Artículo de Investigación
https://doi.org/10.35588/cc.v2i1.4903

Una revisión de la condicionalización bayesiana: propuesta de extensión bayesiana sobre la base de la reconstrucción racional y la condicionalización épsilon

A Review of Bayesian Conditionalization: A Proposal for Bayesian Extension on the Basis of Rational Reconstruction and Epsilon Conditionalization

Rodrigo Iván Barrera Guajardo
rodrigo.barrera@outlook.com
Departamento de Filosofía,
Facultad de Humanidades.
Universidad de Santiago de Chile.
https://orcid.org/0000-0002-3694-4367

Artículo recibido: 22 de abril de 2021
Artículo aceptado: 8 de junio de 2021
Artículo publicado: 31 de julio de 2021

Resumen

La epistemología bayesiana tiene como concepto capital la condicionalización simple. Para comprender de buena forma cómo opera esta regla, se debe dar cuenta de la concepción subjetiva de la probabilidad. Sobre la base de lo anterior es posible esclarecer alcances y límites de la condicionalización simple. En general, cuando esta regla enfrenta una dificultad se hacen esfuerzos por resolver dicha particular cuestión, pero no es usual encontrar propuestas unificadas con la intención de resolver varias de las complicaciones subyacentes al bayesianismo ortodoxo. Por lo mismo, el propósito de esta investigación es justamente proponer aquello: un marco bayesiano ampliado que tendrá por objeto solucionar más de una cuestión particular. Más específicamente, el problema de la incorregibilidad, el problema del surgimiento del cero, el problema de la evidencia antigua y el problema de las teorías nuevas. Lo anterior se logrará adicionando al bayesianismo ortodoxo dos cuestiones, a saber, la condicionalización épsilon y una reconstrucción racional.

Palabras clave: Probabilidad Subjetiva, Grados de Creencia, Dogmatismo Bayesiano, Evidencia Antigua, Nuevas Teorías.

Abstract

Bayesian epistemology has simple conditionalization as its central concept. To understand in a good way how this rule operates, it is necessary to account for the subjective conception of probability. Based on the above, it is possible to clarify the scope and limits of simple conditionalization. In general, when this rule faces a difficulty, efforts are made to resolve this particular issue. Still, it is not usual to find unified proposals to resolve several of the difficulties underlying orthodox bayesianism. Therefore, it is the purpose of this research to propose just that, an extended bayesian framework that will aim to solve more than one issue, more specifically the incorrigibility problem, the zero raising problem, the old evidence problem, and the new theories problem. This will be achieved by adding to orthodox bayesianism the epsilon conditionalization and a rational reconstruction.

Keywords: Subjective Probability, Degrees of Belief, Bayesian Dogmatism, Old Evidence, New Theories.

1. Introducción

La epistemología bayesiana (en adelante EB) considera a un agente racional, el cual se encuentra equipado con grados de creencia en ciertas proposiciones, algunas de las cuales tienen valores de verdad desconocidos para el agente. Si las creencias de dicho agente son consistentes de acuerdo con los criterios establecidos por la teoría, y las creencias se miden en grados por un estándar adecuado y conveniente, se puede demostrar que tienen la estructura matemática de una aditividad finita o una función de probabilidad aditiva.

Si un agente comienza con una probabilidad a priori P_i y este adquiere nueva evidencia, la cual puede ser representada como parte de una declaración probatoria E, entonces, la racionalidad requiere que el agente transforme sistemáticamente las probabilidades iniciales para generar probabilidades a posteriori P_f utilizando una regla de actualización. Pero ¿cuál es la regla que rige esta actualización? En el marco bayesiano ortodoxo, dicha regla es la condicionalización simple (en adelante CS).

Sobre la base de lo anterior, esta investigación se propone: **(1)** determinar si la CS es la regla adecuada para actualizar las creencias de los agentes; **(2)** revisar y superar ciertas dificultades con las que se encuentra esta regla, a saber; **(i)** la incorregibilidad y **(ii)** los problemas del surgimiento del cero. Muy sucintamente, **(i)** indica que la EB no puede dar cuenta de cambios radicales en los grados de creencia y **(ii)** se puede seccionar en tres cuestiones, a saber, el problema del reemplazo, la extensión y el desplazamiento. Todos los problemas mencionados se resolverán empleando un bayesianismo ampliado, que consiste en adicionar a la CS una reconstrucción racional (Lange, 1999) y la condicionalización épsilon (Raidl, 2020), que no es más que una generalización de la CS. No obstante, ¿cuáles son los criterios de inclusión del bayesianismo ampliado aquí propuesto? Para considerar una extensión, ésta debe resolver más de un problema y, además, su inclusión no debe generar problemas adicionales. Por ejemplo, un candidato para pertenecer a esta extensión, pero que no cumplía con los requisitos aquí mencionados, es la condicionalización de Jeffrey (en adelante CJ). La justificación de su exclusión se hará explicita más adelante.

En suma, e incluso aceptando el bayesianismo, cuestión que será el punto de partida en esta investigación, todavía quedan asuntos por resolver. Además, y si bien existen soluciones ad hoc para cada uno de estos asuntos, se pretende proponer una solución unificada —la que se llamará bayesianismo ampliado— que dé cuenta de **(i)** y **(ii)**, pero, a la vez, se haga cargo de problemas típicos de la EB. Hablo, en particular, del problema de la evidencia antigua[1] y del problema de las nuevas teorías.

2. Discusión conceptual

Esta sección tiene por objeto introducir todos los elementos para desarrollar la discusión pertinente a la CS. Para ello, se debe dar cuenta de la EB y los conceptos asociados a ésta, tales como la probabilidad,

[1] Una lista no exhaustiva de diferentes soluciones al problema de la evidencia antigua se encuentra en Hartmann y Fitelson (2015), Sprenger (2015) y Wenmackers y Romeijn (2016).

Vol.2 | N°1 | Pág. 24-54 | Julio 2021 | ISSN-E: 0719-9856
https://doi.org/10.35588/cc.v2i1.4903

sus axiomas y los grados de creencia. Una vez aclarada la cuestión conceptual, se definirá el teorema de Bayes, la condicionalización y el argumento del libro holandés.

Tanto la EB como la epistemología tradicional (en adelante ET) albergan diferentes interpretaciones. En el caso de la ET se encuentran el escepticismo, idealismo y racionalismo, por nombrar algunas. Estas interpretaciones analizan los conceptos de conocimiento y creencia para, posteriormente, revisar las propiedades y límites de estos. Por su parte, la EB se divide principalmente entre objetivistas y subjetivistas, sin embargo, existen dentro de esta división inicial múltiples matices. La EB tiene como propósito el estudio de las propiedades y dinámicas de los grados de creencia, los cuales pueden ser representados a través de probabilidades. Por otra parte, la ET tiene como noción central el conocimiento y, cuando trata la noción de creencia, no la concibe en términos graduales. Por tanto, es válido preguntar tal y como indican Hájek y Hartmann (2010, p.1): ¿cómo se puede hablar entonces de EB? Los autores sugieren una posición moderada, puesto que si bien la EB tiene la capacidad de iluminar problemas particulares de la ET, en ningún caso puede ni reemplazar ni responder todas aquellas cuestiones que son de interés para la ET[2].

Una vez que se fijó un límite entre EB y ET, es preciso indicar los supuestos de la EB. Ésta supone tres cuestiones: **(1)** las creencias admiten grados y estos grados pueden ser representados numéricamente, **(2)** si un agente es perfectamente racional, sus grados de creencia deberían obedecer los axiomas de la probabilidad, **(3)** los grados de creencia se actualizan, a la luz de nueva evidencia, mediante la CS. **(1)** es la afirmación central del bayesianismo; **(2)** indica las condiciones de racionalidad que los grados de creencia deben satisfacer en un tiempo particular, es decir, sincrónicamente. Por último, **(3)** entrega las condiciones sobre cómo los grados de creencia se relacionan en el tiempo o, en otras palabras, diacrónicamente.

2.1. El concepto de probabilidad

Tal y como señala Torretti (2003, p.1), la probabilidad tiene un rol en el diario vivir de los agentes racionales. Ella se encuentra presente en la meteorología, en los diagnósticos médicos, en los mercados de valores y así siguiendo. Por lo mismo, no es sorpresa que la noción de probabilidad esté sujeta a un intenso debate filosófico. Una primera aproximación intuitiva permite señalar que la probabilidad de un evento es una medida numérica que captura qué tan posible es la ocurrencia de dicho evento. A partir de esta intuición es posible realizar un breve recorrido desde la equiprobabilidad hasta la concepción subjetiva de la probabilidad.

La noción de equiprobabilidad retrotrae a la definición de Laplace: la probabilidad de un evento es el cociente entre el número de casos «igualmente posibles» favorables a ese evento y el total de todos los casos igualmente posibles (Torretti, 2003, p.6). Bajo condiciones ideales se considera que al lanzar un dado la probabilidad de conseguir un uno es idéntica a la de conseguir un seis. Se dice, entonces, que los eventos «conseguir un uno» y «conseguir un seis» son equiprobables. Sin embargo, en muchas situaciones

[2] Para una comparación detallada entre ET y EB véase Hájek y Lin (2017).

esto no ocurre y, por lo mismo, esta visión de la probabilidad quedó obsoleta y dio lugar al frecuentismo, donde destacan los aportes de von Mises y Reichenbach entre otros (Gillies, 2000, p.88).

En el frecuentismo se hace referencia a la frecuencia relativa de un evento. Sea A un evento asociado a un experimento aleatorio donde se realizan N ensayos (un ejemplo de esto es lanzar un dado $\mathrm{N} = 100$ veces y A es el evento «conseguir un seis»), la frecuencia relativa de A se define por el cociente entre el número de ocasiones que ocurre A en N ensayos sobre el total de ensayos N. Se dice que, en la medida que N aumenta, la frecuencia relativa se estabiliza y, por tanto, la frecuencia relativa de A cuando N crece indefinidamente corresponde a la probabilidad de A. Formalmente esto se expresa como $\mathrm{P(A)}$. No obstante, el modo en que el frecuentismo define la probabilidad permite generar la siguiente crítica: si $\mathrm{P(A)}$ es igual a la frecuencia relativa de A cuando N tiende a infinito, ¿cómo es posible verificar empíricamente $\mathrm{P(A)}$? Imagine que se ejecutan 10^{10} —o incluso 10^{100}— ensayos para estimar la frecuencia relativa de A y esta converge a p. Esto no permite negar que cuando N tiende a infinito la frecuencia relativa converja a q, siendo q distinto de p.

Por otro lado, la concepción subjetiva de la probabilidad permite sobrepasar la dificultad anterior. Esta investigación adhiere particularmente a la probabilidad subjetiva de de Finetti en su primera etapa, pues, sorprendentemente, esta concepción de la probabilidad fue formulada simultáneamente por Frank Ramsey y Bruno de Finetti. Sin embargo, el primero se aproxima a la probabilidad desde el utilitarismo y el segundo desde las apuestas[3].

Para de Finetti la probabilidad es necesariamente subjetiva, es decir, la probabilidad de un evento A corresponde a un valor numérico, representado por $\mathrm{P(A)}$, que mide la confianza de un agente particular respecto a A en el instante que expresa tal confianza. Observe que todo evento A o bien es verdadero o bien es falso. Sin embargo, cuando existe incertidumbre sobre cuál de estas alternativas es la correcta, la probabilidad es la única forma de cuantificar dicha incertidumbre; de Finetti llamó a esta concepción de la probabilidad previsión. Si bien las previsiones de los agentes son subjetivas, si éstos son coherentes en términos probabilísticos sus previsiones tenderán al acuerdo intersubjetivo. Esto se demuestra con el teorema de representación (Press, 2009, p.239)). Respecto a la pregunta de qué es la probabilidad, de Finneti responde que la misma no existe en términos objetivos. No es una propiedad de la naturaleza ni tampoco de los individuos, sino que describe una relación entre un individuo y la naturaleza, y dicha relación es representada por un grado de creencia[4] (Bernardo, 1997, p.4). De este modo, por ejemplo,

[3] de Finetti fue alejándose del esquema de apuestas para cuantificar la probabilidad hasta llegar a concordar con Ramsey en la aproximación utilitarista (Gillies, 2000, p.57). Esta investigación adhiere a la interpretación original de Bruno de Finetti, es decir, al esquema de apuestas, y entrega una breve justificación del porqué se debiera preferir este último esquema por sobre el utilitarista.

[4] En inglés se utilizan los términos «*credences*» y «*degree of belief*» para hacer referencia a los grados de creencia. Sin embargo, no existe una traducción apropiada para el primer término mencionado, por ello únicamente se hablará de grados de creencia.

 Vol.2 | N°1 | Pág. 24-54 | Julio 2021 | ISSN-E: 0719-9856
https://doi.org/10.35588/cc.v2i1.4903

$\mathrm{P}(\mathrm{A}|\mathrm{E})$ representa el grado de creencia de un agente particular para el evento A bajo condiciones E, que, a su vez, describe la relación entre el agente y la naturaleza[5].

Una característica capital de la concepción subjetiva de la probabilidad propuesta por de Finetti es el operacionalismo. Este sostiene que para cualquier propuesta de interpretación de la probabilidad se debe elaborar una definición operativa adecuada o, lo que es igual, se debe construir un dispositivo competente para medirla. Su propuesta es la siguiente: $\mathrm{P}(\mathrm{A}|\mathrm{E})$ corresponde al máximo precio que un agente estaría dispuesto a pagar en una apuesta realizada bajo condiciones E, donde el agente obtendría el premio si y solamente si A es el caso. En otras palabras, la probabilidad se define en términos de cocientes de apuestas: la probabilidad asignada por un agente a un evento particular debe identificarse con el cociente de apuestas tal que él estaría dispuesto a apostar una cierta suma si el evento es el caso.

