Fraassen, B. (1983), “Calibration: A Frequency Justification for Personal Probability”, en Cohen, R.S. y Laudan, L. (eds.), *Physics, Philosophy and Psychoanalysis*, Dordrecht, Reidel, pp. 295-319.

———(1984), “Belief and the Will”, *Journal of Philosophy*, 81, pp. 235-256.

———(1995), “Belief and the Problem of Ulysses and the Sirens”, *Philosophical Studies*, 77, pp. 7-37.