Previamente se indicó que, si un agente es perfectamente racional, sus grados de creencia deben obedecer los axiomas de la teoría de la probabilidad; esta idea es capturada por el argumento del libro holandés[6]. En términos amplios, este viene a decir que un apostador cuyos grados de creencia no obedecen los axiomas de la probabilidad ganaría o perdería —según el rol que juegue en la apuesta— sin importar los resultados. Si se acepta que un agente racional no va a apostar sabiendo que perderá sea cual sea el resultado, sus grados de creencia deben satisfacer los axiomas de la probabilidad, pues estos aseguran la racionalidad de sus decisiones en torno a las apuestas. Se estudiará con más detalle el argumento del libro holandés más adelante, pues primero se debe precisar qué es el operacionalismo.

El operacionalismo consiste en un sistema de apuestas que permite inferir los valores de las probabilidades. Lo anterior funciona del siguiente modo: considere la proposición $\mathrm{q} =$ «mañana lloverá en Santiago». Un agente asigna a la proposición anterior una probabilidad equivalente a $0{,}1$. Dos preguntas surgen naturalmente: ¿cómo se mide la probabilidad de este agente para tal proposición? ¿Qué significa decir que la probabilidad para la proposición q es equivalente a $0{,}1$? La primera pregunta ya se respondió, indicando que esto se resuelve en términos operacionales o, lo que es lo mismo, con un sistema de apuestas. Respecto a la segunda pregunta; el máximo precio que el agente estaría dispuesto a pagar por un boleto de apuesta que le reportará ganancias de $100\$$, si q es el caso, serían $10\$$. Más generalmente, se puede indicar que la probabilidad de q se mide por el cociente $\mathrm{r} = \frac{\mathrm{b}}{\mathrm{a+b}}$, tal que el agente considera razonable obtener $\mathrm{a}\$$ si q es el caso y $-\mathrm{b}\$$ en caso contrario. El cociente r tiene dos supuestos: (1) el agente no debe presentar ni aversión ni propensión al riesgo, (2) existe independencia entre r y el valor de verdad de la proposición (Huber y Schmidt-Petri, 2008, p.5).

2.2. Axiomas de la probabilidad

Kolmogorov define una estructura matemática abstracta, la cual recibe el nombre de espacio de probabilidad $(\Omega, \mathcal{F}, \mathrm{P})$. Se trata de un conjunto Ω, junto con una colección $\mathcal{F}$ de subconjuntos de Ω, y una

[5] Esta revisión del concepto de probabilidad no pretende ser exhaustiva, sino mostrar que hay múltiples interpretaciones. No se revisó la interpretación lógica, ni la propensista ni la best-system. Un buen resumen de ellas puede consultarse en Hájek (2019). Para un estudio detallado puede consultarse Gillies (2000) y Hacking (2006).

[6] Profundizo acerca de este tema en la sección 2.

función P que asigna números reales a los miembros de $\mathcal{F}$. Sobre lo anterior se definen los axiomas de la probabilidad, siendo estos:

(1) Los valores de P son siempre no-negativos.

(2) $\mathrm{P}(\Omega) = 1$

(3) Para cualquier secuencia de eventos disjuntos[7] $\mathrm{A}_1, \mathrm{A}_2, \mathrm{A}_3, \ldots\ \mathrm{P}(\bigcup_{\mathrm{i}=1}^{\infty} \mathrm{A}_\mathrm{i}) = \sum_{\mathrm{i}=1}^{\infty} \mathrm{P}(\mathrm{A}_\mathrm{i})$.

donde **(1)** corresponde al axioma de no-negatividad, **(2)** es el axioma de normalización y **(3)** el axioma de aditividad contable.

Sobre **(3)** se debe comentar lo siguiente. Nótese que $\mathrm{P}(\bigcup_{\mathrm{i}=1}^{\infty} \mathrm{A}_\mathrm{i}) = \sum_{\mathrm{i}=1}^{\infty} \mathrm{P}(\mathrm{A}_\mathrm{i})$ permite un conjunto infinito de eventos disjuntos A_i, sin embargo, de Finetti —y otros como Ramsey y Savage— señalan que es inapropiado considerar un número infinito de eventos y debiera considerarse un número finito, pues únicamente un número finito de eventos se pueden constatar empíricamente. Luego, según de Finetti, **(3)** debiera definirse como $\mathrm{P}(\bigcup_{\mathrm{i}=1}^{\mathrm{k}} \mathrm{A}_\mathrm{i}) = \sum_{\mathrm{i}=1}^{\mathrm{k}} \mathrm{P}(\mathrm{A}_\mathrm{i})$. En síntesis, de Finetti demuestra que si $\mathcal{F}$ es un algebra, entonces **(1), (2)** y $\mathrm{P}(\bigcup_{\mathrm{i}=1}^{\mathrm{k}} \mathrm{A}_\mathrm{i}) = \sum_{\mathrm{i}=1}^{\mathrm{k}} \mathrm{P}(\mathrm{A}_\mathrm{i})$ son suficientes para que $\mathrm{P}: \mathcal{F} \to \mathbb{R}$ sea considerado una probabilidad. Para una discusión más minuciosa sobre este asunto consultar (Regazzini, 2013).

La terminología entregada hasta ahora tiene su origen en la estadística y no puede aplicarse a la EB. Se procede consecuentemente a indicar una interpretación que sirva para este propósito. Se ha indicado que $(\Omega, \mathcal{F}, \mathrm{P})$ es un espacio de probabilidad. Aquí, un espacio muestral Ω corresponde a un conjunto de mundos posibles —que representan el conjunto de posibilidades epistémicas de un agente— y $\mathcal{F}$ es un σ-algebra que delinea subconjuntos de Ω, y una medida de probabilidad finitamente aditiva P sobre $\mathcal{F}$[8].

Sobre el concepto de posibilidad epistémica, Parrott (2014, p.3), entrega la siguiente definición:

> Se puede pensar que es equivalente a lo que es posible, dado lo que se sabe: las cosas incompatibles con lo que se conoce no son posibilidades epistémicas. Una agente puede creer que es epistémicamente posible que su pareja esté en casa, lo que significa que, sobre la base de los conocimientos de ella, él podría estar en casa. Sin embargo, en tanto la agente sabe que

[7] Se dice que dos eventos A_1 y A_2 son disjuntos si no pueden ocurrir simultáneamente. Formalmente $\mathrm{A}_1 \cap \mathrm{A}_2 = \emptyset$.

[8] Nótese que se adopta una posición conservadora al utilizar una medida de probabilidad finitamente aditiva en lugar de una medida de probabilidad contable. Lo anterior, siguiendo la recomendación de Savage (1954, p.43) quien señala: "gran parte de la teoría matemática moderna de la probabilidad depende del supuesto que las medidas de probabilidad en cuestión son contablemente aditivas, por ello existe la fuerte tentación de asumir la aditividad contable o su equivalente lógico como un postulado. Pero me inclino a estar de acuerdo con de Finetti y Koopman en que, por muy conveniente que sea la aditividad contable, no debería figurar entre los postulados de un concepto de probabilidad personal, a menos que consideremos que su transgresión merece ser calificada de inconsistente. Por lo tanto, parece mejor no asumir directamente la aditividad contable como un postulado, sino reconocerla como una hipótesis especial que da lugar, en su caso, a una gran clase de teoremas útiles". Asimismo, de Finetti (1974, p.119) indica: "nadie ha dado una verdadera justificación de la aditividad contable (aparte de tomarla como una «extensión natural» de la aditividad finita); de hecho, muchos autores también tienen en cuenta casos en los que no se sostiene, pero los consideran por separado, no como absurdos, sino como «patológicos», fuera de la teoría «normal»".

Vol.2 | N°1 | Pág. 24-54 | Julio 2021 | ISSN-E: 0719-9856
https://doi.org/10.35588/cc.v2i1.4903

su pareja trabaja hasta las 20 horas, no es epistémicamente posible que esté en Rusia. En efecto la agente indicaría que lo anterior es falso (Parrott, 2014, p.3)[9].

En resumen, los subjetivistas consideran que la probabilidad es un grado de creencia razonable en una proposición desde un punto de vista individual. Por lo tanto, la probabilidad es una medida subjetiva numérica de un agente particular según su grado de creencia, siempre y cuando sea coherente.

2.3.Grados de creencia

Los grados de creencia representan la confianza de los agentes en la verdad de las proposiciones. Por ejemplo, un agente tendrá confianza en que toda vez que se dirige a una estación de metro en horario hábil, éste estará funcionando. No obstante, el mismo agente tendrá mucha más confianza en que el sol saldrá por la mañana. ¿Cuáles son los *relata* de la relación grados de creencia? Estos son un número, un agente racional en un instante particular y el objeto de la creencia. Esta cuestión es tratada con más detención en Huber & Schmidt-Petri (2008).

Hasta aquí solo se describió el proceso por el cual se cuantifican los grados de creencia. No obstante, es tarea pendiente detallar qué reglas los gobiernan. Antes de formalizar las reglas se ejemplificará cómo la incoherencia probabilística permite construir un libro de apuestas holandés. En este caso la incoherencia se manifiesta pues la probabilidad de un evento más la probabilidad de su negación es superior a uno. Imagine que un agente establece que su grado de creencia con respecto al evento A es igual a $0{,}75$, es decir, $\mathrm{P(A)} = 0{,}75$, y asigna el mismo grado de creencia a la negación de A, es decir, $\mathrm{P}(\neg\mathrm{A}) = 0{,}75$. Esto equivale a decir que el agente está en condiciones de aceptar la siguiente estructura de apuestas:

- Lotería (1): el agente acepta apuestas 3 contra 1 para A (apuesta B_1) y 1 contra 3 para ¬A (apuesta B_2).
- Lotería (2): el agente acepta apuestas 3 contra 1 para ¬A (apuesta B_3) y 1 contra 3 para A (apuesta B_4).

Suponga ahora que usted apuesta contra el agente seleccionando B_2 y B_4 y, consecuentemente, el agente selecciona B_1 y B_3. Si A es el caso, el agente recibe \$1 de la lotería (1) y debe pagar \$3 por la lotería (2) con una pérdida neta de \$2. Si ¬A es el caso, el agente debe pagar \$3 por la lotería (1) y recibe \$1 por la lotería (2). Note que usted gana y el agente pierde \$2 sin importar el resultado. De este modo, usted construyó un libro de apuestas holandés en contra del agente[10].

En suma, ser coherente probabilísticamente equivale a indicar que los grados de creencia están gobernados por los axiomas de la probabilidad.

[9] Para más detalles consultar Easwaran (2011, p.313). Esta y todas las traducciones desde el inglés fueron hechas por el autor.

[10] Note que en el ejemplo entregado se transgrede únicamente el axioma de normalidad. Sin embargo, se puede construir un libro holandés sin importar cuál de los axiomas de la probabilidad se transgreda. Más detalles pueden consultarse en Vineberg (2016).

2.4. Teorema de Bayes

El teorema de Bayes es una consecuencia deductiva de los axiomas de la probabilidad y se define así:

$$P(H_1|e) = \frac{P(H_1)P(e|H_1)}{\sum_{j=1}^{k} P(H_j)P(e|H_j)}$$

Además, se puede interpretar como sigue: un agente considera ciertas proposiciones $H_1, H_2, \ldots, H_k$, pero desconoce cuál de ellas es verdadera. Sobre la base de su conocimiento, el agente puede asignar las probabilidades $P(H_1), P(H_2), \ldots, P(H_k)$. Estas representan la confianza del agente en las proposiciones. Tales probabilidades reciben el nombre de preexperimentales, iniciales o a prioris. Además, el agente puede determinar la probabilidad de observar una evidencia e dada una proposición particular, esto corresponde a la probabilidad condicional $P(e|H_j), j = \{1, \ldots, k\}$. $P(e|H_j)$ que recibe el nombre de verosimilitud. Por último, el teorema de Bayes le permite al agente calcular la probabilidad condicional $P(H_j|e)$ que corresponde a su grado de creencia respecto a H_j después de obtener la evidencia e. $P(H_j|e)$ recibe el nombre de probabilidad post experimental, final o a posteriori[11] (Bhattacharyya y Johnson, 1977, p.93).

El teorema de Bayes proporciona el siguiente corolario, conocido como CS:

$$P_f(H) = \frac{P_i(H)P_i(e|H)}{P_i(e)}$$

La CS[12] permite obtener la probabilidad a posteriori de una proposición H sobre la base de la verosimilitud $P_i(e|H)$ y las probabilidades $P_i(H)$ y $P_i(e)$. La CS es la regla más básica que gobierna los cambios temporales de las probabilidades y sostiene que, si la probabilidad de e sube a uno y $P(H|e)$ no

[11] En adelante solo se utilizarán las expresiones a priori y a posteriori. El subíndice i denotará una probabilidad a priori (P_i) y el subíndice f una probabilidad a posteriori (P_f). Para el mismo propósito —cuando se requiera— se hará referencia a P_i utilizando P_t y a P_f con P_{t+1}.

[12] Se introduce esta nota por sugerencia de uno de los revisores, agradezco sus comentarios.

Una extensión de la CS es la CJ que se define por:

$$P_f(H) = \frac{P_i(e/H)P_i(H)}{P_i(e)} P_f(e) + \frac{P_i(\neg e/h)P_i(H)}{P_i(\neg e)} P_f(\neg e)$$

El problema de aceptar esta regla es que no se puede establecer su método de activación. La CS se activa dadas las condiciones de certeza y rigidez, pero la CJ debe activarse sin apelar a la condición de certeza, quedando así el mecanismo de actualización no establecido de forma unívoca (Bird, 2018; Trpin, 2020). Además, se puede argumentar que si $P_f(e) \neq 1$ se diluye el problema del surgimiento del cero, pero se abre la siguiente interrogante: ¿qué es evidencia? Se debieran aceptar como evidencias proposiciones cuyo grado de creencia sea menor a 1, es decir, evidencia incierta. Son estas últimas cuestiones las que motivan la exclusión de la CJ del bayesianismo ampliado. De todos modos, se debe comentar que la CJ es totalmente compatible con la interpretación subjetiva de la probabilidad. En efecto, Bird (2018, p.9) muestra que la CJ es consecuencia de uno de los axiomas de la probabilidad; la aditividad finita. Por otro lado, Armendt (1980) muestra un argumento diacrónico para la CJ.

Vol.2 | N°1 | Pág. 24-54 | Julio 2021 | ISSN-E: 0719-9856
https://doi.org/10.35588/cc.v2i1.4903

cambia como resultado del cambio temporal en la probabilidad de e, entonces, se debe actualizar $\mathrm{P(H)}$ a $\mathrm{P(H|e)}$. Más formalmente, la CS se debe aplicar si se cumple:

(i) $\mathrm{P_i(e)} < 1$ y $\mathrm{P_f(e)} = 1$

(ii) $\mathrm{P_i(H|e)} = \mathrm{P_f(H|e)}$

Aquí **(i)** corresponde a la condición de certeza e indica que la CS se puede aplicar únicamente cuando la proposición e se tiene por verdadera. Por otro lado, **(ii)**, llamada condición de rigidez, debe su nombre a que $\mathrm{P(H|e)}$ permanece inalterable una vez que $\mathrm{P(e)} = 1$. La conjunción de **(i)** y **(ii)** es la condición necesaria y suficiente para la aplicación de la CS (Jeffrey, 2002, p.46).

En general se dirá que una evidencia e:

(1) Confirma H si y solo si $\mathrm{P(H|e)} > \mathrm{P(H)}$.

(2) No confirma H si y solo si $\mathrm{P(H|e)} < \mathrm{p(H)}$.

(3) Es neutra si y solo si $\mathrm{P(H|e)} = \mathrm{P(H)}$.

Para justificar que la condicionalización es la regla que debe gobernar los cambios de los grados de creencia diacrónicamente, se apela al argumento de Lewis-Teller. Teller —quien dio crédito a Lewis— presentó el argumento diacrónico del libro holandés el cual justifica la aplicación de la CS. Lewis mostró cómo construir un libro holandés contra cualquiera que, al aprender e, cambie su grado de creencia en H a $\mathrm{P_f(H)} > \mathrm{P_i(H|e)}$ o $\mathrm{P_f(H)} < \mathrm{P_i(H|e)}$. Es decir, si un agente actualiza sus grados de creencia de forma inconsistente con la CS, entonces se podrá construir una estrategia de apuestas en la que el agente siempre pierda dinero[13].

3. El problema de la incorregibilidad

Para introducir este tópico, se procederá del siguiente modo: primero, se describirá el problema en cuestión para luego indicar por qué se hace necesaria una solución. Se revisará la propuesta de Lang (1999)[14] —conocida como reconstrucción racional o argumento justificativo— y, a continuación, se explicarán las diferencias entre esta propuesta y la CS. Se explicitará un punto de discrepancia con Lang en torno a las restricciones a las que están sometidas las reconstrucciones racionales. Se indicará —a diferencia de Lang— que estas no debieran estar restringidas únicamente por los axiomas de la probabilidad.

Note que aquel agente cuyos grados de creencia estén gobernados por la CS se encontrará con que sus opiniones respecto a las proposiciones siempre estarán codificadas por su distribución de probabilidad a priori primaria. Esto constituye un problema pues, imagine que un agente decide considerar más

[13] Para ejemplos véase: Earman (1992, p.46), Vineberg (2016) y Talbott (2016). La exposición original se encuentra en Teller (1973) y Lewis (1976).

[14] Por la misma razón dejaré, en lo que resta de esta sección, de citar a Lang (1999). Esto a menos que me refiera a páginas específicas.

seriamente una proposición a la que anteriormente asignó una probabilidad cercana a cero. La CS no puede dar cuenta de este cambio. En otras palabras, llevarlo a cabo implica la transgresión de los axiomas de la probabilidad.

Se define entonces el problema de la incorregibilidad al hecho que un agente se encuentra anclado a su distribución de probabilidad a priori primaria. Lang indica que la CS no debería gobernar los cambios de opinión de un instante de tiempo al siguiente según aparece nueva evidencia, sino que, en lugar de aquello, un agente debe justificar su opinión actual apelando a la evidencia acumulada y cada paso de tal justificación deberá estar gobernado por la CS. Esto ocurriría del siguiente modo: imagine un científico que procede a explicar su grado de creencia en una hipótesis h. Dicho proceso corresponde a una respuesta justificada sobre la base de la evidencia con la que cuenta hasta ahora el científico. Durante el proceso, éste somete a revisión toda la evidencia relacionada con h de la que dispone y, mientras ejecuta el proceso, su grado de creencia en h no se ve afectado. Dicho proceso comienza con la determinación de la distribución de probabilidad a priori primaria por parte del agente. Una vez determinada, este procede a justificar su actual grado de creencia de acuerdo con las distintas evidencias que cuenta; «actualizando» hasta agotar el examen de los elementos del cuerpo de conocimiento. Finalizado este proceso, el agente alcanzará su grado de creencia actual en h. No se debe entender este último «actualizando» como un sinónimo de la CS, sino más bien como la justificación de la actual opinión del científico. Este proceso recibe el nombre de reconstrucción racional, argumento justificativo o, simplemente, justificación.

La propuesta de Lang para superar el problema de la incorregibilidad y a la vez no transgredir los axiomas de la probabilidad opera como sigue: suponga que un agente se encuentra justificando su actual grado de creencia en una proposición h, que equivale a .75. En los primeros pasos de su justificación, el agente toma en consideración una fracción de la evidencia disponible, lo cual lo llevó a un grado de creencia equivalente a .4. A continuación, el agente considera otra porción de evidencia y, consecuentemente, su grado de creencia en h es .6. Finalmente, repite el proceso y termina con un grado de creencia igual a .75. La cuestión es que el cambio en el grado de creencia de un paso a otro en un argumento justificativo no corresponde a una transgresión de la confirmación de la condicionalización. Observe que el agente no está en condiciones de rechazar su grado de creencia anterior y a considerarlo como una conclusión errónea sobre la base de la porción de evidencia considerada en aquel paso. La razón de esto es que las conclusiones anteriores son la base sobre las cuales se construyen las conclusiones posteriores, y un paso posterior no puede considerar el rechazo de un paso anterior como si este fuera una conclusión inapropiada sobre la base de la evidencia considerada en dicho paso. Dicho de otro modo, las opiniones que aparezcan como conclusiones intermedias en un determinado paso del argumento justificativo deben considerarse a lo largo de todo el argumento como si hubiesen sido matizadas por las pruebas tenidas en cuenta en ese paso o antes de él.

Si bien la demostración de que cada transición de un paso a otro en un argumento justificativo debe estar gobernado por la CS es extensa, en lo que sigue se explicará de manera general. Lo primero es esclarecer la noción de calibración[15]. Esta noción manifiesta que sería erróneo transgredir la CS en una

[15] La noción de calibración fue introducida por van Fraassen (1983) y puede entenderse como «estar en sintonía con el mundo». Detalles adicionales pueden consultarse en Hoefer (2012).

Vol.2 | N°1 | Pág. 24-54 | Julio 2021 | ISSN-E: 0719-9856
https://doi.org/10.35588/cc.v2i1.4903

justificación. Se dirá que la opinión de un agente está perfectamente calibrada sólo si su transición de un paso a otro dentro de una justificación está gobernada por la CS (Lang, 1999, p.295).

Dicho lo anterior, considere lo siguiente: un meteorólogo establece que la probabilidad de que llueva mañana es de un 80%. Sostiene lo anterior pues hoy llovió y la temperatura máxima fue de 20°C. Cada vez que se cumplan las mismas condiciones de hoy, es decir, que llueva y la temperatura máxima sea de 20°C, se dirá que mañana es «relevantemente similar» a hoy. Suponga que el pronosticador del tiempo sabe acerca de algún otro día en que, el día anterior al mismo, llovió y la temperatura más alta fue de 20°C. Entonces, el meteorólogo debería tener un 80% de confianza en que lloverá ese día, ya que tiene exactamente las mismas razones en ambos casos para creer que lloverá con la misma probabilidad. Al asignar un 80% de confianza a la afirmación de que lloverá mañana, el pronosticador está perfectamente calibrado si y sólo si llueve exactamente en el 80% de los días que son «relevantemente similares» a mañana. Si el agente está perfectamente calibrado, se dice entonces que su opinión es correcta.

Lang (1999, p.311) sugiere que las conclusiones intermedias en un paso dado de una reconstrucción racional debieran estar perfectamente calibradas, y esto explicaría por qué los agentes se guían por esas conclusiones cuando pasan al siguiente paso teniendo en cuenta algunas evidencias adicionales. Cuando usted justifica su opinión actual, las conclusiones intermedias de cada paso se guían por las del paso anterior. De esta manera, usted considera esas conclusiones intermedias seriamente cuando tiene en cuenta otra evidencia. Si usted cambia de opinión respecto a sus conclusiones intermedias, a las que se debe llegar en un paso determinado, entonces, debe cambiar ese paso ajustando las conclusiones en consecuencia. No puede mantener, en ese paso, las conclusiones intermedias que —ahora considera inadecuadas— y añadir un paso adicional al argumento para negar el efecto de estas conclusiones en pasos posteriores.

No obstante, los agentes pueden cambiar racionalmente sus opiniones de un momento a otro sin necesidad de obtener ninguna nueva evidencia. Cuando el agente establece que sus opiniones anteriores no otorgaron el peso adecuado a ciertas consideraciones, entonces el agente señala que esas opiniones eran incorrectas considerando las pruebas que ya conocía. Incluso cuando no cuenta con evidencia adicional, puede cambiar de opinión reconsiderando cuestiones conocidas.

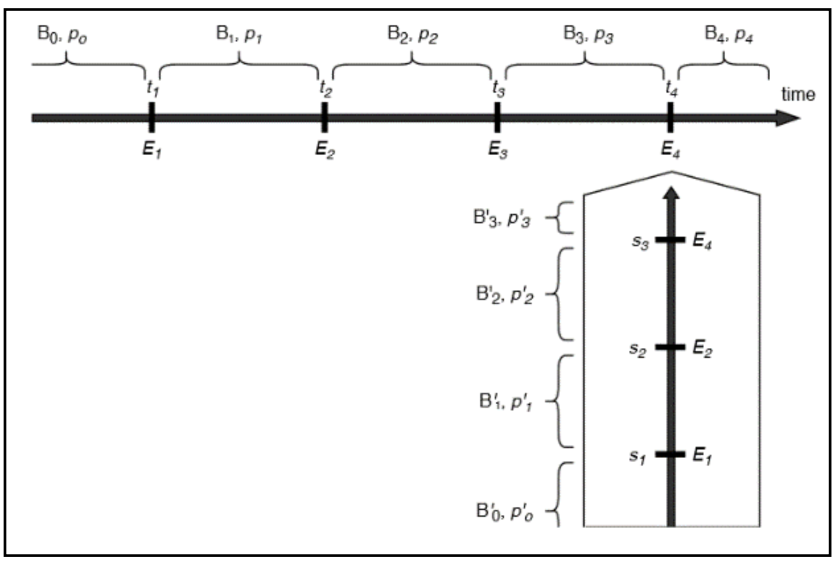
Se considera que las opiniones posteriores son correctas sobre la base de considerar que las opiniones anteriores han sido correctas. Si el agente considera que una conclusión intermedia no era correcta, entonces se debe cambiar ese paso, dando lugar a una nueva justificación. Pero si el agente cambia de opinión sobre si sus opiniones en un momento anterior concedieron el peso adecuado a las pruebas entonces conocidas, no puede volver atrás y cambiar las opiniones que tenía en el pasado.

Llegados hasta acá, resta tratar dos asuntos relacionados con la propuesta de Lang: **(1)** contrastar los argumentos justificativos con la actualización por CS y **(2)** revisar las restricciones que constriñen la justificación.

Suponga la siguiente situación. Un paciente se dirige a un médico pues sospecha que padece una enfermedad E. El médico le indica al paciente que, para comprobar si padece o no la enfermedad, debe indicarle los síntomas y además someterse a tres pruebas, siendo estas P_1, P_2 y P_3. De acuerdo con la CS, el médico establecerá su grado de creencia en la proposición «el paciente padece la enfermedad E». Así, una vez que el paciente indica sus síntomas, el médico, sin contar con ninguna de las tres pruebas, establece un grado de confianza equivalente a $.2$ tras la evaluación de los síntomas indicados por el paciente. Más tarde, el médico recibe los resultados de P_1 y actualiza su grado de creencia a $.5$. Luego, recibe los resultados de P_2, y, nuevamente, actualiza su grado de creencia a $.8$. Finalmente, recibe los resultados de P_3 para definitivamente tener un grado de creencia igual a 1. En cada paso, el médico actualizó su grado de creencia a la luz de nueva evidencia.

El argumento justificativo comienza con una distribución de probabilidad a priori, seleccionando los elementos del cuerpo de conocimientos que el agente considera relevantes. Posteriormente, en cada paso de la reconstrucción racional el agente considera una nueva porción de evidencia y debe reiterar el proceso.

Para esclarecer cómo opera la justificación, sitúe al médico en el instante t_4. En ese momento su cuerpo de conocimiento está formado por los síntomas indicados por el paciente y las pruebas P_1, P_2 y P_3. Para determinar su probabilidad a priori, el médico puede elegir cualquiera de las pruebas o los síntomas indicados por el paciente. Además, puede considerar que los síntomas indicados por el paciente eran falsos e incorporar dicha cuestión al argumento. Misma cuestión puede ocurrir con las pruebas.



Comparación gráfica de la actualización por condicionalización (→) y la reconstrucción racional (↑) (Canson, 2015, p.97).

Note que, en el caso de la reconstrucción racional, B'_0 corresponde al cuerpo de conocimiento que el agente considera relevante al fijar su a priori primario. p'_0 corresponde a la distribución de probabilidad sobre la base de B'_0. Aquí es el agente quien selecciona el orden en el que considera las evidencias, toda vez que su opinión se sustenta sobre la evidencia acumulada no existe restricción de orden.

Ahora que está clara la diferencias entre la CS y la reconstrucción racional, queda por revisar las restricciones a las que está circunscrita esta última. Lang menciona que los grados de creencia deben respetar los axiomas de la probabilidad y, junto con ello, los a priori deben ser no-sesgados [*unbiased*] e

Vol.2 | N°1 | Pág. 24-54 | Julio 2021 | ISSN-E: 0719-9856
https://doi.org/10.35588/cc.v2i1.4903

imparciales [*impartial*] [16]. Determinar a qué restricciones están sujetos los a priori se conoce como el problema de los priori. Esta cuestión divide a bayesianos subjetivos (esta investigación adhiera a estos) y objetivos. Los primeros requieren únicamente que los a priori sean coherentes en términos probabilísticos, mientras que los segundos exigen restricciones adicionales.

Los diferentes agentes bayesianos pueden comenzar con opiniones muy divergentes sobre algún tema, pero rápidamente se acercan a un acuerdo de opinión si observan, en gran medida, las mismas evidencias. Así pues, se afirma que los a priori se «limpian»[17] con la acumulación de evidencia. Ahora bien, lo anterior no afecta del mismo modo a la actualización de los grados de creencia a la luz de nueva evidencia y a la reconstrucción racional propuesta por Lang. La razón de aquello es que ambas no están influenciadas por la historia epistémica. Solo la condicionalización lo está y las justificaciones son reducidas —en términos temporales— en comparación con la historia epistémica. En otras palabras, el efecto de la «limpieza» de los a priori en la reconstrucción racional es marginal en comparación con la condicionalización. Es por esta razón que la propuesta de Lang requiere restricciones adicionales.

Sea J una justificación de n pasos. Para que una justificación sea satisfactoria, cada paso de esta, incluyendo la etapa primaria, podría haber sido un paso cualquiera de la justificación. Es decir, el i-ésimo paso podría ser el paso $1, \dots, i-1, i+1, \dots, n$. En otras palabras, el orden en que se considere la evidencia no tiene importancia. Además, como ya se mencionó, el agente determina las piezas de evidencia que utilizará en la justificación y dichas piezas deben estar ancladas a la realidad. Un ejemplo de una pieza de evidencia no anclada a la realidad es apelar al mes de nacimiento de un paciente para indicar si este puede sufrir una determinada enfermedad. Adicionalmente, la evidencia utilizada para construir una teoría no puede usarse para justificar la teoría. Por lo tanto, el único grado de creencia admisible para e, si e fue utilizado para crear una teoría, es uno[18].

Hasta acá se explicó la razón por la cual se debe proponer una solución al problema de la incorregibilidad y se revisó una solución. Además, se indicó que el problema de los a priori no afecta del mismo modo a la condicionalización y a la justificación, particularmente para la última se requieren restricciones adicionales. Se debe recordar que esta investigación tiene por objetivo establecer si la CS es la regla adecuada para actualizar la creencia de los agentes racionales. Para ello, se debe indicar si existen situaciones en las cuales se debe abandonar dicha regla. Para el caso de los cambios radicales en los grados de creencia, se argumentó que no es necesario abandonar la CS, pues los pasos de la reconstrucción racional están gobernados por ésta.

[16] Lang no es específico en indicar qué entiende por a priori no-sesgado o por a priori imparcial. Además, él comenta que la reconstrucción racional enfrenta el mismo problema de los a priori que el Bayesianismo ortodoxo y este último no hace referencia a las dos restricciones previamente señaladas. Una interpretación de las restricciones de no-sesgo e imparcialidad pudiera ser la siguiente: los a priori deben ponderar la evidencia objetivamente.

[17] Una presentación más formal de «la limpieza de los a priori» se puede consultar en Earman (1992, p.141) y Joyce (2011, p.445).

[18] Esta idea es capturada por el criterio de Zahar-Worrall y viene a decir que e puede dar soporte confirmatorio a S solo si e no es un constituyente de la situación problemática que precedió y resultó en la construcción de S (Koertge, 1978, p.268). Para una discusión más extensa consultar Iseda (1999).

4. El problema del surgimiento del cero

El bayesianismo ortodoxo suscribe, como ya se ha indicado, a la actualización de los grados de creencia a través de la CS. No obstante, los agentes de mente abierta no pueden ser gobernados por esta regla. Un agente de mente abierta es aquel con la capacidad de considerar posibilidades que previamente fueron ignoradas. Por tanto, un agente de mente abierta no tiene la certeza de que sus opiniones convergerán a la verdad. Por ejemplo, considere que cierto fenómeno científico puede ser explicado o bien por la teoría U o bien por la teoría V. ¿Puede asegurar la comunidad científica que estas son las dos únicas posibles explicaciones? Muy probablemente no, o, al menos, no se puede descartar con total certeza la posibilidad que exista otra teoría que pueda dar cuenta del fenómeno. Es válido preguntar qué razones habría para aquello, es decir, que motivos llevan a un agente a considerar posibilidades que antes no contempló. Si bien más adelante se entregarán ejemplos y se formalizará cómo esto ocurre, una de las razones principales es el falibilismo. Por lo mismo, antes de comenzar la discusión en torno al surgimiento del cero, es apropiado comenzar por discutir algunos aspectos de este.

En términos amplios, el falibilismo consiste en aceptar el carácter incierto del conocimiento. Si se sostiene la compatibilidad entre CS y falibilismo, se admite el siguiente compromiso: todos los agentes deben asignar una probabilidad distinta de uno a todas las proposiciones que no son una verdad lógica. Este compromiso se sustenta en el siguiente hecho: si una proposición q tiene $\mathrm{P(q)} = 1$, no hay nada en el aparato bayesiano ortodoxo que pueda cambiar aquello. En otras palabras, la proposición q es —desde una mirada bayesiana— infalible y tal y como comenta Shaffer (2008, p.219) esto constituye una idealización excesiva. Dicho de otro modo, suponer que los agentes asignarán probabilidad uno solamente a proposiciones que constituyen una verdad lógica es una idealización excesiva. En este sentido, Howson y Urbach (2006, p.259) comentan "no hay nada que impida asignar probabilidad uno a cualquier proposición”.

Ahora bien, ¿qué ocurre en particular con las proposiciones observables? ¿son éstas infalibles? Dicho de otra forma, sea q una proposición observable tal que $\mathrm{P(q)} = 1$. ¿Lo anterior sostiene la infalibilidad de q? Howson y Urbach (2006, p.104) indican que no. Las proposiciones observables son falibles y no existen grados de falibilidad ni nada por el estilo. Los informes de observación que son admitidos como evidencia son aceptados como resultado de una decisión o acuerdo y, en esa medida, son convenciones.

Sobre la base de lo anterior, se tiene lo siguiente: el que una proposición tenga grado de creencia 1 no equivale a sostener su infalibilidad, por lo mismo es lícito que un agente, sobre la base de su experiencia hasta el tiempo t, sostenga para una proposición q que $\mathrm{P(q)} = 1$ y en $\mathrm{t'}$ que $\mathrm{P(q)} < 1$, con $\mathrm{t} < \mathrm{t'}$. Sin embargo, previamente se indicó que la CS no permite este cambio. En efecto, la única forma de lograr este cometido es transgrediendo los axiomas de la probabilidad y, por tanto, el bayesianismo ortodoxo y el falibilismo son incompatibles.

Raidl (2020, p.9) comenta que la evidencia puede ser falible —tal y como comentan Howson & Urbach — y lo mismo es el caso para los agentes, dado que estos pueden acceder a evidencia falsa o pueden sostener probabilidades a priori irregulares. Dicho de otro modo, pueden asignar probabilidad cero a

Vol.2 | N°1 | Pág. 24-54 | Julio 2021 | ISSN-E: 0719-9856
https://doi.org/10.35588/cc.v2i1.4903

eventos distintos del conjunto vacío y, consecuentemente, asignar probabilidad uno al complemento de dichos eventos.

Aceptar que la evidencia y los agentes puedes ser falibles equivale a aceptar el bayesianismo de mente abierta, el cual admite defectos de precisión, siendo estos:

(1) Irregularidad inicial. El a priori inicial del agente no es regular. Una función P es regular si $\mathrm{P}(\mathrm{A}) = 0$ implica que $\mathrm{A} = \emptyset$.

(2) El agente comete un error de cálculo.

(3) En algún punto el agente obtiene evidencia falsa.

Un intento por conciliar falibilismo y bayesianismo es llevado a cabo por Weintraub (1993) a través de una taxonomía de proposiciones y haciendo uso de la CJ[19]. La autora comienza señalando que, si la probabilidad respecto a una proposición requiere cambiar una vez un agente le asignó probabilidad uno, esta tendrá que cambiar por un medio distinto a la CS, siendo una alternativa la observación. Si se diferencia entre proposiciones observables y teóricas, se podría argumentar que las primeras se pueden modificar directamente y las segundas únicamente de forma inferencial (o, lo que es lo mismo, mediante la condicionalización). Habría aquí espacio para el dogmatismo ya que, en el caso de las proposiciones teóricas, una vez alcanzada la certeza no sería posible cambiar aquello.

Las proposiciones también se podrían diferenciar entre aquellas que hacen referencia a particulares y aquellas que lo hacen a universales. Las probabilidades de las primeras podrían ser modificadas directamente y las probabilidades de las segundas de forma inferencial.

Alternativamente, se podría indicar que la distinción entre proposiciones teóricas y observables corresponde a la diferencia entre proposiciones falibles e infalibles. En efecto, Lewis (1981, p.14) dice "nada debiera ser tan seguro como todo aquello [con probabilidad 1], con la posible excepción del testimonio de los sentidos". Si se conjunta la idea que las proposiciones observables pueden ser modificadas directamente con lo indicado por Lewis, se tiene que: **(1)** las proposiciones infalibles se pueden modificar directamente y **(2)** las proposiciones teóricas se pueden modificar a través de la condicionalización.

Lo indicado por Lewis está en claro conflicto con la postura falibilista; en adición Raidl (2020) —al igual que lo hacen Howson y Urbach —admite que las proposiciones observables son falibles pues afirma que la evidencia puede serlo y las proposiciones observables pueden ser evidencia.

Sin embargo, si se acepta **(1)** y **(2)** —so pena de contradecir tanto a Raild como a Howson y Urbach — se tiene un falibilismo parcial. Para sobrepasar esto, Weintraub menciona que:

> Considerando que la asignación de probabilidades extremas a proposiciones falibles puede (en ocasiones) ser revocada, la certeza bayesiana, se debe concluir, no genera dogmatismo. Por supuesto, si se produce un cambio en las probabilidades —a través de la regla de Jeffrey—

[19] Recuerdo al lector que 'CJ' está por 'Condicionalización de Jeffrey'.

> desde la reducción de la probabilidad de una proposición a la que previamente se le asignó probabilidad 1, la definición de las probabilidades debe ser extendida a casos donde el denominador de la probabilidad condicional sea equivalente a cero (Weintraub, 1993, p.254)

Varias son las críticas que se pueden aplicar a la propuesta de Weintraub. Una de ellas guarda relación con la regla de Jeffrey. Levi indica que esta regla es una traición al falibilismo. Más precisamente, dice:

El punto de vista de Jeffrey es una traición al falibilismo, no una defensa de éste. En la medida en que nuestros compromisos epistémicos se reflejan únicamente en juicios de probabilidad no extrema, no nos arriesgamos a cometer un error. Si Jones asigna una probabilidad de 0.999 a E, entonces no importa el valor de verdad que tenga E, la «creencia» de Jones (si se puede llamar así) no es falsa. Por lo tanto, si solamente a lo indudable se le debe conceder la probabilidad 1 y sólo se deben creer tales proposiciones (aceptadas como verdaderas), el resultado no es el falibilismo sino un severo escepticismo basado en normas racionalistas de valor epistémico (Levi, 1967, p.209)

Howson y Urbach (2006, p.287) comentan, al igual que Levi, que sostener el grado de creencia 1 en una proposición no equivale a sostener la infalibilidad de dicha proposición, sino que el agente le asigna tal grado de creencia dada su experiencia actual.

En suma, la propuesta de Weintraub no cumple su objetivo, pues entra en conflicto con el falibilismo al sostener la infalibilidad de las proposiciones observables. Además su solución considera la condicionalización de Jeffrey que, tal y como la interpreta Levi, constituye una traición al falibilismo.

4.1. ¿Cómo surgen los ceros?

El surgimiento del cero se puede clasificar en: (1) el problema del reemplazo, (2) el problema de la extensión y (3) el problema del desplazamiento. Más específicamente:

Sea $\mathbb{P}$ el conjunto de funciones de probabilidad sobre un álgebra[20] $\mathcal{A}$, entonces, una función de revisión proposicional es una función $*: \mathbb{P} \times \mathcal{A} \rightarrow \mathbb{P}$, en forma resumida $*(\mathrm{P}, \mathrm{A})$. La CS (también formalizada como $\mathrm{P}_{\mathrm{A}}^{+}$) es una revisión proposicional particular y enfrenta:

- El problema del reemplazo, pues es imposible que $\mathrm{P}(\mathrm{A}) = 0$ y $\mathrm{P}_{\mathrm{A}}^{+} > 0$.
- El problema de la extensión, ya que no es posible que $\mathrm{P}(\mathrm{A}) = 1$, $\mathrm{P}(\mathrm{B}) = 0$ y $\mathrm{P}_{\mathrm{A}}^{+}(\mathrm{B}) > 0$. Note lo siguiente: $\mathrm{P}(\mathrm{B}) = 0$ si y solo si $\mathrm{P}(\neg\mathrm{B}) = 1$. Así, el problema de la extensión no es otra cosa que el problema de la certeza bayesiana.

[20] Para comprender de mejor forma el concepto de σ-algebra, considere lo indicado por Athreya y Lahiri (2006, p.1): imagine un campo abierto $\mathbb{S}$ y una noche de nieve. Al amanecer, usted va al campo a medir la cantidad de nieve en tantos subconjuntos de $\mathbb{S}$ como sea posible. Supongamos ahora que usted tiene las herramientas para medir la nieve exactamente en una clase de subconjuntos, como triángulos, rectángulos, formas circulares, elípticas, etc., sin importar cuán pequeños sean. Es natural tratar de aproximar regiones de formas extrañas mediante combinaciones de estas «formas estándar», y luego utilizar un proceso de limitación para obtener una medida para las regiones de formas extrañas y alcanzar algún límite para tales conjuntos. Sea $\mathcal{B}$ la clase de subconjuntos de $\mathbb{S}$ cuya medida es obtenida de esta manera y sea $\lambda(\mathcal{B})$ la cantidad de nieve en cada $\mathrm{B} \in \mathcal{B}$. Si se cumple con (a) $\emptyset \in \mathcal{B}$, (b) $\mathrm{A} \in \mathcal{B} \Rightarrow \mathrm{A}^{\mathrm{c}} \in \mathcal{B}$ y (c) $\mathrm{A}_1, \mathrm{A}_2, \ldots \in \mathcal{B} \Rightarrow \bigcup_{\mathrm{i}} \mathrm{A}_{\mathrm{i}} \in \mathcal{B}$, entonces $\mathcal{B}$ es una σ-álgebra.

Vol.2 | N°1 | Pág. 24-54 | Julio 2021 | ISSN-E: 0719-9856
https://doi.org/10.35588/cc.v2i1.4903

- El problema del desplazamiento, pues es imposible que $\mathrm{P(A)} \in (0{,}1)$, $\mathrm{P(B)} = 0$ y $\mathrm{P}_{\mathrm{A}}^{+}(\mathrm{B}) > 0$. Suponga que las teorías T, U y V son mutuamente excluyentes. Si un agente considera solo T y U como aquellas teorías que pueden explicar un fenómeno con grados de creencias p y $1 - \mathrm{p}$ respectivamente y, más tarde, el agente descubre que T es falsa y, además, V tiene la capacidad de explicar el fenómeno y pretende éste condicionalizar sobre ambas, es decir, aplicar una función de revisión sobre «T o V», lo anterior no logrará que el grado de creencia en V sea distinto de cero.

Considerando que la CS, o, lo que es lo mismo, el aparato bayesiano ortodoxo, no cuenta con los mecanismos necesarios para sobrepasar los tres problemas descritos, se requiere dar cuenta de una revisión proposicional alternativa. No obstante, antes de definirla, se indicará qué características de la CS permiten la aparición de los problemas previamente señalados.

Recuerde que la condicionalización se define por:

$$\mathrm{P(B|A)} = \mathrm{P}_{\mathrm{A}}^{+}(\mathrm{B}) = \frac{\mathrm{P(B \cap A)}}{\mathrm{P(A)}}$$

De este modo:

1. Si $\mathrm{P(A)} = 0$, entonces $\mathrm{P}_{\mathrm{A}}^{+}$ se indefine.
2. Si $\mathrm{P(A)} = 1$, entonces $\mathrm{P}_{\mathrm{A}}^{+} = \mathrm{P}$.
3. Si $\mathrm{P(A)} \in (0{,}1)$ y $\mathrm{P(B)} = 0$, entonces $\mathrm{P}_{\mathrm{A}}^{+} = \mathrm{P}_{\mathrm{A \setminus B}}^{+}$

donde **(1)** corresponde a la indefinición, **(2)** a la vacuidad y **(3)** a la redundancia. Todas propiedades atribuibles a la CS.

Sometiendo a escrutinio estas tres propiedades se hace evidente que el problema del reemplazo se produce por la indefinición de la CS, pues toda división en la que el denominador es cero es indefinida. El problema de la extensión surge ya que la CS es vacua. En otras palabras, sea A una proposición que tiene grado de creencia equivalente a 1, someterla a revisión nunca cambiará este hecho. Adicionalmente, cualquier revisión proposicional —representada por $\mathrm{P}_{\mathrm{A}}^{*}$— que conserve la certeza enfrentará el problema del desplazamiento.

Se entiende por conservación de la certeza el hecho que $\mathrm{P(B)} = 1$ implica que $\mathrm{P}_{\mathrm{A}}^{*}(\mathrm{B}) = 1$, siempre que $\mathrm{P(A)} > 0$. Expresado de otra forma $\mathrm{P(B)} = 0$ implica $\mathrm{P}_{\mathrm{A}}^{*}(\mathrm{B}) = 0$, si $\mathrm{P(A)} > 0$. Esto se conoce como la conservación del cero.

Una regla de revisión es parcial[21] si $*(\mathrm{P, A})$ está definida si y solo si $\mathrm{P(A)} > 0$. Por otro lado, una regla de revisión es total[22] si y solo si es una función total sobre $\mathbb{P} \times \mathcal{A}^{-}$, donde $\mathcal{A}^{-} = \mathcal{A} \setminus [\emptyset]$. La

[21] Una función parcial es una relación que asocia elementos de un conjunto — el dominio— con, como mínimo, uno de los elementos de otro conjunto —que puede ser el mismo—, llamado contradominio.

[22] Una función es total si está definida para todos sus valores de entrada. Sea $\mathrm{f{:}\,A \rightarrow B}$, f está definida para un elemento $\mathrm{a} \in \mathrm{A}$ si existe un par $(\mathrm{a, f(a)}) \in \mathrm{f}$.

condicionalización es parcial y la parcialidad implica el problema del reemplazo si P no es regular. P es regular si $P(A) = 0$ implica que $A = \emptyset$. La CS no es regular, pues existen elementos, distintos al conjunto vacío, para los que su probabilidad es igual a cero.

4.2. Condicionalización épsilon equívoca

Dadas dos probabilidades P y Q sobre un algebra $\mathcal{A}$, sea $P\epsilon Q = (1-\epsilon)P + \epsilon Q$ su combinación convexa dada por ϵ que pertenece al intervalo $[0,1]$[23]. Observe que ϵ y $(1-\epsilon)$ son ponderadores de las probabilidades Q y P respectivamente, donde ϵ puede interpretarse como el grado de apertura de mente y $(1-\epsilon)$ como el grado de conservadurismo. Note que ambos términos pertenecen al intervalo $[0,1]$, por tanto, se cumple el axioma de normalización.

Para un álgebra finita $\mathcal{A} = \wp(\Omega)$, el equívoco sobre $A \in \mathcal{A}^{-}$es:

Ecuación 1

$$P_{=}^{A}(w)\begin{cases}\dfrac{1}{|A|}, & \text{si } w \in A\\ 0, & \text{eoc.}^{24};\end{cases}$$

$|A|$ corresponde al número de elementos de $A \in \mathcal{A}^{-}$. $\mathcal{A}^{-}$ equivale a $\mathcal{A}$ excluyendo $\emptyset$. La probabilidad Q contiene las proposiciones previamente ignoradas, ya sea que no se consideraron o que, siendo consideradas, se les asignó probabilidad cero.

Imagine la siguiente situación: suponga que tendrá lugar una elección presidencial donde, hasta el tiempo t, solo participan la candidata A y el candidato B. Según las encuestas de opinión la candidata A saldrá electa con una probabilidad p y, consecuentemente, el candidato B será electo con una probabilidad $1 - p$. En el instante $t + 1$ surgen dos nuevas candidatas, siendo estas C y D. Sea C’ la proposición «la candidata C es electa» y D’ la proposición «la candidata D es electa». En efecto, C’ y D’ corresponden a proposiciones previamente ignoradas y pertenecen a Q. En adelante se hará referencia a este ejemplo como el problema de la elección presidencial.

Para el caso de la condicionalización épsilon equívoca, Q es el equívoco sobre las nuevas posibilidades y representa una suposición indiferente frente a éstas, pues si $|A| = n$, cada nueva posibilidad será evaluada con la razón $\frac{1}{n}$ y posteriormente ponderada por ϵ. En otras palabras, las nuevas posibilidades se establecen equiprobables. En el problema de la elección presidencial $|A| = 2$, por tanto, las proposiciones C’ y D’ bajo la condicionalización épsilon equívoca serán evaluadas de la forma $\frac{1}{2} \times \epsilon$. Ahora bien ¿cómo se establece ϵ? La única restricción es que debe pertenecer al intervalo $[0,1]$, por tanto, un valor posible

[23] Sea $S = \{x_1, x_2, \dots, x_k\}$ un subconjunto de $\mathbb{R}^n$ y sean $\lambda_1, \lambda_2, \dots, \lambda_k$ escalares; $x = \lambda_1 x_1, \lambda_2 x_2, \dots, \lambda_k x_k$. Así, x es una combinación lineal sin importar los valores de λ_i y x es una combinación convexa si $\sum_{i=1}^{k} \lambda_i = 1$ y $\lambda_i > 0 \ \forall\, i \in (1, \dots k)$.

[24] La sigla «eoc» es la forma resumida de «en otro caso».

Vol.2 | N°1 | Pág. 24-54 | Julio 2021 | ISSN-E: 0719-9856
https://doi.org/10.35588/cc.v2i1.4903

para épsilon es $\frac{1}{8}$. Si este es el caso, el grado de creencia para C’ equivaldría a $\frac{1}{2} \times \frac{1}{8} = \frac{1}{16}$; misma cuestión es el caso para D’. Es natural que el lector se pregunte ¿cómo se interpreta ϵ? Lo primero es advertir que si $\epsilon = 0$ se obtiene la CS y además $\epsilon^{-} = (1 - \epsilon)$ será necesariamente igual a 1. Recuerde que se propone una revisión proposicional alternativa a la CS con el propósito de modelar cambios en los grados de creencia para agentes de mente abierta, por tanto, ϵ representa el grado de apertura de mente del agente y $\epsilon^{-} = (1 - \epsilon)$ representa el grado de conservadurismo. Un agente bayesiano esencialmente conservador, es decir, aquel agente que no esté dispuesto a considerar un ϵ distinto de cero, se encontrará siempre gobernado por la CS. Por consiguiente, toda vez que dicho agente enfrente una instancia de los problemas del surgimiento del cero se encontrará con que la única forma de dar cuenta de las posibilidades previamente ignoradas es transgredir los axiomas de la probabilidad.

Sea P una probabilidad sobre un álgebra finita $\mathcal{A}$ y $\mathrm{A} \in \mathcal{A}^{-}$. P’ es un paso de condicionalización épsilon de P por $\mathrm{A} \in \mathcal{A}^{-}$ si y solo si existe un ϵ que pertenece al intervalo $[0,1)$ y una probabilidad Q con $\Omega(\mathrm{Q}) \cap \Omega(\mathrm{P}) = \emptyset$ tal que $\mathrm{P}' = (\mathrm{P}\epsilon\mathrm{Q})^{+}_{\mathrm{A}}$[25].

La condicionalización épsilon equívoca, sobre $\mathcal{A}^{-}$ se puede expresar por:

Ecuación 2

$$\mathrm{P}^{\epsilon}_{\mathrm{A}} \begin{cases} (\text{caso } 1)\mathrm{P}^{+}_{\mathrm{A}}, & \text{si } \mathrm{A} \subseteq \Omega(\mathrm{P}) \\ (\text{caso } 2)\mathrm{P}^{\mathrm{A}}_{=}, & \text{si } \mathrm{A} \cap \Omega(\mathrm{P}) = \emptyset \\ (\text{caso } 3)(\mathrm{P}\epsilon\mathrm{P}^{\mathrm{A}\setminus\Omega(\mathrm{P})}_{=})^{+}_{\mathrm{A}}, & \text{eoc.} \end{cases}$$

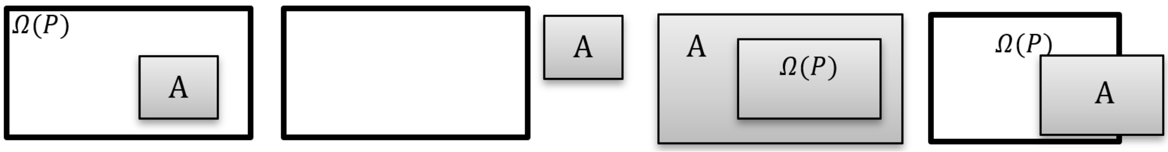
La condicionalización épsilon equívoca es la revisión resultante de una extensión equívoca y conservadora de P a $\mathrm{P}\epsilon\mathrm{P}'_{=}$ para algún ϵ perteneciente al intervalo $[0,1)$. Una extensión de la probabilidad P es una probabilidad R tal que $\Omega(\mathrm{R}) \supseteq \Omega(\mathrm{P})$. Esa extensión es conservadora si y solo si conserva las razones de $\Omega(\mathrm{P})$, y es equívoca si y solamente si todas las nuevas posibilidades sugeridas por la nueva evidencia reciben el mismo valor. Se entiende por nuevas posibilidades aquellas que no se encontraban en $\Omega(\mathrm{P})$; formalmente $\mathrm{A} \setminus \Omega(P)$. Retornando al problema de la elección presidencial, las nuevas posibilidades son C’ = «la candidata C es electa» y D’ = «la candidata D es electa».

$\mathrm{P}^{\epsilon}_{\mathrm{A}}$ se interpreta así: para el caso 1 la CS se encuentra definida. En el caso 2 la condicionalización se indefine y genera el problema del reemplazo. Como se quiere condicionalizar sobre la nueva evidencia A y la intersección del espacio muestral $\Omega(\mathrm{P})$ con A resulta en el conjunto vacío y ni la evidencia aceptada ni los a priori permiten distinguir entre las nuevas posibilidades entonces el a posteriori tampoco lo puede hacer. Por aquello, se debe proceder según $\mathrm{P}^{\mathrm{A}}_{=}$. El caso 3 genera los problemas de desplazamiento y el de la extensión. Para el caso 3, $\mathrm{P}^{\epsilon}_{\mathrm{A}}$ determina que el agente aumente uniformemente las nuevas posibilidades

[25] La colección de todos los posibles resultados de un experimento recibe el nombre de espacio muestral y se denota usualmente con Ω. Un experimento corresponde al proceso de recolectar datos para un fenómeno que exhibe variaciones en sus resultados (Bhattacharyya y Johnson, 1977, p.62). Piense en el experimento lanzar un dado, el espacio muestral será $\Omega = \{1, 2, 3, 4, 5, 6\}$. En el problema de la elección presidencial —a partir del instante $\mathrm{t} + 1$— el espacio muestral equivale a $\Omega = \{\mathrm{A}', \mathrm{B}', \mathrm{C}', \mathrm{D}'\}$.

sugeridas $A \setminus \Omega(P)$, en algún grado leve —es decir, para un valor pequeño de ϵ, y que en consecuencia disminuya los valores de probabilidad de las otras posibilidades, y luego revise sobre A con la CS (Raidl, 2020, p.16).

Una representación gráfica de los casos es la siguiente:



Donde, de izquierda a derecha, las figuras representan para P_A^ϵ: la primera el caso 1; la segunda el caso 2, que genera el problema del reemplazo; la tercera corresponde al caso 3, que genera el problema de la extensión y la cuarta figura también corresponde al caso 3, que genera el problema del desplazamiento.

Con el propósito de aportar claridad, considere los siguientes ejemplos. Suponga que un científico considera que un fenómeno puede ser explicado únicamente por dos teorías, siendo estas U y V. Dicho científico asigna a la teoría U un grado de creencia equivalente a q y $(1-q)$ a la teoría V tal que q pertenece al intervalo $(0,1)$. En el instante $t+1$ aparece una pieza de evidencia que confirma V. Este escenario puede ser modelado sin problema por la CS. No obstante, no ocurre lo mismo con el problema de la elección presidencial, pues este no puede ser modelado por la CS, sino que se debe apelar a la condicionalización épsilon, particularmente al caso 2. Por último, una instancia del caso 3, particularmente del problema del desplazamiento es la siguiente:

Asuma un escenario idealizado donde únicamente se puede dar una y solo una de las siguientes cuestiones: o bien todos los cuervos son negros, o bien todos los cuervos son blancos o exactamente la mitad de los cuervos son negros y la otra mitad blancos. Dado lo anterior, suponga las siguientes hipótesis:

- H_1: «todos los cuervos son negros» $P_i\,(e|H_1) = 1$
- H_2: «ningún cuervo es negro» $P_i\,(e|H_2) = 0$
- H_3: «la mitad de los cuervos son negros» $P_i\,(e|H_3) = 0{,}5$

Sea $e =$ «observar un cuervo negro». Entonces $P_i\,(e|H_1)$ se interpreta del siguiente modo: la probabilidad de observar un cuervo negro, dado que acepta la hipótesis H_1. Como $H_1 =$ «todos los cuervos son negros», entonces $P_i\,(e|H_1) = 1$. $P_i\,(e|H_2)$ corresponde a la probabilidad de observar un cuervo negro, dado que se acepta la hipótesis que indica que ningún cuervo es negro, por tanto $P_i\,(e|H_2) = 0$. $P_i\,(e|H_3)$ corresponde a la verosimilitud de observar un cuervo negro dado que la mitad de los cuervos son negros, así $P_i\,(e|H_3) = 0{,}5$[26].

[26] Advierta que en el ejemplo de los cuervos la $\sum_{i=1}^{3} P_i\,(e|H_i) > 1$ y lo anterior no constituye una transgresión de los axiomas de la probabilidad. Para aportar claridad en este punto, es necesario señalar la diferencia entre probabilidad y verosimilitud [*likelihood*]. La probabilidad se vincula a los resultados posibles; la verosimilitud se vincula a las hipótesis. Los resultados posibles son mutuamente excluyentes y exhaustivos. Suponga que se le pide a un sujeto que prediga el resultado de

Vol.2 | N°1 | Pág. 24-54 | Julio 2021 | ISSN-E: 0719-9856
https://doi.org/10.35588/cc.v2i1.4903

Suponga además que $P_i(H_1) = 0{,}1$, $P_i(H_2) = 0{,}9$ y $P_i(H_3) = 0$. En otras palabras, para un agente dado la probabilidad de que H_1 sea cierta, antes de observar e, es igual a $0{,}1$; la probabilidad de que H_2 sea cierta, antes de observar e, es igual a $0{,}9$ y la probabilidad de que H_3 sea cierta, antes de observar e, es igual a 0.

Recuerde que los axiomas de la probabilidad exigen que $\sum_{k=1}^{3} P_i(H_k) = 1$, entonces, como $P_i(H_1) = 0{,}1$ y $P_i(H_2) = 0{,}9$, concluimos que $P_i(H_3)$ debe necesariamente ser igual a cero.

Para aplicar la CS se debe calcular $P_i(e) = \sum_{k=1}^{3} P_i\,(e|H_k) \times P_i(H_k) = 1 \times 0{,}1 + 0 \times 0{,}9 + 0{,}5 \times 0 = 0{,}1$. Así, $P_i(e) = 0{,}1$. Imagine que se observa un cuervo negro. Así, $P_f(H_1) = \frac{P_i\,(e|H_1)P_i(H_1)}{P(e)} = \frac{1\times 0{,}1}{0{,}1} = 1$. Consecuentemente, el axioma de normalización requiere que $P_f(H_2) = P_f(H_3) = 0$. Bajo estas circunstancias $P(e) = 1$. Es decir, desde el punto de vista del agente la probabilidad de observar un cuervo negro es 1. Sin embargo, ¿qué ocurre si ahora, en contra de todo pronóstico, el agente observa un cuervo blanco? Lo primero es advertir que con la CS no se puede dar cuenta de lo anterior. No obstante, ¿cómo puede la condicionalización épsilon actualizar los grados de creencia bajo estas circunstancias?

Primero se debe establecer ϵ. Sea $\epsilon = 0{,}2$. Así, se tiene que $P_i\,(H_1) = 0{,}8$ y que $P_i(H_3) = 0{,}2$. Ahora se debe recalcular $P(e) = 1 \times 0{,}8 + 0{,}5 \times 0{,}2 = 0{,}9$. Antes de continuar se debe indicar lo siguiente: e es evidencia para H si y solo si $P(H|e) > P(H)$. Luego, en este ejemplo e es evidencia para H_1 y $\neg e =$ observar un cuervo blanco, es evidencia para H_3. Se sabe que $P(\neg e) = 1 - P(e) = 0{,}1$ y que $P(\neg e|H_k) = 1 - P(e|H_k)$. De esta manera:

- H_1: «todos los cuervos son negros» $\quad P_i\,(\neg e|H_1) = 0$
- H_2: «ningún cuervo es negro» $\quad P_i\,(\neg e|H_2) = 1$
- H_3: «la mitad de los cuervos son negros» $\quad P_i\,(\neg e|H_3) = 0{,}5$

Ahora es posible calcular $P_f(H_1) = \frac{P_i\,(\neg e|H_1)P_i(H_1)}{P(\neg e)} = \frac{0\times 1}{0{,}1} = 0$. Este resultado es esperable, pues como se observó un cuervo blanco, es natural asignar una probabilidad cero a la hipótesis que señala que todos los cuervos son negros. Asimismo, $P_f(H_3) = \frac{P_i\,(\neg e|H_3)P_i(H_3)}{P(\neg e)} = \frac{0{,}5\times 0{,}2}{0{,}1} = 1$.

Retornando a la condicionalización épsilon equívoca, note que ésta asigna a todas las nuevas posibilidades la misma probabilidad. La razón de lo anterior radica en el supuesto de ausencia de información adicional. Véase el problema de la elección presidencial, donde a las candidatas C y D se le

cada uno de los 10 lanzamientos de una moneda. Sólo hay 11 resultados posibles, siendo estos; ningún acierto, 1 acierto, 2 acierto, …, 10 aciertos. El resultado real siempre será uno y solo uno de los resultados posibles. Por lo tanto, las probabilidades que se adjuntan a los posibles resultados deben circunscribirse al axioma de normalización, es decir $P(0\text{ acierto}) + P(1\text{ acierto}) + \cdots + P(10\text{ aciertos})$ debe sumar uno. Lo anterior no aplica a las verosimilitudes, pues estas no son ni mutuamente excluyentes ni exhaustivas. Una revisión más detallada a esta materia puede consultarse en Hacking (1972) y Edwards (1984).

asignó el equívoco $\frac{1}{16}$. En otras palabras, el agente no cuenta con información adicional que le permita ponderar a las candidatas de forma distinta, por ejemplo, estableciendo que $\mathrm{P}(\mathrm{C}') > \mathrm{P}(\mathrm{D}')$ o que $\mathrm{P}(\mathrm{C}') < \mathrm{P}(\mathrm{D}')$. Sin embargo, sería totalmente válido que un agente quisiera diferenciar entre ambas. Si esto fuera el caso, se debe apelar a la condicionalización épsilon libre o a la condicionalización épsilon binaria, pues ambas pueden dar cuenta de información adicional[27].

En virtud de que ya se ha expuesto en detalle la condicionalización épsilon, vale comentar una última cuestión respecto al parámetro épsilon. Imagine un agente que se encuentra en pleno aprendizaje de una materia particular. Se espera que, dado que el agente está en un terreno hasta cierto punto desconocido, su ϵ esté distante de cero. Del mismo modo, un agente experto en una particular cuestión debiera contar con un ϵ cercano a cero. En otras palabras, durante el proceso de aprendizaje se produce una caída continua del parámetro ϵ producto de la acumulación de evidencia. Así $\epsilon^{-} = (1 - \epsilon)$ representa la confianza que el agente asigna a la fuente que proporciona la evidencia.

Ahora bien, recuerdo al lector que una cuestión que quedo pendiente es justificar la exclusión de la CJ del bayesianismo ampliado. Lo primero es precisar que tanto la CS como la CJ son casos particulares del axioma de aditividad. Además, la CJ, si bien tiene un carácter más amplio que la CS, enfrenta sus propias dificultades, a saber, ¿cuándo se inicia el proceso de actualización? Se puede indicar que este proceso debe iniciarse cuando $\mathrm{P}(\mathrm{e})$ es distinta de 1. Pero esto no da cuenta del proceso racional que inicia la actualización. Así, la inclusión de la CJ abre una cuestión que no tiene clara respuesta. Por consiguiente, no cumple con los requisitos del bayesianismo ampliado para su inclusión. No obstante, vale la pena indicar que esto no constituye una pérdida de generalidad para el bayesianismo ampliado, pues para constituirse esta pérdida debe indicarse un problema que pueda ser solucionado por la CJ y no por el bayesianismo ampliado.

Otra crítica que se podría hacer a esta investigación es que no considera el bayesianismo cuántico ya que, según este, cualquier agente puede evaluar sus expectativas según su propia experiencia personal. En otras palabras, existe compatibilidad entre esta interpretación del bayesianismo y el subjetivismo. Y, en efecto, hasta aquí parece no existir conflicto alguno para que el bayesianismo ampliado incorpore al bayesianismo cuántico, pero se debe comentar lo siguiente: los modelos de actualización probabilísticos cuánticos están basados en lógicas no clásicas y por esto $\mathrm{P}(\mathrm{H}|\mathrm{E})$ no es invertible. Lo anterior no socava los modelos cuánticos de actualización, sino que manifiesta que no son compatibles con el bayesianismo ampliado aquí propuesto, puesto que este está regido por la CS y esta considera invertir probabilidades condicionales[28].

Recapitulando, recuerde que uno de los objetivos de esta investigación es dar cuenta de un aparato bayesiano ampliado que tenga la capacidad de sobrepasar problemas clásicos que afectan al bayesianismo ortodoxo. Si bien hasta aquí solo se han resuelto de manera explícita el problema de la incorregibilidad y

[27] Para más detalles sobre la condicionalización épsilon libre y binaria véase Raidl (2020).

[28] La discusión en torno a la condicionalización se Jeffrey y el Bayesianismo cuántico se incluyó por sugerencia de los revisores. Agradezco nuevamente sus comentarios. Para más detalles sobre bayesianismo cuántico consulte Khrennikov (2016), Garola (2018) y DeBrota et al. (2020).

 Vol.2 | N°1 | Pág. 24-54 | Julio 2021 | ISSN-E: 0719-9856
https://doi.org/10.35588/cc.v2i1.4903

el conjunto de problemas provocados por el surgimiento del cero, esto ya constituye un avance significativo. A continuación, se explicitará cómo este mismo aparato puede resolver el problema de la evidencia antigua y el problema de las teorías nuevas.

5. Evidencia antigua y teorías nuevas

Glymour (1980) presentó a los bayesianos el siguiente problema: suponga que se propone una hipótesis H que explica cierta evidencia ya conocida, representada por e. ¿Puede H ser confirmada por e? Intuitivamente, la respuesta es sí, e puede confirmar H. En otras palabras, aquella evidencia que se conoce con antelación puede dar soporte a una nueva hipótesis. Como ejemplo, considere el problema de explicar el desplazamiento del perihelio de mercurio. Durante mucho tiempo la mecánica newtoniana (en adelante MN) no dio cuenta de este fenómeno y postuló hipótesis auxiliares (e.g. la existencia de otro planeta dentro de la órbita de Mercurio); hipótesis que no fueron confirmadas. Einstein se dio cuenta en la década de 1910 que su Teoría General de la Relatividad (en adelante TGR) daba cuenta del cambio del perihelio. De acuerdo con la mayoría de los físicos, explicar esta «evidencia antigua» (en el sentido de los datos conocidos anteriormente) confirió un grado sustancial de confirmación a la TGR, quizás incluso más que algunas piezas de evidencia nuevas, como las observaciones del eclipse solar de Eddington de 1919 (Sprenger & Hartmann, 2019, p.131).

La cuestión es que la CS no puede dar cuenta de la situación anterior, pues como e es conocida —es evidencia antigua—, formalmente $\mathrm{P(e)} = 1$:

$$\mathrm{P(H|e)} = \frac{\mathrm{P(H)P(e|H)}}{\mathrm{P(e)}} = \mathrm{P(H)P(e|H)} \leq \mathrm{P(H)}$$ [29]

En otras palabras, e no puede confirmar H. Muchos esfuerzos se han realizado por solucionar este problema. Uno de ellos —que es de interés para esta investigación— es llevado a cabo por Howson (1991). Su propuesta resumida es la siguiente: primero, se debe notar que los datos no constituyen una evidencia a favor o en contra de una hipótesis aislada, sino que la evidencia e se encuentra en conjunción con información contextual. Por ejemplo, el encontrar un diccionario en cierta avenida de Santiago no es evidencia para esclarecer si Pedro mató a Diego, o una encuesta de opinión sobre los efectos del confinamiento obligatorio de la población debido a una pandemia no es evidencia para establecer quién será el próximo presidente de Chile. Luego «e es evidencia para» denota una relación tripartita entre la evidencia, la hipótesis y un cuerpo K de información. Una condición para aplicar la condicionalización es que $\mathrm{e} \notin \mathrm{K}$. De este modo Howson (1991, p.548) señala que el problema de la evidencia antigua no es en realidad un problema, sino más bien un recordatorio que si $\mathrm{e} \in \mathrm{K}$, entonces esa pieza de evidencia debe ser eliminada del cuerpo de conocimiento K.

[29] Note que se tiene la igualdad $\frac{\mathrm{P(H)P(e|H)}}{\mathrm{P(e)}} = \mathrm{P(H)P(e|H)}$ pues $\mathrm{P(e)} = 1$. Luego, se sabe que $\mathrm{P(H)P(e|H)} \leq \mathrm{P(H)}$ ya que $\mathrm{P(e|H)} \in [0,1]$.

La solución de Howson se puede resumir en dos pasos. Primero, se debe eliminar e del cuerpo de conocimiento (el cuerpo de conocimiento una vez se elimina la evidencia e, es representado por $K\backslash e$). Luego, se debe calcular $P^c(e)$[30]. Frente a esta propuesta se pueden esgrimir ciertas objeciones, por ejemplo, imagine que únicamente α y β pertenecen a K para un agente dado. Lo anterior, se puede axiomatizar como $\{\alpha, \beta\}$ o $\{\alpha, \alpha \rightarrow \beta\}$. Imagine que α constituye evidencia antigua. Se debe entonces eliminar α, sin embargo, aquello resultará en $\{\beta\}$ y $\{\alpha \rightarrow \beta\}$. ¿Cuál es el criterio de elección con el que cuenta el agente? ¿Cómo determina si seleccionar $\{\beta\}$ o $\{\alpha \rightarrow \beta\}$?

Howson y Urbach comentan que Suzuki (2005) entrega la solución indicando:

> Suzuki ha demostrado que hay funciones de contracción probabilísticas consistentes que representan la eliminación de e de K en relación con condiciones límite plausibles en tales funciones (estas condiciones son proporcionadas por la conocida teoría AGM (Alchourron-Gardenfors-Makinson) (Howson y Urbach, 2006, p.299)[31]

En efecto, Suzuki indica:

> El resultado de eliminar de K «todo aquello que depende de la evidencia antigua e» implicará al menos el resultado de eliminar de K todo lo lógicamente dependiente de e […]. Asumo que el resultado de eliminar de K todo lo lógicamente dependiente de la evidencia antigua e es el resultado de la eliminación de algunas proposiciones de K tal que no sea posible deducir e (Suzuki, 2005, p.39).

Aceptada la solución de Suzuki, resta con indicar cómo se calcula $P^c(e)$. Es decir, cual es el grado de creencia en e si se obvia que e es el caso. Se sabe que $P^c(e) = \sum_{i=1}^{k} P^c(e|H_i) \times P^c(H_i)$, y por las leyes de la probabilidad se tiene:

$$P^c(e) = \sum_{i=1}^{k} P^c(e|H_i) \times P^c(H_i) = P^c(e|H_1) \times P^c(H_i) + \cdots + P^c(e|H_k) \times P^c(H_k) + P^c(e|\neg(H_1 \vee \ldots \vee H_k)) \times P^c(\neg(H_1 \vee \ldots \vee H_k)).$$

Salmon señala (véase Curd y Cover (1998, p.643)) que el término $P^c(e|\neg(H_1 \vee \ldots \vee H_k)) \times P^c(\neg(H_1 \vee \ldots \vee H_k))$ es intratable[32], pues el agente nunca agotará todas las hipótesis posibles que pueden explicar un fenómeno y, más importante aún, el agente no puede determinar qué proporción de todas las posibles explicaciones se están considerando. Sin embargo, desde el bayesianismo subjetivo se puede tomar la siguiente postura: simplemente indicar que $P^c(\neg(H_1 \vee \ldots \vee H_k)) = 0$. Es natural que el lector se pregunte cómo se interpreta $\neg(H_1 \vee \ldots \vee H_k)$. Imagine que para cierto fenómeno se consideran solamente

[30] $P^c(e)$ corresponde al grado de creencia contrafáctico en e.

[31] Para más detalles del modelo AGM véase Hansson (2017).

[32] Algunos autores (Wenmackers y Romeijn, 2016; Tesic, Eva, y Hartmann, 2017) han llamado a $\neg(H_1 \vee \ldots \vee H_k)$ *catch-all hypothesis*. Para una solución del problema de la evidencia antigua mediante la introducción del *catch-all hypothesis*, véase Wenmackers y Romeijn, (2016).

Vol.2 | N°1 | Pág. 24-54 | Julio 2021 | ISSN-E: 0719-9856
https://doi.org/10.35588/cc.v2i1.4903

dos hipótesis H_1 y H_2, en aquel caso $k = 2$ y $\neg(H_1 \vee H_2)$ representa una hipótesis alternativa tal que ni H_1 es el caso ni H_2 es el caso.

Resumiendo, si se establece que $P^c(\neg(H_1 \vee ... \vee H_k)) = 0$ se tiene:

$$P^c(e) = \sum_{i=1}^{k} P^c(e|H_i) \times P^c(H_i) = P^c(e|H_1) \times P^c(H_i) + \cdots + P^c(e|H_k) \times P^c(H_k).$$

Sin embargo, queda una tercera cuestión problemática con la propuesta de Howson, pues se requiere que $P(e)_{t+1} < 1$ toda vez que $P(e)_t = 1$. Una forma de abordar esta situación es apelar a la reconstrucción racional propuesta en Lang (1999), quedando del siguiente modo: el agente puede construir un argumento en el cual e confirme H. Esto ocurre si el agente toma como su conjunto primario de proposiciones relevantes un subconjunto de K que contiene la hipótesis H y todas las proposiciones auxiliares que dan soporte a H, incluyendo e. Sobre este conjunto, el agente construye una distribución de probabilidad tal que su grado de creencia en e es menor que 1. En este momento, se aplica la propuesta de Howson para calcular $P^c(e)$. Esto es análogo al grado de creencia contrafáctico señalado previamente, es decir, el agente define una distribución de probabilidad sobre el conjunto de proposiciones que establecería como relevantes tal que e no fuera el caso. En el siguiente paso se introduce $P(e) = 1$ y luego actualiza su distribución de probabilidad, de acuerdo con la CS.

Por ejemplo, sea $K = \{e_1, e_2, e_3, e_4\}$. Además, e_1, e_2 y e_3 confirman la MN, sin embargo e_4 no es explicada por la teoría y se considera una anomalía. Se tiene además que $P(e_i) = 1, i \in \{1,2,3,4\}$. Luego, en t_1, aparece la TGR. La TGR es confirmada por e_5 y e_6, evidencias conocidas después de t_1. Pero la TGR también puede dar cuenta de K. Si se pone la atención en e_4, que no era explicada por la MN, se debe proceder del siguiente modo[33]:

Paso 1 ➔ $P(MN) = 0{,}3; P(TGR) = 0{,}7; P^c(e_4) = 0{,}7$

Paso 2 ➔ $P(e_5) = 1; P(e_6) = 1; P(MN) = 0{,}3; P(TGR) = 0{,}7; P(e_4) = 1;\ P_i(e_4|MN) =; P_i(e_4|TGR) = 1;\ P_i(e_5|MN) = 0; P_i(e_5|TGR) = 1;\ P_i(e_6|MN) = 0; P_i(e_6|TGR) = 1$

Paso 3 ➔ $P(MN) = 0; P(TGR) = 1$

Es apropiado señalar que esta estrategia no es la única para superar el problema de la evidencia antigua. Una lista no exhaustiva de soluciones es la siguiente: Sprenger (2015), Hartmann y Fitelson (2015) y Wenmackers y Romeijn (2016). Ahora bien, sobre la base de la propuesta explorada en esta investigación se cree que ha quedado demostrado que el bayesianismo ampliado puede dar cuenta de la evidencia antigua.

El problema de las teorías nuevas Talbott lo resume así:

> Suponga que hay una teoría H_1 que generalmente se considera altamente confirmada por la evidencia disponible E. Es posible que simplemente la introducción de una teoría alternativa

[33] En este ejemplo idealizado e_1, e_2 y e_3 confirman tanto la MN como la TGR.

H_2 pueda conducir a una erosión del soporte de H_1. Es plausible pensar que la introducción de Copérnico de la hipótesis heliocéntrica tuvo este efecto en la astronomía ptolemaica, centrada en la tierra, que no había sido cuestionada anteriormente. Este tipo de cambio no puede explicarse por la CS (Talbott, 2016).

En relación con el mismo punto, Strevens señala lo siguiente:

> El problema de la evidencia antigua bien podría llamarse el problema de las nuevas teorías, pues si no existieran teorías nuevas o, lo que es lo mismo, si todas las hipótesis estuvieran disponibles antes de que se obtenga cualquier pieza de evidencia el problema no existiría (…) sin importar el nombre que se le dé, el problema es considerado por la mayoría de los bayesianos como uno que urge solucionar (Strevens, 2006, p.15).

Sobre lo que indica Strevens es conveniente introducir un matiz. Si todas las hipótesis para un fenómeno particular son conocidas —a pesar de la alta improbabilidad de aquello— el problema tampoco desaparece, ya que si una teoría conocida se tiene por falsa, no hay nada en el aparato bayesiano tradicional que pueda cambiar aquello. En efecto, esta situación fue señalada previamente como el problema del desplazamiento. Por otro lado, hay una estrecha relación entre la nueva evidencia y las nuevas teorías o, dicho de otro modo, las nuevas piezas de evidencia motivan la aparición de nuevas teorías. En suma, la solución previamente expuesta para solucionar el problema de la evidencia antigua no permite solucionar en toda ocasión el problema de las teorías nuevas.

Suponga la teoría T_1 y que a tal teoría un agente asigna el grado de creencia 1. Luego, que aparece una teoría nueva T_2 y que el agente quiere asignarle una probabilidad mayor que cero. Esto evidentemente no es posible, pues equivale al problema de la certeza tratado previamente. Ahora bien, y desde otra perspectiva, si existen dos teorías, siendo estas T_1 y T_2 con grados de creencia p y $(1-p)$ respectivamente y otra teoría T_3 que se consideró falsa, la CS tampoco podría dar cuenta de tal situación. Esto se conoce como el problema del desplazamiento.

A continuación, se indicará como la condicionalización épsilon equívoca puede solucionar el problema de la certeza bayesiana[34].

Recuerde que la condicionalización épsilon equívoca está definida por:

$$P_A^{\epsilon}\begin{cases}(\text{Caso }1)P_A^{+}, & \text{si } A\subseteq\Omega(P)\\ (\text{Caso }2)P_{=}^{A}, & \text{si } A\cap\Omega(P)=\emptyset\\ (\text{Caso }3)(P\epsilon P_{=}^{A\setminus\Omega(P)})_A^{+}, & \text{eoc.}\end{cases}$$

El problema de la certeza bayesiana, tal y como se describió con antelación corresponde al caso 2 de la función P_A^{ϵ}. Observe que se tienen dos teorías T_1 y T_2, además $P_i(T_1)=1$ y $P_i(T_2)=0$. Note que $\Omega(P)=\{T_1\}$ y $\Omega(Q)=\{T_2\}$, entonces $\Omega(Q)\cap\Omega(P)=\emptyset$. Recuerde que se indicó que, dadas dos

[34] Una solución al problema del desplazamiento fue ejemplificada en la sección 3.2.

Vol.2 | N°1 | Pág. 24-54 | Julio 2021 | ISSN-E: 0719-9856
https://doi.org/10.35588/cc.v2i1.4903

probabilidades P y Q sobre un álgebra $\mathcal{A}$, $\mathrm{P}\epsilon\mathrm{Q} = (1-\epsilon)\mathrm{P} + \epsilon\mathrm{Q}$ denota una combinación convexa tal que ϵ pertenece al intervalo $[0,1]$. La combinación convexa exige que ϵ y $(1-\epsilon)$ sean mayores que cero y que $\epsilon + (1-\epsilon) = 1$. Además, el número de elementos de A, representado por $|A|$ es igual a 1. Como el agente quiere reducir el valor de $\mathrm{P}(\mathrm{T}_1)$, puede asignar, por ejemplo, un $\epsilon = 0{,}1$. Si lo anterior es el caso, entonces $(1-\epsilon)\mathrm{P} + \epsilon\mathrm{Q} = (1-0{,}1) \times \mathrm{T}_1 + 0{,}1 \times \mathrm{T}_2$. Nótese que este es el caso más simple, sin embargo, si en lugar de considerar solo una nueva teoría el agente considerase dos, representadas por T_2 y T_3 con $\epsilon = 0{,}1$, se tiene que $(1-\epsilon)\mathrm{P} + \epsilon\mathrm{Q} = (1-0{,}1) \times \mathrm{T}_1 + 0{,}1 \times \frac{1}{2} \times \mathrm{T}_2 + 0{,}1 \times \frac{1}{2} \times \mathrm{T}_3$, así $\sum_{k=1}^{3} \mathrm{P_f}(\mathrm{T_k}) = 1$.

En este punto una objeción posible es que, si en lugar que un y solo un agente se enfrente a esta situación lo hacen n agentes a la vez, nada impide que estos agentes asignen valores muy distintos a ϵ. En efecto, no hay nada que impida lo anterior ni tampoco hay necesidad de tal impedimento. En relación con esto, Hawthorne dice:

> Si la inducción bayesiana va a producir acuerdo intersubjetivo entre los agentes con respecto al apoyo inductivo de las hipótesis, entonces la evidencia debe producir de alguna manera una convergencia para concordar entre las probabilidades posteriores de diferentes agentes bayesianos, a pesar de sus desacuerdos iniciales sobre las plausibilidades de las hipótesis. Por lo tanto, los defensores de la inducción bayesiana han investigado las circunstancias en las que las pruebas podrían «lavar» los efectos de los antecedentes subjetivos y hacer que los agentes converjan en sus probabilidades posteriores (Hawthorne, 1994, p.241).

Los teoremas de convergencia bayesiana establecen condiciones bajo las cuales la acumulación de evidencia induce a los agentes a converger al acuerdo intersubjetivo. Esta materia es explorada en detalle por Friedman (1997), Grünwald (2012), Belot (2013) y Huttegger (2015).

En términos amplios, se puede aseverar que sin importar qué tan diversos sean los valores de ϵ entre los agentes, sus opiniones —entendidas como grados de creencia— convergerán en la medida que se vean expuestos a las mismas evidencias.

6. Conclusión

Recuerde que esta investigación suscribe al bayesianismo subjetivo y se esbozó una breve crítica al frecuentismo, mas no se realizó una discusión exhaustiva. Se puede argumentar que en ciertos contextos —por ejemplo, en simulación— el frecuentismo es la interpretación adecuada, pues aquí los parámetros poblacionales son conocidos y por tanto la principal crítica esgrimida en su contra se desvanece. En consecuencia, la validez del subjetivismo y del frecuentismo pareciera ser contextual, y tales contextos deben ser explicitados.

Adicionalmente, se mostró un mecanismo capaz de lidiar con el falibilismo y con los problemas que éste genera para el bayesianismo, a saber, el dogmatismo bayesiano y el surgimiento del cero. Este mecanismo es el bayesianismo ampliado. Este, a su vez, se encuentra dotado con una reconstrucción racional que permite a los agentes cambiar sus grados de creencia sin apelar a nueva evidencia, cuestión que el bayesianismo ortodoxo no está en condiciones de hacer, pues se indicó que la condicionalización simple se activa a la luz de nueva evidencia.

Se argumentó, además, por qué debe excluirse la condicionalización de Jeffrey. Sin embargo, la versión del bayesianismo ampliado aquí expuesto no pretende ser una versión inmutable, sino que todo lo contrario. Si se llegara a encontrar una dificultad que no pueda ser superada por lo aquí expuesto, se debiera buscar una extensión que, sin reemplazar ni la CS ni la condición épsilon, permita superarla.

Referencias

Armendt, B. (1980). "Is there a Dutch book argument for probability kinematics?" *Philosophy of Science* 47(4): 583-588. doi: https://doi.org/bjrwpm

Athreya, K. B. y Lahiri, S. N. (2006). *Measure Theory and Probability Theory*. Nueva York: Springer Science & Business Media.

Belot, G. (2013). "Bayesian orgulity". *Philosophy of Science*, 80(4): 483-503. doi: https://doi.org/gpkp

Bernardo, J.M. (1997). *Bruno de Finetti en la Estadística Contemporánea*. José-Miguel Bernardo (página personal, Universitat de València). URL: https://www.uv.es/~bernardo/publications.html. Consultado el 13 de julio del 2020.

Bhattacharyya, G., y Johnson, R. (1977). *Statistical Concepts and Methods*. USA: John Wiley & Sons.

Bird, A. (2018). Inference to the Best Explanation, Bayesianism, and Knowledge. En McCain, K. y Poston, T. (2018) *Best Explanations: New Essays on Inference to the Best Explanation*. Oxford: Oxford University Press. doi: https://doi.org/gprb

Curd, M., y Cover, J. A. (1998). *Philosophy of science: The central issues.* New York: W.W. Norton & Company.

de Finetti, B. (1974). *Theory of Probability: A Critical Introductory Treatment*. London: John Wiley and Sons. doi: https://doi.org/gprd

DeBrota, J. B., Fuchs, C. A., Pienaar, J. L. y Stacey, B. C. (2020). "The Born Rule as Dutch-Book Coherence (and only a little more)". arXiv preprint arXiv:2012.14397. URL: https://arxiv.org/abs/2012.14397

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

Easwaran, K. (2011). "Bayesianism I: Introduction and arguments in favor". *Philosophy Compass*, 6(5): 312-320. doi: https://doi.org/cwv9c9

Edwards, A. W. (1984). *Likelihood*. Cambridge science classics. Cambridge University Press.

Elga, A. (2016)." Bayesian humility." *Philosophy of Science*, 83(3): 305-323. doi: https://doi.org/gg432b

Friedman, R. D. (1997). “Towards a (Bayesian) convergence?” *The International Journal of Evidence & Proof, 1(special):* 348-360. doi: https://doi.org/gprg

Garola, C. (2018). “An Epistemic Interpretation of Quantum Probability via Contextuality”. *Foundations of Science*, 25(1), 105-120. doi: https://doi.org/gprh

Gillies, D. (2000). *Philosophical theories of probability*. USA. Psychology Press.

Glymour, C. (1980). *Theory and Evidence*. Princeton, New Jersey: Princeton University Press.

Grünwald, P. (2012). The safe bayesian. En International Conference on Algorithmic Learning Theory (pp. 169-183). Berlin: Springer Heidelberg.

Hacking, I. (1972). “Likelihood”. *The British Journal for the Philosophy of Science*, 23(2): 132-137. doi: https://doi.org/dn48vp

Hacking, I. (2006). *The emergence of probability: A philosophical study of early ideas about probability, induction and statistical inference*. Cambridge University Press. doi: https://doi.org/gprk

Hájek, A. (Fall 2019). *Interpretations of Probability*. The Stanford Encyclopedia of Philosophy, Edward Zalta (ed.): URL: https://plato.stanford.edu/archives/fall2019/entries/probability-interpret/

Hájek, A. y Hartmann, S. (2010). *Bayesian epistemology*. Oxford: Oxford University Press

Hájek, A. y Lin, H. (2017). “A tale of two epistemologies?”. *Res Philosophica*, 94(2): 207-232.

Hansson, S. O. (Winter 2017). *Logic of Belief Revision*. The Stanford Encyclopedia of Philosophy, Edward Zalta (ed.). URL: https://plato.stanford.edu/archives/win2017/entries/logic-belief-revision/

Hartmann, S. y Fitelson, B. (2015). “A new Garber-style solution to the problem of old evidence.” *Philosophy of Science*, 82(4): 712-717. doi: https://doi.org/gprn

Hawthorne, J. (1994, January). “On the nature of Bayesian convergence”. En *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* (Vol. 1994, No. 1, pp. 241-249). Philosophy of Science Association. doi: https://doi.org/gprp

Hoefer, C. (2012). “Calibration: Being in Tune with Frequencies.” *Dialéctica, 66(3),435-452*. doi: https://doi.org/gprq

Howson, C. (1991). “The 'old evidence' problem”. *The British Journal for the Philosophy of Science*, *42*(4): 547-555. doi: https://doi.org/bk6kb2

Howson, C. y Urbach, P. (2006). *Scientific reasoning: the Bayesian approach*. Open Court Publishing.

Huber, F. y Schmidt-Petri, C. (2008). *Degrees of Belief*. Springer Science & Business Media. doi: https://doi.org/cwqwh6

Huttegger, S. M. (2015). “Bayesian convergence to the truth and the metaphysics of possible worlds”. *Philosophy of Science, 82*(4): 587-601. doi: https://doi.org/gprr

Iseda, T. (1999). "Use-Novelty, Severity, and a Systematic Neglect of Relevant Alternatives". *Philosophy of Science,* 66: 403-413. doi: https://doi.org/d5v9wb

Jeffrey, R. (2002). "*Petrus Hispanus Lectures: I. "After Logical Empiricism" II. "Radical Probabilism"". En* Works by and about Richard Jeffrey (página personal, Universidad de Princeton). URL: http://www.princeton.edu/~bayesway/. Consultado el 12 de mayo del 2020.

Joyce, J. M. (2011). "The devolepment of subjective Bayesianism". En Gabbay, M. y Hartmann, S. (eds.) *Handbook of the History of Logic* (Vol. 10, pp. 415-475). USA: North-Holland.

Khrennikov, A. (2016). "Quantum Bayesianism as the basis of general theory of decision-making." *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*. *374*(2068) doi: https://doi.org/f3rnrj

Koertge, N. (1978). "Towards a new theory of scientific inquiry". En Cohen, R. Davidson, S , Nuchelmans, G., Salmon,W. (1974) *Progress and rationality in science* (pp. 253-278). Holland: Reidel Publishing Company. doi: https://doi.org/d2swcr

Lange, M. (1999). "Calibration and the Epistemological Role of Bayesian Conditionalization." *The Journal of Philosophy*, 96(6): 294-324. doi: https://doi.org/bzw33s

Levi, I. (1967). Probability Kinematics. *The British Journal for the Philosophy of Science*, 18 (3): 197–209. doi: https://doi.org/b3k2t5

Lewis D. (1976). "Probabilities of Conditionals and Conditional Probabilities". En Harper W.L., Stalnaker R., Pearce G. (eds) IFS. The University of Western Ontario Series in Philosophy of Science (A Series of Books in Philosophy of Science, Methodology, Epistemology, Logic, History of Science, and Related Fields), vol 15. Dordrecht: Springer. doi: https://doi.org/gprt

Lewis, D. (1981). "Causal Decision Theory". *Australasian Journal of Philosophy*, 59(1): 5–30. doi: https://doi.org/bs3w2t

Parrott, M. (2014). "Bayesian models, delusional beliefs, and epistemic possibilities." *The British Journal for the Philosophy of Science*, 67(1): 271-296. doi: https://doi.org/f3r4vz

Press, S. J. (2009). *Subjective and Objective Bayesian Statistics: Principles, Models, and Applications.* New Jersey: John Wiley & Sons.

Raidl, E. (2020). "Open-Minded Orthodox Bayesianism by Epsilon-Conditionalization." *The British Journal for the Philosophy of Science*, 71(1): 139–176. doi: https://doi.org/gprv

Regazzini, E. (2013). The origins of de Finetti's critique of countable additivity. En Jones,G, Shen, X. (2013) *Advances in modern statistical theory and applications: a Festschrift in honor of Morris L. Eaton* (pp. 63-82). Institute of Mathematical Statistics.doi: https://doi.org/gprw

Savage, L. J. (1954). *The Foundations of Statistics*. New York: John Wiley and Sons

Shaffer, M. J. (2008). "Bayesianism, convergence and social epistemology." *Episteme,*5(2): 203-219. doi: https://doi.org/bskckz

Sprenger, J. (2015). “A novel solution to the problem of old evidence.” *Philosophy of Science*, 82(3): 383-401. doi: https://doi.org/gprx

Sprenger, J., y Hartmann, S. (2019). *Bayesian philosophy of science*. Oxford University Press. doi: https://doi.org/gprz

Strevens, M. (2006). The Bayesian Approach to the Philosophy of Science. En D. M. Borchert (ed.), *Encyclopedia of Philosophy, second edition*. Macmillan Reference (pp. 495-502.).

Suzuki, S. (2005). “The old evidence problem and AGM theory.” *Annals of the Japan Association for Philosophy of Science*, 13(2):105-126. doi: https://doi.org/gpr2

Talbott, W. (Winter 2016). "Bayesian Epistemology". The Stanford Encyclopedia of Philosophy, Edward N. Zalta (ed.). URL: https://plato.stanford.edu/archives/win2016/entries/epistemology-bayesian/

Teller, P. (1973). “Conditionalization and observation”. *Synthese*, 26(2): 218-258. doi: https://doi.org/dt3n5j

Tesic, M., Eva, B., & Hartmann, S. (2017). *Confirmation by explanation: A Bayesian justification of IBE*. Open Access LMU. URL: https://epub.ub.uni-muenchen.de/41934/. doi: 10.5282/ubm/epub.41934

Torretti, R. (2003). “El concepto de probabilidad”. *Diálogos*, 38(81): 407-448.

Trpin, B. (2020). “Jeffrey conditionalization: proceed with caution”. *Philosophical Studies*. 177(10): 2985-3012. doi: https://doi.org/ghc9dh

van Fraassen, B. C. (1983). “Calibration: A Frequency Justification”. En Cohen, R. y Laudan, R. (eds), (1983), *Physics, Philosophy and Psychoanalysis*. Dordrecht: Springer.

Vineberg, S. (Spring 2016). "Dutch Book Arguments". *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL: https://plato.stanford.edu/archives/spr2016/entries/dutch-book/

Weintraub, R. (1993). “Fallibilism and rational belief”. *The British journal for the philosophy of science*, 44(2): 251-261. doi: https://doi.org/c26twh

Wenmackers, S., & Romeijn, J. W. (2016). “New theory about old evidence”. *Synthese*, 193(4):1225-1250. doi: https://doi.org/gfjw5r